

Naomi Nagy*, Michael Iannozzi and David Heap
**Faetar null subjects: a variationist study
of a heritage language in contact**

<https://doi.org/10.1515/ijsl-2017-0040>

Abstract: Faetar is an under-documented variety descended from Francoprovençal and spoken in two isolated Apulian villages in southern Italy as well as in the emigrant diaspora, especially in the Greater Toronto Area. Speakers use two series of subject pronouns (*strong* and *weak* pronouns), producing sentences with zero, one or two overt subject pronouns. The status of the overt forms as subject pronouns, emphatic pronouns, left- or right-dislocated pronouns, clitics, or affixes is not clear. Contrary to the predictions of the Null Subject Parameter hypothesis (Perlmutter 1971, *Deep and surface structure constraints in syntax*. New York: Holt, Rhinehart and Winston; Chomsky 1981, *Lectures on government and binding*. Dordrecht: Foris), these grammars have subject pronoun paradigms that are variable and conditioned by a number of linguistic factors (including person, tense, information status and subject type). This article delineates which aspects vary diachronically, spatially, or between individuals – a necessary prerequisite to constructing a theoretical model that accounts for this variation. By comparing the patterns of use in France, Italy, and Toronto, and using sources that span nearly a century, we see that despite the very small size of its speech community, Faetar shows little sign of accommodating to English’s virtually categorical presence of subject pronouns, nor to Italian’s high null subject (hereafter \emptyset -subject) rate, nor to the conditioning effects found in those languages.

Keywords: Faetar, heritage language, pro-drop, null subject, Toronto

1 Introduction

The geolects of the Gallo-Italo-Romance linguistic continuum show considerable variation in subject pronoun usage (Heap 2000; Kristol, this issue). Contrary to the predictions of the Null Subject Parameter hypothesis (Perlmutter 1971;

*Corresponding author: Naomi Nagy, Department of Linguistics, Sidney Smith Hall, 4th floor 100 St. George St., Toronto, ON M5S 3G3, Canada, E-mail: naomi.nagy@utoronto.ca

Michael Iannozzi: E-mail: miannozz@uwo.ca, David Heap: E-mail: djheap@uwo.ca, Inter-faculty Program in Linguistics Arts & Humanities Building Rm 2R34, 1151 Richmond Street, London, ON N6A 3K7, Canada

Chomsky 1981: 240), these grammars have subject pronoun paradigms that are not only *partial* (i.e. some, but not all grammatical persons have pronoun usage which approaches categorical frequencies) but also *variable* and conditioned by a number of linguistic factors. Even within one subfamily, Francoprovençal (FP), Diémoz (2007) finds a surprising amount of variability in subject pronoun usage between neighboring villages and across linguistic contexts.

The variation examined in this article concerns these subject pronouns, specifically alternation between overt and \emptyset -subjects of finite verbs in declarative main clauses, in Faetar. Faetar is a Francoprovençal variety spoken in Faeto and Celle, two villages in the Foggia region of Apulia in Italy, as a result of migration from a FP-speaking region of France in about the fourteenth century (see Zulato et al. [this issue] on Faetar's ethnolinguistic vitality, demographics and revitalization efforts). Faetar has a variable phonology and lexicon in which FP-origin forms compete with cognate forms more characteristic of surrounding Italian varieties (Nagy and Reynolds 1997; Nagy 2011a). After examining the variability in the variety spoken in Apulia, we examine data from speakers living in the Greater Toronto Area, Canada (hereafter, Toronto), the largest group of diaspora Faetar speakers. Speakers from Faeto and Celle emigrated in the 1950–1970s, with several hundred settling in the Toronto area. This community has maintained regular usage of Faetar into the second generation. Currently speakers range in age from about 30 to 100 years old.

Because Faetar is a minority language both in Italy and in Canada, we might expect its morphosyntax, like its phonology and lexicon, to show results of contact with neighboring varieties. Our data support the hypothesis that Faetar subject pronoun morphosyntax varies between FP-like (subject pronoun usually present) and Italian-like (subject pronoun usually absent), but do not reflect ongoing contact-induced change. Faetar has a variable subject pronoun system which is typologically more similar to that found in other FP varieties than to neighboring southern Italian \emptyset -subject varieties.

We present our findings of a historical change in subject pronoun usage: both by comparing contemporary Faetar speakers of different ages and by contrasting the patterns found in Jaberg and Jud's (1928–1940) *Sprach- und Sachatlas Italiens und der südschweiz (AIS)* atlas data with data recorded in the 1990s in Faeto and the early 2010s in Toronto. Multivariate statistical analysis indicates that linguistic factors including grammatical person, type of nominal subject (noun, demonstrative pronoun, or empty), tense, information status of the subject (new or old), and presence of preverbal clitics have significant effects on subject pronoun usage, as in other Gallo-Italo-Romance varieties (Heap 2000).

