

VARIABILITY IN INTERROGATION AND NEGATION

IN SPOKEN FRENCH

Aidan Benedict Coveney

Thesis submitted for the degree of PhD

University of Newcastle upon Tyne

Department of French Studies

September 1989

NEWCASTLE UNIVERSITY LIBRARY

287 54978 4

h2.5
- . 5 80

CONTENTS

ABSTRACT.....	x
ACKNOWLEDGEMENTS.....	xii
INTRODUCTION.....	1
CHAPTER 1: GRAMMATICAL VARIABILITY.....	4
1.1 Introduction.....	4
1.2 A simple typology of grammatical alternations.....	5
1.3 The problem of the infrequency of certain forms.....	9
1.4 The case against grammatical variables.....	12
1.4.1 Grammatical variables are inextricably linked to variable rules.....	12
1.4.2 Grammatical variables investigated have not exhibited social differentiation (or stylistic variation).....	13
1.4.3 Grammatical variables involve meaning-bearing units....	16
1.4.4 The concept of grammatical variables violates the "one form: one meaning" axiom.....	17
1.4.5 The meaning of a grammatical form is the reflection of all its uses.....	20
1.4.6 Cases of structural mismatch cannot be analysed in terms of grammatical variables	22
1.5 The case for grammatical variables.....	24
1.5.1 Evidence from other- and self-repairs in conversation..	24
1.5.2 Evidence from correction, repetition, interpretation and translation.....	27
1.5.3 Evidence from linguistic change.....	29
1.6 Approaches to the problem of equivalence.....	31
1.6.1 Historical data.....	31
1.6.2 Intonational variability.....	32
1.6.3 Non-variationist quantitative studies.....	33
1.6.4 Weak complementarity.....	33
1.7 Two problematic cases of "equivalence".....	35

1.8 Criteria for identifying grammatical variables.....	38
CHAPTER 2: A CORPUS OF SPOKEN FRENCH.....	42
2.1 Introduction.....	42
2.2 The Somme and the Picard dialect.....	44
2.3 Centres de vacances.....	50
2.4 Fieldwork procedure.....	53
2.5 The interviews.....	56
2.6 Characteristics of the speakers.....	62
2.7 Transcription.....	67
2.8 Computerisation of the corpus.....	71
2.9 Use of the Oxford Concordance Program.....	73
CHAPTER 3: THE VARIABLE OMISSION OF THE NEGATIVE PARTICLE....	78
3.1 Defining the variable and its variants.....	78
3.2 Meaning and (ne).....	80
3.3 The socio-stylistic value of the variants.....	81
3.4 History.....	82
3.5 Geographical diffusion of the omissibility of <u>ne</u>	87
3.6 Negation in Picard.....	88
3.7 Previous quantitative studies of (ne).....	89
3.7.1 Data, informants and methodology.....	89
3.7.1.1 Gougenheim <u>et al</u> (1964).....	90
3.7.1.2 Ashby (1976) and (1981).....	90
3.7.1.3 Sankoff & Vincent (1980).....	91
3.7.1.4 Diller (1983).....	91
3.7.1.5 Lüdicke (1982).....	91
3.7.1.6 Moreau (1986).....	92
3.7.2 A comparison of the overall frequency of <u>ne</u> in previous surveys and in the present study.....	92
3.7.3 Children's use of (ne).....	94

3.8 Non-tokens and categorical tokens.....	95
3.8.1 General exclusions.....	95
3.8.2 Excluded occurrences of the negative particle.....	97
3.8.3 Excluded occurrences of the negative items.....	98
3.9 Editing the files of the concordances.....	101
3.10 Linguistic constraints.....	103
3.10.1 Introduction.....	103
3.10.2 The subject of the verb.....	104
3.10.3 Negative items.....	108
3.10.4 The presence or absence of a non-subject preverbal clitic.....	110
3.10.5 Phonological environment.....	112
3.10.6 Collocations or "preformed sequences".....	114
3.11 Interpersonal variation.....	121
3.12 Style-shifting.....	129
3.13 Conclusion.....	132
CHAPTER 4: INTERROGATIVE STRUCTURES.....	134
4.1 Defining the variables and their variant structures.....	134
4.1.1 Direct and indirect interrogatives.....	134
4.1.2 Interrogative comment clauses.....	135
4.1.3 Yes/no interrogatives and WH interrogatives.....	136
4.1.4 Variant structures occurring in the corpus.....	136
4.1.5 Clitic inversion and NP inversion.....	137
4.1.6 The [Q=S V] structure.....	138
4.1.7 Structures involving a dislocated NP coreferential with subject clitic.....	139
4.1.8 Structures not occurring in the corpus.....	140
4.1.8.1 The interrogative particle <u>ti</u>	140
4.1.8.2 Structures said to be 'rivals' of <u>est-ce que</u>	141
4.1.8.3 The structure [QEV NP].....	142

4.1.8.4 The structure [Q=S V-CL].....	142
4.1.8.5 The structure [E NP V-CL].....	143
4.2. The socio-stylistic evaluation of the variant structures by grammarians.....	143
4.3. The history of interrogative structures in French.....	145
4.3.1 Introduction.....	145
4.3.2 Old French: before c. 1300.....	146
4.3.3 Middle French: from c. 1300 to c. 1600.....	147
4.3.4 Classical French: from c. 1600 to c. 1800.....	148
4.3.5 Modern French: post-1800.....	149
4.4 Interrogatives in Picard dialect (as spoken in the Somme).....	150
4.4.1 Yes/No interrogatives.....	150
4.4.2 WH Interrogatives.....	151
4.4.3 Other Structures.....	151
4.4.4 Picard structures and the informants in the present study.....	152
4.5 Previous quantitative studies.....	153
4.5.1 Data, informants and methodology.....	153
4.5.1.1 Introduction.....	153
4.5.1.2 Gougenheim <u>et al</u> (1964).....	153
4.5.1.3 Pohl (1965).....	154
4.5.1.4 Terry (1970).....	155
4.5.1.5 Soll (1982).....	156
4.5.1.6 Behnstedt (1973).....	156
4.5.1.7 Ashby (1977).....	157
4.5.1.8 Lefebvre (1981).....	157
4.5.1.9 Lafontaine & Lardinois (1985).....	158
4.5.2 Results of previous quantitative studies.....	158
4.5.2.1 Reorganisation of the data.....	158
4.5.2.2 Yes/No interrogatives.....	159

4.5.2.3 WH interrogatives.....	161
4.6 Interrogatives in the present corpus: obtaining sufficient data and identifying the tokens.....	164
4.6.1 Obtaining sufficient data.....	164
4.6.2 Identifying occurrences of the variants: general problems.....	166
4.6.3 Problems in identifying tokens of (WHQ).....	168
4.6.4 Problems in identifying tokens of (YNQ).....	171
4.7 A preliminary comparison of interrogatives in the corpus with previous quantitative studies.....	171
4.8 The problems of acceptability and equivalence.....	174
4.8.1 The need for judgements of acceptability and equivalence.....	174
4.8.2 The equivalence of variants.....	176
4.8.3 Solutions to the problems of acceptability and equivalence.....	176
CHAPTER 5: THE COMMUNICATIVE FUNCTIONS OF INTERROGATIVES.....	180
5.1 Introduction.....	180
5.2 A set of distinctive features for categories of communicative function.....	181
5.2.1 Features concerning the relationship between speaker, utterance and addressee.....	181
5.2.2 Features concerning the author's knowledge, beliefs and assumptions.....	183
5.2.3 Features concerning propositional content.....	187
5.3 Exemplification and discussion of categories of communicative function.....	191
5.3.1 Introduction.....	191
5.3.2 Request for information.....	192
5.3.3 Request for opinion.....	194
5.3.4 Check (on the knowledge of the addressee).....	195
5.3.5 Request for clarification.....	196
5.3.6 Request for advice.....	198

5.3.7 Request for reminder.....	199
5.3.8 Suggestion.....	199
5.3.9 Offer.....	200
5.3.10 Request for permission.....	201
5.3.11 Request for action.....	202
5.3.12 Tentative assertion.....	204
5.3.13 Emphatic assertion.....	205
5.3.14 Exclamatory assertion.....	210
5.3.15 Sub-topic-introducing question.....	212
5.3.16 Pre-announcement.....	214
5.3.17 Post-announcement.....	217
5.3.18 Echo question.....	219
5.3.19 Indeterminate communicative function.....	221
5.4 Conducive questions.....	226
5.4.1 Terminology.....	226
5.4.2 Four types of conducive questions.....	227
5.4.3 Affirmative questions expecting a positive answer.....	228
5.4.4 Affirmative questions expecting a negative answer.....	235
5.5 Negative questions.....	238
5.5.1 Logical structure.....	238
5.5.2 History.....	239
5.5.3 Negative questions expecting a negative answer.....	240
5.5.4 Negative questions expecting a positive answer.....	244
5.6 Conclusion.....	249
 CHAPTER 6: CONSTRAINTS ON INTERROGATIVES.....	 251
6.1 Introduction.....	251
6.2 Coding.....	251
6.3 Categorical tokens.....	253

6.3.1 The exclusion of categorical tokens in variationist methodology.....	253
6.3.2 Categorical tokens of (YNQ).....	254
6.3.3 Categorical tokens of (WHQ).....	261
6.4 Quantification of non-categorical tokens.....	267
6.5 Categorical linguistic constraints.....	268
6.5.1 [V-CL] and [QV-CL].....	268
6.5.2 [QSV].....	273
6.5.3 [QESV].....	274
6.5.4 [QV NP].....	275
6.6 Categorical pragmatic constraints.....	282
6.6.1 Yes/No interrogatives.....	282
6.6.2 WH interrogatives.....	284
6.7 Variable linguistic and pragmatic constraints.....	286
6.7.1 Preliminary comments.....	286
6.7.2 [SV] and [ESV].....	290
6.7.3 [V-CL] and [QV-CL].....	294
6.7.4 [Q=S V].....	302
6.7.5 [QV NP].....	303
6.7.6 [SVQ].....	307
6.7.7 [QSV].....	320
6.7.8 [QESV].....	324
6.8 The effect of speaker identity and audience design.....	327
6.8.1 A quantitative analysis.....	327
6.8.2 Yes/No interrogatives.....	328
6.8.3 WH interrogatives.....	330
6.8.4 Further considerations.....	334
6.9 A socio-pragmatic interpretation of the variability.....	338
6.9.1 Problems with a quantitative treatment of the data.....	338
6.9.2 Yes/No interrogatives.....	339

6.9.3 WH interrogatives.....	342
6.10 Variability in interrogative comment clauses.....	344
6.11 Concluding remarks.....	346
SUMMARY AND CONCLUSION.....	347
APPENDIX 1.....	352
APPENDIX 2.....	354
APPENDIX 3.....	358
REFERENCES.....	361

ABSTRACT

In this thesis, a set of defining properties of grammatical variables is proposed, taking particular account of the precise extent to which variants should be required to be equivalent, semantically and pragmatically. These principles are then applied in a variationist analysis of negation and interrogation in spoken French, with data from a corpus from the Somme, northern France. Computer-assisted data-handling techniques are employed, notably the Oxford Concordance Program.

For the (ne) variable, a large proportion of the data is analysed in terms of preformed sequences, which strongly favour the omission of the negative particle. There is also evidence that age is the most important extra-linguistic constraint, but this is interpreted as being a case of age-grading rather than of change in progress, as has sometimes been supposed. It is suggested that the negative particle has all but disappeared from northern French vernacular styles.

To check the pragmatic equivalence of variant interrogative structures, a taxonomy of communicative functions is set up, drawing from research on speech acts, conversational structure and communicative grammar. The interrogatives in the corpus are then classified in terms of this taxonomy.

In Yes/No interrogatives, clitic inversion is found to be completely absent from the corpus, and the minority use of est-ce que is shown to be motivated by pragmatic and socio-pragmatic factors, ie it is often used when the speaker does not expect an

answer from the addressee, or to encode politeness.

WH interrogatives constitute one of the most complex grammatical variables studied so far, with six variant structures occurring in the corpus, and the choice among them being constrained by a large number of linguistic, discoursal and pragmatic factors. In order to take account of the unacceptability of some structures in certain contexts, the notion of "semi-variable" tokens is proposed. This is reflected in the method of calculating each variant's relative frequencies, as these exclude those contexts where the variant would be unacceptable, or non-equivalent to the structure actually used. The productive use of clitic inversion in the corpus is seen to be minimal, and the choice of the WH-final structure (as opposed to a WH-fronted one) is shown to be motivated overwhelmingly by discoursal considerations. The female informants are found to favour the est-ce que structure (partly, again, for politeness), whereas the male speakers use rather more of a non-standard variant.

ACKNOWLEDGEMENTS

I should like to express my thanks to the following:

- Lesley Milroy, who has supervised the later stages of this research, especially for her invaluable comments on the draft chapters; Peter Trudgill, who supervised the project in the earlier stages, for his advice and support;
- the speakers whose language is described here, and all those who made the fieldwork possible, especially the directors and staff of the camps for their warm welcome;
- Geoff Hare and Tony Lodge, for their unstinting support and practical help in many ways; the former, in particular, for his guidance on French socio-professional categories and social classes, and the latter for making possible the computerisation of the corpus; John Green, for timely advice and encouragement;
- my French colleagues and former colleagues, who lent me their ears and their intuitions on many occasions: Sophie Baillarguet, Jacqueline Hollaar, Veronique Leygue, Catherine Ogus, Anne Soulié, and, especially, Thérèse Butler;
- Judy Baker, Judith Hunter and John Law for their advice and assistance in handling "la machine infernale" and the Oxford Concordance Program;
- the Inter-Library Loans staff of the University Library;
- Patricia Sears for typing earlier drafts of chapters 1 and 2, and for her help in many other ways; Carmel Mayfield and Tessa Havelock for typing parts of chapter 4;

- many other colleagues, in this country and abroad, who have discussed this research with me, in person or by correspondence, and who have sent me copies of their own work: especially, William Ashby, Adrian Battye, Michel Blanc, Jenny Cheshire, Viv Edwards, Martin Harris, Marie-Anne Hintze, Vi Howe, Beatriz Lavandera, Bob Le Page, Geoffrey Leech, Jim Milroy, Gillian Sankoff, Pierrette Thibault and Sebastião Votre.
- my students (especially those on my sociolinguistics courses), from whom I have learned a great deal;
- Mike and Thérèse Butler, Graham McGregor, Ian Mason and Miranda Stewart, for many enjoyable hours of conversation, during which I have picked their brains;
- Helena and Jonathan Sears for sharing with me their penetrating insights into the nature of language; Patricia and Christopher Sears for buying a second-hand car from an ex-fieldworker;
- and finally, all my family, but especially my parents, for their unfailing support in every way.

INTRODUCTION

This thesis sets out to investigate the variability in negation and interrogation in a corpus of French, in order to reveal the linguistic and other constraints affecting the variation, as well as any interpersonal differences among the speakers whose speech is studied here. Negation and interrogation, although related and at times interconnecting areas of grammar, nevertheless pose contrasting problems in a variationist analysis, and differ also in the relative complexity of their constraining factors, and in the richness of their social differentiation - at least in a corpus of a relatively modest scope. The initial motivation for this research lay in the belief that the teaching of French as a foreign language could benefit from more sociolinguistic information about the contemporary spoken language, but it is also hoped that the present study will be of interest to linguists who do not always have ready access to this type of information. Even today, it is probably fair to say that spoken French (especially spoken metropolitan French) is a relatively neglected area of linguistic inquiry.

In chapter 1 we review the problems involved in the quantitative analysis of grammatical variability, and, in particular, address the arguments for and against the extension of the sociolinguistic variable to this level of language.

Chapter 2 describes the corpus of spoken French upon which the study is based: how, and from whom, the recordings were obtained, the speakers' socio-geographical and dialectal background, and the characteristics of the interviews recorded. Finally, we

outline the transcription process, the computerisation of the corpus, and the use of the Oxford Concordance Program to extract the data required for the present study.

The variable omission of the negative particle ne is the subject-matter of chapter 3. The geographical and historical progress of this variable is sketched briefly, before we consider certain linguistic factors which affect the likelihood of ne being present or absent. We also examine the extent to which the individuals and groups of speakers in the study differ in their use of this variable.

The last three chapters of this thesis are concerned with variability in direct interrogation. The various structures to be found in spoken French are described in chapter 4, followed by a brief outline of their socio-stylistic evaluation and their history since Old French. We then review a number of previous quantitative studies in this area, and make an initial comparison with the data of the present study. The discoursal/pragmatic sources of the interrogatives are discussed, together with a consideration of the problems of equivalence and acceptability as these affect a variationist analysis of this area.

Chapter 5 is devoted to examining the communicative functions of the interrogatives in the corpus. Firstly a taxonomy of the relevant communicative functions is sketched out, based on a set of pragmatic and semantic features, and then each of the communicative functions realised by interrogatives in the corpus is illustrated and discussed.

Finally, in chapter 6 we consider the complex web of constraining factors affecting the speaker's choice of interrogative structure, linguistic, pragmatic and discoursal, before looking at the evidence for social differentiation in this area. The chapter closes with a brief discussion of variability in interrogative comment clauses (eg tu vois?), which had been excluded from the main analysis.

CHAPTER 1: GRAMMATICAL VARIABILITY

1.1 Introduction

Sociolinguists have been analysing grammatical variability since at least 1968 (Labov et al., 1968), but the enterprise has proved a good deal more controversial than the analyses of phonological variation which established the methodological and theoretical principles of this particular current of sociolinguistic research (Labov, 1966). The beginning of the debate on the problems involved can be said to date from Lavandera's 1978 article, 'Where does the sociolinguistic variable stop?', in which she challenged Sankoff's assertion in 1973 that the extension of the sociolinguistic variable from phonology to syntax was not a conceptually difficult step to take. The arguments on both sides will be considered presently, but it may be useful to clarify two points from the outset.

Firstly, it has been suggested that the sociolinguistic variable is simply an analytical tool, rather than a concept forming part of a theory of language, and that therefore the precise definition of variables is unimportant (Hudson, 1980:157). This view is rejected unequivocally here. The variable is indeed part of a sociolinguistic theory of language, and it interlocks with other concepts which form part of this theory. Clusters of sociolinguistic variables serve to distinguish related varieties from each other, and the attitudes of speakers towards the variants are one of the defining characteristics of the speech community (Labov, 1966:7). Moreover, the precise definition of a variable and its variants is of fundamental importance in variationist methodology. However, it is also important to

distinguish between variables which have yet to be fully investigated, and those which have already been shown to be socially differentiated in some speech community. It may be useful to make the distinction explicit by referring to the former as alternations, rather than variables as such.

A second, but less contentious, point which needs to be made here is that throughout the present study the broader term grammar will be used in preference to (but to include) syntax. This is simply because the latter is sometimes assumed to cover only the connections between elements of sentence structure, and might therefore be taken to exclude some of the phenomena which we shall need to discuss occasionally, such as pronouns.

1.2 A simple typology of grammatical alternations

In order to discuss some of the problems associated with the analysis of grammatical variability, it will be helpful to be able to refer to different types of grammatical variable, or "alternations", to use a more neutral term. One way of classifying these alternations is into three types, defined purely on formal grounds: omissible items, alternating items and alternating structures. The illustrative examples will be from spoken French, and it remains to be seen whether this tripartite typology could accommodate grammatical alternations in other varieties, or even a wider range from French itself. No such claims are being made here. Some, though not all, of these alternations have been analysed as linguistic variables, in France and/or Quebec, and have been found to be socially differentiated. All of them are likely to be constrained by

linguistic factors, and, in some cases, by discorsal and pragmatic ones too. In certain instances, these may indeed prove to be the only influences on the variability, if research reveals that there is no social or stylistic differentiation. Their inclusion here is not intended to imply that they have already been shown to be socio-stylistically significant.

