

SUBJECT PRONOUN VARIATION IN FAETAR AND FRANCOPROVENÇAL

David Heap, University of Western Ontario
Naomi Nagy, University of New Hampshire

ABSTRACT

We compare 15 Gallic and one southern Italian Francoprovençal (FP) dialects, examining the considerable synchronic and diachronic variation in pronoun presence, in order to trace the historical trajectory of FP pronoun usage. Contrary to the predictions of the Null Subject Parameter hypothesis (Perlmutter, 1971, Chomsky, 1981), these grammars have subject pronoun paradigms that are not only "partial" but also variable, and conditioned by a number of linguistic factors. To construct a theoretical model that fully accounts for this variation, it is necessary to determine the range of variation and factors which correlate to it and to note which aspects vary diachronically. We report distributions representing diachronic and synchronic variation. Significant effects of grammatical person, tense, information status, subject type, and negation are shown.

1. INTRODUCTION

The geolects of the Gallo-Italo-Romance linguistic continuum show considerable variation in subject pronoun usage (Heap 1997). Contrary to the predictions of the Null Subject Parameter hypothesis (Perlmutter 1971, Chomsky 1981:240), these grammars have subject pronoun paradigms that are not only "partial" (some but not all grammatical persons are regularly used) but also variable, conditioned by a number of linguistic factors.

The Faetar speech community, a Francoprovençal (FP) isolate spoken in Southern Italy for six centuries (since its original settlers emigrated from the Gallic FP region) has a highly variable phonology and lexicon in which FP-origin forms compete with cognate forms more characteristic of surrounding Italian and Apulian varieties (Nagy 1996). Faetar also has a variable subject pronoun system which is typologically more similar to the FP varieties of southeastern France, western Switzerland and northwestern Italy than to its southern Italian "null subject" neighbours. The data in this study support our hypothesis that its subject pronoun morphology also varies between FP-like (subject pronoun usually present) and Italic types (subject pronoun usually absent).

Statistical analysis (using GoldVarb 2.0) indicates that linguistic factors including grammatical person, type of nominal subject (noun, demonstrative, or empty), type of clause, subject's position in clause,

and presence of object and/or reflexive proclitics have significant conditioning effects on subject pronoun usage, as they do in other Gallo-Italo-Romance varieties (Heap 1997). We also show a historical change in subject pronoun usage: both by comparing contemporary Faetar speakers of different ages and by contrasting the patterns found by Jaberg & Jud (1928-40) in the early part of this century with patterns established by fieldwork in this decade. The significance of these linguistic and social variables will need to be taken into account in any theoretical model of variable subject pronoun usage, such as the one we are currently constructing.

2. DATA

We analyse the FP subject pronoun system using data from the sources listed below. The 70-year difference between Jaberg & Jud's and Nagy's fieldwork allows us to explore the shift in subject pronoun usage in real time.

- (1) Linguistic atlas data (Gilliéron & Edmont, 1902) compiled for Heap (1997). 15 survey points in the French *départements* of Ain and Isère, the area where the Faetar settlers are believed to have originated (Nagy 1996), are examined. There are 1,317 records from 100 maps.
- (2) Linguistic atlas data from Jaberg & Jud (1928) for Faeto (AIS point 715) is examined. 90 maps are examined.
- (3) Recent fieldwork data (Nagy 1996) are used to obtain a dynamic portrait of subject pronoun usage across the contemporary Faetar speech community. Speakers range in age from 11 to 77 and are balanced according to sex and degree of contact with Italian. Faetar data are from a sample of 2,092 sentences from 20 speakers.
- (4) Two contemporary Faetar speakers provided translations of Italian sentences, mostly verb paradigms, via oral cues, during fieldwork conducted in 1994 by Francesca Giuliani and Naomi Nagy. Transcriptions of these elicitations are examined as they provide data more comparable to the Jaberg & Jud data.

The atlas data selected are from atlas points with (quasi-)categorical subject pronoun presence listed in Table 1. The numbers in the rightmost column indicate the number of grammatical persons which normally have subject pronouns present, as determined by approximately 100 samples (atlas maps) per town. The last line provides information for Faeto.