Our findings show that the factors constraining subject pronoun usage in Faetar do so variably, rather than categorically, and the effect size of each variable

differs among samples. Careful comparison of intra- and inter-speaker variation can play an important role in language documentation. Meyerhoff (2015: 78) motivates simultaneously pursuing language documentation and variationist sociolinguistics, noting “that a dual focus can benefit both enterprises”. In particular, she notes that such approaches can “provide a positive framework for the local community to evaluate synchronic variation and change”, an important asset for language revitalization. While variable, the patterns documented across samples of Faetar reflect no accommodation to Italian or English patterns of \emptyset -subject usage. The significance of these linguistic and social variables will need to be taken into account in any theoretical model of variable subject pronoun usage.

2 Comparison of subject pronoun systems in varieties in contact

English speakers produce overt subjects virtually categorically, even in casual speech. A rate of 2% \emptyset -subjects was documented for conversational Toronto English by Nagy et al. (2011: 139). Subject pronoun absence was predominantly in contexts where the subject has the same referent as the previous clause (necessarily “old” information) and where the clause is conjoined to the previous clause. No other factors significantly correlated with the presence versus absence of \emptyset -subjects in a sample of 400 finite clauses from Toronto English (Nagy et al. 2011: 140).

Italian has a much higher rate of \emptyset -subjects. Rates of \emptyset -subjects around 80% have been documented for Heritage (Calabrese) Italian in Toronto (Nagy 2014), Heritage Italian in Germany (Schmitz et al. 2016), and Italian in southern Italy (Nagy 2015; Schmitz et al. 2016). As in English, Italian \emptyset -subjects are favored in contexts where the subject shares a referent with the previous clause’s subject. Additionally, for both Homeland Italian in Calabria ($n = 748$) and for Heritage Italian in Toronto ($n = 1,047$), \emptyset -subjects are favored with plural more than singular subjects. For Heritage Italian, presence of a proclitic favors \emptyset -subjects, as do past perfect clauses (versus present and past imperfect). There is no evidence of a change in progress with respect to the rate of use of \emptyset -subjects in either Heritage or Homeland Italian (Nagy 2015).

FP, like Italian, has alternation between \emptyset and overt subjects, but with little consistency found in the contexts that favor overt versus \emptyset -subjects across regional varieties (Diémoz 2007). Heap (2000) finds some fairly consistent patterns based on grammatical person.

Faetar has two pronoun types, strong and weak, in addition to \emptyset -subjects. Examples (1–5), produced by speakers in Faeto, illustrate the five possible subject

types. All options may occur in phrases with or without [+ Argument] subjects, though strong pronouns are rare in sentences with nouns as subjects (compare to a Swiss FP variety, Kristol [this issue]).

- (1) **No overt subject pronoun**
/ɛ lu dʒóɾe Ø stav a la kaz/
'and that day, [Ø = I] was at the house' (F1F79)
- (2) **Weak pronoun**
/e i stávo vakánt/
'and it was vacant' (F1M92A)
- (3) **Strong pronoun**
/no íɫɛ sta tú:tə/
'No, he was always...' (F1F79A)
- (4) **Strong + Weak pronoun**
/íɫɛ i e lu me prefɛrí:tə/
'She-strong she-weak is my favorite' (F1F79A)
- (5) **Expletive pronoun**
/o súndə kúntə k e pa luá:/
'There are tales that are not true' (FOM30A)

This article focusses not on the specific forms of Faetar pronouns, but rather on the alternation between overt and Ø-subjects in these several varieties. We compare the probabilistic patterns of variation between different samples of Faetar and its source (FP) and adstrate (English, Italian) varieties. If FP in contact with Italian (Homeland Faetar) shows a pattern of variation which is similar to the neighboring Italian vernacular, or if Heritage Faetar shows a pattern of variation which shifts towards either Toronto English or Toronto Italian, then we may have evidence for contact-induced change. These similarities and differences in factors governing the variability also shed light on the behavior of a minority language in contact with majority varieties.

3 Methodology

We analyze the Faetar subject pronoun system using data from three sources: AIS atlas data from the 1920s and two corpora of spontaneous Faetar speech

(Faeto in the 1990s, Toronto in the 2010s). The 70-year difference between Jaberg and Jud’s 1920s atlas elicitation and Nagy’s 1990s fieldwork allows us to explore the shift in subject pronoun usage in real time, but the two spontaneous speech corpora (1990s, 2010s) constitute our primary focus here. The speakers recorded in Toronto are 25 years older, on average, than those recorded in Faeto, and the latter were recorded nearly 25 years earlier, making the two data sets comparable in terms of generations. A sample of finite clauses from all three sources are coded for the presence or absence of subject pronouns and factors conditioning this variation. These data provide a dynamic portrait of subject pronoun usage across the Faetar speech community. See Figure 1.