The first category involves the presence or absence of a grammatical item, and is exemplified by the negative particle ne, as in (1), and by the empty or "dummy" subject pronoun il when preceding certain impersonal verbs, notably y avoir and falloir. The presence and absence, in rapid succession, of ne and il is exemplified in (1) and (2), respectively:

- (1) j' Ø étais pas très partie prenante ça ce n'est qu'après /
(2422)
(2) il faut qu' Ø y en ait un qui soit capable de prendre sur lui son agressivité ... (2115)

(Throughout the present study, examples taken from the corpus to be described in chapter 2 will be followed by a reference number, which serves to locate the extract in the corpus.) A third, though undoubtedly more controversial, candidate for membership of this category, is a third person subject pronoun, coreferential with a preceding NP, within or "dislocated" from the same clause:

- (3) donc euh - a priori la révolution c'est quelque chose de néfaste enfin ou des événements comme ceux-là / Ø sont néfastes et on les rejète en bloc. / (2420)

The second type of grammatical variable consists of alternations between two or more items from a sub-system of the grammar, such as pronouns, verb paradigms or prepositions. Examples of this type seem, in French at least, to be more numerous than the other

two. In (4) the indefinite personal subject pronoun is realised as t (the elided form of tu), but as on in (5):

- (4) mais t as aussi des jeunes qui arrivaient qui voulaient faire un stage d'animateur qu'étaient jamais allés en colo. / (2513)
(5) s'i y a un canoë d cassé j dois êt - si je s si jamais je peux pas le faire / j dois être capable de dire comment on répare un canoë là ... (2586)

A second case also involves on, here as the first person plural pronoun, in alternation with nous:

- (6) alors à ce moment-là on explique aux parents "voilà / euh nous nous ne vérifions ni euh comment - ce que écrivent les enfants / et on ne / on - ne vérifie pas combien de lettres ils écrivent. / (2321)

Thirdly, the simple future tense alternates with the periphrastic future form, aller + infinitive, to refer to future time:

- (7) et j pense que ça va plus durer quand même euh très longtemps hein. / (21378)
(8) mais je crois que on fera encore euh [LAUGH] quelques années. / (2645)

Another, less well-known, alternation involving verb paradigms is the expression of past punctual meaning by means of the choice between the perfect and the pluperfect tenses, as in (9) and (10) respectively:

- (9) et j'ai fait mon bilan de form - oui - mon bilan de formation oui / (4143)
(10) mais euh je l'avais d'ailleurs marqué dans mon bilan de formation ... (4169)

The third type of alternation is between two or more structures, which may differ not only in word order, but also in the grammatical items they involve and their accompanying prosodies. An example which does consist of just a difference in word order is the alternation between pre- and post-nominal position for many adjectives:

- (11) et malgré tout faire un léger bénéfice. / (3470)
(12) et bon ils ont établi les prix avec un bénéfice euh léger. / (3472)

A second example involves the position, in certain subordinate WH

clauses, of the subject NP: pre-verbal in (13), but post-verbal, or inverted, in (14):

(13) les trousses de pharmacie / que les groupes emmènent. /
(1117)

(14) pour euh revenir sur les problèmes relationnels euh le petit
coup que - nous a fait euh Valérie hier - (2132)

A third example concerns the choice in WH interrogative clauses from a range of structures, including one where the WH word is in post-verbal position, as in (15), and others where it is fronted and there is either no other marker of interrogativity, or inversion of the subject clitic, or one of various "reinforcers" of the WH word, as in (18-20), respectively:

(15) oui on en était où là? / (4671)

(16) ou j'en étais là? / (28569)

(17) ou en étais-je? / (21608)

(18) ou est-ce qu'il était? / (2410)

(19) qui c'est qui voulait une poire? / (said by a non-informant
at 2417)

(20) quel âge que t as? / (example noted, but not recorded)

Whereas the three categories of variables proposed here are based solely on the formal characteristics of the alternating forms, the typology proposed by Romaine (1984:418-20) includes as criteria for the different types, their linguistic and social/stylistic conditioning factors. The first two types in Romaine's typology are "pure" phonological and morphophonemic, and are not of direct concern to us here. Her morphosyntactic type appears to correspond to what we have called omissible items (p.420), and presumably her morpholexical type is equivalent to our alternating items, although this is not entirely clear. On the other hand, what Romaine calls "pure" syntactic variables no doubt refers to the same phenomena that we have termed alternating structures. As an example of this last type, Romaine cites the "agentless passive", investigated by Weiner & Labov

(1983), and this example led her to conclude that "pure" syntactic variables are not conditioned by social or stylistic factors. However, in order for such generalisations to have some validity, they clearly need to be based on more than one or two examples, and indeed Winford has pointed out that in creole continua there are many "pure" syntactic variables which are involved in social differentiation (1984:272). We shall not attempt here to refine or develop the typologies put forward by Romaine and Winford, but merely add that ultimately such typologies may need to take account of other factors, such as the subjective evaluation of the forms or structures involved, and, perhaps more importantly, they may also need to be based on a more detailed grammatical analysis than the one we have adopted here. A consequence of this latter development might well be that certain alternations would be reclassified. For example, in considering the phrase structure of the alternating structures in (3), we might wish to consider the first NP to be a "topic" rather than a straightforward subject (with c' being treated as an agreement marker). In that case, the alternation would be better considered as involving two different structures rather than simply an omissible item.

1.3 The problem of the infrequency of certain forms

One difficulty which is encountered in attempting to quantify grammatical variability, which is not generally a problem at the phonological level, is the rarity of certain grammatical forms. The reason for this, as has been pointed out by Milroy (1987:143-4) for example, is clearly that phonological systems

involve a smaller number of forms than do grammatical systems, and therefore the phonological items involved are virtually certain to occur reasonably frequently in any stretch of discourse. (In fact, the infrequency of tokens can be a problem with phonological variables too, when they occur in small sets of lexical items: for example the intervocalic (th) variable studied by Milroy, 1980:119.) Furthermore, since grammatical variables can generally be said to involve meaning-bearing units, the frequency of certain grammatical forms will depend very much on what the speaker is talking about. Omissible items, somewhat paradoxically, seem to be reasonably frequent in discourse: at least, this is the case with the three mentioned in 1.2, as well as the complementiser que in Montreal French. But alternating items and alternating structures may be seriously affected by this problem, and it is in the study of certain verb forms that it has proved most acute. Wolfram & Christian (1976), for example, obtained only 65 occurrences of the perfective aspect marker done in Appalachian English (according to Chambers & Trudgill, 1980:90), and Harris (1984:316) reports on the elusiveness of the "hot news" past verb form in his Hiberno-English data, especially in interviews as opposed to more spontaneous interaction.

One partial solution to this problem is described by Rickford (1983:101). In his study of Black English bin, he supplemented his recordings by noting down on small record cards every occurrence which he heard during his participant observation fieldwork. The beginning of a second solution to the problem lies in an increased understanding of the communicative function of

the grammatical forms in question, as was pointed out by Labov (1972a:204-5). This is, in any case, important for a satisfactory analysis of the variable. Once such an understanding has been acquired, it is possible to attempt to elicit the forms by creating a communicative need for the informant to use them. In interviews this is most simply done by gearing the questions to a particular function, as Lavandera demonstrated extremely effectively in her study of variation in the verb forms used in conditional clauses in Spanish: she obtained 1,418 tokens from her 105 informants, or about 14 each (Lavandera, 1975). However, there are other techniques which can be used, even within the confines of an interview, as we shall see in chapters 2 and 4, with regard to interrogatives.

The communicative functions of grammatical forms has, for some time, been a central concern for language teachers and speech therapists alike, for very practical reasons in both cases, and both these groups have developed techniques for eliciting certain grammatical forms by creating communicative needs for them (cf Klippel, 1984, for example). Some of these techniques could probably be adapted profitably by sociolinguists for use in their fieldwork. It should be mentioned, however, that whilst the functions of some items and structures may be fairly transparent (eg the future tense, interrogatives), those of certain others may be far more difficult to specify (eg the subordinate WH clauses in which there is variation in the position of the subject NP, as in (13) and (14) above).

1.4 The case against grammatical variables

1.4.1 Grammatical variables are inextricably linked to variable rules

Let us now move on to consider the principal objections which have been raised to the concept of the grammatical variable. Firstly, it has been argued that the latter is inextricably linked to the concept of the variable rule, which, in turn, is rejected as unhelpful, or even incoherent (cf García, 1985:193; Harris, 1984:310-1; Romaine, 1984:414).

A number of the French alternations referred to in 1.2 either have been, or could be, described in terms of generative rules. In particular, the omissible items ne and impersonal il might be handled by deletion rules, whereas the alternating structures involving attributive adjectives and the subject NP in some WH clauses would be candidates for treatment as movement rules, or transformations. Indeed the second of these two alternations was accounted for by Kayne (1972) and Kayne & Pollock (1978) by the optional transformation which they called stylistic inversion.

There are other alternations, however, which either would require several such rules, or cannot be described in terms of rules at all. The several different WH interrogative structures available in French, for instance, which are widely recognised as constituting a sociolinguistic variable, have to be handled, in transformational theory, by a number of rules at various levels of the grammar (Lefebvre & Maisonneuve, 1982:201-4). Even more

starkly, it seems unlikely that even the most imaginative transformationalist would nowadays wish to relate by rule the two or more grammatical items involved in the second type of alternation mentioned in 1.2, such as nous and on used as first person plural subjects.

Since many grammatical alternations cannot be described in terms of (variable) rules, it is clear that the concept of the linguistic variable is in no way dependent on that of the variable rule. Moreover, many studies of phonological and grammatical variability have been carried out without having recourse to variable rules (cf especially, the studies from various research projects reported in Trudgill, (ed) 1978, and Romaine, (ed) 1982). In addition, it is quite clear that using the VARBRUL computer program for the statistical analysis of variable data does not imply any commitment to the variable rule concept. Many quantitative studies have used this program without even implying that any rule is involved in the alternation, let alone actually drawing up a variable rule (eg Laberge, 1977; Sankoff & Thibault, 1977; Thibault, 1983).

1.4.2 Grammatical variables investigated have not exhibited social differentiation (or stylistic variation)

A second objection to the concept of grammatical variables is basically an empirical one: that few, if any, grammatical variables have in fact been shown to be differentiated socially (let alone stylistically), and therefore they do not qualify as sociolinguistic variables. Lavandera (1978:172-3) pointed out that Sankoff (1973) had demonstrated only linguistic constraints

on several grammatical variables: the position of the Tok Pisin future marker bai, the omissibility of the complementiser que and the alternation between on and tu or vous as indefinite subject pronouns in Montreal French. But the reasons for this seem to have been simply (a) that Sankoff wished to show that a variationist analysis of grammatical features could reveal interesting linguistic constraints (which had not previously been suspected), and (b) that, certainly in the case of the on - tu/vous variable, the analysis was still in its early stages, and evidence for social differentiation was not yet available. Laberge (1977, 1983) later showed quite clearly that, not only was the indefinite subject pronoun a sociolinguistic variable, but also that it was a change in progress. Warren (1988) has similarly reported very clear social conditioning of the presence or absence of que in Montreal. There is, then, ample evidence that at least some variables of the omissible item and of the alternating items types are socially differentiated (cf also Kemp, 1981, for an overview of several variables in Montreal French).

What has proved more elusive, however, is a clear case of socially differentiated alternating structures, or what Romaine (1984) called "pure" syntactic variables. One candidate was the English agentless clauses involving an alternation between agentless passives and actives with a generalised subject, as in the following:

- (21) The liquor closet got broken into.
- (22) They broke into the liquor closet. (Weiner & Labov, 1983:341)

But Weiner & Labov found that the sex of the speaker had no

effect on the choice of structure, and that age and social class exerted only a very modest influence. The principal conditioning factor was a discorsal one: there was a strong tendency to use the passive when the NP to be fronted had already occurred at the start of a preceding clause.

A second candidate was the alternation between "reduced" and "complete" coordinate structures in Montreal French:

- (23) Puis eux-autres ils font des couvertes tu sais puis ils font
des bijoux.
(24) Il parle l'anglais puis Ø Ø le français. (Emirikian &
Sankoff, 1979:79)

In their analysis, Emirikian & Sankoff found that there was a slight tendency for working-class speakers to use complete forms rather more than middle-class speakers. But far more significant was the effect of linguistic constraints, such as the type of coordinate structure involved.

Both of these studies, then, concluded that social factors do not greatly affect the speaker's choice between alternating structures. It might well be supposed that many similar alternations involving structures which have been traditionally recognised as paraphrases and related by transformations (eg extraposition, clefting) would yield similar results, ie that there would be substantial evidence of linguistic, discorsal and/or pragmatic constraints but only the faintest hint of social differentiation. It seems likely, then, that socially differentiated alternating structures are less common than the two other types of grammatical variable. This observation cannot, however, be taken as evidence that such variables do not exist,

since only a relatively small number of different varieties have, as yet, undergone a thorough variationist analysis of many areas of their grammar: Montreal French (Kemp, 1981; Sankoff, 1980), American Black English Vernacular (Labov, et al., 1968), non-standard Reading English (Cheshire, 1982), British Black English (Edwards, 1986). Moreover, in chapter 4 evidence will be provided from several quantitative studies to show conclusively that Yes/No and WH interrogatives in French are socially and stylistically differentiated.

1.4.3 Grammatical variables involve meaning-bearing units

Huspek (1986) has argued that even phonological variables involve meaning, albeit of a symbolic rather than a conceptual kind. However, such meaning is highly subjective in nature, and clearly does not affect either the propositional content or the pragmatic force of an utterance. It is generally agreed that the units involved in phonological variation are not bearers of meaning at all, in any reasonably narrow sense of the word. Indeed, it could be claimed that phonological variants are not "different ways of saying the same thing", because, in themselves, they do not say, or mean, anything at all (cf Winford, 1984:269). In contrast, Lavandera (1978:175) and Romaine (1984:411) have argued that one of the problems with extending the domain of the sociolinguistic variable beyond phonology is that the units involved are, necessarily, meaning-bearing.

Although this point is basically valid, it is not quite as straightforward as it might seem at first sight. Most cases of

alternating items do indeed involve meaningful forms, for example the alternations between different pronouns or verb paradigms which we saw in 1.2. However, not all grammatical items are equally meaningful: for example, à and de, which alternate as infinitival complementisers following certain verbs (such as continuer), have a purely grammatical function, without, apparently, contributing anything to the propositional content of an utterance, or indeed any other aspect of its meaning.

The issue is even less clear-cut when we come to the two other types of grammatical alternations. Firstly, items such as ne and impersonal il are only omissible because they are communicatively redundant, ie convey no meaning which is not also expressed by some other element in the utterance. It could be argued, therefore, that they bear little, if any, meaning at all. The third type of grammatical alternation involves structures (not utterances as such), which could be represented in terms of the appropriate category symbols: N + Adj alternating with Adj + N, for example. Such abstract structures can only be said to bear very limited meaning in themselves in comparison with the meaning that is conveyed by the exponents of these structures, ie utterances in which the skeletal structures are filled out with lexical items.

1.4.4 The concept of grammatical variables violates the "one form: one meaning" axiom

A third, and possibly more serious, objection to grammatical variables is that the assumption upon which they are based, that two or more items or structures can mean the same thing, is in

conflict with a fundamental principle of language, viz that one form corresponds to one meaning. This view has been most cogently argued for by Bolinger, who has stated his belief that: "(...) in syntax there is no such thing as two different surface structures with the same deep structure (that is, with the same meaning)" (1977:4). Bolinger was in fact arguing against a tendency within formal linguistics to assume too readily a paraphrase relationship between pairs of sentences which were then to be related by transformations. However, his point needs to be answered also by variationists working on syntax, since they too make the same assumption in analysing pairs of structures as variants of a variable. Bolinger backed up his claim with evidence from several alternations in English, including the omissibility of the complementiser that, as in (25), and the variable position of the verbal particle in (26) and (27):

(25) I thought (that) you might need some help.

(26) They hauled in the lines but didn't get them in.

(27) *They hauled the lines in but didn't get them in. (Bolinger, 1977:11, 17)

Although his data appear to be confected rather than taken from naturally occurring spoken language, Bolinger nevertheless argues convincingly that in certain circumstances, pairs of utterances involving these alternations do not mean the same thing, or even are not equally acceptable (cf (26) and (27)).

What he does not demonstrate, however, is that such differences in meaning or acceptability affect all, or even many, instances of such alternations. Indeed, perhaps one reason why his examples are so intriguing is precisely because they are somewhat unusual. He himself admits that in some cases the differences in meaning

or use are either very infrequent or extremely subtle (1977:9, 13). Incidentally, both that-deletion and particle movement, as in (25) and (26-27) respectively, have been analysed as variables in Philadelphia by Kroch & Small (1978), who found that not only were the variants given different values in "prescriptive judgement tests", but they were also differentiated socially.

Referring again to the one form: one meaning axiom, Bolinger has also stated "If a language permits a contrast in form to survive, it ought to be for a purpose." (1977:19). Whether this ought to be the case is surely a rather subjective matter, and an appropriate object of investigation only for a prescriptive approach to language. What can be properly explored within linguistics is whether it is in fact the case. Synchronic and diachronic evidence from phonology would suggest that Bolinger's statement is too general. It is well known that many, possibly most, phonological variables are stable as opposed to being changes in progress, and certain forms are known to have been in variation for centuries, eg English (h) and (ng) (Strang, 1970:79-81) and French (l) (Sankoff & Cedergren, 1971:64). Further evidence against the general validity of the one form: one meaning/function axiom is provided by Milroy & Milroy (1985a:349) who have reported that there is considerable overlap in the realisations of the vowels /a/ and /ɛ/ in Belfast vernacular. The axiom is quite probably more valid in standard varieties than in non-standard ones, but there is strong evidence that even in standard languages it is by no means respected consistently. To the extent that the axiom has validity in standard languages, it is, in some measure at least, an artifact

of codification, and if it is held up as a rational principle which languages can be expected to follow, then it is little more than a tenet of prescriptivist ideology.

1.4.5 The meaning of a grammatical form is the reflection of all its uses

Sankoff (1988:153-4) has referred to a claim (unattributed) that all the semantic and pragmatic nuances acquired by an item or structure in the various contexts in which it occurs are retained whenever the item or structure is used. Since it is almost certain that no two items or structures have identical contextual distributions, it follows that they cannot be semantically equivalent either.

Sankoff does not refer to the source of this argument, but it appears that a similar line of thinking underlies the approach of the Groupe Aixois de Recherches en Syntaxe, who have stated their belief that pairs of utterances such as the following:

- (28) ... pour qu'il tombe pas.
- (29) ... pour pas qu'il tombe.

involve semantic differences which are more than mere nuances (1980:307). They point out that in certain contexts one of the variants is excluded, for example pour que ...pas in the following:

- (30) Je fais ça pour pas que plus on l'énervé moins il parle.

And they also comment that, in Quebec, (31) implies a positive response, whereas (32) is neutral:

- (31) Tu passeras ce soir à la maison?
- (32) Passeras-tu ce soir à la maison?

(They add, however, that this distinction does not seem to hold for other French speakers.) Apart from these two examples, however, the GARS do not elaborate on the precise nature of these differences, and simply state that they accept that a consequence of their position is to recognise that since the different structures are often associated with particular social groups, it follows that different groups are meaning different things.

The two examples discussed above do not, in reality, pose any obstacle to a variationist analysis, as they would simply be treated as categorical tokens in which only one variant is possible, and would be excluded from the quantification of variable tokens.

As for the proposition that whenever a form or structure is used, it somehow conveys the meanings from all the contexts in which it can be used, Sankoff (1988:153-4) responds that "distinctions in referential value or grammatical function among different surface forms can be neutralized in discourse". We could go a step further and say that forms and structures derive their meaning or function only from their use in discourse, and the idealised meaning or function which is imagined when considering forms or structures out of context is nothing but the imperfect recall of the totality of their occurrences in discourse.

