Writing a Sociolinguistic Grammar of Faetar

Naomi Nagy

Writing a Sociolinguistic Grammar of Faetar

Writing a Sociolinguistic Grammar of Faetar

Naomi Nagy

1 Introduction

Traditionally, reference grammars for little known languages represent the language as a homogeneous entity, ignoring variation in favor of conciseness. In contrast, sociolinguists focus on linguistic variation and its correlation to culturally relevant distinctions among speakers. We are interested in the issue of representing such variation *in the grammar*. (Note that here' grammar' refers both to a theoretical mental model of language and a book describing a language.) When a sociolinguist is confronted with the task of writing a grammar, the traditional methods of homogeneous grammar preparation clash with sociolinguistic goals.[1]

The biggest difference between the goals of the two types of undertakings relates to the scope of the data. Grammars describe many (ideally, all) parts of the language but are normally based on data from a small number of speakers. In contrast, publications in the field of sociolinguistics usually address only a very narrow part of the language, but they do so using data from many speakers. Because more speakers' forms are included, information may be provided regarding culturally relevant distinctions such as sex, age, or attitude of the speaker toward the language. In addition, because more speakers' forms are included, sociolinguistics publications describe more than one way to say a certain thing, while a grammar may prescribe (or describe) only one form per function. For these reasons, sociolinguistic publications are not meant to be used as reference texts—they are not meant as language-learning aids. Grammars may be made for this purpose. The aim of the two types of publications also differ: a grammar attempts to be theory neutral, providing data that could be lent in support of various formal linguistic theories. A sociolinguistic paper argues in favor of some particular model of linguistic form and/or social structure.

To illustrate the difficulties of combining these two types of enterprises, I discuss my recently completed grammar of Faetar (Nagy 2000a) as a case in point. I describe the structure of this grammar and discuss

---

some of the problems I have encountered in developing it. This grammar is meant as a reference text for linguists: it is published (by Lincom Europa) in a series of grammars of endangered languages. It includes competing variants for many aspects of the language and provides information about the speakers and contexts that produced each datum.

2 Background on Faetar

Faetar is a Francoprovençal (FP) dialect indigenous to Faeto, a village in southern Italy.[2] It was settled by people from the Francoprovençal region of France sometime in the 13th-15th century. There are currently about five hundred residents of Faeto, nearly all of whom speak Faetar and Italian. The people of Faeto are all cognizant of the historical difference between their language and the Italian dialects of the surrounding region. The language is referred to variously as [lu frantʃáj] 'French', [lu provɛnsál] 'Provençal' and [lu fajdár] 'Faetar'.

Faetar has not been codified. While there are several people who occasionally write brief texts in Faetar, there are no readers who are not dependent on the accompanying Italian translations. There are some texts that provide phonetic transcriptions of the language and/or present aspects of Faetar within the formalizations of linguistic theory: Castielli 1975, Fino 1970, Gallucci 1988, Hoffman 1968, Giuliani 1995, Kattenbusch 1982; however none are extensive descriptions published in English and none contain (much) information about the sociolinguistic variation found in the language.

3 Motivation for the Grammar

The need to examine non-standardized vernacular languages, such as Faetar, as part of the enterprise to reconcile grammatical theory and empirical observation was addressed by (Chambers 2000). There are several reasons to produce a grammar of an endangered language. As stated in the following quotes, these range from the scientific interest of seeing what a particular language can tell us about linguistic theory to providing a tool for cultural preservation.

> At this point in the history of linguistics, at least, *each* language
> offering testimony for linguistic theory brings something impor-

---

[2] A similar variety is spoken in the neighboring village of Celle San Vito. However, all data discussed here is from Faeto.

tant, and heretofore not known or not yet integrated into the theory. In many cases, data from a 'new' language forces changes in the developing theory, and in some cases, linguistic diversity sets an entirely new agenda (Hale 1998:194).

Documentation of languages that are near extinction will insure that these languages can contribute to scientific inquiry and to the cultural knowledge of those who are losing their ancestral language... As language represents an important component of any culture, the loss of a language can result in the loss of cultural identity (Goebl et al. 1996:659).

The contribution of this grammar is to provide a *variable* description of the language: one that describes the varied ways that the language is used to express meaning. It also is meant to serve as a form of documentation of a language whose number of speakers has been declining rapidly.[3]

Speakers have been predicting Faetar's demise for decades and lamenting the "imperfect form" of their present vernacular, which they attribute to contact with Italian. While it is clear that Italian and the local vernaculars have influenced Faetar, some of the changes are the inevitable result of language change, which is always in progress. I hope that by providing documentation of the language in a formal grammar format, speakers may recognize Faetar as a complete linguistic system, rather than an "ex-language" that has been steadily chipped away at by Italian. A standardized orthographic system could be an important tool in providing the necessary status to Faetar to make speakers accept it as a "real" language. This in turn may slow the decline of the language. I say this because my observations suggest that many children are hesitant to speak Faetar because their elders repeatedly tell them that they do not speak it correctly. If it were possible to validate multiple ways of speaking by including them all in a grammar, the effect of this behavior might be diminished.