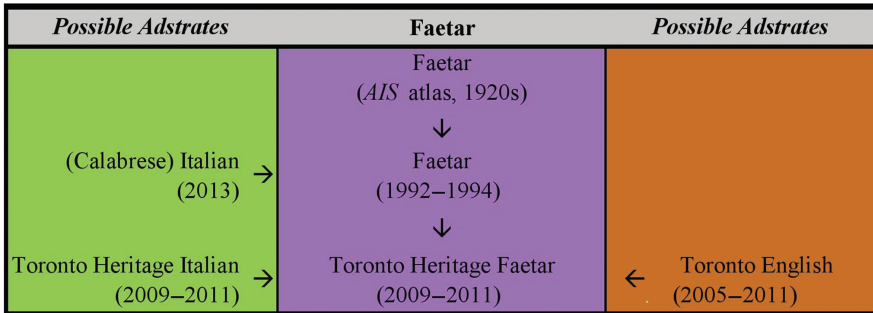


Figure 1: Comparisons included in this study, chronologically arranged. (Arrows indicate direction of possible influence, from larger to smaller languages. Data is from conversational speech except where otherwise indicated. Dates indicate recording period.)

3.1 Data

The Homeland Faetar sample consists of 1,650 sentences from 20 speakers age 11–77 at the time of recording. Recordings were made in interviewees’ homes with a (non-native speaker) interviewer. Conversations are casual and cover a range of topics, following standard sociolinguistic interview techniques (Labov 1984). Some spontaneous speech was prompted by pictures in a children’s book (Amery and Cartwright 1987).

Similar fieldwork conducted in 2009–2011 produced conversational data for Heritage Faetar in Toronto (Nagy 2011b). The Heritage Faetar data constitute a sample of 959 sentences from 13 speakers, of which six are first generation (immigrants from Faeto or Celle), age 70 to 92 at the time of recording, and seven are second generation, age 32 to 58. See Table 1.

Table 1: Faetar speaker distribution.

Generation	Males	Females	Total
Homeland	10	10	20
Heritage Generation 1	4	2	6
Heritage Generation 2	4	3	7
Total	18	15	33

All conversational speech was transcribed in ELAN (<http://tla.mpi.nl/tools/tla-tools/elan/>, Wittenburg et al. 2006) between 2009 and 2014. Also in ELAN, each finite declarative main clause was marked for one of the types of subject listed in (1–5), a dependent variable later simplified to just overt versus \emptyset -subjects. In Faetar, both the strong and the weak form may appear adjacently, without emphatic effect. In addition to the [+human] pronouns, we consider generic pronouns: including [ki] ‘who’, which is used as a generic [+human] pronoun in Faetar, and the expletive pronoun [o] used in weather expressions like/o piówe/ ‘it’s raining’ and the existentials/o aját/‘there is’ and/o ajánde/‘there are’.

To prepare for multivariate analysis that determined the effects of contextual and social factors, we coded additional characteristics of each token’s context for factors (independent variables) predicted to condition the presence or absence of the \emptyset -subject. Coding these contextual factors ensures that differences between speakers or groups of speakers that are due to differences in the distribution of the types of sentences they produce (e.g., referring to new or old information, first versus third person reference) do not preclude accurate comparison of \emptyset -subject rates. Coding was done by two of the authors in consultation with native speakers of Faetar. For the Heritage Faetar data, 100 tokens were coded per speaker, or as many as available in shorter recordings. For Homeland Faetar, the sparser available data was coded exhaustively, producing 18 to 376 tokens per speaker.

Ninety maps were examined from AIS point 715, which is Faeto. All relevant responses, consisting of sentences containing finite verbs, in the AIS questionnaire were coded as above. It is most likely that these sentences were elicited using translation tasks reminiscent of Gilliéron and Edmont’s prompt *Comment dites-vous en patois ...?* ‘How do you say in local speech...?’. While this methodology produces a different type of data, not ideal for comparison to the other samples, it is the only comparator available from an earlier period. As in Gilliéron and Edmont’s FP atlas data (1902–1910), we see a system with quasi-categorically overt subject pronouns in most if not all grammatical persons (Heap 2000).

3.2 Independent variables

Independent variables were selected for analysis based on previous findings regarding \emptyset -subject variation in Romance languages (see extended discussion in Heap and Nagy 1998; Nagy 2014). Six linguistic factors were coded for each token (sentence) considered. Sentences were coded for (1) person, number, and gender of their subject; (2) presence of a [+Argument] subject (noun or a demonstrative pronoun); (3) information status of the subject (new versus old information); (4) tense (simplified to past versus non-past as other distinctions were not significantly correlated to the rate of \emptyset -subjects); (5) presence or absence of a (post-verbal) negation marker; and (6) presence of a proclitic other than a subject (reduced to the binary [presence/absence]). Factors (1) and (2) were combined to avoid collinearity: only 3rd person subjects may co-occur with a noun or a demonstrative pronoun.

To see if Faetar is undergoing change either in Italy or in Toronto, the social variables coded for the conversational speech data were: age group, binned into three categories that are comparable, in real time, across the Homeland and Heritage samples; and generation (for Heritage speakers). Individual speaker was coded as a random effect to determine if any individuals were behaving in an unusual way, and to eliminate excessive influence of any outlier speakers or differences due to different sample sizes.