Table 1: FP data points examined (adapted from Heap, 1997:159)

Town	Dépt.	Point	System Type
Clonas	Isère	829	6
Monestier-de-Clermont	Isère	849	3
S.-Priest	Isère	912	6
Villars-en-Dombes	Ain	913	6
Lent	Ain	915	5
Reptonges	Ain	917	6
S.-Jean-de-Bourmay	Isère	921	6
Morestel	Isère	922	6
Forcieu	Ain	924	6
Brion	Ain	926	5
Charavines	Isère	931	6
Surjoux	Ain	935	5
Sassenage	Isère	940	5
Theys	Isère	942	6
Le Bourg-d'Oisans	Isère	950	6
Faeto	Apulia	715	6

3. LINGUISTIC FACTORS

3.1. Dependent variable: Subject pronoun presence

There are several options for the form of subject pronouns in both FP and Faetar. In addition to having a null-subject present, there may be one subject pronoun present, which may be either the strong or the weak form. In Faetar, both the strong and the weak form may appear adjacently, without emphatic effect. In addition to the [+human] pronouns, we consider some generic pronouns (including [ki] 'who', which is used as a generic [+human] pronoun in Faetar). Here are examples of the possible forms of subject pronoun: the dependent variable in our study (the first three types exist in the FP data):

no subject pronoun	[mmdʒ]	'I eat'
weak pronoun only	[dʒə mmdʒ]	'I eat'
[o], [u]	[o piówə]	'It's raining'
strong + weak pronoun	[dʒi dʒə mmdʒ]	'I eat'
strong pronoun only	[dʒi mmdʒ]	'I eat'
[ki] in main clause	[ki mmdʒ læ dʒəlat]	'One eats ice cream'

3.2. Independent linguistic variables

Seven linguistic factors are coded for each token (sentence) considered. (1) Each sentence was coded according to person,

number, and gender of the subject. In this analysis, only [+human] subjects are considered. (2) Each sentence is coded for tense (future, present, other) to determine whether tenses with more syncretism would behave differently. (3) The information status of the subject was coded as new or old. Subjects of embedded clauses were coded separately as a group. This factor could not be reliably coded for the atlas data. (4) Presence of a nominal subject, either a noun or a demonstrative pronoun was coded. (5) The position of the subject was coded into five categories: beginning of the main clause elsewhere in the main clause, beginning of the embedded clause, elsewhere in the embedded clause, or in a relative clause.¹ (6) Sentences were coded as negative or affirmative. In these dialects, negation is post-verbal. (7) The presence of a proclitic other than the subject was coded as follows: reflexive, "en"-type partitive, direct object, indirect object, or none.

3.3. Independent social variables

In the analysis of the Faetar contemporary data, social variables are considered: age, sex, and amount of contact with Italian. (No analysis of social factors is possible for the atlas data because there is only one speaker per dialect.) The speakers range in age from 11 to 77 and are divided into 4 groups, roughly by generation: <20, 21-40, 41-60, 61+. Four means of determining the amount of contact with Italian were also considered: amount of schooling (in Italian), language/place of work, presence of non-Faetani in the family, and residence elsewhere in Italy. These intensity of contact factors were not found to have significant effects and so will not be discussed further.

4. RESULTS

4.1. Overall patterns

Before examining the effects of the linguistic factors, overall distributions showing differences between the dialects considered, are presented. For the FP data as a whole, 19% (251/1317) of the sentences examined have \emptyset -subjects. In contrast, in the Faetar data, 50% (1056/2092) of the sentences have \emptyset -subjects. The older Faetar data (Jaberg & Jud, 1928-40) resembles the FP atlas data (Gilliéron & Edmont 1902-10): 18% (16/88) of the sentences have \emptyset -subjects. The distribution of types of pronoun forms present in Faetar is shown in Table 2.

¹Sentences with null subjects are coded as having the subject in initial (canonical) position in the FP data, but were not coded for this factor group in the Faetar data.