Currently, both young and old speakers say that the young people don't use the language properly, or as their ancestors did. Here are a few quotations, which I have translated, from Italian or Faetar to English. The first is from the introduction to a school project in which students interview their grandparents and transcribe what they have said in Faetar and translate it to Italian. As many citizens see this booklet, this written state-

---

[3] The population has dropped from approximately 5,000 speakers in the mid 20th century to less than 500 speakers today.

ment may well have influenced the other speakers who provided the spoken statements below it.

(1) There is ... the phenomenon of the gradual loss (especially in the younger generations) of the most archaic vocabulary and phonetic and morphological structures which are most typical of our language, because these are too far from today's predominant reality: the Italian mass-media. There is the phenomenon of Italianization of our Francoprovençal, that is, the addition of the inflections and colorings from Italian (*Ricerca*... 1991:11).

(2) Our Faetar is already a bit Italianized. We don't speak a real Provençal. Even my parents don't really speak it. They never did... They don't teach the children, so they make mistakes... so now we have a different language. (Speaker F32A, Tape 10A)

(3) Faetar isn't spoken like it used to be. It's more a bastardization now. (Speaker M77, Tape 18A)

(4) Only an imperfect form of the language is learned now. (Speaker F80, Tape 17B)

These changes are not imagined by the speakers. Several variables I have investigated, such as variable subject pronoun use and variable deletion of post-tonic segments, show significant differences across the generations (see Nagy & Heap 1988, Nagy & Reynolds 1997). Some of these changes in progress, though not all, can be attributed to influence from Italian.

As part of their effort to fend off the perceived death of Faetar, several speakers have encouraged me to write a book describing their language. John Carosielli, a Philadelphian who grew up speaking Faetar with his Faeto-émigré parents, first suggested this to me in 1992. I hope that the grammar will serve him (and other English-speaking Faetani) and that the next version, to be written in Italian, will be of use to people in Faeto.

As discussed in (Nagy 2000b), designing the second (pedagogical) grammar highlights a significant hurdle: the lack of training in applied linguistics provided to students of theoretical linguistics. Developing pedagogical skills for teaching people about their language is generally not a part of the training received in a graduate program in linguistics, nor an activity that is rewarded in the academy. This has been succinctly pointed out by (Craig 1998:155-6) in a discussion of the potential dissonance between the demands of the field and the demands of an academic career:

The point to realize is that there is no division of labor in the field, that the linguists, with their formal education are the main—supposedly expert—resource for whatever project is wanted, from literacy programs to bilingual education programs, to revitalization programs, to translation of legal texts.

Additionally, there is the challenge of efficiently and effectively codifying this non-written language without misrepresenting the facts of language change and variation. To this end, the following elements integral to the construction of a sociogrammar are addressed:

- types of data
- organizational procedures for representing variable data
- development of orthography
- value judgments regarding data and speakers
- coordination with other grammars

4 Challenge 1: Selecting the Types of Data to Include

There is a well known trade-off between the ease of collecting 'artificially elicited' data such as recitations of verb conjugations (which lack any element of natural or vernacular speech) and the difficulty of procuring naturalistic speech, especially in a language in which the researcher lacks fluency (Fuller 2000). While a sociolinguist can often determine ways to elicit samples of the variable under study in various 'naturalistic' ways, it is not possible to construct such methodologies for every aspect of a language to be described in a grammar. Nor is it efficient to transcribe every recorded utterance and search for samples of, e.g., each form of each tense of a verb. Therefore, I used a variety of methods to collect the speech data that is described. Because of the trade-offs between accuracy and efficiency, each example included in the grammar is annotated according to how it was collected. Readers may determine the representativeness of a particular example. Figure 1 illustrates the relative representativeness of each type of data, the trade-off with how easy it is to collect and analyze, and the annotation scheme which I used in the grammar to represent the type of data each example represents.

Figure 1

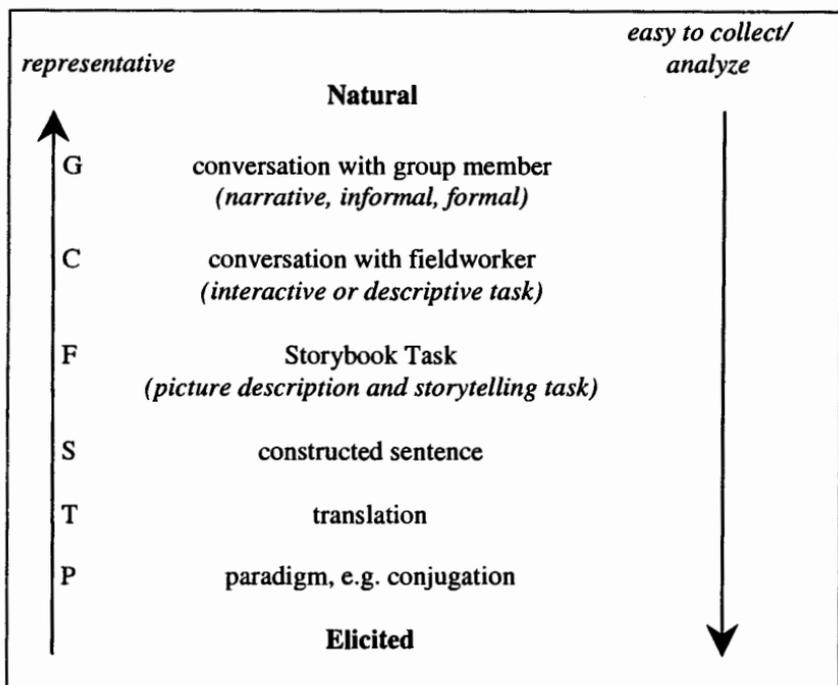| | | *easy to collect/* |
representative		*analyze*
	Natural	
G	conversation with group member	
	(narrative, informal, formal)	
C	conversation with fieldworker	
	(interactive or descriptive task)	
F	Storybook Task	
	(picture description and storytelling task)	
S	constructed sentence	
T	translation	
P	paradigm, e.g. conjugation	
	Elicited	