Based on analyses of phonological variation (Nagy and Reynolds 1997), we anticipated that younger speakers in Faeto would show more Italian influence than older speakers, resulting in higher rates of \emptyset -subjects. We predicted influence from English in younger speakers in Toronto, leading to lower rates of \emptyset -subjects and different constraints conditioning the variation. Additionally, for Homeland speakers, an effect of the amount of contact with Italian was sought by comparing rankings of degree of self-reported contact with Italian to rates of \emptyset -subject usage. Four types of contact with Italian were considered: amount of schooling (in Italian), language/place of work, presence of non-Faetani in the family, and residence elsewhere in Italy. These contact factors had no significant effects and will not be discussed further. As sex had been shown not to have a significant effect in pilot studies, it is not considered here.

3.3 Methods of distributional and logistic regression analysis

Once the coding was completed in ELAN, the time-aligned tiers of transcription and annotation were exported to a text file as the input for distributional analysis (Section 4.1) and multivariate analysis (Section 4.2) (cf. Nagy and

Meyerhoff 2015 for methodological details). Rbrul (Johnson 2014) was used to conduct mixed-effects logistic regression models with speaker as a random effect. The Homeland, Heritage, and AIS samples were examined separately. The binary dependent variable in all instances was overt versus \emptyset -subject, and the application value was \emptyset -subject (as opposed to an overt subject pronoun, as listed in [1–5]). Social factors were analyzed for the conversational data but not the atlas data, where there is only one speaker. Tokens that were unclear (semantically or phonetically) were excluded from analysis as were tokens in contexts that were too rare to be included in statistical analysis (e.g., second person plural subjects).

4 Results

We first compare the rates (proportions) of \emptyset -subject use in each sample. We next compare the effects of the linguistic constraints to see whether there is evidence of change in the grammar and, if so, whether that can be appropriately attributed to contact with either Italian (in Italy and Toronto) or English (in Toronto). Comparison of constraints is a more robust method of documenting contact-induced change than comparison of rates, being less sensitive to stylistic differences (cf. Poplack and Levey 2010).

4.1 Overall rate of \emptyset -pronoun usage

The overall distributions show differences in rate of subject pronoun usage among varieties. The older Faetar atlas data (Jaberg and Jud 1928–1940) replicates the low rates of \emptyset -subjects seen in FP atlas data for France (Gilliéron and Edmont 1902–1910): 19% (17/90) of the sentences for Faeto (AIS point 715) have \emptyset -subjects.

The distribution of types of pronoun forms used in conversational Faetar is shown in Table 2.

The rate of \emptyset -subjects is much higher in the Homeland and Heritage samples than in the atlas data. Whether this is due to a change over time or to the difference in speech style (atlas data is elicited in isolated phrases, often via translation) is unclear. However, as both the data collection in Faeto and Toronto consist of conversation of the same type with the same interviewer, we can see a difference in rate of \emptyset -subject use: lower in the Heritage sample than the Homeland. Figure 2 shows the distribution of subject types, by age, in the two samples.

The distribution across ages may reflect ongoing change in the Homeland sample but seems remarkably stable in the Heritage sample. In all cases the

Table 2: Distribution of types of subject pronouns in conversational Faetar.

Subject form	Homeland Faetar		Heritage Faetar	
	%	n	%	n
no subject pronoun	55 %	902	39 %	373
weak form of pronoun only	38 %	627	43 %	410
strong + weak pronoun	2 %	14	4 %	35
[o] (generic, existential)	5 %	77	2 %	19
strong pronoun only	2 %	30	13 %	122
TOTAL		1,650		959

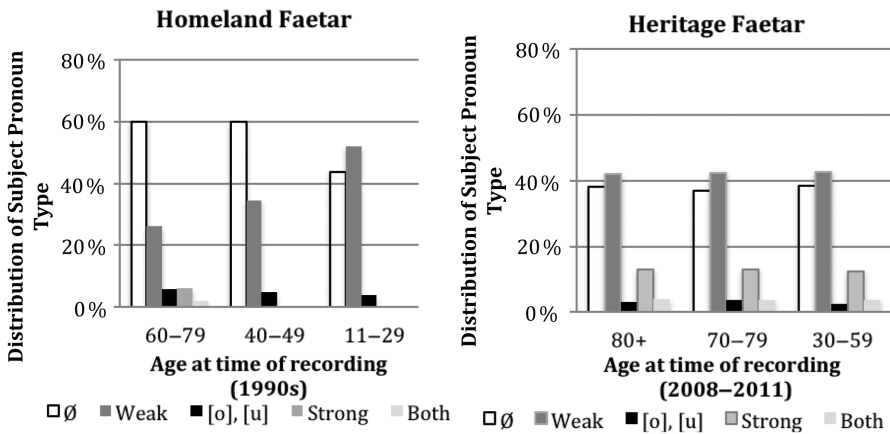


Figure 2: Distribution of subject types by age group in Homeland (n = 1650) and Heritage Faetar (n = 959).

rates of Ø-subjects are lower than those observed for Italian (~80 %, Nagy 2015; Schmitz et al. 2016) but considerably higher than the 2% observed for English (Nagy et al. 2011).