1.4.6 Cases of structural mismatch cannot be analysed in terms of grammatical variables

Let us now consider the final objection which has been raised against the use of the linguistic variable at the level of grammar: that the notion is inapplicable when there is a structural mismatch between grammatical sub-systems of two varieties in contact. Harris (1984) has discussed certain verb forms in Hiberno-English vernacular (HE) expressing four distinct temporal/aspectual meanings, which in Standard English (StE) would all be expressed by the perfect tense. He demonstrates quite clearly that, as the HE tense system has more forms than StE in this area, there is a "structural mismatch" between the two varieties, in the sense that one form in StE maps onto (or fulfils the functions of) several forms in HE. Harris concludes that the sociolinguistic variable cannot be used to analyse the variation between the HE and StE forms (p.314).

It would be unwise to presume to discuss the semantic distinctions made in a variety of which one has very little knowledge, especially in an area as subtle as tense and aspect, but a certain number of general points can be made. Firstly, the fact that two forms can be analysed as belonging to two different systems does not, in itself, mean that they cannot be used by speakers to express the same meaning. Just as, on the phonological level, the fact that two segments belong to the different, and frequently "mismatching", systems of distinct varieties does not prevent us from recognising that they may have the same function in a given lexical set, and can be quantified

as variants of a variable.

Secondly, Harris appears to assume that a variationist analysis would require the setting up of a single variable in this area, effectively ignoring the structural mismatch (cf also Winford, 1984, who discusses a similar problem involving the tense systems of varieties within a creole continuum). But if it is true that any temporal-aspectual meaning in this area which can be expressed in HE can also be expressed in StE (and Harris does not appear to contradict this), then the variation in verb forms could indeed be quantified by establishing a separate variable for each of the four meanings which are identified. For each variable, the variants would, of course, be the HE and StE verb forms which are used to express that particular meaning. There may well be methodological problems in adopting such an approach: Harris mentions the relative infrequency of tokens, and the task of assigning the appropriate semantic category to each token and the need to take account of semantic, discoursal and pragmatic factors (p.316). However, these do not seem to be insurmountable difficulties, and they certainly do not constitute fundamental problems of principle for a variationist analysis.

What emerges from this discussion is that several apparently theoretical objections to the use of the variable in analysing grammatical variation can be reinterpreted as methodological problems. In particular, as Sankoff (1988:154) has also pointed out, the meaning of each token must be inferred, and a systematic procedure has to be arrived at for handling cases where the meaning cannot be so determined, for example, because of

ambiguity. This means that the analysis of grammatical variability may well be more labour-intensive than that of phonological phenomena, and it may also mean that the analyst needs to have been present when the speech was recorded.

1.5 The case for grammatical variables

1.5.1 Evidence from other- and self-repairs in conversation

Let us now move on to consider some of the arguments which serve to justify the extension of the sociolinguistic variable to grammatical and syntactic phenomena. An area which provides some evidence that speakers themselves recognise two forms or structures as semantically and pragmatically equivalent is that of conversational repairs. Speakers may repair a previous utterance of theirs because, for one reason or another, it has not been understood by their addressee. One such incident, noted but not recorded, occurred during a residential course in France which I attended. During conversation over the breakfast table one morning a woman asked a younger male the following question:

(33) Tu es d'où?

Although her addressee said nothing, it was apparent from his perplexed expression that he had failed to understand what she had said. The woman provided an identical repetition of her question, but was again not understood. There is, of course, more than one possible explanation for this, but the one which seemed most likely at the time was that the addressee had failed even to identify the utterance as a question (possibly because it was not spoken with the rising intonation which would have unambiguously

characterised it as an interrogative), and had perhaps even interpreted the sequence of segments as the homophonous
(34) Tu es doux.

which would certainly not have been a relevant contribution to the conversation at the time, even if it were a truthful one. Finally, the woman switched to an alternative WH interrogative structure:

(35) D'où tu es?

and was, to general relief, at last understood by her addressee. It seems clear that, in this situation at least, the two structures provided different ways of saying precisely the same thing. Very similar incidents can be heard almost daily on French radio phone-in programmes, when the animateur poses the following (highly predictable) question to a listener:

(36) Vous nous appelez d'où?

Another example, recorded just before the beginning of an interview with one of the informants in this study, shows a similar switch from the WH-final structure to a WH-fronted one:

(37) child: (tu veux que j) - j peux en faire combien? /
B: (eh ben) / pourquoi "combien" ? /
child: combien j peux en faire? /
B: ben - avec un t en as pas assez? /

In this extract, B seems to have understood the child's first question, but asks why the child should want to make more than one of the objects she was making at the time. The child, however, appears to interpret B's first query as signalling that he has not understood her question, and she therefore provides the paraphrase of it.

In other cases, speakers self-repair spontaneously, perhaps motivated by a feeling that a form they have just used is

non-standard, or inappropriate in some way, and replace it with a standard, or other, equivalent:

(38) tout c qui est mis euh comptabilité c'est Paul qui l fait
hein / parce qu'on pense qu'i faut pas être euh deux sur la
comptabilité sinon tu / on sait plus qui / problème des sorties
(et) les rentrées faut qu y ait une personne quoi. / (22210)

Here the speaker switches from the non-standard tu to the standard on as the indefinite subject pronoun. Ironically, however, on is somewhat ambiguous in this context, as it could conceivably have first person plural reference, as in the previous occurrence at the beginning of the same line. However, another example, from a child, shows that not all cases of self-repair involve a switch from non-standard to standard variants:

(39) Alors, qui est-ce, qui c'est qui gagne? (Soll, 1982:50)

In this instance, the speaker switches mid-question from the standard structure with est-ce (qui) to the non-standard alternative with c'est qui.

In some cases, of course, a repair triggered by a misunderstanding may entail a switch to a form or structure which we would probably not wish to consider as equivalent to the first form/structure used. In the following extract, for example, it is difficult to know whether the switch from present to imperfect tense was motivated by a wish simply to make the utterance a little more polite, or by a desire to actually modify the temporal reference of the verb:

(40) B: et qu'est-ce que tu veux me faire dire pour ça? /
A: hein? /
B: qu'est-ce que tu voulais me faire dire? /
A: rien / (491820)

1.5.2 Evidence from correction, repetition, interpretation and translation

Similar evidence can be gleaned from incidents when speakers correct another speaker (typically a child), when they feel the item or structure used was inappropriate. Ervin-Tripp (1973:41) has given a nice example of this, involving the non-standard English completive aspect marker done:

(41) 6 year-old child: She done ate up all of my potato chips.
Mother: "Done ate!" She has ... have ate up all my potato chips.

The mother here apparently believes that has/have ate is a more correct way of saying the same thing as done ate. It is interesting to note that, in their study of Appalachian speech, Wolfram & Christian (1976) did not find it possible to treat done as a variant of a variable, because it was not thought to be "in alternation with other forms that could be considered to be equivalent ways of doing or saying the same thing" (Chambers & Trudgill, 1980:90). Yet the mother in (41) surely provides us with firm evidence that done is equivalent to, and in variation with, at least one other form in some contexts, if not all.

This ability of speakers to "correct", or translate, from one variety to another has been tapped by several sociolinguists who have investigated the grammatical features of divergent, but related, dialects. Labov et al. (1968:311) had their Black adolescent informants repeat back various sentences, some of which had Standard English grammar and some the grammar of Black English Vernacular, as in the following well-known pair:

(42) There is no cat that can get in any coop.
(43) It ain't no cat can't get in no coop. (p.324)

Labov et al. reported that some of the Standard English sentences were understood and repeated back instantly, but in BEV rather than Standard English. This appears to demonstrate that the informants considered the BEV and Standard English sentences to be equivalent, even if they did not have sufficient active command of the latter's grammar to be able rapidly to repeat the sometimes complex sentences. In similar fashion, Rickford (1987:161-2) administered "correction tests" to his Guyanese Creole informants, involving translation of sentences both from Creole into English and vice versa. The informants' ability to perform this task provides us with further evidence that speakers are quite familiar with the concept and reality of equivalent grammatical forms.

These last two cases involved variation between forms or structures from two highly divergent varieties which happen to be in contact, but, whilst it seems very likely that such situations are particularly rich in grammatical variables, evidence of this kind can also be found in speech communities which do not apparently involve contact. Labov (1975:50-2) reported the "Jay-walking experiment" whereby passers-by were asked their opinion about a (fictitious) incident, and in replying, revealed their interpretation of one of the following four forms which had been used by their interlocutor:

- (44) He was arrested to test the law.
- (45) He was arrested.
- (46) He got arrested.
- (47) He got arrested to test the law.

The results of the experiment showed that many of the informants interpreted got arrested in a causative sense when it was followed by the purpose clause, as in (47), but that without the

purpose clause (as in (46)), far fewer people interpreted got arrested causatively - indeed only the same number as those who interpreted (45) in this way. This important finding shows that the get-passive and the be-passive can be interpreted (and presumably used) equivalently, ie that the distinction between them is neutralised in certain contexts. And secondly, it contradicts the suggestion, referred to in 1.4.5, that a given form retains an ideal, unique meaning in whatever contexts it occurs.

For a final, anecdotal, example we return to the variation in WH interrogatives in French. (I am indebted to Geoffrey Hare for this example.) A woman who was at Roissy airport to meet some friends on their arrival was overheard telling these friends what she had done when their arrival time was past and they had still not appeared: she went up to arriving passengers as they emerged from customs, and asked several of them either (48) or (49):

(48) Vous venez d'où?

(49) D'où venez-vous?

Although these utterances did not involve repair, or indeed, repetition to the same person (apparently), it seems clear that the contexts in which the two structures were used were identical in all respects, except for the interlocutor, and that the woman's meaning, in the widest sense, was the same.

1.5.3 Evidence from linguistic change

The evidence that has been reviewed so far has been intended to demonstrate that speakers themselves show an awareness of the

equivalence of certain alternating grammatical forms and structures in their ability to manipulate such alternations for various communicative purposes. However, the reality of equivalence does not, of course, prove that such alternants are socially or stylistically differentiated. Whether they are or not is clearly an empirical matter, and can only be known definitively by carrying out a variationist analysis on the alternation. As was mentioned in 1.4.2, a fairly substantial number of cases of sociolinguistic grammatical variables have already been documented in the literature, although those of the alternating structure type do appear to be rarer than omissible and alternating items.

Perhaps the most compelling piece of evidence that there must indeed be grammatical variability is the undeniable existence of grammatical change. With regard to linguistic changes over a long time-span, we do not, of course, have native-speaker informants to provide us with evidence of the semantic and pragmatic equivalence of parallel items or structures from, say, Old and contemporary French. However, there is probably a consensus among scholars (not to mention laypersons) that it is generally possible to translate from one historical variety to another, and that the only really intractable obstacles are lexical rather than grammatical. (According to Robins, 1980:23, even translation between two distinct languages is always possible, although usually difficult.) Translatability between two historically connected varieties would seem to imply that any grammatical items or structures which have been lost from the earlier variety have been replaced by new forms (or new uses of old forms) in the

later variety, to fulfil the corresponding functions. Now, since it seems inconceivable that a grammatical change could, even in the speech of one individual, run its course overnight, from innovation to completion, it must be concluded that all change is reflected in synchronic variation between two or more items or structures. This point is not, of course, new to historical linguists, and has been made explicitly before, by Rydén (1980:38), for example. The forms in alternation during the period of synchronic variation which reflects, and eventually effects, the change, will necessarily be socially (and stylistically) differentiated as the change spreads through the speech community, and these variants must be sufficiently equivalent to enable communicative needs to be satisfied while the change is in progress. These alternating forms will be, therefore, sociolinguistic grammatical variables. Moreover, it seems reasonable to assume that there are other grammatical variables which are diachronically stable, and that, just as in phonology, not all variables reflect ongoing change.

1.6 Approaches to the problem of equivalence

1.6.1 Historical data

The extent of equivalence of alternating forms, and how to deal with this, are problems which have faced linguists working in several different areas. Firstly, historical linguists have adopted various approaches to these problems, ranging from a variationist-type approach, where the meaning of each token is explored thoroughly, to a reliance on the criterion of the

syntactic distribution of forms, with little or no regard for differences of meaning between the alternants. Often, indeed, it is not explicitly stated that the form in question is equivalent to any other form, although this does seem usually to be implicit nevertheless. This appears to be the process which lies behind many of the remarks to be found in linguistic histories concerning general trends in a language in earlier periods, such as the following: "This use of aller + INFINITIVE is first found, sporadically, in the 14th century, and by the 17th century is frequent." (Price, 1971:240). Such an approach seems eminently justified with historical data in particular, since it is far more difficult to make judgements of acceptability and equivalence, firstly because there are no native-speaker informants to question, secondly because the quantity of observable language is far less than in the contemporary language, and thirdly because all the data are written rather than spoken.

1.6.2 Intonational variability

A second area in which such an approach to quantification has been found necessary is in the study of intonational variability. Local (1982:90) has demonstrated convincingly that there is significant social differentiation in the ratios of occurrences of one tone to another in Tyneside English. Moreover, he argues "It is not possible to delimit (in anything other than an ad hoc way) the functions of particular (co-occurrences of) intonational features in a given variety until we have information concerning the distribution and status of those features within those

varieties." (p.103). In this case, then, the justification for quantification without assuming equivalence lies in the fact that not enough is yet known about the meaning or function of the forms in question.

1.6.3 Non-variationist quantitative studies

A similar approach was adopted by many linguists, working before the development of the variationist paradigm, who studied the ratios or relative frequencies of alternating grammatical items or syntactic structures (cf for example, Gougenheim et al., 1964, on interrogatives, and Ross, 1974, on adjective position). In so far as such studies were seeking to discover the overall frequencies of such forms (often with important applied objectives in mind), the relative disregard of semantic or pragmatic differences could be justified, even though, with hindsight, it is possible to argue that this may have, to some extent, undermined the validity of the results. The value of such work is, nevertheless, considerable, not least to subsequent researchers, who can turn to these earlier studies for comparative evidence regarding linguistic change.

1.6.4 Weak complementarity

The most controversial approach to the problem of equivalence is probably that proposed by D Sankoff & Thibault (1979), involving the notion of "weak complementarity". They argue that their method is appropriate as a first step in the analysis of alternations which are differentially distributed in the speech

community, but where full semantic/pragmatic equivalence of the alternants has not been established, and/or where there is considerable uncertainty over which contexts are categorical (ie where only one variant is possible). In this approach, the analyst calculates the rate of occurrence of the two (or more) alternants per 1,000 lines of transcribed speech. Sankoff & Thibault stress that this procedure is a preliminary to a more rigorous analysis: "(...) once weak complementarity has been proven, the linguistic variable, or incipient variable, should be reanalyzed in terms of variant proportions." (pp.9-10). They acknowledge too that the method rests on the assumption that all speakers need to express certain basic "discourse functions" at similar rates. Clearly, this assumption would be invalid if the discourse genre and topics were not comparable for all speakers. But in Sankoff & Thibault's data they were indeed comparable, on the whole, since the speakers were all responding to the interview questionnaire of the Montreal survey (cf Sankoff & Sankoff, 1973:56-8). Although the weak complementarity approach does not give a definitive or complete picture of the variability, it has proved useful as a preliminary to a more rigorous analysis, and it provides a simple and rapid method of discovering whether forms are indeed socially differentiated.

The objectives of the present study, however, include the investigation of the factors which constrain the variability in interrogation and negation in French, and since some of these factors may well be discorsal or pragmatic, neither the weak complementarity approach nor the straightforward counting of frequencies will be appropriate. Instead, it will be necessary to

examine the context, meaning and function of each token, and to be aware of a wide range of contextual features which may be of relevance (cf the discussions by Harris, 1984:316, and Sankoff, 1988:154). It will then be possible to distinguish between categorical and variable tokens, and to proceed to the quantitative analysis of the latter in order to discover what factors constrain the variability, and whether there is any evidence of social differentiation.

1.7 Two problematic cases of "equivalence"

Before attempting to draw up a set of defining criteria for grammatical variables, it may be instructive to consider two groups of sentences, discussed by Romaine (1984), which are in some sense equivalent, but which would certainly not qualify as sociolinguistic variables. Although these examples are decontextualised sentences, as opposed to the utterances in context which form the data of actual studies of variability, they will help us to clarify our criteria for identifying grammatical variables. The first set of sentences is as follows:

- (50) Because it was cold, I closed the door.
- (51) I closed the door, because it was cold.
- (52) It was cold, therefore I closed the door.
- (53) *I closed the door, therefore it was cold. (Romaine, 1984:416)

Romaine suggests that we might wish to consider (50) and (51) as variants of the same syntactic variable, because the truth value of the two sentences is the same, and it is only the word order which varies. Most instances of word order variation, however, involve either words within a phrase, or of phrases within a clause, and it is significant to note that in (50) and (51) it is

the order of the two clauses with respect to each other which is varying. Romaine suggests, not unreasonably, that the choice between (50) and (51) is probably motivated, not by speaker characteristics such as age, sex and social class, but rather by a pragmatic constraint involving foregrounding and backgrounding. It does appear unlikely in the extreme that instances of clause-order variation would prove to be sociolinguistic variables, even though the evidence is at present limited, and could only be negative, in any case. But for that reason, we might well prefer to narrow the scope of the third type of alternation described in 1.2, by specifying that the alternating structures must be infra-clausal. Looking now at (52) in addition to (50) and (51), Romaine suggests that "Intuitively, one would like to say that these are variants of the same variable (...)" (p.416). But, again intuitively, it also seems improbable that such cases should turn out to be variants of sociolinguistic variables. Moreover, they would be excluded a priori if we required that exponents of the alternating structures in a variable of the third type should include the same lexical items: not only are therefore and because different lexical items, they also have quite different senses, at least in (51) and (52). In the former, because could be glossed as "the reason (for this) being", whereas in the latter, therefore means "which is the reason why". With regard to the ungrammatical (53), Romaine points out that a construction similar to this was possible in Old English (p.417). At first sight, this would appear to pose a serious challenge to the approach to syntactic variability which is being adopted here. It was argued in 1.5.3 that grammatical change takes place via socio-stylistically differentiated

grammatical variables, and this could be taken to imply that a grammatical item or structure only disappears from a variety if it is being replaced. But the intended meaning of therefore in (53) seems to be the same as that of because in (51), and so it would appear that one of the functions of the Old English equivalent of therefore is fulfilled in contemporary English by because. (Of course, (53) would be perfectly grammatical if the therefore-clause were intended as a post hoc deduction, rather than as the consequence of the first clause, as in (52).)

The second set of data discussed by Romaine is (54-58), which she considers to be all ways of broadly saying the same thing, ie of asking a person to close the window:

- (54) It's cold in here.
- (55) I'm cold.
- (56) Are you cold?
- (57) Would you close the window?
- (58) Close the window! (Romaine, 1984:422-3)

Once again, these appear intuitively to be a most unpromising set of candidates as variants of a sociolinguistic variable, but on what precise grounds can they be excluded? The first point is that the senses of (54-6) are different from each other and different from (57) and (58), since they include different lexical items. (57) and (58), although lexically closer than the others, nevertheless differ in that (58) does not have either the overt subject or the modal verb of (57). Secondly, there are the syntactic and semantic differences among the sentences: (54) and (55) are declaratives and statements, (56) and (57) are interrogatives and questions, while (58) is syntactically imperative and semantically directive (following the terminological distinctions of Huddleston, 1984:351). (54-57)

could all, of course, be said to be (more or less) indirect requests for action: we shall return to the problems posed by such indirect speech acts in chapter 5.

1.8 Criteria for identifying grammatical variables

To conclude this review of some of the problems inherent in the analysis of grammatical variability, let us now attempt to draw up a list of criteria for the identification of sociolinguistic grammatical variables.