Figure 1: Types of data collected

Especially in cases where the researcher is not fully fluent in the language under study, naturally occurring speech is more difficult to elicit, record, transcribe, and translate. This is partially due to the unnaturalness of native speakers speaking their in-group language to an outsider. Also, the function of each utterance is not as easily matched to its form.

To mitigate these difficulties, structured tasks can be used. For example, the researcher can ask speakers to describe pictures or recount something that is already known to the researcher, such as their daily routine or the days of the week. However, this method is still problematic. To illustrate, Table 1 lists the forms that were produced in response to a request to describe a picture of a child's dress. The codes in the second column indicate which speakers produced each form in response to the picture. In the context, all these words meant 'dress.' However, later requests to speakers F26A and F32 for clarification of the different shades of meaning represented by the different forms in the second column produced the definitions in the third column. The discrepancies illustrate that even such a

method as picture description for controlling the functions of the utterance is not foolproof.

Utterance	Speaker	Post hoc gloss
vést(ə)	F83	*dress*
la vestít(ə)	F9B, F26A, F27B, F40, M25A, M27B, M28A, M34A	*man's suit*
la vestín(ə)	F6, F27A, F32, F56B, F65A, F80, M32, M78A	*little dress*
lo vestətín(ə)	F30, F88	*little suit for boy or girl*
la vestarél	F65A	*cute dress*
la vestiʃwól(ə)	M72A, M12, M44A, M76, M77	*cute suit for little boy*
vestətíəl(ə)	M47C	*suit for little boy*

Table 1: Defining diminutives

Translation from a shared language to the language under study is often used to quickly elicit many forms. One obvious default of this method is the influence that the shared language might have on the forms produced. A second problem is that, here, too, the data is not fully predictable. The following exchange illustrates this. Italicized forms are the Italian prompts and forms transcribed in IPA are *expected* to be the Faetar translations.

(5) NN: *Diamo il libro al ragazzo.*   *Let's give the book to the boy.*
     M81: [denə lu lívərə a lu kwattrá]   Give the book to the boy.
     NN: *Hai datto il libro alla mama?*   *You gave the book to the mom?*
     M81: [ʎo]   Yes.   (Tape 43A)

Another technique is to have speakers produce paradigms such as verb conjugations or all possible combinations of a preposition + article. Here the data is likely to be influenced by any languages studied formally, as the speaker must draw from somewhere the knowledge of each paradigm's structure. That is, if a speaker studied French in school, and those studies included memorization and recitation of verb conjugations, these patterns are likely to be drawn on when the speaker is asked to recite Faetar conjugations. A further shortcoming of this method is that it works only with educated speakers and thus cannot be used to represent the full range of speakers' forms.

A final problem with asking for translations, conjugations, or other extensive paradigms is the effects of fatigue or confusion on the speakers. For example, when I was collecting examples of the use of deictics meaning 'this' (proximal) and 'that' (distal), speakers varied greatly in the degree of correlation in the pairing of "sVt-" forms to proximal prompts and "sVl-" forms to distal prompts. I do not feel comfortable reorganizing this section of the grammar to include all "sVt-" forms as proximal examples and all "sVl-" forms as distal examples, although it is possible that that is the general pattern and deviations from it were caused by fatigue or confusion during recording sessions.

A balance of different types of data resolves the types of problems noted here. The formally elicited forms can be checked against the naturally elicited data for 'vernacularness' and accuracy. The formally elicited forms can, in turn, be used to determine the meaning and function of the naturally elicited data. However, once the language is understood by the researcher, more problems abound.

5 Challenge 2: Organizational Procedures

The traditional format of a grammar does not allow for representation of differences among forms used by different speakers, or multiple forms used by a single speaker. Even grammars that provide multiple examples do not usually indicate the type of speaker or context that produced each form. The Faetar grammar is a description of speech that was collected over an eight-year period from about eighty of the 500 residents of Faeto and it is important to know what sort of speaker and context produced each form. The following subsections address the organization of data from multiple speakers, in multiple styles, and relating to multiple variables in the language.

5.1 Multiple Speakers

Rather than placing all of the information about the speaker and context in the text adjacent to each utterance, I relegated most of the information to an appendix containing of speaker and context information and annotated each utterance with the minimal information necessary for cross-reference. This annotation includes a speaker code that indicates the sex of the speaker (M or F) and the speaker's age at the time of the first recording, a single letter indicating the type of data (see Figure 1), and the number of the tape on which the recording is archived. Extracts from that table are shown in Table 2. This table identifies the speaker in column 1 and

shows where they reside and their occupation in columns 2 and 3. The date of the interview, a unique identifying number for the interview, and the tape on which the interview is archived are shown in the next columns. The final column indicates what sort of data was recorded in broad terms (conversation, picture description task, or grammatical elicitations). Indices showing the amount of contact with Italian the speaker has are also provided.