Table 2: Distribution of types of subject pronouns in the Faetar data

no subject pronoun	50%	[ki] as main clause subject	1-2%
weak form of pronoun only	41%	strong + weak pronoun	1-2%
[o], [u] (generic, existential)	5%	strong pronoun only	1-2%

4.2. Effects of linguistic variables

Table 3 compares the effects of the linguistic variables for the FP (first two columns) and Faetar data (last two columns). Factor weights are from a Goldvarb 1-level binomial analysis. Each factor group is labeled as significant or non-significant (for each data set), as determined by Goldvarb step-up/step-down binomial analysis. The factor groups which are significant in both the FP and the Faetar data are discussed below.

Table 3: Factor weights (Input value = \emptyset -subject pronoun)

Factor	FP (Input = 0.100)		Faetar (Input = 0.230)	
	Weight	Freq.	Weight	Freq.
1: Person	SIG.		SIG.	
1 sg.	0.859	0.18	0.037	0.04
2 sg.	0.216	0.01	0.200	0.17
3 sg. masc.	0.385	0.34	0.563	0.36
3 sg. fem.	0.035	0.37	0.677	0.46
1 pl.	0.638	0.10	0.029	0.07
2 pl.	0.525	0.03	no data	
3 pl. masc.	0.234	0.44	0.494	0.22
3 pl. fem.	no data		0.456	0.23
2: Tense	NON-SIG.		SIG.	
future	0.467	0.24	0.105	0.22
present	0.492	0.24	0.524	0.34
any other tense	0.513	0.19	0.246	0.25
3: Information Status²	N/A		SIG.	
new			0.672	0.41
old			0.394	0.25
embedded clause			0.106	0.14

² It is interesting to note that this factor apparently has a significant counter-functional relationship to subject pronoun presence.

Table 3: Factor weights (continued)

Factor	FP		Faetar	
	Weight	Freq.	Weight	Freq.
4:Nominal subject		SIG.		SIG.
none	0.231	0.09	0.309	0.25
noun	0.998	0.88	0.716	0.30
demon. pronoun	0.983	0.73	0.896	0.72
5:Subject's position		NON-SIG.		SIG.
start of MC	0.553	0.24	0.213	0.30
elsewhere in MC	0.506	0.27	0.211	0.20
start of EC	0.337	0.05		0.00
relative clause	0.247	0.07	no data	
else in EC	0.581	0.97	0.619	0.11
6:Negation		SIG.		SIG.
affirmative	0.491	0.22	0.486	0.33
post-verbal negative	0.852	0.26	0.729	0.39
7:Non-subject proclitics		SIG.		SIG.
reflexive	0.668	0.09	0.843	0.57
none	0.442	0.22	0.467	0.31
"en"-type partitive	0.580	0.50		N/A
direct object	0.608	0.18	0.862	0.67
indirect object	0.820	0.29	1.000	

For both the FP and the Faetar data, there is a strong effect of subject person. However, the two dialect groups pattern rather differently. In FP, the 3rd persons (singular and plural, masculine and feminine) all disfavour \emptyset -subjects (although only 3rd singular feminine disfavour enough to have a factor weight lower than the input value). In contrast, Faetar exhibits a pattern of favouring the \emptyset -subject in exactly the opposite conditions: 1st and 2nd person rank below the input value and all 3rd person forms rank above it. Weights, in rank order, are given in Tables 4 and 5.

Table 4: Effect of person: Rank ordering for FP data

Factor weights	Persons in rank order	Effect
0.035 0.100 0.216 0.234 0.385	3rd sg. f. INPUT VALUE 2nd sg. 3rd pt. (f. & m.) 3rd sg. m.	all disfavour Ø-subject
0.525 0.638 0.859	2nd pt. 1st pt. 1st sg.	all favour Ø-subject

Table 5: Effect of person: Rank ordering for Faetar data

Factor weights	Persons in rank order	Effect
0.029 0.037 0.200 (no data)	1st pl. 1st sg. 2nd sg. 2nd pl.	all disfavour Ø-subject
0.230 0.456 0.494 0.563 0.677	INPUT VALUE 3rd pt. f. 3rd pt. m. 3rd sg. m. 3rd sg. f.	all favour Ø-subject

Heap (1997:160) found a pattern similar to this FP pattern in Table 4 exhibited by the overall data for the Gallo-Italic Romance dialect continuum throughout southeastern France and northern Italy, as shown in Table 6.