(i) The parallel utterances which include realisations of the variants must have the same propositional content. (This is the same criterion which Jacobson, 1980:23, referred to as descriptive synonymy, and other terms which are sometimes used for the same type of meaning are conceptual, cognitive or denotative meaning, and sense.)

(ii) The parallel utterances must have the same communicative function. (Communicative function will be discussed further in chapter 5, as it is a particularly important aspect of the meaning of interrogative utterances. For the moment it can be taken to mean illocutionary force. This criterion is somewhat narrower than what Jacobson, 1980:26, among others, seems to have envisaged by "functional" comparability.)

(iii) The parallel utterances must have (or be able to have) the same lexical items as each other. (This is, in effect, a partial, formal reflex of the semantic criterion (i).)

(iv) The variants must be to some extent formally related. (This criterion does not apply to the omissible item type of variable, since in a structuralist framework we would not wish to say that \emptyset is formally related to any item at all. With regard to the two other types, the precise nature and extent of the relationship will depend upon the variable involved. It will often mean that the variants are of the same syntactic category, such as pronouns, Noun Phrases, subordinate clauses. But whether or not they are relatable in terms of the formalism of any particular grammatical theory is immaterial.)

(v) The variants must be socially differentiated within the speech community. (But stylistic differentiation is not a necessary criterion, just as not all phonological variables are "markers", cf Labov, 1972a:237. In practice, however, it might well be that most, if not all, grammatical variables will prove to be involved in style-shifting, since speakers could be expected to be more aware of the alternating forms involved than is sometimes the case in phonological variation: cf the suggestion, in Chambers & Trudgill, 1980:84-8, that markers are variables of which speakers are more aware, for various reasons.)

(vi) The variants must occur in the vernacular style of at least some speakers in the speech community. (This might well effectively rule out syntactic alternations where one of the variants occurs only in formal styles, such as in the alternation between extraposed and non-extraposed sentences, where the latter seem not to occur in the vernacular. There are, of course, many

such alternations where one of the structures is characteristic of the "elaborated" syntax of written language: cf Biber, 1988, for an exhaustive study of this kind of basically asocial variation.)

(vii) The alternation must be a (relatively) language-specific one, rather than a (near) universal. (Thus, an alternation between interrogative structures will be a possible sociolinguistic variable, since most languages seem to have alternatives in this area only if they have very clear functional motivations (cf Sadock & Zwicky, 1985). However, an alternation between active and passive sentences is most unlikely to prove to be socially significant, since both constructions are widespread in the world's languages (Keenan, 1985).)

A number of significant exclusions from the above criteria may be pointed out. Firstly, it is not necessary that the variants should occur as possible and equivalent alternatives in all contexts. Categorical contexts will be excluded from the quantitative analysis, just as certain phonological variables are contextually restricted: for example, the variation between [t] and [ʔ] in English is not found word-initially. Secondly, the variants need not have the same expressive meaning (cf Lyons, 1981:54), a term which here is intended to cover not only connotative and affective meaning, but also emphasis and politeness (cf Leech, 1981:12-6; Lyons, 1981:54). Emphasis and politeness are not excluded from the analysis, however, since they may be considered as factors motivating the speaker's choice of variant in some cases. But the highly subjective nature of

relative judgements of these aspects of meaning makes it undesirable to categorise one variant as being more or less emphatic or polite than another regardless of the circumstances. Moreover, there is ample evidence that change (and hence synchronic variation also) overrides such differences in expressive meaning. Thirdly, thematic ordering is also omitted from our list of criteria, since it is by no means certain that an apparent contrast at this level rules out the possibility of social differentiation between the variants.

Finally, to return to the distinction made at the beginning of this chapter between the variable as a methodological tool, or procedure, and the variable as a concept within the sociolinguistic theory of language, it is suggested here that sociolinguistic grammatical variables must meet criteria (i-vii) above. However, the variable (or variationist) procedure may be profitably adopted to investigate the factors constraining other alternations also (as was clearly demonstrated by Weiner & Labov, 1983), and in such cases it is only the semantic and formal criteria which need be applied, ie (i-iv) above.

CHAPTER 2: A CORPUS OF SPOKEN FRENCH

2.1 Introduction

This chapter is concerned with describing the corpus which provides the data for the subsequent variationist analysis of negation and interrogation. It was decided to focus on the informal speech of (mostly) young adults from northern France, since it was supposed that this would be of particular relevance to teachers and learners of French as a foreign language. As no suitable corpus appeared to be already available, it was necessary to carry out fieldwork specifically for this project. In any case, given the research objectives, it was particularly advantageous, if not essential, for the fieldwork to be carried out by myself, so that the subsequent interpretation of the informants' speech would be facilitated by a knowledge of relevant contextual and situational factors, which can be crucial in an area such as the pragmatic interpretation of interrogatives.

However, it must be recognised that the non-native fieldworker is clearly at some disadvantage in comparison with "insider" fieldworkers who operate within their native speech communities (eg Trudgill, 1974, and the fieldworkers on the Montreal project, cf Sankoff, & Sankoff, 1973), or even with native-speaker "outsiders" who have nevertheless lived in or near the community for some time (eg Cheshire, 1982; Milroy, 1980). The non-native fieldworker's background knowledge of the community, its members and its culture is likely to be more restricted than that of

native fieldworkers. And when the data consist of interviews with the fieldworker, there is clearly a risk that even a fluent non-native will fail to elicit an informal style of speech from the informants. In spite of this, there are several cases in the sociolinguistic literature of successful fieldwork being carried out by non-natives, eg Gardner-Chloros' study of code-switching in Strasbourg (1985a, b), the surveys of multilingual communities in Belize and St Lucia (Le Page & Tabouret-Keller, 1985), Russell's (1981) study of Mombasan Swahili and Lennig's (1978) research into variation and change in Parisian French vowels.

Early sociolinguistic research was often based on random or "quasi-random" samples of speakers from a given speech community, eg Labov (1966), Sankoff & Sankoff (1973), and Trudgill (1974). This had the considerable advantage of enabling the researcher to make generalisations about the speech of an entire community, and to observe significant patterns of differentiation in various sub-groups. Other studies, however, have chosen to sacrifice representativeness and breadth, to some extent, in order to put a higher priority on obtaining larger quantities of vernacular speech, eg Labov et al. (1968), Cheshire (1982), Milroy (1980), and, in Paris, Laks (1977, 1983). There is some evidence to suggest that securing interviews from a random sample of speakers in France is more difficult than has been found in Montreal, Norwich and New York, for example. The Orléans survey encountered a failure rate of over 75%, with refusals being most frequent among working-class speakers (Blanc & Biggs, 1971). According to Lennig (1978:9-10), an attempt to obtain interviews from a random sample in 1968 in the Paris suburb of Villejuif met with almost

total failure.

Since the present study is not, in any case, attempting to make a definitive characterisation of the linguistic behaviour of a speech community, it was not necessary to obtain a random sample of speakers. Moreover, given the non-native and non-resident status of the fieldworker, it was important to minimise the socio-cultural distance between the latter and the informants. For these, and other practical reasons, it was decided to select as a setting for the fieldwork children's holiday centres - an area of French life of which I already had substantial experience, as an "animateur" (youth leader). Although there is at present little evidence regarding the influence of geographical origin on grammatical variability in metropolitan French, it seemed important to restrict the scope of the research to speakers from a single department, in this case, the Somme.

2.2 The Somme and the Picard dialect

The department of the Somme is located half-way between Paris and Calais, and the river from which it takes its name runs east-west across the department, through the small towns of Ham, Peronne and Corbie, before reaching Amiens (the geographical and administrative centre), passing through the medium-size town of Abbeville, and finally emerging into the Channel at St Valéry-sur-Somme. Together with the other two departments which form the region of Picardy (the Aisne and the Oise), the Somme can be characterised as an area of modest urbanisation with many villages and mostly medium and small towns (Moreau, 1981). In

1982 (the year of the fieldwork), the Institut National de la Statistique et des Etudes Economiques (l'INSEE) recorded the population as 543,600, of which about 131,000 lived in Amiens, and another 25,000 in Abbeville, the second largest town. The proportion of the Somme's inhabitants which lives in rural areas is significantly higher than the average French department: 42% compared with 27% in 1975 (INSEE, 1981). Between 1975 and 1982, the populations of Amiens and Abbeville registered a slight decline, attributable to the phenomenon of "périurbanisation", ie of people going to live in neighbouring rural communes, and commuting to work in the towns (Le Courrier Picard, 16.7.82). Several indicators point to Amiens' socio-economic dominance in the Somme: it is a commercial centre for virtually the whole department, its industries attract manual workers from a radius of over 30 km and its newspaper, Le Courrier Picard, is distributed throughout the department (but not beyond). Only the far east of the Somme escapes the influence of Amiens to some extent, due to its proximity to St Quentin, the largest town in the neighbouring department of the Aisne (Centre Régional de Documentation Pédagogique, 1981:15). As a result of its relative lack of urbanisation, and its higher than average unemployment rate, the Somme has attracted few foreign workers: in 1979 there were only about 12,000 (in comparison with some 4 million in France as a whole).

The inhabitants of the Somme appear to see their department as being at the heart of Picardy, and the historical and linguistic evidence lends support to this belief. The contemporary region of Picardy is only about thirty years old, but what is now the Somme

also accounted for the major part of the pre-revolutionary province of Picardy (which also included the coastal strip from Calais southwards and the north of the present-day Aisne). Moreover, the whole of the Somme is situated within the Picard dialect area, which also covers most of the Nord Pas-de-Calais region, plus a part of Belgium (Carton, 1981:16).

Picard is one of the dialects of the langue d'oïl group which covers northern France, and is, unsurprisingly, quite close linguistically to Standard French. Although it is generally reckoned that Picard, along with the other northern dialects, had been assimilating to French for several centuries, it was only from about the 17th century that it came to be associated with social inferiority (Crampon 1981:556). After Grégoire's language survey in 1794 claimed that about 12 million French citizens (out of a total population of some 26 millions) had little or no command of spoken French, the State adopted an unequivocal policy of promoting French and of eradicating dialects and regional languages (cf de Certeau et al., 1975). The principal instrument of this policy through the 19th century was state primary education, although this only became free in 1881 and obligatory in 1882 (Price, 1987:316-7). Teachers were required to suppress the use by schoolchildren of any variety other than French, and made frequent use of ridicule and punishment in their attempts to do so. To what extent they were successful is unclear, but it seems certain that the language policy, together with other factors such as urbanisation, led to a change in attitudes among dialect-speakers themselves, and heralded, in the long term at least, the definitive decline of dialects such as Picard. Several

commentators have suggested that the 1914-18 war had a decisive impact on the use of dialects (eg Martinet, 1962:114). The furthest advance of the German army came in September 1914 and took in the entire eastern half of the Somme, including Amiens and the northern town of Doullens. However, by December the front had been forced back eastwards some 30-40 kilometres, and remained there for most of the war. By November 1918, over 50% of the buildings in the eastern Somme had been destroyed and hundreds of thousands of hectares of land were devastated (CRDP, 1981:59). In his description of the Picard spoken in Mesnil-Martinsart (5 km north of Albert, near the departmental border with the Pas-de-Calais), Flutre commented that the dialect had not survived the 1914-18 war: the two villages had been rased to the ground, the inhabitants were evacuated and dispersed, and many did not return after the war (1955:VIII). The population of Picardy took until 1962 to return to its 1914 level, whereas France as a whole had reached this point by 1931 (CRDP, 1981:58). From 1940 to 1944, the whole of the Somme was again occupied by the German army, and this led to a massive exodus: by July 1940, the population of Amiens had shrunk from 90,000 to less than 7,000. Although there is little doubt that the two wars accelerated the process of dialect shift in the Somme, this would certainly have been taking place anyway. Even in the far south of France, there has been an inexorable trend in this direction in the 20th century, as is reported by Eckert for a Gascon-speaking community in the Pyrenees (1980:181).

Today all observers appear to agree that Picard is disappearing as a first language (eg Crampon, 1981:560), though it seems to

have fared better in those parts which were less severely affected by the 1914-18 war, such as the Vimeu area along the coast south of St Valéry (Carton, 1981:16). From this side of the Channel, it may seem surprising that dialect should have survived so long in an area so close to the nation's capital. The principal reason for this is probably to be found in the fact that the population is far less urbanised than in English counties located a comparable distance from London, and that traditional patterns of rural employment have endured much longer in the Somme.

Since 1945 the language policy of the French State has been liberalised considerably. In 1951 the loi Deixonne was passed to permit the teaching of certain regional languages in secondary schools, and in 1985 the government established a Conseil National des Langues et Cultures de France (Libération, 8.8.85). The plurality indicated within the title of this body would have been inconceivable before 1940. However, these developments have generally been limited to varieties which are linguistically remote from French itself, thus excluding Picard and other northern dialects. Nevertheless, they have perhaps contributed indirectly to a shift to more positive attitudes vis-à-vis Picard, at least among certain sections of society. There are now popular writers, singers, rock groups, theatre troupes and puppeteers who write and perform in Picard in preference to French. But most strikingly, perhaps, the oppressive language policy of the Third Republic has been thrown into reverse in primary schools. Since about 1977, trainee teachers in the Ecoles Normales in Amiens have been given lectures on Picard, in order to foster more positive attitudes and to help prepare them to

cope with the children who, in some areas, still arrive at school speaking a variety which, if not "pure" dialect, has nevertheless many Picard features (Crampon, 1978). A number of books on Picard have appeared which are specifically aimed at teachers (eg Emrik (1978), and even a Picard language course complete with (scripted) recordings (Debrie, 1983).

These developments found an echo among the informants in the present study, several of whom were teachers or trainee-teachers. Those informants who reported that Picard was still spoken suggested that it was to be heard mostly among the elderly and in the north and west of the Somme (and still more so in the Pas-de-Calais). A few claimed to have (or to have had) a modest active command of the dialect, including one 18 year-old, who added, however, that her brothers used it far more than she did, because it was "natural" for them, but not for her. If such a pattern were general, it would coincide with Gal's (1979) finding that women made the shift to the more prestigious variety well before men in the same community. One of the young male informants recorded for the present study produced two or three "party piece" sentences in Picard. But several other young people present at the time were quite unable to understand what he said. A feature which was frequently mentioned as being characteristic of Picard (and which seems to be a stereotype), was the use of che or ch' as the definite article. Indeed one informant was heard to use this feature occasionally in his speech, though he did not do so at all on the recording. However, the same speaker reported that he could not understand his grandmother when she spoke Picard. Several informants commented that "real" Picard was

no longer spoken, but referred in often uncomplimentary terms to mixed varieties (cf Eloy, 1987, on Picard/French co-occurrence restrictions in Amiens speech). Another informant summarised the contemporary sociolinguistic^{situation} in Picardy as follows: his grandparents' generation (born in the late 19th century) spoke only Picard, his parents' generation (born around the 1920s) understood the dialect and spoke it to some extent, but his own generation (born in the 1940s and 1950s) understood but did not speak it. During the fieldwork period, I did not hear any Picard produced spontaneously, by either children or adults. Nevertheless, the dialectal background will be borne in mind as a possible influence on the variability in negation and interrogation to be analysed in later chapters.

2.3 Centres de vacances

In France, many children between the ages of 4 and 14 spend part of their summer holidays in residential holiday centres in the company of between 30 and 200 other children. There are also camps d'adolescents for the 15-18 age group. In 1983 the number of children and teenagers taking part in these group holidays was 1.3 million (Verdié, 1985:135). The centres are situated in (former) châteaux, boarding schools, camp sites or purpose-built centres away from the children's home area, in the countryside, the mountains or at the coast. They are organised by a variety of institutions: local councils, companies, nationalised enterprises, churches, charitable organisations and some small firms which exist solely to run such centres. They are usually financed by the above organisations and by varying

degrees of parental contributions. Centres de vacances were originally set up (initially in the late 19th century, but increasingly during the Front Populaire government of 1936-38) to enable underprivileged children from the cities to spend the summer in a more healthy environment. In the 1960s and 1970s, however, they began to cater also for the more affluent classes, and there are now very few that conform to the stereotype of near-slum conditions evoked in Pierre Perret's light-hearted song, "Les jolies colonies de vacances". Almost all centres are now mixed, and the children may come from a limited catchment area (a town, department or region), or they may be from several different areas of France. There are also some binational and international centres. It is widely acknowledged that centres can make a very significant contribution to children's personal and social development, and French people are often astonished to learn that centres exist only marginally in Britain.

Each centre, which normally lasts between three and four weeks, is administered by a director, aided by one or more assistant directors, a bursar, a health assistant or nurse, and kitchen and cleaning staff. There are also a number of animateurs (at least one for every twelve children) to look after the children and organise their leisure activities. These staff are employed, with generally very modest remuneration, only for the duration of the centre. Directors and assistant directors are often teachers, and the animateurs are typically students or final-year lycéens (cf Gaudard, 1982).

Centres must conform to regulations laid down by the Ministry of

Youth and Sport, and are subject to frequent inspection.

Directors and animateurs are normally required to have completed at least one training course prior to their first centre. These courses are generally residential, last for several days and are similarly subject to ministerial guidelines and inspection.

Directors and animateurs who successfully complete the whole training programme receive a diploma from the Ministry of Youth and Sport. The courses are run by various non-profit-making associations which exist to train staff, develop resources and generally to support the centres de vacances movement.

Until 1984 the Central Bureau for Educational Visits and Exchanges arranged annually for over fifty young people from the U.K. (usually students of French) to follow a training course and work as animateurs in France. For a student of French, the linguistic benefits alone of this experience are enormous. When it comes to learning everyday spoken language, children aged between 7 and 14 are probably the best models one could wish for, because of their clarity of articulation, their moderate speech rate, their distinctive intonation patterns and their uninhibited expressiveness. The animateur is engaged in face-to-face conversation with children for about twelve hours per day, or some 300 hours over the duration of the centre. I took advantage of the CBEVE scheme myself as a student, and continued to work as an animateur for several summers following, subsequently becoming involved in helping to run training courses for both British and French animateurs. When it was necessary to arrange fieldwork for the present study, it seemed natural to make use of the experience I had accumulated of this small, but

influential area of French life, which is, moreover, a haven of vernacular speech.

2.4 Fieldwork procedure

Having obtained from various associations based in the Somme a large number of addresses of directors of centres, I wrote to them explaining that I was intending to make a study visit of several centres in order to learn more about them myself, and to assemble an audio-visual documentation (slides and recordings) which could be used for the preparation of British animateurs wishing to work in France. I asked the directors to complete and return a questionnaire requesting details of the centre, and of the number and geographical origin of animateurs and other persons working in the centre. I also asked whether the directors would be willing for me to stay at their centres for three or four days as a paying guest. In view of the fact that accommodation is often quite limited, it would probably not have been reasonable to stay much longer than this, and in any case it was necessary to visit a fairly large number of centres over the two months in which they run (July and August) in order to maximise the number of potential informants. The response rate to this questionnaire was excellent, and the respondents were unanimously positive, and often most enthusiastic about the project. With the information thus obtained, it was possible to arrange an itinerary to visit fifteen centres, some in Picardy itself, some in various other parts of France. As time was relatively short, and several tape recorders and a large stock of tapes needed to be taken along too, it was decided not to follow

in the cycle tracks of Edmond Edmont, but to purchase a second-hand car to travel the 2,500 miles.