4.2 Conditioning effects

Table 3 compares the effects of the independent variables in the three samples. For conversational data, factor weights are calculated by binomial logistic regression analyses with Ø-subject as the application value, comparing the rate of Ø-subjects to all types of overt subject pronouns combined. Factor weights are provided for factor groups (constraints), showing the likelihood of

Table 3: Logistic regression models of factors affecting the likelihood of \emptyset -subjects in conversational data for Homeland and Heritage Faetar. Distributional data from AIS is provided for comparison.

Factor group	AIS (19 % \emptyset) n = 90		Homeland (57 % \emptyset) n = 1,580, Input = 0.37			Heritage (38 % \emptyset) n = 942, Input = 0.40		
	N	% \emptyset	FW	N	% \emptyset	FW	N % \emptyset	
Grammatical person* [+ /-Argument] subject (87)							(71)	
Generic	8	75	0.94	537	91	0.82	47 68	
+ Arg 3rd sg	9	0	0.61	419	51	+ Arg 3rd m 0.82	117 68	
-Arg 3rd sg	13	8	0.51	321	37	-Arg 3rd m 0.67	339 52	
+ Arg 3rd pl	7	0	0.47	51	39	+ Arg 3rd f 0.74	38 61	
-Arg 3rd pl	3	33	0.36	136	21	-Arg 3rd f 0.53	71 35	
1st pl	6	67	0.07	31	7	0.32	46 20	
1st sg	25	8	[]	62	3	0.11	192 6	
2nd pl.	7	43			<i>no data</i>		<i>no data</i>	
2nd sg.	12	0			<i>no data</i>	0.12	92 7	
Preverbal clitics							(42)	
Present	18	39	0.71	18	72	[0.57]	53 60	
None	72	13	0.29	1477	58	[0.43]	889 37	
Age group							(27)	
Oldest		<i>N/A</i>	0.62	379	67	[0.60]	196 48	
Middle			0.53	646	62	[0.45]	483 36	
Youngest			0.35	470	47	[0.44]	263 35	
Tense							(22)	
Non-Past	58	21	0.61	1426	59	[0.53]	474 37	
Past	32	16	0.39	69	42	[0.47]	468 40	
Information Status^a							(20)	
New	30	20	0.60	1185	65	0.55	648 42	
Old	7	29	0.40	310	34	0.45	294 30	
Negation								
Negative	4	25	[0.54]	52	60	[0.51]	97 35	
Affirmative	86	19	[0.46]	1443	59	[0.49]	845 39	
Speaker	<i>N/A</i>		0 (std. dev.)			0.21 (std. dev.)		

The significance cut-off level is $p < 0.05$. The range, or relative effect size (in parentheses for each factor group), is calculated by subtracting the least favoring context's factor weight from that of the most favoring context. The 1st sg context is excluded from the model for the Homeland sample because it is so close to categorical.

^aFifty three tokens were not possible to code for this factor in the atlas data.

a \emptyset -subject in each context. Non-significant factors' weights appear in square brackets. Factor groups are listed in order of decreasing range in the Homeland model: the factor group listed first has the biggest effect on the likelihood of a \emptyset -subject appearing. Within each factor group, the factors are listed in decreasing rank of favoring \emptyset -subjects.

Grammatical person has the strongest effect in all data sets and information status has the smallest effect that still surpasses the threshold for significance in both models. Three conditioning factors that are significant in the Homeland model, but with small ranges, are not significant in the Heritage model. The direction of effect suggested by the percentages is, however, similar across samples. Individuals behave significantly differently from each other in the Heritage sample, but this random effect has a standard deviation of 0 in the Homeland sample, suggesting more similarity among Homeland speakers.

We now look at the favoring versus disfavoring contexts within each factor group. Grammatical person has the strongest effect for both groups of speakers, with 3rd person, particularly generic subjects, favoring \emptyset -subjects, and first and second person disfavoring. Cross-linguistically, different effects have been reported for grammatical person of the subject (Jenkins 2000; Travis 2007). Heap (2000: 115) reports that in atlas data for the Gallo-Italic Romance dialect continuum throughout southeastern France and northern Italy, 3rd persons (singular and plural, masculine and feminine) are among the least likely to have \emptyset -subjects, with 0–20 % \emptyset -subjects in 49–55 % of varieties sampled. In contrast, Table 3 shows that *AIS* data for Faetar, like conversational Faetar (Homeland and Heritage), favors \emptyset -subjects in the 3rd person [-Argument]-subject contexts, a pattern suggested though not quantified by Marzys (1981: 52–53). One interesting difference between the Homeland and Heritage varieties is that, for Homeland, the best fitting model splits the 3rd person tokens by number (like Italian) while for the Heritage variety, 3rd person is split by gender.

A surprising finding is that pronouns with a new-information referent are more likely to be \emptyset than those with old-information referents, in the Homeland and Heritage samples. This relatively small effect appears to contradict the same-versus switch-referent effect reported in most studies of \emptyset -subjects, cross-linguistically: continuity of referent (a subset of the old-information category) favors \emptyset -subject occurrence in Spanish (Cameron 1993; Silva-Corvalán 1982; Travis 2007); Egyptian Arabic (Parkinson 1987), and in English (Harvie 1998).