Speaker	Home	Occupation	Date	IV	Tape	Content[4]
M6	Faeto	elementary	9/2-4/93	69	22a, b	conv
M7	Faeto	elementary	08/06/94	123	26a	FW
M13C	Faeto	student	03/29/00	75	41b	conv
M23B	Faeto	student	08/21/93	33	20a	FW
M23B	Faeto	student	08/23/93	48	20b	conv
M28C	Faeto	civil servant	04/03/00	89	NR	grm
M30	Faeto	soldier	03/29/00	78	42a	conv
M44A	Faeto	engineer	09/04/93	22	16b	conv
M74	Faeto	merchant	03/29/00	77	42a	grm
M81	Faeto	ret. farmer	03/30/00	80	43a	grm

Table 2: Speaker information

Although this table provides a fair amount of information about each speaker, it may not necessarily provide the factors that turn out to be relevant in accounting for all observed variation. Therefore, it is necessary to provide information so that one may return to the original data for more information about the speakers and/or context of the utterance. This absolves the author from needing to exhaustively present all information about each speaker's group membership(s).

   There have been some previous attempts to include social information in a grammar, but I am not aware of any that attribute every example to a speaker so that correlations to social factors could be explored for any aspect of the language. However, some grammars do make some mention of which types of speakers favor which variants. For example, the following types of descriptive statements are embedded in the Phonotactics chapter

---

[4] The abbreviation used in this column are:

conv   Conversation (with me or a native speaker)
FW   Storybook Task. Elicited by asking speakers to describe pictures in Amery & Cartwright's *First 100 Words* to me.
grm   Grammar (elicitation by paradigm and/or translation)

of *A Grammar of Tamambo: The language of Western Malo, Vanuatu* by Dorothy Jauncey. It reports mostly age effects, but also sex and education.

> Occasionally, older speakers articulate the prenasalised voiced stop with a slight [r] off-glide... – optionally. (Jauncey 1997:25)

> Fricative /β/ is often realised by younger speakers (<35) as [v] as in /βeᵐbe/ → [vɛmbɛ] 'butterfly'. (Jauncey 1997:26)

5.2 Multiple Styles

Having explored some of the ways that speaker variation can be reported in a grammar, I turn now to style variation. Again, the best that can be done at the outset of the project is to carefully report where each datum came from so that information regarding stylistic or contextual features can be retrieved when determined relevant.

As a non-native speaker and out-group member, there is a limit to the types of styles I observed. Even when I was able to record 'real' conversation, transcribing, translating, and coding is difficult. For this reason, I relied on elicited data to supplement the naturalistic data to a greater degree than is common in sociolinguistics research. Another reason for eliciting forms is to collect more reliable information on the types of variants available to the speakers. It is not clear how much natural speech data one must sift through before concluding that only one (or any other number) form exists for a certain structure or concept. By targeting a particular function, one may more efficiently determine the range of possible variation. One troublesome aspect of this is that a grammar that reports more than one form for a particular function may be making (or be read as making) an implicit claim to have represented all possible forms in a way that a traditional grammar does not.

5.3 Multiple variables

The final element of complexity is that, unlike most sociolinguistic articles, a grammar addresses the full range of the language and finds variation in many parts of it. It is not necessarily possible or desirable to conduct quantitative analyses of each variable. Therefore, it is not clear which linguistic, stylistic, or social variables are correlated to the variation of any linguistic dependent variable. Even without that added complication, there are difficulties in organizing the different parts of the language in a clear and systematic way.

I initially attempted to model my grammar on Stich's (1998) grammar of FP. Having collected different types of data than he did, this was not feasible. Additionally, in the process of trying to organize my data into his paradigms, I realized that one might lose valuable material by attempting to make one language fit into the description of another. As pointed out by (Hale 1998) (quoted above) and (Chambers 2000) with particular emphasis on documenting vernacular speech, one of the purposes of documenting more languages is to discover new linguistic features and structures. If one only reports elements that correspond to previously described elements, one descreases the likelihood of making significant new contributions. This may be clearly seen in thinking about Nick Evans' finding that tense is marked on nouns in Kayardild, described in (Wuethrich 2000:1156). If Evans had described that language in the format of some existing description of another language, there would have been no place to note that.

Combining Stich's format, that recommended by the publisher, and my own design, I organized my grammar in a traditional format. There are chapters on Phonology, Lexical Morphology, Grammatical Morphology, Derivational Morphology, and Syntax, as well as three transcribed texts and an appendix containing information about each speaker. Variation is noted if it was noticed, but no exhaustive attempt was made to collect data from all types of speakers in all contexts for all variables. (It is no coincidence that '!' is the sign for both factorial combination and exasperation.) The following section expands on this combinatorial problem.