Staying at the centres themselves proved a considerable advantage, as it was possible to get to know people quite rapidly, and to become, in effect, a temporary member of these temporary communities. In conversation with potential informants who were native-speakers of French, I asked where they were from (taking care to avoid the Tu es d'où? formulation). If they lived and/or worked in the Somme, I would at some stage ask if they were willing to record an interview to talk about their experience of centres de vacances. A number of potential informants (all female) made excuses or simply refused this request. The reason seemed in most cases to be a reluctance to be recorded rather than any antipathy towards the fieldworker, since they often talked with me at some length "off the record". The same problem has been encountered in other studies, even when the fieldworker was female also (eg Farhat Khan, personal communication). There appears to be some evidence to suggest that, on average, females are more sensitive than males about having a permanent record made of part of their self-image (cf Milroy's comparison with photographing human subjects, 1987:88). In contrast with the refusals, however, most of the informants appeared to enjoy the interview and the opportunity it provided to talk at some length about matters of considerable interest to both themselves and the fieldworker.

Another problem in securing interviews was the fact that many of the people working in the centres were extremely busy and had

very little spare time in which to simply sit down and talk. Of course, a lot of spontaneous speech is produced in the centres, but it tends to be in circumstances which are not conducive to recording, for example in large, noisy refectories, or during games outdoors. Even indoor activities proved difficult to record successfully, as the rooms in which they took place invariably had very poor acoustics. More sophisticated recording equipment would have helped to overcome some of these problems, but this was unfortunately not available: a simple cassette recorder was used, together with either a clip-on neck-tie microphone or a hand-held or free-standing microphone. Rechargeable batteries enabled interviews to be recorded in various situations, including a number out of doors (in quiet areas with plenty of vegetation) and in one case, whilst coming down a mountainside. Another interview was made in a local bar, during the informant's evening off, but most were recorded in one of the offices, or some other room in the camp. Various other, more spontaneous recordings were also made, but the quality of these is very uneven, and they comprise mostly small amounts of speech from a large number of different individuals, especially children. Consequently, these supplementary recordings do not constitute a suitable corpus for the present study. In addition to the recordings, some three hundred slides were taken of various aspects of the centres visited, and a selection of these was later combined with some of the recordings for a tape-slide sequence which has been shown in this country to future animateurs.

2.5 The interviews

It has been recognised for some time in sociolinguistics that informal speech can best be obtained in spontaneous, preferably peer-group interaction, rather than in interviews (cf Milroy, 1987:39-67). Nevertheless, for this research project, interviews were seen to have two distinct advantages: firstly, in most cases, they provide a guaranteed means of recording a fairly large quantity of speech from each informant (which is particularly important in the study of grammatical as opposed to phonological variability), and secondly they allow the fieldworker to attempt to elicit certain grammatical forms by means of specially framed questions and other techniques.

As I stayed in each centre for several days, I was not a total stranger to the informants when the interviews were recorded. Moreover, because of my contacts in the centres de vacances movement, we often had some acquaintances in common, and the number of these naturally increased as the fieldwork progressed. Several of the interviews included sections where mutual acquaintances were the topic of conversation. Another factor which increased the informality of the interviews was the reciprocal tutoiement which had already been established between myself and the majority of the informants. The tu/vous alternation cannot be characterised as a sociolinguistic variable, due to the fact that usually the choice of pronoun with which to address any given person is a more or less permanent one, determined by considerations of politeness (cf Brown & Levinson, 1987:22-5, 198-204). But it is possibly the most

significant sociolinguistic feature of contemporary French, partly because of the subtle and variable nature of the norms governing the alternation. In general terms, reciprocal tutoiement is very closely associated with vernacular speech, and in many centres de vacances it is universally practised, even between a child and the director. Among the thirty informants, there were six exceptions to the general pattern of reciprocal tutoiement in the interviews. Three of these informants were women in their fifties, who would not therefore normally exchange tu with a 27 year-old male unless he were a member of their family. Another was a male informant aged 37, who I interviewed shortly after my arrival at the centre (since he was not going to be available later), and with whom I had not yet established reciprocal tutoiement. The two other exceptions were 18 year-old females who were still at school, and who managed to avoid using a pronoun of address at all in the interview, except on one or two occasions (when they used vous). I, in fact, used tu to them, as I always had done with animatrices in centres de vacances or on training courses. In retrospect, this may not have been the best strategy, particularly as neither of these two informants had previously experienced a training course or centre, but it has since become clear that non-reciprocal usage is by no means uncommon among native-speakers in situations of uncertainty.

The starting point, and in most cases the principal topic, of the interviews was the informant's experience of centres de vacances (both as a child, and as an adult) and of any training courses they had followed. This generally led in to more specific questions about aspects of the organisation of everyday life in

centres, including matters such as the groups to which the children were allotted, their pocket money and mail etc. Games, outdoor pursuits, craftwork and other leisure activities were the topic of another set of questions. With informants whose work in the centre involved planning and preparing meals, food was also a major topic. Questions about the geographical origin of the children often preceded some discussion about Picardy and Picard dialect. All informants were also asked about their life outside the summer holiday period, their work or studies, their own leisure activities and their plans. Not all of these topics were systematically covered with all the informants, especially when they spoke at length on one particular topic. The questions were formulated more or less spontaneously (accompanied by some degree of hesitation and reformulation), though a note of the topics to be covered was kept at hand. Questions which have clearly been phrased in advance, or which are even read aloud, seem inevitably to increase the formality of an interview. Similarly, most of the questions concerned the direct experience of the informants or their opinions and attitudes, rather than knowledge of a more general kind, and they could thus be answered quite readily. The value of this approach lies in the fact that it enables the informant to feel in control of their subject-matter, which is probably essential for the production of informal speech.

My own background, together with French people's considerable curiosity about Britain, enabled me to go beyond the normal rôle of an interviewer by talking a little about my experience of centres de vacances and about myself and Britain more generally. This often caused the interview to evolve into something

approaching a normal conversation in which all participants have, in principle, an equal right to control the topic and to ask questions.

A number of questions were included partly in an attempt to elicit particular grammatical variables which might otherwise have occurred only infrequently. The questions about games, craft activities and local recipes often led on to descriptions of how to go about these various processes, and these produced large numbers of tokens of the generalised indefinite subject pronouns on and tu or vous. Initially, it was possible for me, as only a semi-insider, to ask these questions in genuine ignorance of the answers, but after hearing several similar accounts of a game, for example, this ignorance had to be feigned somewhat. A useful variant was to explain that I did not quite see the difference between, say, the jeu de l'oie and the jeu de piste. Questions about the informant's plans, both for their participation in centres de vacances and more generally, unsurprisingly succeeded in eliciting instances of the simple and periphrastic future forms of verbs. Both of these grammatical variables are to be the subject of later studies. However, more formal elicitation techniques (such as tests of various kinds) were not attempted, since they would have been quite out of place in the context of the holiday centres, and would have been virtually impossible to justify to informants.

In some, but not most, of the interviews, one or more other persons were present for some of the time, and there were also a number of briefer interruptions, particularly in the interviews

with camp directors and other senior personnel. The duration of the interviews obtained was of course variable, but was on average 36 minutes, giving a total corpus of 18 hours for the thirty informants. This is modest indeed when compared with the 270 hours of the Ottawa-Hull corpus (Poplack, to appear) or the 300 hours of the Orléans project (Blanc & Biggs, 1971), but for the purposes of the present study it has generally proved adequate.

The recording quality of the interviews was also variable, due partly to the different settings in which they took place, but was quite satisfactory on the whole. Four early recordings made with the neck-tie microphone were seriously marred, however, by interference which, in some cases, was apparently the result of static created by the informant's clothing. The result was that portions of these interviews were quite inaudible. Consequently, the neck-tie microphone was abandoned and later recordings were made exclusively with the hand-held omnidirectional microphone.

Milroy & Milroy (1985b:136) characterise interviews as having a clear two-part structure in which the fieldworker elicits and the informant replies, and as typically having a slow pace with pauses between turns and no interruptions. They liken this discourse pattern to classroom interaction and test situations, and add that, although the structure may be modified as the interview progresses, the initial stages of all interviews conform to this pattern. Many of the "interviews" in the present corpus, however, do not in fact resemble this pattern at all. Whilst it is true that the vast majority of the questions are

asked by the fieldworker, there are very few of the minimal, elliptical answers characteristic of the interview extract analysed by Milroy & Milroy (pp. 132-3). This may well be because the questions asked did not tend to be a series of specific factual ones which could be answered in this way, but included many which requested opinions or which necessarily required fairly lengthy responses. Care was also taken to avoid, as far as possible, embedding questions within a request (eg "Could you tell me where you were born?"), since these seem to characterise more polite and more formal interviews, and elicit the same kind of speech in response. The speaking rate of most informants in the interviews was not noticeably slower than at other times. Indeed in certain cases it was so rapid that subsequent transcription proved quite difficult. Moreover, the recordings exhibit a fair amount of overlapping speech and interruptions, and very few long pauses between turns. Several of them begin with the informant asking questions of the fieldworker, and in most of them the informant quite quickly starts producing lengthy turns, often consisting of narratives. Although two or three informants showed signs of nervousness, they were very much in the minority. The speech of the informants is not at all characteristic of what one would expect in a test situation, and indeed, given Milroy & Milroy's defining characteristics of interviews, it might well be more accurate to describe the recordings as elicited conversations, rather than interviews as such. For the sake of clarity and brevity, however, we shall continue to refer to them as interviews.

Information concerning the speakers' residence, occupation,

education and age was, in many instances, obtained outside the interview, but when this had not been done, the questions were generally inserted individually at relevant points in the conversation, rather than as a sequence. Sometimes it was possible to prepare an appropriate context for such questions: for example, having asked younger informants about their intended job, I would ask whether their parents were in a similar occupation, and the informant would usually either confirm this or, if it were not the case, say what they did anyway. Providing information about myself, such as my year of birth, sometimes prompted informants to offer similar information about themselves in exchange.

2.6 Characteristics of the speakers

With regard to the sample of speakers for this study, it had been hoped to record a sufficient number of older speakers to enable some comparison to be made with the younger adults who were to form the focus of the research. This was perhaps unrealistic in retrospect, and I soon discovered that older people (especially men) tend not to choose to spend their summer holidays working in centres de vacances. In the event, just three older female informants (in their fifties) were recorded, the oldest males being only 37 years of age. Apart from the three older women, then, the informants were aged between 17 and 37, and comprised similar numbers of women and men. Given the relative numbers of animateurs and of more senior personnel in the centres, it was of course easier to find potential informants in the 17 to 22 age group than older ones. Although no specific quotas were set, an

attempt was made to record speakers from different social backgrounds, particularly any who were neither students nor teachers, since it was suspected that these two groups would be particularly numerous in the centres. Consequently, three broad social classes are reasonably represented among the younger speakers, but in the 24 to 37 age group, there is just one person from the working class. As with the lack of older male informants, this is not surprising: few people in employment choose to spend their four weeks of summer holidays working in a children's camp. Most of those who do work in centres tend to be school-teachers, who benefit from the same holidays as the children, ie about two months.

It is well known that the concept of social class is a controversial one and that different groups of factors have been taken into account in various sociolinguistic studies to construct social class indices (cf Guy, 1988:39-49; Milroy, 1987:97-101). Despite some fundamental criticisms and certain very real problems in classifying speakers satisfactorily, no superior alternative system of a comparable scale appears to be available. The Linguistic Market Index developed by Laberge (1977; cf also Sankoff & Laberge, 1978) perhaps succeeds in achieving an intuitively more satisfying classification for speakers such as teachers and secretaries, whose work generally requires them to use more standardised speech than other occupations within a similar social class. However, the descriptions of the individuals' economic backgrounds upon which the index scores were based, were not comparable for all speakers, as Guy has pointed out (1988:45). Even more seriously,

perhaps, it could be said that there is a certain circularity in asking judges (sociolinguists who know the speech community well) to study the economic history of individuals and to give them scores according to how important it is for them to have a command of the standard language, and then to correlate those scores with linguistic behaviour. In a sense, the judgements amount to guessing how standard the linguistic behaviour of the speakers is likely to be, and it is hardly surprising if the judges are found to have guessed correctly.

Several other social variables which have proved to have considerable explanatory power in the analysis of linguistic variation often require a long-term participant observation methodology: for example, the strength of social networks and the position within the peer-group, as studied by Milroy (1980) and Laks (1983), respectively. The centres de vacances where the fieldwork for the present study was carried out were all temporary communities, and the extent to which the adults working in them knew each other previously varied enormously. Some groups included several members of the same family, or friends who had known each other for many years and saw each other frequently throughout the year, whereas others consisted mostly of strangers, and many included people from various parts of France.

It has often been suggested that a speaker's education (eg as measured by the highest educational qualification obtained) is likely to correlate with, and perhaps even explain, socially differentiated variation in language. Ashby (1981, and personal communication) classified his Tours informants into three groups

based principally on educational achievement. In the present study, however, it was decided to classify speakers according to occupation rather than education, partly because the two variables are to some extent connected anyway (ie certain occupations are open only to people with a given level of education), and partly because level of educational attainment is dependent on age in various respects - on the one hand, the school-leaving age has risen progressively this century, and on the other, most of the younger informants had not yet completed their studies. Most importantly, it was suspected that, for this group of speakers at any rate, education would shed less light on the linguistic variability than a straightforward indication of social class based on occupation.

Each informant was therefore allotted to the appropriate socio-professional category (as specified by the INSEE, eg in Desrosières, 1984), and then to one of the three broad classes widely used in French social research (cf Marceau, 1977:6-9). These are supérieure, moyenne and populaire, which will be glossed here as the upper, intermediate and working classes. The upper class includes, amongst others, secondary-school teachers and liberal professions, the intermediate class covers primary-school teachers, technicians and shopkeepers, whilst the working class takes in, for example, manual workers in industry and agriculture. As is customary (if not entirely satisfactory), students have been classified according to the occupation of the head of their household and married women according to their husband's occupation, unless their own was of a higher status (cf Mullineaux & Blanc, 1982:11; Trudgill, 1974:38-9). Three

informants could not be classified in terms of class because the information obtained about their parents', or their own occupation, was insufficiently precise for them to be allotted to a socio-professional category.

The other two characteristics according to which the informants are classified in this study are sex and age. Whilst biological age is, in itself, totally objective, what is less clear is the basis on which age groups are to be established. Not only is age a continuous variable, but also a given biological age may have a varying significance from one culture to another. One sociolinguist who has tackled this problem is Thibault (1983), who set up five groups among the Montreal informants according to their shared experience of important socio-economic developments in Montreal, as well as their progress through significant landmarks in their own lives. Thibault's two younger groups, les jeunes (15-23 years) and les jeunes adultes (24-35 years), correspond very closely to the two principal age groups in this study. 23 years of age seemed to be an appropriate dividing line since none of the informants below that age was either married or in permanent full-time employment, whereas most of the informants in the 24-37 age group possessed both these characteristics, and there are grounds to believe that both marriage and employment have a significant effect on the individual's social network pattern, and consequently on their linguistic behaviour also. The three women in their fifties (all from the intermediate social class) clearly needed to be considered as a group apart. Table 2.1 shows the distribution of the speakers in terms of age, sex

and social class.

Table 2.1 Distribution of the informants according to age group, sex and social class

		Male	Female
17-22 years	working class	2	2
	intermediate class	3	2
	upper class	1	2
	class unknown	2	0
<hr/>			
24-37 years	working class	0	1
	intermediate class	3	2
	upper class	4	2
	class unknown	1	0
<hr/>			
50-60 years	intermediate class	0	3
<hr/>			
	Totals	16	14
<hr/>			

2.7 Transcription

It is not always realised that many, perhaps most, quantitative studies of phonological variation have involved little or no phonetic transcription. Once the variants have been identified, the analyst listens to the recordings and simply notes, for each occurrence of the variable, the variant used (Peter Trudgill, personal communication). An exception to this was the procedure

adopted in the Sociolinguistic Survey of Multilingual Communities (McEntegart & Le Page, 1982:110-1), whereby narrow phonetic transcriptions were made, but as this was so time-consuming, just 400 words of each speech style were transcribed per speaker. In the Belfast survey, on the other hand, a detailed phonetic transcription of the speech of a number of individuals was carried out as part of the pilot study, but subsequently just the lexical item in which the occurrence of the variable was located was transcribed (Milroy, 1987:117, and personal communication).

For the analysis of grammatical variables, it is clear that a larger context needs to be transcribed if linguistic constraints are to be taken into account. For many variables and their constraints, this can clearly be done selectively, ie without a transcription of the entire corpus (cf Cheshire, 1982:27). However, when discoursal and pragmatic constraints and/or a large number of different variables are to be analysed, it has often been considered necessary to transcribe all the data. In the projects on the French of Montreal and of Ottawa-Hull, funds were available to pay for full-time research assistants to type full orthographic transcriptions directly into computer files (Sankoff & Sankoff, 1973; Poplack, forthcoming). Similarly, though on a much smaller scale, in the ESRC-funded study of Black English in Dudley, the transcription and coding (in preparation for computerisation) of the 23 hours of data were carried out by a team of research workers (Edwards, 1986:136).

For the present study, in order to allow for the investigation of potential discoursal and pragmatic constraints on the variability

(and also to facilitate later studies of a wide range of linguistic features), a detailed orthographic transcription of the whole corpus was made. It is estimated that five minutes of recorded speech required at least one hour to be transcribed, with pauses, hesitations and overlapping speech also indicated. When the quality of the recording is poor, or the speech is difficult to decipher for some other reason (such as a very rapid delivery), the time required is correspondingly longer. Although a good deal was learnt by transcribing the entire corpus, researchers on future projects should probably be discouraged from undertaking transcription on this scale, unless assistance is available.

The transcription was made in conventional orthography, with some modifications, in order to reflect in particular the elision which affects many frequent items, such as clitics and certain discourse markers. The words most often elided are shown in table 2.2.

Table 2.2 Words most frequently transcribed in elided form

Elided form/Full form		Elided form/Full form	
j	je	d	de
t	tu, te	n	ne
i	il, ils	ors	alors
c	ce	ben	bien
m	me	fin	enfin
s	se	m	mais
l	le	qu	qui, que
es	les		

Although the use of these abbreviated forms gives a more faithful representation of the pronunciation, it has the effect of creating more homographs, which proves a disadvantage when the corpus is computerised. Limited use was made of conventional punctuation: the apostrophe, question mark and quotation marks having their conventional values, and the full stop being used occasionally to indicate a falling tone, where it seemed important to show this. Pauses were indicated by an oblique and very short pauses by a hyphen. Overlapping speech was enclosed by asterisks (for both speakers), and uncertain transcription by parentheses. Short sections of inaudible speech were represented by full stops (one per syllable) within parentheses. Longer stretches were indicated in upper case enclosed by square brackets, as too were various comments on non-verbal contextual features, such as laughter, noise or the arrival or departure of a third person. (Many extracts from the transcripts will be presented in later chapters: where it seems important, the fact

that an extract is preceded and/or followed by speech from the same person will be indicated by three suspension marks.) The speech of different individuals was written on separate lines, since this facilitates visual inspection of the turn structure of the interviews. Capital letters were used only for proper names, all of which have been altered, in order to ensure the anonymity of the speakers.

2.8 Computerisation of the corpus

Despite the attractions of full transcripts of the recordings, it must be said that these do not simplify the task of examining the occurrences of the variables which are to be analysed. In order to identify which contextual factors might prove to be constraints on the variability, it is convenient, if not essential, to have tokens in their context arranged sequentially on paper, ie concordances. To produce these manually, when dealing with thousands of occurrences and substantial context is a task of appalling repetitiveness and tedium, fit only for a computer. An attempt to produce such manual concordances was abandoned after several weeks. Fortunately, it later became possible to have the transcripts typed into computer files which, after editing, could then be processed with the help of the Oxford Concordance Program to give lists of all words occurring in a given transcript, and concordances of selected items.