We turn next to the factors that test our hypothesis that other material in the surface subject position will decrease the likelihood of an overt subject pronoun. As hypothesized, subject pronouns are less likely to surface when other material fills the preverbal space in linear surface order, even with different syntactic roles. The presence of preverbal clitics has this effect in both Homeland and

Heritage datasets: \emptyset -subjects are more frequent in clauses with preverbal clitics. This hypothesis is also supported by the effects of [+ Argument] subjects. In both samples, there are fewer \emptyset -subjects in clauses with a demonstrative pronoun or noun subject. It is important to note that subject pronouns may follow both noun and demonstrative subjects, suggesting that these pronouns are, in such cases, clitics.

Negation affects \emptyset -subject presence in languages with pre-verbal negation like Spanish and Italian (Geeslin and Gudmestad 2011; see Renzi and Vanelli 1983: 130 for northern Italian dialects). However, as negation is post-verbal in Faetar, it does not interact significantly with subject pronouns (cf. Heap 2000: 130–131; Diémoz 2007: 4).

For the Heritage Faetar data, models containing either the factor generation or age group were compared.¹ We report the model with age group – the stronger predictor (although a non-significant effect) – in Table 3. The similarity of the age effect in the two communities, shown in Figure 3, is striking, particularly the parallelism in the slope showing \emptyset -subject usage decreasing in apparent time.

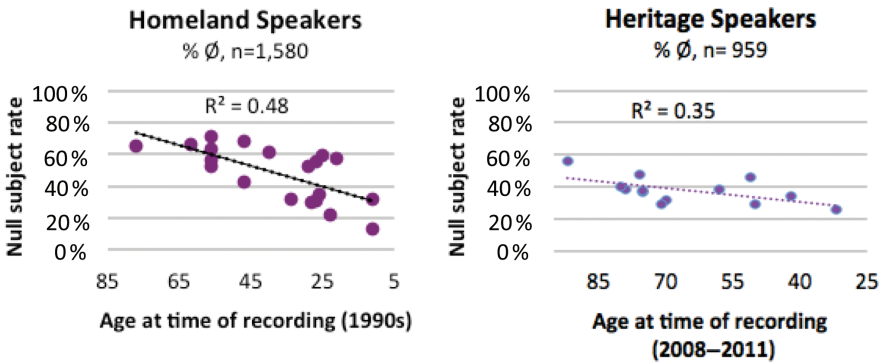


Figure 3: Individual rates of \emptyset -subject use in Homeland and Heritage Faetar.

5 Discussion: change versus stability

Comparing the atlas and the conversational Faetar data sets appears to suggest that, over time, a shift toward more \emptyset -subjects has taken place: the percentage

¹ These two factors classify the speakers in similar but not identical fashion, and so cannot be included simultaneously in any model.

of \emptyset -subjects rose from 19% in the atlas data to 57% for Homeland conversational Faetar. This would not be an unexpected effect given that Faetar has been in contact with Italian, a \emptyset -subject language, for some 600 years (while the Gallic varieties of FP have been in contact with French, a non- \emptyset -subject language). Much of the difference can more plausibly be attributed to the different types of data compared: sentence elicitation/translation in the atlas data versus conversational speech.² Turning then to the more reliable conditioning effects, we see consistency between the AIS and Homeland samples in the direction of effects.

Comparison of Heritage and Homeland varieties suggests a decrease in the rate of use of \emptyset -subjects. However, some of this rate difference must be attributed to a different distribution of tokens in the two conversational data sets. We note the comparatively large number of tokens in the Homeland sample (versus the Heritage) with generic reference, reporting new information and/or with non-past temporal reference, all contexts that favor \emptyset -subjects. Thus, again, the overall rate difference must be taken with a grain of salt. As discussed above, the conditioning effects run in the same direction in the Homeland and Heritage models, though fewer reach significance in the latter. This decrease in significant conditioning factors can be interpreted as an ongoing change from the Homeland to the Heritage variety.

The pattern of younger speakers in both communities using fewer \emptyset -subjects (Table 3, Figure 3) confirms this interpretation and suggests a shift away from a more Italian-like grammar. The motivation of marking a distinctive identity is plausible both in Faeto and in Toronto, where Faetar speakers may feel overshadowed by the large Italian-speaking community. Italian was, until the 2011 census, Toronto's most common mother tongue after English (Statistics Canada 2014). The lack of a generational difference in the rate in Heritage Faetar indicates that it would be unwise to attribute the change to contact with English: the second generation, born and raised in Toronto, have more contact with English, but do not have a significantly different \emptyset -subject rate from Generation 1 speakers. Rather, the Heritage speakers are clearly continuing a trend established in the Homeland sample, where English influence is implausible.

We next consider the possibility of influence from Italian and English. The change in the effect of grammatical person from FP to Faetar may be due to

² To overcome this inequivalence, a small corpus of Faetar elicitation/translation data, recorded in 1994, was compared. This data set shows virtually categorical usage of subject pronouns: only 2 of 271 sentences have a \emptyset -subject, suggesting that stylistic rather than diachronic variability better accounts for the difference (Heap and Nagy 1998: 298).

contact from Italian (the person hierarchy in the AIS data is less clear but may be transitional between FP and later Faetar). In Italian, as in the conversational Homeland and Heritage data for Faetar, \emptyset -subjects are most likely to be found in 3rd person (Schmitz et al. 2016). There is no change in the effect of person from the Homeland to the Heritage sample that could be attributed to English contact: the person hierarchy remains stable, while no statistically-significant person effect is reported for English \emptyset -subject variation (Harvie 1998: 20–22; Nagy et al. 2011).