Table 6: Rank ordering in Gallo-Italo-Romance

Rank	Pronoun type	Effect
1st 2nd 3rd	2nd person sg. 3rd person pl. 3rd person sg.	all disfavour Ø-subject
4th 5th 6th	1st person pl. 1st person sg. 2nd person pl.	all favour Ø-subject

The other significant linguistic variables relate to the presence of other preverbal material, i.e., other arguments in subject position. The hypothesis is that subject pronouns will be less likely to surface when other material fills the same part of the linear surface order. This view

is supported by the effects of noun subjects and object pronouns (proclitics), which precede the verb.

The presence of a [+Argument] subject (noun or demonstrative pronoun) has a significant effect in both data sets: doubling of overt NPs with subject pronouns is disfavoured, while subject pronouns are favoured where there is no overt NP. This is particularly true in FP where sentences with noun and demonstrative pronoun subjects have factor weights of 0.99 vs. 0.23 for sentences without a [+Argument] subject. Less strongly, the same pattern is found in Faetar: 0.89 (for demonstrative pronouns) and 0.72 (for nouns) vs. 0.31. That is, some doubling is permitted in Faetar but is virtually categorically proscribed in Gallic FP.

In both FP and Faetar, the absence of (object) proclitics favours subject pronoun usage, and any proclitic form favours the use of zero pronouns. Again, this supports the hypothesis that subject pronouns surface less frequently when their argument position is otherwise occupied.

The non-significance of negation in FP also supports the hypothesis that preverbal material tends to block subject pronoun usage (Heap 1997: 208). Since all the negation considered here is post-verbal, we expect it to be "invisible" to the subject pronoun. We have no explanation for the slightly significant effect of this factor in contemporary Faetar.

4. A change in progress: but which way?

Comparing the atlas and the main Faetar data sets suggests that, over time, a shift toward more \emptyset -subjects has taken place, as the percentage of \emptyset -subjects rose from 18% to 50%. This is not an entirely unexpected effect given that Faetar has been in contact with Italian, a [+Null Subject] language for some 600 years (while the Gallic varieties of FP have been in contact with French a [-Null Subject] language). The degree of change over the past 70 years is, however, surprising. Furthermore, this apparent effect is contestable.

The problem lies in the differences in types of data compared: sentence elicitation/translation for FP and conversational speech for Faetar. To overcome this inequality, a small corpus of contemporary Faetar elicitation/translation data was compared to the FP atlas data. This data set shows virtually categorical usage of subject pronouns: only 2 of 271 sentences have a \emptyset -subject, indicating a change in the opposite direction for Faetar: toward a [-Null Subject] language.

This more surprising pattern is supported by the age-grading found in the Faetar conversational data. A pattern of younger speakers using fewer \emptyset -subjects suggests an ongoing shift away from an Italian-like grammar. Within most age groups, males (right side) use more subject pronouns than females (left side), indicating the

typical type of change from above: one in which females lead in acquiring the surrounding prestige form. The sex difference is, however, small and may not be significant. (Age and sex were combined in one factor group, which is significant.) See Figure 1.