The transcripts required considerable editing before being passed on to the typists, because the original versions included much minor detail which was found to result in an "overcrowded" typed

version, and sometimes also confused the typists. Therefore, many "back channels" (cf Brown & Yule, 1983:92) from the interviewer were deleted, when they seemed to have no other significance than to encourage the informant to keep talking. Also deleted were asterisks indicating overlapping speech where these occurred in the middle of words. As conventional accents were not available, the plus sign was used to represent the acute accent which occurs at the end of past participles, in order to facilitate searches for these items in future studies. It was not felt necessary to represent accents in any other cases. In contrast, several additional characters were used to flag items which could not otherwise be easily retrieved. Particularly relevant to the present study is the flagging of the negative adverb plus, and of the restrictive adverb que, which are both loci for the (ne) variable, and have very frequent homographs. Also flagged, for future purposes, were left- and right-dislocated NPs and non-dislocated subject NPs. For flagging, it was simply necessary to insert a distinctive character at the end of each occurrence of the item in question.

As some of the typists had no knowledge of French, it is quite understandable that a fair number of errors were made in some of the computerised transcripts. The more serious ones, together with any which would directly affect the present study, were corrected, but a good deal of editing remains to be done before the files can be used for the study of a wider range of linguistic features.

2.9 Use of the Oxford Concordance Program

The computerised transcripts were then prepared for use with the Oxford Concordance Program, by ensuring that the beginning of each conversational turn was marked in accordance with the "Cocoa" format references (cf Hockey & Marriot, 1982). Thus the beginning of a turn by speaker A is indicated in the text by <S A>. These references enable the user to instruct the computer to select items from the speech of one or more specified speakers, whilst ignoring those occurring in the speech of others. However, the references themselves do not appear in the concordances produced, which means that the beginning and end of the speaker's turn is not indicated. Such an indication is clearly essential in the analysis of linguistic, discoursal and pragmatic constraints, and so it was necessary to follow the Cocoa references by a letter to represent the new speaker: <S A> A. The capital letter outside the angled brackets would then appear in the concordances produced.

The Oxford Concordance Program is thoroughly and clearly documented for users and the basics can be mastered fairly quickly even with very little background knowledge of computing. The user simply needs to be able to edit and manage files. Version 1.4 of OCP was used for the present study, but version 2 is now available, together with a new user's manual which includes exemplification from a variationist study of Tok Pisin (Hockey & Martin, 1988).

To produce a concordance with OCP it is necessary to create an

empty file (to accommodate the concordance) and a command file, in which a set of commands is typed, taken from the wide range available, and which specify to the computer which items are to be selected from the data file (the computerised transcript) and from which speaker(s), how much context is to be provided for each occurrence, and how the concordance is to be laid out in the (formerly empty) results file. Once these three files are prepared, the user types in the following: run *ocp
1=commandfilename 2=emptyfilename 5=datafilename. (The three underlined file names are, of course, replaced by the actual names which have been given to the files.) OCP is now available for use on microcomputers, but for the present study it was run within the Michigan Terminal System of the mainframe computer at the University of Newcastle upon Tyne. OCP uses relatively substantial amounts of time of the computer's Central Processor Unit, but nevertheless, a concordance of several hundred occurrences of a given item is produced in seconds, and printed in minutes, rather than the hours it would take to be done manually.

For the analysis of the (ne) variable, it was decided to produce two concordances from each transcript, one with all occurrences of ne, or its reduced variants n' and n, and the other with all tokens of the negative words pas, jamais etc., whether or not these were accompanied by ne. The command file to produce concordances of the various negative words is shown in (1):

```

(1) 1 *INPUT
    2 text 1 to 90 and comments "<" to ">".
    3 references cocoa.
    4 select except where (S = "A", "X").
    5 *WORDS
    6 alphabet "A=a B=b C=c D=d E=e F=f G=g H=h I=i J=j K=k L=l
    7 M=m N=n O=o P=p Q=q R=r S=s T=t U=u V=v W=w X=x Y=y Z=z ^".
    8 diacritics "+".
    9 punctuation "@ % * # . :: ? [ ] ( ) : " - ' /".
10 *ACTION
11 do concordance.
12 maximum context left 70 letters and right 30 letters.
13 pick words "**^ PAS JAMAIS NI RIEN PERSONNE NUL* AUCUN* GUERE".
14 *GO

```

(The line numbers at the extreme left do not need to be inserted into the file and are included here only for ease of reference.)

This command file consists of three sections. The first, *INPUT, defines the parts of the data file which are to be scanned. Line 2 specifies that the maximum width of the file is 90 columns, and that material inside the angled brackets is to be treated as "comments" rather than as part of the text to be scanned. Line 3 confirms that the references conform to the "Cocoa" conventions. (Cocoa was apparently an ancestor of OCP.) Line 4 tells the computer to ignore the speech of A (the fieldworker) and X (third persons). The second section, *WORDS, specifies the characters which are used in the data file and their functions. The alphabet command lists all the characters which are used to form words, and indicates that upper and lower case letters are to be treated as equivalent. A significant addition is the hat symbol (^), which is used to post-flag those occurrences of plus and que which are loci for the (ne) variable. The plus sign is declared as a diacritic on line 8, and on line 9 the various punctuation symbols which are to be interpreted as word separators are listed. These include three characters used in the transcripts for flagging NPs in various syntactic positions: @, % and #.

Notice also that the colon and the quotation marks are preceded by the escape symbol (:), as they are "reserved characters" in OCP. The final set of commands, *ACTION, specifies what the user wants to be copied from the data file into the empty results file. Line 11 requests a concordance (as opposed to an index or wordlist), and line 12 indicates how much preceding and following context is required for each occurrence. Here, a fairly substantial amount of preceding context is needed in order to capture long NP subjects of negated verbs, and to cater for instances where the negative item or que occurs several words after the actual locus for (ne). Some following context is also necessary, however, in order for past participles and other significant elements to be shown. On line 13, the words to be selected for the concordance are listed in upper case. The asterisk here enables the user to specify all words beginning nul- and aucun- (thus taking in feminine and plural forms, as well as nullement and aucunement), and all words ending in the hat symbol, ie all flagged occurrences of que and plus. *GO on line 14 closes the command file and signals to the computer to begin execution of the commands.

A similar command file was prepared in order to produce concordances of the negative particle ne in the computerised transcripts. In this file the computer was instructed to pick all instances of ne and its elided forms, and line 12 of the command file was modified so as to give a smaller amount of preceding context and more following context.

To produce concordances of interrogatives, the question mark was

taken out of the punctuation characters, and placed among the alphabeticals. A concordance of all direct interrogatives could then be obtained by specifying all words ending in ? in the PICK command. However, if the question marks in the text were immediately preceded by a word, with no intervening space, then the concordance would list the interrogatives in the alphabetical order of those words. By inserting a space before every question mark in the data files, the tokens were presented in the concordance in the order in which they occurred in the transcript. Since the interrogative structures themselves are much larger than the occurrences of the (ne) variable, and discursal and pragmatic factors need to be considered as potential constraints, a substantial amount of context is required in their concordances. This was easily specified in the MAXIMUM CONTEXT command. However, when the context required exceeds one line, it is necessary to add a further set of commands to the file, under the heading *FORMAT, which enable the user to state the number and length of the lines over which the context for each occurrence is to be given.

The concordances produced in this manner greatly facilitated the subsequent analysis of the variability in negation and interrogation which forms the core of the present study, and which will be described in the next and following chapters.

CHAPTER 3: THE VARIABLE OMISSION OF THE NEGATIVE PARTICLE

3.1 Defining the variable and its variants

This chapter is concerned with what is possibly the best known sociolinguistic variable in contemporary French, the presence or absence of the negative particle ne. In cross-linguistic studies of sentential negation, French is the standard exemplar of an "embracing" device comprising the preverbal clitic ne and one of several post-verbal negative "auxiliaries" (Payne, 1985:224). These auxiliaries have traditionally been classified as adverbs (pas, jamais, plus, nulle part, guère, nullement), pronouns (rien, personne) or determiners (aucun, nul) (Grevisse, 1986:1477). In view of the virtual disappearance of ne in certain varieties, the term "auxiliary" seems something of a misnomer, and so we shall refer to pas, rien etc as the negative items, and to ne as the negative particle. In addition to the various negative items, ne is variably present with the restrictive adverb que:

- (1) ... ça a duré que quinze jours / ... (17447)
- (2) ... nos grands-parents ne parlent / que le patois ... (14920)

(Although the logical structure of such utterances does not include the negative operator, we may note that one corresponding expression in English is nothing but.)

The negative particle ne is standardly pronounced as [nø] preconsonantly, and as [n] prevocally. In rapid speech the /n/ may be realised merely as the nasalisation of a neighbouring vowel, but since this is a superficial phonetic phenomenon, and

is still a gesture of nasality, there does not seem to be a strong case for treating it as a distinct variant. A more serious candidate, perhaps, is the frequent allegro pronunciation of ne as [n] preconsonantly, and which might be considered an intermediate variant between [nθ] and θ. There is perhaps a parallel here with the contraction and deletion of the copula in American Black English (Labov, 1972b, chapter 3), where the variants are [iz] and θ, together with the intermediate variants [s] (following voiceless consonants) and [z] (elsewhere). However, in some cases at least, it could be argued that treating the full, contracted and deleted variants as a single variable may obscure more than it reveals. Chambers & Trudgill (1980:74-5) discuss some data from Black and White speakers from the Mississippi Delta, in which it is clear that Blacks delete the copula far more frequently than do Whites, whereas the latter contract it much more than do Blacks. But if an overall score is calculated by weighting the variants and treating the contracted form as the intermediate variant, Blacks and Whites have very similar scores and the differences are obscured. With these data at least, more would be revealed by treating contraction and deletion as two distinct variables. Similarly, it seems reasonable to consider the realisation of the French negative particle as [n] to be a result of the variable process of schwa deletion, and not as a distinct variant of the grammatical variable (ne).

In this study therefore, we shall follow previous researchers in analysing (ne) in terms of two variants, ne (including elided forms) and θ. We shall not, of course, be concerned with the

pleonastic or "expletive" ne, which, in formal styles, occurs optionally in certain contexts:

(3) Je crains qu'on ne me trompe.

(4) Avant que Louis ne parte. (both from Grevisse, 1986:1492)

3.2 Meaning and (ne)

It is, of course, well known that the scope of the negative operator in sentential negation is by no means always the whole sentence (Gaatone, 1970:54-5). Indeed some writers have used the terms predicate or VP negation on the grounds that the subject is often excluded from the scope of the negation. However, Payne (1985:199) points out that, equally, the scope frequently does include the subject.

Generally, the presence or absence of ne in a negative clause cannot be said to change the meaning, except perhaps in terms of emphasis (which is not, in any case, part of propositional meaning). Grevisse (1986:1492) suggests that speakers tend to reintroduce the negative particle for emphasis or to repair an utterance which has been misunderstood, as is illustrated in the following extract from the corpus:

(5) B: ... [QUIETLY] ça m'avait pas plu. /

A: comment? /

B: ça ne m'avait pas plu ... (35)

Désirat & Hordé (1976:157) go so far as to say that, in spoken French, ne can be considered as simply an emphatic particle without any negative value of its own.

One area in which the presence or absence of ne may convey a difference in the scope of the negative operator is in certain cases involving a modal verb, as in (6):

(6) On peut pas aimer ce film. (Diller, 1983:174)

Without context or intonation, this sentence is ambiguous, corresponding to both (7), where the negation is internal, and to (8), where it is external:

(7) On peut ne pas aimer ce film. [poss [neg on aime ce film]]

(8) On ne peut pas aimer ce film. [neg poss [on aime ce film]]

In an actual utterance, (6) would very probably be disambiguated prosodically, with stress on peut for the internal reading, and on pas for the external interpretation. In one potentially ambiguous instance from the corpus, the presence of ne avoids the ambiguity:

(9) alors tu vois c'est sp sous ce prétexte-là tu peux laisser tu
peux ne rien faire pendant deux mois quoi. / (4165)

A type of structure which has attracted considerable attention is negative "raising" or "transportation" (cf Prince, 1976), where, as in English, the negative markers are transferred to the verb in the main clause from the subordinate clause:

(10) Il faut qu'il ne périsse pas.

(11) Il ne faut pas qu'il périsse. (Grevisse, 1986:1488)

However, there does not appear to be any evidence that the presence or absence of the negative particle ne affects the meaning of such sentences.

3.3 The socio-stylistic value of the variants

According to Désirat & Hordé (1976:155), educated speakers feel that they are speaking badly when they omit ne, and this is due principally to the fact that ne is obligatory in written French. Contemporary grammars do not generally condemn the omission of ne as characteristic of vulgar or uneducated usage, but simply note

that it occurs very frequently in "popular" and colloquial speech (Chevalier et al, 1964:429; Mauger, 1968:374; Wagner & Pinchon, 1962:397). The latest edition of Le bon usage adds that ne disappears almost systematically, not only in working-class speech, but also in colloquial styles in Quebec, Paris and the Berry (Grevisse, 1986:1491). It appears that textbooks in France, however, prefer not to mention this deviation from the written norm (Désirat & Hordé, 1976:155). Whilst ne can clearly be regarded, then, as the more prestigious variant, its absence from informal styles does not seem today to be subject to stigmatisation.

3.4 History

The history of negation in French is well documented and will be sketched only very briefly here (cf Price, 1962, 1971:252-7). In Old French the principal marker of negation was ne itself, but it was optionally reinforced by one of several expressions which, etymologically, denoted small quantities: pas, mie, point, gout(t)e etc. It is thought that these were originally used for emphasis (Mohren, 1980), in the same way as in contemporary French pas is often reinforced by adverbial expressions of degree such as du tout and absolument. However, these items are reckoned to have lost this emphatic value quite early within the Old French period, and came to have negative meaning in their own right. This is most clearly shown by their use without ne in elliptical contexts. The frequency of these items in negative clauses increased through the Middle French period, and by the early 17th century they were more or less obligatory, in

non-archaic styles, in those contexts where they are used today. Contemporary written French, of course, still reflects this classical norm.

The most recent development in this area is, of course, the one with which we are directly concerned: the omissibility of ne. As is probably the case with most changes "from below", where the innovation carries less prestige than the older variant, it is difficult to locate the beginning of the change in time and space - all the more so in the case of ne omission, since this change certainly began several centuries ago. Although Pohl (1975:18) has reported that examples of ne deletion are to be found in certain medieval Flemish texts, the most generally held view seems to be that the process was only clearly under way in the Renaissance period - a number of cases from L'Heptaméron by Marguerite de Navarre (1492-1549), are cited by Gougenheim (1973:242). There is fairly substantial evidence that the deletion of ne had advanced considerably by the 18th century: Désirat & Hordé (1976:156) quote several instances from texts by Vadé (1720-57) which represent the speech of workers in the fish-market (presumably in Paris, though Vadé himself was born in the Somme), and Roelandt (1975, cited by Sankoff & Vincent, 1980) reports a fairly high rate of deletion in popular speech in 1730. However, Pohl (1968:1344) has commented that the fact that some relatively uneducated speakers (who are, presumably, less affected by the conservatism of written French) still have fairly high rates of use of ne today would suggest that the dramatic demise of the negative particle has taken place only quite recently. He points out that Balzac (1799-1850) did not include

ne omission in his representations of popular speech, unlike later 19th century novelists, such as Zola and Maupassant. Pohl concludes that a significant acceleration in this change took place in about 1820-50 (1975:22).

The causes of this cycle of changes in negation in French have been the subject of considerable controversy recently, in particular over whether they can be explained in terms of typological change (Harris, 1978; Ohkado, 1989; Schwegler, 1983; Venneman, 1974). However, a discussion of the complex issues involved in this area is beyond the scope of the present study. Instead we shall limit our brief discussion here to the questions most relevant to the variable (ne), in particular to how and why the negative particle has come to be variably omitted. Several factors contributing to the demise of ne have been proposed. Firstly, since about the 14th century, French has had phrasal rather than word stress, and this has meant that pas and the other negative items have typically been stressed (and still are), whereas ne has been unstressed (Harris, 1978:26; Posner, 1985:171). Secondly, with the loss of schwas from the 16th century onwards, ne has normally been realised as a single segment, [n] (Pohl, 1968:1344). However, as Posner (1985:177) has pointed out, such lack of phonetic substance need not lead to disappearance, or even to reinforcement by an auxiliary, witness the case of similarly lightweight negative particles in south central Italian dialects. A third factor which has been suggested is that the preverbal clitics of Modern French are comparable to bound prefixes, and that their bond with the verb has led to the exclusion of ne, which was the only item able to separate the

subject clitic from the rest of the verbal complex (Harris, 1978:26,119). In support of this suggestion, Posner points out that north-western Italian dialects, in common with many French dialects, use obligatory subject clitics, and that they too have lost the preverbal negative particle (1985:188). However, Occitan has likewise ousted ne, but it has never used obligatory subject pronouns (Schwegler, 1983:309). In addition, it is perfectly possible for a language to have preverbal subject prefixes which are separated from the verbal stem by a negative infix: this is the case in certain negative verb forms in Swahili, for example (cf section 3.10.2 below). A fourth point which has been proposed recently as a possible cause of the demise of ne is its homophony, in some dialects, with the "adverbial pronoun" en (Posner, 1985:189-92). This receives some indirect support in data from Montreal, where en is realised as [nã] (Sankoff & Vincent, 1980:298-9), and where ne has disappeared almost completely. It is not hard to see how, in this variety, a preverbal clitic sequence of ne + en, [nənã], would readily be reduced to [n:ã] and then [nã].

Despite counter-arguments, it seems likely that most of the above four factors contributed to the loss of the negative particle. But surely the most significant point is that once the auxiliaries had themselves taken on negative meaning and had become obligatory, ne became redundant, and this, together with the fact that it was both phonetically lightweight and separated by the verb from its erstwhile auxiliaries, led to its progressive disappearance. An almost identical "negative cycle" has occurred in English and other Germanic languages during the

last millenium (cf Ohkado, 1989; Strang, 1980:151,209-10,312). This cycle can be interpreted in terms of the communicative needs of speakers roughly as follows. Speakers seem to generally feel the need to emphasise negative sentences, possibly because these are more marked and less frequent than positive ones, and because they are generally used to deny positive propositions which either have been expressed in the discourse or which the speaker believes might be assumed by the hearer. The reinforcing expressions which speakers use often become grammaticalised as negative auxiliaries, increasing in frequency, shedding the lexical and syntactic restrictions of their etyma and becoming semantically negative (cf Horn, 1978:148-51). Short, unstressed grammatical items, such as negative clitics, often undergo contraction, and, if they become communicatively redundant, ultimately tend to be deleted entirely (cf Meillet, 1982:139-41). These processes can be seen as responses to the need for a language to be both rhetorically expressive and "quick and easy" to speak (Slobin, 1977).

Before leaving the history of negation, let us note that Pohl (1975:23) has suggested, in addition, that socio-economic factors may have led to the sudden acceleration in the loss of ne which he believes to have occurred between 1820 and 1850. Specifically, he points to the multiplication of social contacts in the 19th century, particularly as a result of the development of railways. We might add that this was also the beginning of widespread migration to towns and cities from rural areas, and that this was one of the factors which led to a shift from regional languages and dialects to French, and consequently to a significant

increase in the number of speakers of French, initially as a second, and later as a first, language.

3.5 Geographical diffusion of the omission of ne

It is well known that linguistic changes do not generally affect simultaneously all areas where a language is spoken. but usually diffuse geographically from the place(s) of origin of the innovation (often the capital city, or at least a major centre of population or influence) outwards to other towns and neighbouring rural areas, and eventually to more remote areas.