The other significant feature of \emptyset -subject variation in Faetar, that subjects with the same referent as the previous subject disfavor \emptyset -subjects compared to those with a different referent, is a further clue that Faetar is maintaining its distance from both Italian and English patterns, which both show the opposite pattern. Its grammar appears immune to contact effects in this respect. The effect of preverbal material is consistent across varieties, providing no clues about contact effects.

Thus, while there does appear to be a change in progress in Faetar toward a lower rate of \emptyset -subjects (based on comparison of the Homeland and Heritage samples, keeping in mind the caveat related to differences in elicitation method affecting the AIS data), there is little change in its conditioning constraints. Most importantly, the small differences between the three samples of Faetar that are compared here cannot be attributed to contact with either Italian or English.

6 Conclusion

We have shown that a number of factors, both linguistic and social, govern the presence of subject pronouns in FP dialects spanning both time and space. The effects of these factors are variable, rather than categorical. The diachronic and synchronic patterns provide information that will be helpful in constructing a theoretical account of this part of the grammar and its various forms over time, one that can account for the variable presence of both strong and weak pronouns, in contexts both with and without a [+Argument] subject.

Faetar is resisting pulls in all directions from its FP roots, evident when we examine both the constraint system and the rates of \emptyset -subjects. In the Homeland data, the direction of change (to lower rates of \emptyset -subjects among younger speakers) is opposite of what would be predicted for an assimilatory effect of contact with Italian. Yet, the apparent-time decrease in \emptyset -subject rate also cannot be attributed to English as neither age nor generation has a significant effect in the Heritage data. In contrast, age does affect the variation in the Homeland data, collected at a time when English was not used in Faeto.

As comparison of constraints is considered a more robust method of documenting contact-induced change than comparison of rates (cf. Poplack and

Levey 2010), we may safely say that Faetar retains much of its FP grammar. Heritage Faetar remains similar to Homeland Faetar: both exhibit a trend toward fewer \emptyset -subjects and similar directions of effects. In this way, our study contributes to the growing body of literature showing systematic behavior in endangered varieties, and integrating variationist description with documentation of lesser-studied languages, aiming for the mutual benefit to both fields noted in Meyerhoff (2015: 78). As discussed in Zulato et al. (this issue), members of the Faetar community are proud of their connection with Gallic languages. We hope that this indication of surviving evidence of Gallic influence can further bolster the status of Faetar. We have shown that, in both Italy and Canada, where the tiny Faetar communities are in intense contact with larger English and/or Italian communities, this contact does not appear to affect subject pronoun usage.

Acknowledgements: This research would not be possible without generous contributions of time and knowledge from our speakers; valuable discussions about pronouns with Julie Auger, Miriam Meyerhoff, Terry Nadasdi, Gillian Sankoff and Vera Richetti Smith; research assistants Tonia Djogovic, Rick Grimm, Tiina Rebane; and financial support from: SSHRC Standard Research Grant 410-2009-2330, Salvatori Research Grants in 1992, 1993, and 1994 to the first author; SSHRC Doctoral Fellowship 752-91-2167 to the third author; and support from Keren Rice's Canada Research Chair, the Faculty of Arts and Science at the University of Toronto and the School of Arts and Sciences at the University of Pennsylvania. We thank the *Chicago Linguistic Society* editors for permission to adapt sections of Nagy and Heap (1998) in this article.

References

- Amery, Heather & Stephen Cartwright. 1987. *First 100 words*. London: Usborne.
- Cameron, Richard. 1993. Ambiguous agreement, functional compensation, and nonspecific *tú* in the Spanish of San Juan, Puerto Rico, and Madrid, Spain. *Language Variation and Change* 5(3). 305–334.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- Diémoz, Federica. 2007. *Morphologie et syntaxe des pronoms personnels sujets dans les parlers francoprovençaux de la Vallée d'Aoste*. Tübingen: A. Francke.
- Geeslin, Kimberly & Aarnes Gudmestad. 2011. Using sociolinguistic analyses of discourse-level features to expand research on L2 variation: Native and non-native contrasts in forms of Spanish subject expression. In Shu-Ju Diana Tai & Zhi Li, eds. *Selected proceedings of the 2009 Second Language Research Forum*, 16–30.
- Gilliéron, Jules & Edmond Edmont. 1902–1910. *Atlas Linguistique de la France*. Paris: Champion.