5.4 Multiple Tiers x Multiple Forms

Grammars are organized in many-tiered hierarchies. For example, one might choose to look at the second person plural (2p) form of the indicative imperfect past of a Type 1 verb in the "Verbs" section of the "Morphology" chapter within the "grammar" part (as opposed to the lexicon or sample texts) of the grammar book. Each other form in the book differs in some way on at least one level of that hierarchy. A decision tree representing this process is shown in Figure 2.
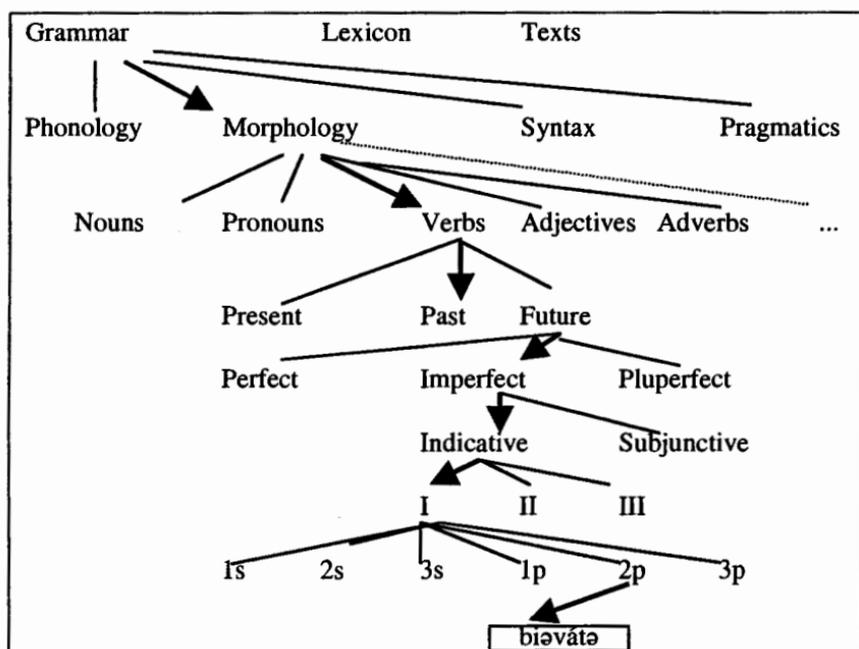
Figure 2: Hierarchy in a traditional grammar

A variable grammar requires at least one more dimension in order to show each such form for each cross-section of the relevant social and contextual factors. In addition to showing the possible forms for each cell of this many-dimensioned array, a sociolinguist is tempted to show the relative frequency of each form and which factors best correlate to it. Doing this would involve a structure of the type shown in Figure 3, all fitting into the single box at the bottom of Figure 2.[5] This tree shows the decision paths followed to arrive at the second plural imperfect indicative forms of a Type I verb produced by three speakers from the middle age group, but differing in sex and occupation. (Any other combinations of appropriate social factors could also be used.)

---

[5] This diagram is actually a simplification as it does not include a style axis and does not make allowances for multiple forms produced by one (type of) speaker.
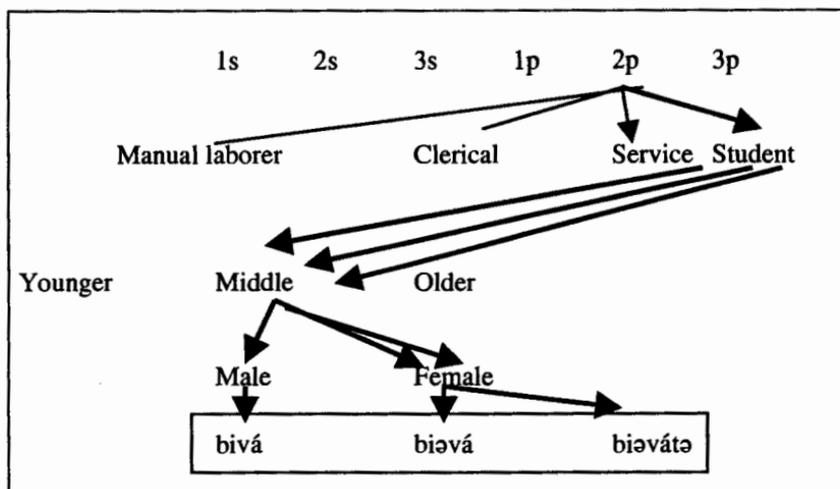
Figure 3: Hierarchy in a variationist grammar (group variation)

To simplify the grammar, I make no explicit claims about which forms are favored by which groups or styles. Rather, I annotate each example to indicate the speaker and context. Readers may draw their own conclusions about which aspects of the speaker's identity and the context were relevant in selecting a particular variant. Thus, the structure of the Faetar grammar is that shown in Figure 4 (combined with Figure 2): a variable but not sociolinguistically prescriptive grammar.
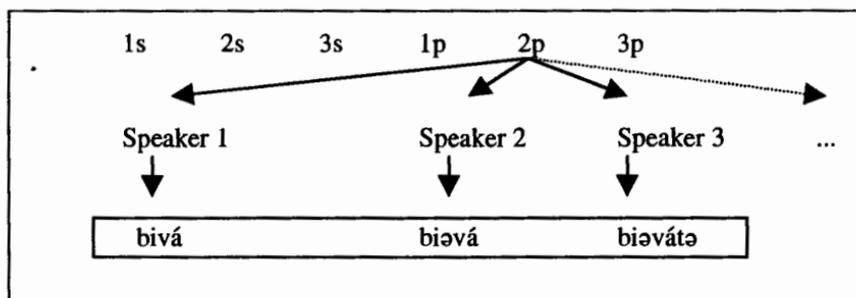