Some evidence for the geographical diffusion of ne omission has been extracted by Lüdicke (1982:50-2) from the Atlas linguistique de la France (Gillieron & Edmont, 1903-10). For eleven sentences used in the ALF questionnaire, the number of negative particles omitted by each informant in northern France was plotted on a map. The map shows that between 7 and 11 omissions were made by the informants in Paris and a large area including the Berry, the Centre, the Vendée and Poitou-Charente. Between 3 and 6 negative particles were deleted by the informants in areas around the edge of the main zone of omission, including most of the south of Ile de France, by several in Normandy and also by small clusters of informants in various parts of the north and east. Finally, two or fewer omissions were made by most of the informants in the whole of north-eastern France, francophone Belgium, Picardy and the far north, Brittany and much of Normandy. Clearly the evidence is very limited, being based on just eleven tokens per informant, in elicited translations into dialect from French

sentences (in which ne appears to have been systematically retained) and with only a handful of informants per département. However, the findings seem to suggest that, in the 19th century, the omission of ne had advanced furthest in Paris and in central and western dialects, and this is corroborated by other evidence.

On the basis of a number of small-scale quantitative studies, Pohl has concluded that "the loss of ne has progressed more in France than in Belgium, in Paris than in the provinces, in towns than in the countryside, and among monolinguals than among bilingual or bidialectal speakers (1968:1352), which is precisely the type of diffusion pattern we would expect in contemporary France.

In the south, Occitan had begun to use negative emphasisers by at least the 16th century, with the first cases of ne deletion appearing already in the 17th century. By the 19th century the postverbal negators occurred alone in most instances (Schwegler, 1983:308). According to Posner, ne was lost more rapidly in Occitan because of the absence of conservative pressure from a written norm (1985:193). The same could probably be said of the non-standard varieties of French in Canada.

3.6 Negation in Picard

We have noted already that most of the Picard informants in the ALF survey retained all eleven instances of the negative particle, and neither Debrie (1983) nor Flutre (1955) appear to mention any tendency to omit ne in Picard. According to Debrie

(1983:35-6), predicate negation is by means of preverbal ène accompanied by one of the following postverbal negative items: po, jamwé, pu, mi, pwin, which correspond to French pas, jamais, plus, mie and point (the last two being archaic in French). Flutre comments that mie is more emphatic than point, and also that the two negative items rien and personne are reinforced by an immediately preceding point, mie or pas. Similar cases of double negatives can occasionally be found in French literature of the 19th and early 20th centuries (Grevisse, 1986:1485), but today they would be considered non-standard. The phenomenon is also found in Montreal French, though apparently less so among younger speakers (Lemieux, 1985:102-36). The following example of double negation in the present corpus was produced by a 17 year-old male:

- (12) ... dans l bilan général j'ai pas eu aucune difficulté. /
 ... (23566)

If the speaker's intention here was to emphasise the negation, then the fact that he preferred to use a second negative item rather than ne would suggest that, for certain speakers at least, ne is indeed now semantically void.

3.7 Previous quantitative studies of (ne)

3.7.1 Data, informants and methodology

For a description of contemporary usage, we are fortunate now to have a number of quantitative studies, most of which focus on a particular speech community. Before discussing the overall rates of ne retention in these, let us consider briefly the characteristics of the various corpora on which they are based.

3.7.1.1 Gougenheim et al (1964)

Most of the data for the Français Fondamental project were obtained, in the early 1950s, from 275 informants, of diverse geographical origins, and of whom just eleven were children. Of the adult informants, 104 were from professional backgrounds (especially teachers and students), but a wide range of other occupations was also represented. Gougenheim et al did not carry out a specific study of negation as such, but Pohl (1968:1343) retrieved from their frequency counts the number of occurrences of pas and ne, which were, respectively, the eighth and the nineteenth most frequent words in the corpus. The frequency of ne retention will therefore be only approximate, as some cases of pas will not have been loci for (ne), and some instances of ne will have been with other negative items. Pohl (1975) also provides omission rates of ne for a wide range of individuals and groups, but he does not generally say the number of tokens on which these were based, and so these figures will not be included in this review.

3.7.1.2 Ashby (1976) and (1981)

In his first quantitative study of the negative particle, Ashby used Malécot's (1972) corpus of 25 hours of speech from fifty middle-class Parisian adults. His (1981) study, however, was based on interviews conducted by himself with 35 informants from Tours, equally divided among males and females, various socio-educational backgrounds, and two age groups, 14-21 and

51-64.

3.7.1.3 Sankoff & Vincent (1980)

Sankoff & Vincent's investigation into the productive use of ne in Montreal drew data from interviews with sixty informants, of both sexes and diverse ages and social classes, from the Sankoff-Cedergren corpus, recorded in 1971.

3.7.1.4 Diller (1983)

In a fairly small-scale study, Diller used interviews with twelve older adults, from two socio-economic classes, from an unspecified speech community in southern France.

3.7.1.5 Lüdicke (1982)

In addition to her investigation into the distribution of ne in the Atlas linguistique de la France, Lüdicke presented data from recordings with five female informants of diverse ages, occupations and geographical backgrounds. In view of the geographical diversity in particular, we shall not in fact present an overall frequency for these speakers. However, it is interesting to note that the rates of omission of ne varied from 91.3% for a 22 year-old Parisian nurse to just 12.9% for a 56 year-old housewife from Alsace (Lüdicke, 1982:48).

3.7.1.6 Moreau (1986)

Finally, Moreau, who was interested in the linguistic constraints on (ne) rather than social or stylistic differentiation, drew her data from interviews on Belgian radio with thirty male celebrities from France.

3.7.2 A comparison of the overall frequency of ne in previous surveys and in the present study

Let us now compare the global results from these surveys with each other and with those of the present study. Table 3.1 shows the overall rates of retention of ne found by the various researchers, together with the number of tokens from which these rates have been calculated. The frequencies are for each corpus as a whole, and are not an average of the scores of individual speakers. (As will be explained later, a significant number of tokens have been excluded from the total for the present study on the grounds that they do not allow the speaker an effective choice of variants.)

Table 3.1 Relative frequencies for the retention of ne
in seven quantitative studies

Variety	Researcher(s)	<u>ne</u> retained	N=
Montreal	Sankoff & Vincent	0.5%	c10,000
France	Gougenheim <u>et al</u>	61.9%	5,308
Paris, middle class	Ashby (1976)	55.8%	1,020
Tours	Ashby (1981)	37%	2,818
Southern France lower class	Diller	58%	641
upper class	"	78%	
French male celebrities (radio)	Moreau	50.25%	3,158
Somme	present study	18.8%	2,932

The most striking feature of this set of results is surely the major division between France and Montreal, where ne is retained in only a minute number of cases. Sankoff & Vincent (1980:301) reported that 45 of the 60 speakers did not use ne at all in their interview (which lasted an hour or more). The speaker with the highest rate of use was a 61 year-old woman from a wealthy business family, who had studied in France for several years. Even she, however, used ne only 8% of the time. Of the surveys conducted on metropolitan French, it is the present study which reveals the lowest overall rate of retention of the negative particle. In contrast, we see that the older speakers studied by Diller, especially the upper-class speakers, have very high scores, as do the numerous, and demographically heterogeneous, speakers who contributed to the Gougenheim corpus recorded in the

1950s. Ashby's middle-class Parisian informants have an overall rate of ne retention a little higher than that of the male celebrities, and fairly substantially higher than that of his Tours informants from diverse social backgrounds. It is interesting to note that, in both his studies Ashby found massive interpersonal variation. Among the middle-class Parisians, two speakers retained ne 100% of the time (they were both older females), but another used ne in only 8% of cases. In Tours one speaker (a 15 year-old apprentice waiter) omitted ne consistently, whilst another (a male final-year lycéen from an intermediate socio-educational background) had the astonishing retention rate of 94% (Ashby, 1981:676-7, 1988:696).

3.7.3 Children's use of (ne)

One section of society which is represented rather poorly in these surveys is children, and yet the systematic omission of ne is a well-known characteristic of children's speech. Pohl (1968:1351-2) reports on the usage of two 5 year-old twins, from a middle-class family from Mons: the girl's rate of retention was 28%, while the boy's was just 4%. Moreover, there was evidence that on the occasions when the boy did use ne, he did so apparently haphazardly and without intending it to convey negation. Pohl also comments that the use of the negative particle increases once the child begins school, and at the age of 12, the twins were using ne 39.4% of the time (Pohl, 1975:20). Labelle (1976, cited in Lemieux, 1985) found that almost all of twenty 5 year-old informants from Paris and Montreal omitted ne consistently. The 9 year-olds, however, had an average retention

rate of 9.9%. (It is not known in what situations the recordings were made.) Lemieux (1985:95-6) concludes, with Pohl, that children in both Paris and Montreal appear to begin using the negative particle after they have started school.

3.8 Non-tokens and categorical tokens

3.8.1 General exclusions

Let us now move on to the analysis of the variability of (ne) in the present corpus. It will be recalled that two concordances were produced, one containing all occurrences of ne, n' and n, and the other with all instances of the various negative items, aucun, jamais, ni, pas, personne, plus, que and rien. The next step was to scrutinise these concordances in order to eliminate from the subsequent analysis all occurrences which were either not tokens at all of the variable (ne), or which were judged to be categorical, ie where the speaker effectively had no choice as to which variant to use.

Following Sankoff & Vincent (1980) and Ashby (1976, 1981), it was decided to exclude tokens where on is followed by a vowel, since it does not seem possible systematically to distinguish between a liaison [n] and a [n] representing ne:

(13) mais on (n') avait pas faim. / (3166)

More generally, all instances where the word immediately preceding the locus of the variable ends in a [n], and where it was impossible to determine whether ne was present, were excluded from the quantification:

- (14) rien (n') est ouvert quoi. / (2312)
(15) nous sur la troisième chaîne y en a jamais. / (25574)

John Green (personal communication) has in fact suggested that the liaison [n] and the ne [n] may be pronounced slightly differently, and there is no doubt that, in slower more careful styles, it is quite possible to emphasise a n' prosodically with additional volume and/or duration, or by pronouncing a preceding word-final schwa, as in the following examples:

- (16) On n'en sait rien.
(17) Personne n'est venu.

It is true that, in quoted utterances, speakers often produce the variant which seems more appropriate to the original "author" of the utterance and the situation in which it was produced, as in (18), where the (hypothetical) author was a camp director speaking to a parent on the telephone, and (19), where the speaker is quoting very young children:

- (18) ... on explique aux parents "voilà / euh nous nous n vérifions ni euh (comment) - ce que écrivent les enfants / et on ne / on - on ne vérifie pas combien de lettres ils écrivent" ...
(23212)
(19) ... au début les enfants [disent] "non je sais pas" ...
(15414)

Nevertheless, speakers clearly retain at least a measure of freedom as to which variant to choose, witness the following quotations, very similar to the last two, where the other variants are used:

- (20) ... on explique aux parents "votre gamin est libre il il écrit quand il le veut ... c'est pas à nous de régler ça ..." (23303)
(21) ... pour d'autres [enfants africains] euh ça va très loin jusqu'à des attitudes physiques / qui sont par exemple euh "je n sais pas m'asseoir à table" ... (144524)

In view of the latitude which speakers have when making quotations, then, it was decided to include such tokens among those to be quantified.

Repaired utterances, however, were excluded when the repaired material was repeated identically, since, in such cases, it seems that the speaker is, in effect, deleting the "false start" which they have just made. In the following example, therefore, the first c'est pas was not counted:

- (22) ... on essaie d leur dire que c'est pas - c'est pas la meilleure solution que de fumer ... (2228)

Cases of repair where the other variant was used, however, were retained (cf example (5) above).

3.8.2 Excluded occurrences of the negative particle

Among the occurrences of the negative particle excluded from the quantitative analysis were certain expressions where it is used on its own, without the support of an auxiliary. Although these expressions clearly involved negation originally, it is arguable that they have since lost this, witness their closest equivalents in English: anything (or no matter what), nevertheless, if only:

- (23) "n'importe quoi mais je veux pas manger chez les Anglais" (13980)
 (24) mais n'empêche que / quand un curé venait nous dire ... (13161)
 (25) ça m'a pas laissé un très bon souvenir. / si c n'est le souvenir du fric ... (495)

In more common equivalents of the expression in (25), ne is supported by que, as in the versions in (26) and (27), which also require the presence of ne (and are therefore categorical tokens to be excluded):

- (26) ... à partir du du petit déjeuner plus copieux disons un peu plus / et anglais même bon / si c n'est que de mettre du bacon ... (21025)
 (27) y a certaines règles qu'i faut respecter / hein ne serait-ce qu de se lever de bonne heure / ... (11420)

Finally, there is just one occurrence in the corpus of ne being used alone to negate the verb savoir, and this could perhaps be regarded as a fixed locution, functioning as a compound quantifier, and roughly equivalent to beaucoup:

- (28) ... qui n'étaient pas allés au cinéma depuis je n sais
combien de temps quoi. / (4456)

3.8.3 Excluded occurrences of the negative items

The number of non-tokens in the concordances of the negative items was substantially larger than in those of the negative particle. Firstly, there were instances of these items being used with a non-negative sense, for example, pas in (29), aucune and rien in (30) and jamais in (31) (one of many such instances of its use as a reinforcer of si):

- (29) y a un pas à franchir qui est pas qui est pas facile quoi. / (3666)
(30) j veux dire sans aucune moquerie sans rien ... (2939)
(31) "... tout est en bois / euh si jamais ça crame là-dedans qu'est-ce qui va s passer? " / (2235)

Although aucun, jamais and rien can also be found in affirmative interrogatives, as in (32) and (33),

- (32) Y a-t-il aucune raison pour ça?
(33) Avez-vous jamais rien entendu de si absurde? (both examples from Byrne & Churchill, 1986:431)

there are no such occurrences in the corpus, and informal observations suggest that speakers are now tending to avoid this use, perhaps because the negative value of these items is so much more common than their positive sense, and there is a risk of them being interpreted as negators in interrogatives.

Secondly, cases where these items were used without a verb were

similarly excluded, since they do not, of course, provide a locus for (ne). All the negative items can negate other constituents in such contexts, and aucun, jamais, rien and personne can occur as one-word utterances:

- (34) on fait des réunions préparatoires (au deb) - fin pas
préparatoires mais des réunions de début d camp ... (2277)
(35) et y a un moniteur / mais que pour les dortoirs ... (5129)
(36) mais non le stage euh / rien - je trouve que c'est une
histoire de fric ... (573)

A third, and numerically quite significant, set of non-tokens involves the collocation pas mal. Very often it is quite clear that this is used as a colloquial synonym of beaucoup (which itself, of course, derives historically from a collocation), as in the following examples, where we see that the literal meaning becomes progressively less plausible:

- (37) Bertrand a fait euhm - m'avait pas mal expliqué ... (3143)
(38) ça nous avait pas mal plu ... (3164)
(39) à la maison j'achète pas mal de céréales ... (3218)
(40) j commençais à paniquer pas mal ... (2540)

Additionally, we may note that the insertion of ne would produce syntactically acceptable sentences in examples (37) and (38), since pas mal is functioning as a complement of the verb, whereas in (39) it is a quantifier within the NP. In (40), on the other hand, pas mal is an adverbial modifying the infinitival complement, and since the pas is not located after a tensed verb, the insertion of ne is impossible syntactically in the contemporary spoken language. In other cases, pas mal functions as an equivalent of bien, sometimes quite unambiguously, thanks to the presence of déjà, which cannot occur in negative contexts:

- (41) l'avis satisfaisant c'est déjà pas mal / ... (201243)
(42) *Ce n'est déjà pas mal.

In other cases, the fact that pas retains its negative value is conveyed by the use of the reinforcing expression non plus, which

is, naturally, restricted to negative contexts:

- (43) mais d'un autre côté c'est pas mal non plus qu'ils se qu'ils
 changent de de personnes qu'ils voient d'autres gens ... (6122)
(44) *C'est bien non plus que ...

Occurrences such as (43) were, then, retained for the quantitative analysis, as too were ambiguous cases where pas mal could be interpreted as equivalent to either pas mauvais or bien:

- (45) moi je trouve que c'est pas mal ce truc ... (3217)

A fourth set of occurrences clearly involves negation, unlike some of the cases discussed so far, but was excluded from subsequent analysis on the grounds that the speaker did not have a choice of variant. The most numerous occurrences in this category involve an omitted or ellipted subject, where the following verbal group begins with a consonant (but not a semi-vowel), and where consequently the absence of ne is obligatory, as in the following examples:

- (46) ...fin je veux dire être euh / faut pas être mou quoi ...
 (18461)
(47) j sais pas très bien / sais pas / ... (1553)
(48) bon mais - peut-être qu'avec les enfants que l'on a là euh /
 sont pas tellement motivés ... (11302)
(49) donc c'est vraiment ma première puisque pouvais pas en faire
 avant ... (2316)

(It is assumed here that the speaker's decision to omit the subject precedes the choice regarding the negative particle.) It seems that cases such as these have, in previous studies, been included as occurrences of ne deletion, but since in the present corpus there are some forty such tokens, the decision to include or exclude them is not a trivial one for the subsequent quantification. Verbs in the imperative do not, of course, fall into this category, since it is quite possible to omit or retain the negative particle even when the verbal group begins with a consonant:

-)
- (50) "mm t'inquiète pas ..." (21395)
 (51) eh ne recassez pas de boîtes hein / [SHOUTING TO CHILDREN]
 (4815)

(If, on the other hand, the verbal group begins with a vowel, ne seems far more likely to be retained.) Rather more controversially perhaps, it was also decided to exclude the small number of examples of the non-standard structure pour pas que, where pas can be said to have been "transported" from the subordinate clause to the subordinating conjunction which introduces it. Native-speakers have suggested that ne is not possible in this context:

- (52) le centre leur a donné - leur a donné un peu d'argent quoi.
 / pour pas qu'ils soient trop / (23693)

However, it must be pointed out that examples of this structure with ne retained are attested in literature representing non-standard speech (Grevisse, 1986:1489), and even more significantly, Pohl has in fact reported two such occurrences in speech (1968:1355).

In addition to these four categories of non-tokens and categorical tokens, it was, of course, necessary to delete from the files containing the concordances of the negative items all those tokens where ne was retained.

3.9 Editing the files of the concordances

For the next stage of the analysis, the two files of occurrences of ne on the one hand, and of tokens of the various negative items, on the other, were edited so that OCP could be run on them in turn, to produce wordlists of the words most closely involved with the variable. Each token of the variable was already on a

separate line of the file, but it was also necessary to delete any extraneous linguistic material which would invalidate the figures obtained on the wordlists: for example, any verb other than the negated verb in question, any additional instances of the negative items or the proclitics, and also any items which in a wordlist would be indistinguishable from such items, eg the demonstrative determiner ce. Many instances of expressions which are written as two or more separate words were modified so as to appear as a single item in the wordlists (eg en plus as enplus, tout à fait as toutafait, (il) y a as (il) ya), and certain words were given a diacritic to distinguish them from significant homographs, eg à was represented as a+, as accents were not available. Negated infinitives and present participles were tagged with an initial \$, so that they would appear as a group at the end of the wordlist. NP subjects, it will be recalled, had already been tagged with a following %.

(53) and (54) show extracts from the edited files of tokens where ne is present and absent, respectively:

- (53) 141 de n pas \$donner
 186 qu'i n faut pas laisser //
 1264 qui n'est pas ferme
 (54) 15 j'avais aucune formation
 1100 / i faut jamais laisser
 144 / j suis peutetre pas toutafait

(The initial number on each line serves to locate the token in the corpus.) The subsequent production of OCP wordlists of these files had the considerable advantage of providing automatic counts of the various words in the files, which could then be used in the study of linguistic factors which constrain the variability.

3.10 Linguistic constraints

3.10.1 Introduction

There is, of course, a very large number of contextual factors which could conceivably exert an influence on the variable presence or absence of ne. Indeed, Ashby (1976, 1981) has examined many such factors, including the type of clause (independent/dependent; if dependent, then relative/infinitive/conditional etc), type of verb (transitive/intransitive/impersonal), role of the negated verb (auxiliary/modal/main/non-finite), presence or absence of one or more object proclitics, preceding and following phonological environment, the accompanying negative item (pas, jamais etc) and the subject of the verb.