- Harvie, Dawn. 1998. Null subject in English. *Cahiers linguistiques d'Ottawa* 26. 15–25.
- Heap, David. 2000. *La variation grammaticale en géolinguistique: Les pronoms sujet en roman central*. München: Lincomm Europa.
- Heap, David & Naomi Nagy. 1998. Subject pronoun variation in Faetar and Francoprovençal. *Papers in sociolinguistics. NWAVE-26 à Université Laval*, 291–300.
- Jaberg, Karl & Jakob Jud (eds.). 1928–1940. *Sprach- und Sachatlas Italiens und der südschweiz*. Zofingen: Ringier.
- Jenkins, Lyle. 2000. *Biolinguistics: Exploring the biology of language*. Cambridge: Cambridge University Press.
- Johnson, Daniel Ezra. 2014. Rbrul (Version 2.2) [Computer software]. <http://www.danielezrajohnson.com/rbrul.html> (accessed 23 May 2014).
- Labov, William. 1984. Field methods of the project on linguistic change and variation. In John Baugh & Joel Sherzer (eds.), *Language in use: Readings in sociolinguistics*, 28–53. NJ: Prentice-Hall.
- Marzys, Zygmunt. 1981. Les pronoms personnels sujets dans le parler francoprovençal de Faeto et Celle. *Vox Romanica* 40. 48–58.
- Meyerhoff, Miriam. 2015. Turning variation on its head: Analysing subject prefixes in Nkep (Vanuatu) for language documentation. *Asia-Pacific Language Variation* 1(1). 78–108.
- Nagy, Naomi. 2011a. Lexical change and language contact: Faetar in Italy and Canada. *Journal of Sociolinguistics* 15(3). 366–382.
- Nagy, Naomi. 2011b. A multilingual corpus to explore geographic variation. *Rassegna Italiana di Linguistica Applicata* 43(1). 65–84.
- Nagy, Naomi. 2014. A sociolinguistic view of null subjects and VOT in Toronto heritage languages. *Lingua* 164B. 309–327. doi:10.1016/j.lingua.2014.04.012 (accessed 15 March 2016).
- Nagy, Naomi. 2015. Cross-dialect vs. cross-linguistic contact in Southern Italy. Paper presented at *ICLAVE 8*, University of Leipzig, 27–29 May.
- Nagy, Naomi, Nina Aghdasi, Derek Denis & Alexandra Motut. 2011. Null subjects in heritage languages: Contact effects in a cross-linguistic context. *University of Pennsylvania Working Papers in Linguistics* 17(2). Article 16. <http://repository.upenn.edu/pwpl/vol17/iss2/16> (accessed 1 February 2016).
- Nagy, Naomi & David Heap. 1998. Francoprovençal null subject and constraint interaction. In M. Catherine Gruber, Derrick Higgins, Kenneth Olson & Tamra Wysocki (eds.), *CLS 34: The Panels*, 151–166. Chicago: Chicago Linguistic Society 34.2.
- Nagy, Naomi & Miriam Meyerhoff. 2015. Extending ELAN into variationist sociolinguistics. *Linguistic Vanguard* 1(1). 271–281.
- Nagy, Naomi & Bill Reynolds. 1997. Optimality theory and variable word-final deletion in Faetar. *Language Variation and Change* 9(2). 37–56.
- Parkinson, Dilworth. 1987. Constraints on the presence/absence of optional subjects pronouns in Egyptian Arabic. In Keith Denning, Sharon Inkelas, Faye C. McNair-Knox & John R. Rickford (eds.), *Variation in language: NWAV-XV at stanford*, 348–360. Stanford: Stanford University Department of Linguistics.
- Perlmutter, David. 1971. *Deep and surface structure constraints in syntax*. New York: Holt, Rhinehart and Winston.
- Poplack, Shana & Stephen Levey. 2010. Contact-induced grammatical change: A cautionary tale. In Peter Auer & Jurgen Schmidt (eds.), *Language and space: An international handbook of linguistic variation*, 391–419. Berlin/New York: De Gruyter.

- Renzi, Lorenzo & Laura Vanelli. 1983. I pronomi soggetto in alcune varietà romanze. In Paola Benincà (ed), *Scritti linguistici in onore di G.B. Pellegrini*, 121–145. Pisa: Pacini.
- Schmitz, Katrin, Laura DiVenanzio & Anna-Lena Scherger. 2016. Null and overt subjects in Italian and Spanish heritage speakers in Germany. *Lingua*. doi:10.1016/j.lingua.2016.04.004 (accessed 12 July 2016).
- Silva-Corvalán, Carmen. 1982. Subject expression and placement in Mexican-American Spanish. In J. Amastae & L. Elias-Olivares (eds.), *Spanish in the United States: Sociolinguistic aspects*, 93–120. Cambridge, NY: Cambridge University Press.
- Statistics Canada. 2014. Mother tongue of person. <http://www.statcan.gc.ca/concepts/définitions/language-langue01-eng.htm> (accessed 27 April 2015).
- Travis, Catherine. 2007. Genre effects on subject expression in Spanish: Priming in narrative and conversation. *Language Variation and Change* 19(2). 101–135.
- Wittenburg, Peter, Hennie Brugman, Albert Russel, Andreas Klassmann & Han Sloetjes. 2006. ELAN: A professional framework for multimodality research. In *Proceedings of LREC 2006, fifth international conference on language resources and evaluation*, 1556–1559.