Figure 4: Hierarchy in a Variationist Grammar (individual variation)

6 Challenge 3: Codifying an Oral Language

Faetar is an oral language. A few people write occasional short texts in Faetar. These are always accompanied by an Italian translation, which

provides essential assistance in reading the Faetar text—even the authors
of the texts are not fluent readers. Each has developed a different tran-
scription system, and the battle is heated between proponents of repre-
senting etymological patterns vs. current pronunciation, which translates
to using a French-like or an Italian-like spelling system. The alternative of
using a system of phonetic symbols has been rejected on the grounds of
difficulty of typing as well as reading (for non-linguists). To this, I add the
difficulty of constructing an orthography of a language whose pronuncia-
tion varies greatly. There are variable realizations of word-initial gemi-
nates due to a sandhi process as well as highly varying amounts of post-
tonic deletion. One must decide whether to represent each utterance as it
was produced (phonetically, to some extent) or to choose an invariant
spelling for each word, which is the norm in codified languages.

In the grammar intended for linguists, I use IPA symbols to represent
each datum as it was produced in the instance for which it was coded.
However, a less technical orthography is necessary for the second version
of the grammar, meant for the people of Faeto. Thinking that the intui-
tions of native speakers would provide a useful starting point for the de-
velopment of an orthography, I produced the following Italian sentence,
asked someone to say it in Faetar, and then to write it down.

(6) Conosco la figlia/donna che studia/parla il faetano.
     'I know the girl/woman who studies/speaks Faetar.'
     [dʒi dʒə kwannáj la fiʎ/fen k i stúdjə/ parlə lu fajdár]

The responses in (7) were written. There is no word that everyone spells
alike, although there was agreement on pronunciation. Such complications
are multiplied as different pronunciations are introduced.

(7)	G	cuanej	la	figlj	ch	i	studj	lu	faitar	M25B
	G	Quanai	n	fenn	k	i	studij	lu	faitar	M23C
	Ji ge	quanaij	la	fenne	que	ij	studje	lu	faitare	F24B
	Je	cuana'j	la	figl'y	k	i'	studi'y	lu	faitar	F20A
	Je	cuanij	la	'ffenn '	ch'	i	studij	lu	fajtar	F50E
	J	Quanaj	le	figl'	ch	i	parl		faitar'	M28B
	Je	Quhnaj	le	fen	chi		parl	lu	faitar	M20B

7 Challenge 4: Value Judgments

Although no attempt was made in this grammar to attribute correlation of linguistic variants to certain aspects of speakers' identities, I found it difficult to represent the different variants without indicating that one was more highly valued than another. The spatial organization of the book forces the author to act on certain value judgments, and these should be made explicit.

Grammars that provide one variant in the text and then list others parenthetically or in footnotes implicitly indicate a preference for the first variant listed. To avoid constructing a prescriptive grammar, one must present the variants in as egalitarian a manner as possible. No variant should be presented in a way that make it seem like an' alternative' form if one does not have prescriptive goals in mind. (The range of data types included in this grammar is described above in Section 4.)

7.1 Judging Types of Speaker

I wished to avoid valorizing any type of speaker over any other. Data from all types serve a purpose. Valuable work has been done by including semispeakers to broaden the pool of speakers that can provide data for small languages (Dorian 1981: Ch. 4; Sankoff et al. 1997), so they should not be excluded. Assuming the validity of the Apparent Time Construct (Bailey et al. 1991) data from older speakers provide evidence of older forms of the language that may be approaching obsolescence. Similarly, data from young speakers provides valuable clues about new directions the language may be taking. Data from educated speakers is valuable in part due to its ease of collection by translation or grammatical elicitation. Speakers who have studied languages can also provide valuable input about the structure of the language. Un- (or less-) educated speakers, on the other hand, provide data that is valuable because it is untainted by the influence of formal study of language. Monolingual speakers provide similarly valuable data, while multilingual speakers can provide parallels from other languages that may help the researcher understand the forms or functions better, and they can participate in translation tasks. And so on. Thus, I have collected data from anyone who self-identifies as a Faetar-speaker and can speak in a relatively fluid manner.

Accepting data from a wide range of speakers forces the issue of organizing it. If one lists the variants alphabetically, the shorter forms (in Faetar, more FP-like, more typical of older speakers) get listed first. And that is only if one can construct an alphabetization schema for words tran-

scribed in IPA. If one lists variants alphabetically by speaker code, or
some such, one unwittingly valorizes the first speakers listed (in my cod-
ing method, that would be the youngest girls, whose language is most
Italian-like). Any linear organization of forms implicitly suggests that the
first variant listed is the' best' one, and that others are alternatives that
may not be acceptable to all. One must explicitly state otherwise. How-
ever, that may not be sufficient.

I have chosen to represent the forms in the order that I randomly
came across them in my data. Where several speakers provide the same
variant, I attributed the form to the first few speakers I came across, with
an aim to presenting some of the social and stylistic breadth of the form.
However, due to time constraints, I sometimes chose to look through
elicitation or translation data first which means that I list first forms pro-
vided by educated speakers.