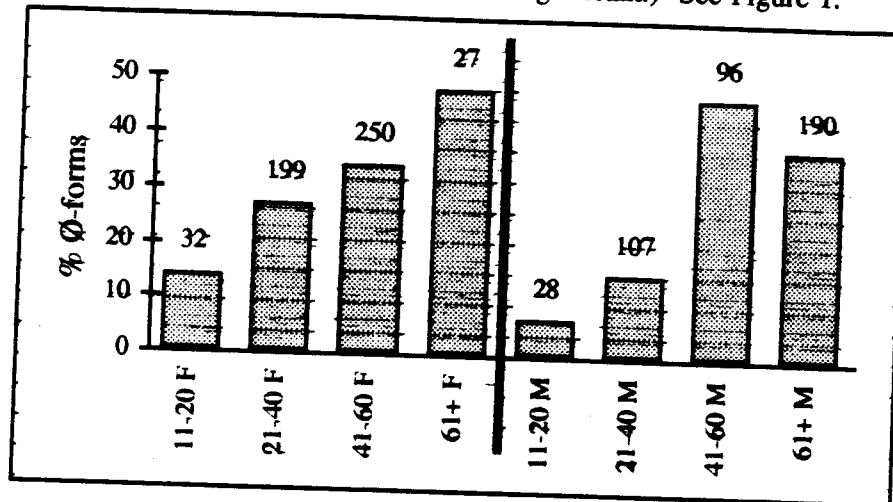


Figure 1: Percent Ø-subject sentences according to age and sex of speaker (20 Faeto speakers). The numbers above each bar represent the number of sentences considered for that group of speakers.

5. SUMMARY

We have shown that there are a number of factors, both linguistic and social, which govern the presence of subject pronouns in numerous dialects of FP spanning both time and regions. Contrary to Marzys (1981), the effects of these factors are variable, rather than categorical: grammatical person as a category is a significant factor in both Faetar and the other FP varieties, but the relative effect of each grammatical person varies between geolects. Similarly, the other linguistic factors considered do not have a categorical effect on subject pronoun presence/absence. By considering the effects of subject person and of other material in preverbal condition, we are able to account for a great deal of the variation. This variation has proved valuable in tracing the history of subject pronouns in FP: 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.

The results of this research, along with other studies of the development of Romance pronoun systems (Auger 1994, Nadasdi 1995, Heap 1997) and other studies of contact-induced variation in

Faetar (Nagy 1994, Nagy & Reynolds 1997), provides data for the long-term cross-linguistic project of determining which social and linguistic factors are most reliable for predicting how a language will change when it comes into contact with another (cf. Nagy 1997).

REFERENCES

- Auger, J. (1994), *Pronominal clitics in Québec Colloquial French: A Morphological Analysis*. Univ. of Pennsylvania PhD dissertation.
- Chomsky, N. (1981), *Lectures on Government and Binding*. Dordrecht: Foris.
- Gillieron, J. & E. Edmont. (1902-10), *Atlas Linguistique de la France*. Paris: Champion.
- Heap, D. (1997), *La variation grammaticale en géolinguistique: Les pronoms sujet en roman central*. Univ. of Toronto PhD dissertation. Toronto Working Papers in Linguistics.
- Jaberg, K. & J. Jud. (ed.) (1928-1940), *Sprach- und Sachatlas Italiens und der südschweiz*. Zofingen: Ringier.
- Marzys, Z. (1981), Les pronoms personnels sujets dans le parler francoprovençal de Faeto et Celle. *Vox Romanica* 40:48-58.
- Nadasdi, T. (1995), *Variation morphosyntaxique et langue minoritaire: le cas du français ontarien*. University of Toronto PhD dissertation.
- Nagy, N. (1994), Language Contact: Italian (?) Geminates in Faetar. *Belgian Journal of Linguistics* 9:111-128.
- Nagy, N. (1996), *Language contact and language change in the Faetar speech community*. Univ. of Pennsylvania PhD dissertation. Philadelphia: IRCS.
- Nagy, N. (1997), Modeling contact-induced language change. *Penn Working Papers in Linguistics* 4.1:399-418. (Papers from *NWAVE* 25).
- Nagy, N. & B. Reynolds. (1997), Optimality theory and variable word-final deletion in Faetar. *Language Variation and Change* 9.2:37-56.
- Perlmutter, D. (1971), *Deep and Surface Structure Constraints in Syntax*. New York: Holt, Rhinehart & Winston.