Providing attribution to more than one speaker makes display un-
wieldy in a two-dimensional text. Several people have suggested a web-
based grammar that would allow for easier navigation through the hierar-
chies, but that remains for future work. For now, hierarchies as shown in
Figures 2 and 4 are utilized. To show the implementation of that format,
the conjugation of a verb in (almost) all its tenses and moods is shown in
Table 3 (opposite). It is considerably longer than that found in a prescrip-
tive grammar, where each tense could occupy only a single line. The
rightmost column lists the speakers who produced the forms on that line.
For this verb, the three singular persons share the same form in all tenses,
so all singular persons are represented in one column, allowing an ac-
ceptably narrow display. Verbs where all persons differ introduce further
formatting problems.

Another type of value judgment problem exists—that of determining
the validity of explicit judgments or grammatical information provided by
speakers. Research has shown that speakers are not fully aware of the
range of ways that they may express a particular function (Labov 1975,
Nagy & Karins 1993, Tillery 2000). In order to present the less common
forms of verbs, such as the conditional and subjunctive, I resorted to ask-
ing educated speakers to conjugate verbs in those forms. Speakers indi-
cated that there is confusion about when to use each of these forms in
Italian, and that evidently carried over into Faetar. The general tendency
was to use the following suffixes.

(8) **Person**   **Conditional**   **Subjunctive**

1s-3s	-írrə	-íssə
1 p	-ərán	-əsián
2p	-ərá	-əsiá
3p	-əránd	-əsiánd

INF	bájə	<u>to drink</u>	*F26A P 88, F29C P 88, F12B P 76*		
	bájˠ		*M74 77*		
PrP	bián	*M28C P 84*	**PP** biáw	*F26A P 88, F29C P 88,*	
					F32A P 79
	1-3s	**1p**	**2p**	**3p**	
PRES	bájə	biún	bií	biúnd	*F26A P 88*
	"	"	bíə	"	*F32A P 79,*
					M23C P 81
	"	bəvún	bəví	"	*F12B P 76*
			bəvé	bəvúnd	*F12B P 76*
	baíj (3s)				*F11B F 139*
IMPF	bívə	biəván	biəvátə	biəvánd	*F26A P 88*
	"	"	biəvá	"	*F32A P 79*
	"	"	bivá	bivánd	*M23C P 81*
PST	bií(t)	biərún	bíər	biərúnd	*F26A P 88*
	"	"	bíítəvə		*F29C P 86*
FUT	birá	birán			*F32A P 79*
CND	bírrə	biərán	biərá	biəránd	*F32A P 79*
SBJ	bíísə	biəsián	biəsiá	biəsiánd	*F26A P 88*
IMP	baí	bián	biíə		*F26A P 88*

Table 3: Sample verb paradigm (partial)

However, in certain cases, speakers provided forms with the suffixes labeled subjunctive in response to a request for the conditional, and vice versa. This was true also for sentences produced in more naturalistic contexts, suggesting that the functions (of indicating conditionality and all the many things that the subjunctive is used for) are not expressed with mutually exclusive sets of forms. I classified all forms elicited with the subjunctive suffixes as subjunctive forms even if the speaker produced them in response to a request for conditionals, and vice versa. I do not feel very

comfortable about this decision, as it requires forcing Faetar into a mold developed for other languages and prioritizing the non-native speaker linguist's analysis over that of native speakers. There is the alternative possibility that, in natural speech, both types of suffixes are used to indicate both subjunctive and conditional. Without knowing *a priori* what contexts require each verb mood, it is not possible to know if that is so.

7.2 Judging Borrowings

Another area of value judgments presents itself. Many aspects of the variety currently spoken in Faeto are similar, to varying degrees, to the neighboring Italian varieties. While it is tempting to exclude all such parts of the language from a grammar of Faetar, one must contend with the fact that FP and Italian have shared many features for years. Therefore, a similarity to Italian does not immediately disqualify a feature of the language from being indigenous to Faetar. Due to the uncertainty of the time and place in which Faetar originated, and the lack of extensive localized documentation of FP from the period of its original settlement, attribution of origins for many aspects of the language remains questionable. Furthermore, the alternation between forms that are indigenous to Faetar and others may well be of interest to future researchers who study, for example, the linguistic representation of local vs. broader identity.

However, excluding or including possibly non-indigenous variants are not the only options. One might choose to include them but mark them as borrowings. As is often the problem with borrowings, one would need a clear protocol for identifying which borrowings had become incorporated as part of the language. It would be extremely time-consuming to try to sort them out due to the phonological similarities of FP and Italian.

8 Challenge 5: Coordination with Other Work

This discussion of Faetar's similarities to neighboring varieties leads to my last point: the coordination of one grammar with descriptions of other languages. To facilitate comparative research, it would be helpful to note which aspects of Faetar resemble other languages. To do this, one must decide whether to compare Faetar to geographically, typologically, and/or genetically close languages. Each such comparison would have certain effects. Highlighting similarities to Italian might be interpreted as suggesting that those aspects of the language are not native to Faetar, but rather a result of Italianization, a process feared by many Faetar speakers (see Section 3). Comparison to non-standard regional Italian varieties,

which have low status themselves, might further stigmatize Faetar. Comparisons to French might highlight the distinctive origin of Faetar in a way that would have a positive impact on the preservation of the language. These comparisons can be made explicitly but also will be implied by the orthographic system selected to represent the language.

9 Conclusion

This grammar is the outcome of a desire to provide a description of an endangered language both for its speakers who may be interested in preservation of the language and to scholars interested in studying vernacular, Romance, oral, and/or contact languages. My training as a variationist sociolinguist influenced my production of the grammar, prohibiting me from representing any aspect of the language as invariant that might actually be variable. This brought to the forefront issues such as:

- the interplay between naturalistic and elicited speech data of various sorts and their differing;
- representing competing forms that carry out the same linguistic function without devoiding them of their social function;
- organizing the many dimensions of linguistic, contextual/stylistic, and social variables;
- codification of an oral language;
- recognizing judgments related to the value of data and speakers; and
- the relation of one grammar to other grammars and how this may construct different relationships among the languages represented.

The goal is for the resulting grammar to provide a picture of Faetar that is more faithful to the ever-changing and socially-situated nature of the actual speakers, rather than some "ideal[ized] speaker-listener in a homogeneous speech community" (Chomsky 1965:3-4).

References

Amery, Henry and Stephanie Cartwright 1993. *First 100 Words*. Oklahoma: EDC.
Bailey, Guy, Tom Wikle, Jan Tillery and Lori Sand. 1991. The apparent time construct. *Language Variation and Change* 3.3: 241-265.
Castielli, Roberto. 1975. *Funtan' d' Fait': Testi integrali dei cori con presentazione traduzione e note*. Faeto, Italy.

Chambers, Jack. 2000. Universal sources of the vernacular. Paper presented at the International Conference on Language Variation in Europe, Universitat Pompeu Fabra, Barcelona.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Mass: MIT Press.

Craig, Colette Grinevald. 1998. Language contact and language degeneration. In *The Handbook of Sociolinguistics*. Florian Coulmas (ed.). Blackwell. 257-270.

Dorian, Nancy. 1981. *Language Death: The Life Cycle of a Scottish Gaelic Dialect*. Philadelphia, PA: University of Pennsylvania Press.

Fino, Carmela. 1970. *Il dialetto francoprovenzale di Celle S. Vito in provincia di Foggia*. Master's thesis. Università di Bari.

Fuller, Janet. 2000. Language choice and speaker identity: The influence of a researcher's linguistic proficiency in interviews. *Southern Journal of Linguistics* 24.1: 91-102.

Gallucci, Concetta. 1988. Provenzale o Franco Provenzale? *Il Provenzale*. Aug. 1988: 5-7.

Giuliani, Francesca. 1995. *Ricerche sul francoprovenzale di Faeto e Celle San Vito in provincia di Foggia*. Master's thesis. Università di Firenze.

Goebl, Hans, Peter Nelde, Zigmund Stary, and Wolfgang Wolck. 1996. In *Contact Linguistics: An International Handbook of Contemporary Research*. New York: de Gruyter.

Hale, Ken. 1998. On endangered languages and the importance of linguistic diversity. In *Endangered languages: Current issues and future prospects*. Lenore Grenoble and Lindsay Whaley. Cambridge, Mass.: Cambridge Press. 192-216.

Hoffman, Robert J. 1968. *The Franco-provençal dialect of Faeto*. Master's thesis. Ohio State University, Columbus, OH.

Jauncey, Dorothy. 1997. *A Grammar of Tamambo: The language of Western Malo, Vanuato*. Doctoral dissertation. Australian National University, Canberra.

Kattenbusch, Dieter. 1982. *Das Frankoprovenzalische in Süditalien. Studien zur synchronischen und diachronischen Dialektologie*. Tübingen: Gunter Narr Verlag.

Labov, William. 1975. *What is a linguistic fact?* New York: Humanities Press.

Nagy, Naomi. 1996. *Language contact and language change in the Faetar speech community*. Linguistics. Doctoral dissertation. University of Pennsylvania, Philadelphia. IRCS Technical Report 96-08.

Nagy, Naomi. 2000a. *Faetar*. München: Lincom Europa.

Nagy, Naomi. 2000b. What I didn't know about working in an endangered language community: Some fieldwork issues. *International Journal of the Sociology of Language* 144:143-160.

Nagy, Naomi and David Heap. 1998. Francoprovencal null subject and constraint interaction. In *CLS 34:The Panels*. Chicago Linguistic Society. 34.2:151-166.

Nagy, Naomi and Krisjanis Karins 1993. Developing an experimental basis for determining grammaticality. *Penn Review of Linguistics* 17: 93-100.

Nagy, Naomi and Bill Reynolds 1997. Optimality theory and variable word-final deletion in Faetar. *Language Variation and Change* 91: 37-56.

Sankoff, Gillian, Naomi Nagy, Pierrette Thibault, Hélène Blondeau, Marie-Odile Fonollosa, and Lucie Gagnon. 1997. Variation and the use of discourse markers in a language contact situation. *Language Variation and Change* 9.2: 191-218.

Stich, Dominique. 1998. *Parlons Francoprovençal: Une Langue Méconnue*. Paris: L'Harmattan.

Tillery, Jan. 2000. The Reliability and validity of linguistic self-reports. *Southern Journal of Linguistics* 24.1: 55-68.

Wuethrich, Bernice. 2000. Learning the world's languages- before they vanish. *Science* 288: 1156-9.

English Department
University of New Hampshire
Durham, NH 03824
ngn@hopper.unh.edu