Whether it is conceivable that such a large number of different factors should simultaneously constrain the variation is a moot point. In the present study, at any rate, only a few of these factors will be examined, and these will be factors where there could be a functional motivation for a differential use of ne.

For this analysis of linguistic constraints, the 2,932 variable tokens from all thirty informants have been grouped together. This has the advantage of ensuring that there are sufficient numbers of tokens for all the factors to be examined. There does not seem to be any theoretical objection to this procedure, since it appears that no study of variability has yet suggested that linguistic constraints operate differentially within a speech

community.

3.10.2 The subject of the verb

Let us first consider the effect of the subject of the verb, which has been found to be a significant constraint in several previous studies. The different rates for the retention of ne according to subject are shown in table 3.2.

Table 3.2 Relative frequencies of the retention of ne
according to subject (variable tokens only)

Subject	ne retained	N=
Noun Phrase	67.2%	174
None (infinitive verb)	62.5%	40
None (present participle) (50%)		4
qui	44.3%	140
nous	(85.7%)	7
elles	(66.7%)	6
vous	38.9%	18
elle	33.3%	54
ça	29.7%	195
ils	27.9%	251
on	15.9%	233
il	12.2%	131
je	11.5%	977
tu	10.4%	58
ce	2.3%	433
—	—	—
All clitics	14.6%	2,363
—	—	—
All variable tokens	18.8%	2,932

(Note: Tokens where the subject is ellipted, and where the segment following the locus for (ne) is a consonant, have been excluded.)

Given that the overall rate of retention was 18.8%, we find some very striking differences here, in particular between those for NPs and qui in comparison with the clitic subjects. The high rate

of retention when there is no subject and the verb is in the infinitive is also noticeable.

The strength of this constraint has already been noted by Ashby (1982) and Diller (1983), among others. Diller pointed out a correlation between rate of ne retention and the semantic weight of the subject, but it is not clear how this could be said to account for the constraint. A different line of explanation starts from the observation that the clitic subjects collocate with negated verbs far more frequently than do NPs, not only because they are about ten times more frequent in discourse, but also because they form a closed class, whereas NPs are, of course, an open, indeed, infinite set. As frequent collocations tend to become eroded phonetically, the communicatively redundant negative particle is clearly a strong candidate for such erosion. The relative pronoun qui, which is also very frequent in discourse, is not a clitic (witness example (55)), even though, like them, it undergoes elision in informal speech, as in (56):

(55) ... qui visiblement n'étaient pas d'accord / ... (28451)

(56) ... y en avait un qu'avait pas reçu de lettre ... (18778)

We may suppose also that, given the greater length and diversity of NP subjects, there is perhaps a tendency for speakers to require more time in producing them and hence for their use to slow down the rate of speech somewhat. This, in itself, may increase the speaker's degree of self-monitoring, which in turn would lead to a tendency to use more formal variants, especially those, like ne, which are associated with written language. With regard to the high rate of ne retention with infinitives, several factors may be relevant. Firstly, as we have seen in section 3.2, the use of ne in such contexts can occasionally be motivated by a

need to avoid ambiguity. Secondly, it may be that since ne is normally adjacent to the negative item in such contexts (items which precede the infinitive, such as pas, rien and plus being far more frequent than those which follow, such as personne, aucun and que), the auxiliary serves to support the ne, just as in Latin ne had merged with the adjacent oenum to give noenum (and then non), rather than disappearing completely (Meillet, 1982:140). Thirdly, it is interesting to note that infinitives and present participles have been conservative contexts previously in the history of negation in French: the use of non to negate the verb in Old French continued longer with them than with other verb forms (Grevisse, 1986:1477). The conservatism of infinitives and present participles can perhaps be attributed to their relative infrequency in discourse, as well as to another factor, to be discussed presently.

Another suggested interpretation of this constraint is that the cliticisation of these pronominal subjects has "squeezed out" the redundant negative particle (Ashby, 1981:681; Harris, 1978:118). There is, of course, ample evidence of the clitic status of these forms. Firstly, they can be separated from the verb only by other clitics: the direct and indirect object pronouns, the reflexive pronouns and ne itself: hence the ungrammaticality of (57):

(57) *Je souvent sors.

Secondly, they often undergo elision and assimilation, as in (58):

(58) [ʃtRuv] for je trouve; [ta] for tu as; [ivwa] for il voit,
or ils voient; [zave] for vous avez.

Thirdly, it has been suggested that there is an increasing tendency for them to occur even when there is a NP subject

(Harris, 1978:119), for example:

(59) et puis la discussion ça marche pas toujours / (18827)

However, it is important to note that the cliticisation of pronominal subjects need not entail the elimination of other preverbal items. In Swahili, subject prefixes coexist with preverbal negative markers in imperative and subjunctive verb forms, as in the following examples:

(60) Wa-si-end-e 'Let them not go' (Ashton, 1947:70)
they + not + go + SUBJUNCTIVE

(61) U-si-mw-ambi-e 'Don't tell her/him' (Polomé, 1967:110)
you + not + to her/him + say + SUBJUNCTIVE

One crucial contrast with French, however, is that in Swahili there is normally no other marker of negation, and therefore the negative infix is communicatively essential, unlike ne in French.

The extremely frequent collocation of ce with est is no doubt the reason for its very low rate of ne retention. But if it is the frequency of certain collocations which favours higher rates of omission of ne, then perhaps we should be taking into account the larger grammatical context more systematically. We shall take up this point again in 3.10.6.

3.10.3 Negative items

As for the negative complements, the differences in the rates of omission of ne are much slighter, as we see from table 3.3. (Only the first negative item is taken into account here, ie instances of, for example, jamais in plus jamais are not counted. There were no occurrences in the corpus of the items associated with formal styles, nul, nullement, point and guère.)

Table 3.3 Retention of ne according to negative item

Negative item	ne retained	N=
que	34.9%	109
personne	33.3%	24
ni	(33.3%)	3
jamais	26.2%	84
plus	25.8%	209
rien	21.2%	146
aucun(e)	21.2%	33
pas	16.4%	2,317

Total		2,925
(All variable tokens		18.8%
		2,932)

(Note: In a small number of occurrences the negative item is inaudible or absent.)

Such differences as there are among the various negative items are more difficult to explain in functional terms, than was the case with the different types of subject. The higher rate of retention for que may be due to a need to reinforce it (albeit at a distance), since, with the elision of the schwa, que often consists of just the one segment [k]. Ashby (1981:678-9), who also found a higher than average retention rate for que (but also for plus), suggested that "que and plus may be less categorically negative than other second negatives". By this, he was presumably referring to the fact that que is a restrictive, not a negative adverb, and that sentences with plus presuppose that the proposition was formerly true. It is not clear, however, why this

should lead to a greater use of ne. On the whole, then, it is difficult to see any very clear pattern in these results, except for a slight tendency for the more frequent items, rien, plus, and especially pas, to have lower rates of retention than the average.

3.10.4 The presence or absence of a non-subject preverbal clitic

As we have already seen, it has been suggested that an important factor leading to the historical loss of ne was the cliticisation of preverbal pronouns, not only subjects, but also reflexives, direct and indirect objects, and the "adverbial pronouns" y and en (Harris, 1978:118; Posner, 1985:188). If this was indeed a major factor, then we might expect to find a similar constraint continuing to operate in synchronic variation today. Table 3.4 shows the retention rates for all tokens where there is at least one preverbal non-subject clitic. For each token, only the first such clitic has been taken into account, so that no token is represented twice in the table. Tokens involving infinitives have been excluded here wherever the negative item (most frequently pas) occurs before the verb, since in such cases ne and the other clitic(s) are not adjacent. Apart from that, however, all tokens have been considered, including those which have a non-clitic subject, such as qui or a NP.

Table 3.4 Retention of ne in tokens with a non-subject proclitic

Proclitic	ne retained	N=
le/l'/la	19.1%	68
les	7.7%	26
lui/leur	(33.3%)	12
y	37.5%	16
en	13.8%	58
me/m'	17.7%	62
te/t'	(11.1%)	9
se/s'	34.6%	81
nous	0%	11
vous	-	0
All proclitics	21.3%	343
All tokens	18.8%	2,932

From table 3.4, it appears that the presence of a non-subject clitic does not particularly inhibit the presence of ne: in fact, the rate of retention of ne is slightly higher in such contexts than in the corpus as a whole. This evidence does not necessarily disprove the theory that a similar constraint operated historically - it is conceivable that the constraint has simply ceased to operate, especially in the case of a group of speakers where the loss of ne is not far from completion. However, these results must be said to call into question the existence of such a historical constraint, and make the need for quantitative data from the past even more apparent. One final, and very minor, point concerning the data in table 3.4 is that ne is not retained in any of the few contexts where nous immediately follows the

locus of the variable: as was pointed out in 3.4, it seems that a potential [nən] sequence within a proclitic array, increases the likelihood of ne being omitted.

3.10.5 Phonological environment

Another factor which certainly merits investigation as a potential constraint is the phonological environment. In particular, it has been suggested (eg by Pohl, 1968:1354) that ne tends to be retained where its presence would prevent a hiatus between two vowels, as in ça n'a pas, qui n'est pas etc.

However, a serious problem is encountered in attempts to carry out a quantitative analysis of this. Ashby investigated phonological environment as a factor affecting (ne) in both the Paris and Tours corpora. In the former, 49/1019 tokens (0.5%) were identified as being intervocalic (1976:29), but in the Tours corpus 894/2818 tokens (31.7%) were so categorised. As there seems to be no reason why the number of intervocalic occurrences should vary so greatly from one body of speech to another, it must be concluded that different defining criteria were applied in each case. This may well be quite justified, but it serves to illustrate that deciding on such criteria is more problematic than might be supposed.

The main problem, and a possible explanation for the discrepancy referred to above, would seem to be as follows. As we have seen, approximately 80% of the occurrences of this variable in the present corpus have clitic subjects, with je and ce being

particularly numerous. Now, if ne is present, the preceding phonological environment in such cases will normally be [ʒə] and [sə]. But if ne is absent, and the following segment is a vowel, the schwa will automatically fall, and the preceding segment will now be a consonant. The transformation of the preceding phonological environment is even more radical when the following verbal group begins with a voiceless consonant, as in (62):

(62) je te vois pas [ʃ tə vwa pa] ; je sais pas [ʃs pa] ;
je suis pas [ʃɥi pa].

That is to say, not only is there a loss of the schwa, but also assimilation, and in the last two examples in (62) the assimilation is coalescent, involving the merging of [ʒ] and [s] to give [ʃ].

In such instances, when the phonological environment is so radically different according to whether the ne is present or absent, it seems more reasonable to say that the grammar is constraining the phonology, rather than vice versa. One solution might be simply to exclude such cases from a quantitative analysis of this constraint. This would, of course, affect a large proportion of all tokens: not only those with je and ce, but also those with tu, qui and ça when these are followed by a vowel, since they too undergo elision, for some speakers, if not all. Instead of proceeding along these lines, it was decided in the present study not to attempt any general quantification of phonological environment, but rather to bear it in mind as a contributing factor affecting certain sub-sections of the data.

3.10.6 Collocations or "preformed sequences"

Lüdicke (1982:43-7) has demonstrated that factors which have sometimes been considered as independent constraints on (ne) may interact very significantly. For example, in negated forms of être or avoir, the segment following ne is usually a vowel (80% of cases), whereas with other verbs, a consonant is far more common (90% of cases). Therefore, the grammatical factor, "nature of the verb" (avoir/être vs others), is not independent of the phonological factor, "following segment" (vowel vs consonant). The real effect of these two factors can only be known, therefore, if a sophisticated statistical analysis is carried out, such as that performed by the VARBRUL program.

Another indication of the interaction of certain linguistic constraints on (ne) is the common observation that ne is omitted extremely frequently in particular collocations: Pohl, for example, has commented that he himself often omits ne in c'est and il y a (1968:1356), and Ashby (1981:678) found that ne was omitted more frequently in il (n') y a pas, je (ne) sais pas, il (ne) faut pas and ce (n') est pas than in all other cases taken together.

Moreau (1986) has taken this a stage further, pointing out that much spoken language consists of more or less set sequences, rather than of unique and novel sentences, as has sometimes been supposed. Observing that, in most of the contextual factors which she analysed (subject, verb, tense, negative item), the ne omission rates tended to be higher for the more frequent

exponents of these categories, she went on to investigate quantitatively the relevance of "preformed sequences" to the variability. The sequences which she selected were the most frequent ones in her corpus, and comprised a singular clitic subject, a form of one of the twelve or so commonest verbs and the negative item pas. Only tokens which had the three elements following immediately after each other were counted as examples of the sequences, so that any tokens with preverbal non-subject clitics or with degree adverbs preceding pas were not included. As Moreau had predicted, the ne omission rate for these sequences was found to be much higher than those of other occurrences.

As Moreau herself acknowledges (1986:151), there are a good number of other candidates as preformed sequences (such as je ne sais plus), and the only reason for not quantifying them also was that they were insufficiently frequent to justify distinct categories. In total, the occurrences which were categorised as sequences accounted for 34.2% of the data.

In adapting Moreau's basic idea to the present study, it was decided to quantify separately all negative sequences comprising a verb form preceded by any subject clitic or qui, which occurred at least nine times in the corpus. Additionally, the following four sequences were included: (il n') y a, (il n') y avait, (il n') y en a, je (n') en sais rien. The major difference from Moreau's approach, then, was the decision not to require the presence of pas immediately following the verb, thus taking in instances of, for example, je (ne) sais plus, and of (il n') y a même pas. By doing this, it was possible to include, as

sequences, a far greater proportion of the tokens of this variable: 1,778/2,932 (60.6%). But quite apart from this advantage, this decision seems justified by three other factors. Firstly, since affirmative clauses are far more numerous than negative ones, it could be argued that it is essentially the subject + verb sequence which speakers operate with, regardless of whether a negative item follows or not. Secondly, there is evidence from pausing and hesitation phenomena that speakers do indeed operate with these subject + verb sequences, as in extracts (63) and (64):

- (63) d'un autre côté c'est - pas mal non plus qu'ils se qu'ils changent de de personnes qu'ils voient d'autres gens ... (6122)
(64) je sais pas si je ferai euh / que qu'une colo au mois d'août ... (4600)

Thirdly, it seems quite likely that pas, especially after a form of the copula, forms part of a collocation with following adjectives or adverbs, all the more so as such collocations themselves occur very frequently as independent utterances: pas mal, pas cher, pas évident, pas possible, pas facile, pas solide, pas tellement, etc. (The first two of these occur so frequently that they might almost be thought of as single lexical items meaning, respectively, "good" or "a lot", and "cheap".)

The sequences which were selected for separate quantification are shown in table 3.5, in descending order of frequency, together with their rates of ne retention.

Table 3.5 Retention of ne in preformed sequences

sequence	ne retained	N=	sequence	ne retained	N=
ce/ça est	3.6%	386	j'étais	5.9%	17
je sais	8.6%	315	elle a	18.8%	16
j'ai	8.7%	172	il a	12.5%	16
(il) y a	2.4%	165	il est	0%	15
c'était	0%	56	qui est	33.3%	15
on peut	5.6%	54	ça va	13.3%	15
ils ont	44.2%	52	je veux	42.9%	14
(il) y avait	0%	44	qui ont	61.5%	13
je suis	14.3%	42	j'en sais rien	0%	13
j'avais	8.1%	37	je crois	9.1%	11
je peux	17.7%	34	on fait	9.1%	11
je pense	18.5%	27	on va	0%	11
ça a	52%	25	ça marche	0%	10
ils sont	9.1%	22	on pouvait	10%	10
ils avaient	30%	20	qui sont	40%	10
je connais	10%	20	je vois	20%	10
il faut	10.5%	19	qui a	55.6%	9
tu as	16.7%	18	j'aime	0%	9
on sait	27.8%	18	tu peux	0%	9
(il) y en a	0%	18			
All the above sequences			10.1%	1,778	
All variable tokens			18.8%	2,932	

The most important point to arise from table 3.5 is the substantial difference between the ne retention rates for the selected collocations taken as a whole and all variable tokens. It is clear that ne is considerably less likely to be retained in frequently occurring collocations than in other contexts. Moreover, this tendency is not solely attributable to the contribution of the handful of extremely frequent sequences, c'est, je sais, j'ai and (il) y a. Although these do indeed have very low ne retention rates, there are others which have similarly low scores, most notably, perhaps, c'était and (il) y avait both with zero.

In looking at the rates of individual collocations, we can, of

course, have less confidence in the frequencies which are produced from a rather small number of occurrences. This said, a number of the sequences are worth closer examination. Firstly, it is surely no coincidence that the three which have by far the highest rates of ne retention (over 50%) are qui ont, qui a and ça a, which all involve a potential hiatus when ne is omitted. Naturally, speakers of French are perfectly capable of producing these sequences without ne - they have to do so in any case in affirmative clauses. But it is generally acknowledged that French phonology tends to avoid hiatus where possible. Many speakers in fact avoid the hiatus by eliding the vowel of the pronoun, as we saw in 3.10.5, but it may be that elision takes place less readily with the three sequences in question than with others such as tu as and qui est, which have lower rates of ne retention. Phonological factors may also be playing a role in two other sequences where the negative particle is used more frequently than the average. Ils ont (44.2%) and ils avaient (30%) are the only sequences in table 3.5 which involve a more or less obligatory liaison if ne is omitted, and it may be that speakers feel that this would highlight the absence of the negative particle, which is still, of course, considered a non-standard feature. Not all the instances of higher rates of ne retention can be accounted for, however: there seems no particular reason why je veux, qui sont and on sait should have scores of well above the average.

Looking now at the sequences in which ne is omitted consistently, it is noticeable that three of these - c'est, c'était and il est - include third person forms of the copula, which are often

followed by other frequent sequences of pas combined with an adjective or degree adverb plus adjective: pas (tellement) évident, etc. Two other collocations which have zero ne retention rates are (il) y en a and (il) y avait. Together with the score of 2.4% for (il) y a, these may again be due in part to a phonological factor: once the impersonal pronoun il has been omitted (as it is extremely frequently), there is a very strong tendency to also omit ne, in order to avoid the rare initial cluster [nj]. (There are very few common words which begin with this (eg nier and nièce), or with the articulatorily similar palatal nasal [ɲ].) Four of the other sequences where the negative particle is deleted consistently in the corpus involve non-standard or markedly colloquial features which we would expect to find collocating with Ø rather than with ne. Firstly, there is j'en sais rien, which generally functions as an emphatic equivalent of je (ne) sais pas:

- (65) A: ... est-ce que les gens en général vont à l'église dans ta région? / est-ce qu'ils sont pratiquants * ou? * /
 B: * moi j'en * sais rien du tout. / ... (27365)

(12 of the 13 instances of j'en sais rien were produced by male informants.) A second such sequence, ça marche, is a colloquial equivalent of ça va or ça fonctionne:

- (66) ... parce que ça marche pas bien mes études. / (23191)
 (67) ... et puis la discussion ça marche pas toujours / (18827)

Thirdly, in several of the occurrences of on va, aller is used as an auxiliary verb, and on has definite reference, ie as an equivalent of nous:

- (68) et puis on s'est dit "on va pas rester sur une mauvaise impression" on est revenus. / ... (3350)

And finally, in the majority of the examples of tu peux, tu has indefinite rather than definite reference:

(69) ... puisque pour travailler avec des gosses handicapés faut quand même un petit peu les connaître tu peux pas les envoyer en centre comme ça tout seuls c'est pas possible. / ... (4134)

The use of aller as an auxiliary is a somewhat colloquial feature, and the use of tu instead of on is non-standard (cf Emirkanian & Sankoff, 1985; Laberge, 1977, 1983), and both would therefore be expected to co-occur with the absence rather than the presence of ne. (It is interesting to note that Moreau found je (ne) vais pas to be the sequence in which ne was omitted most frequently in her corpus (1986:152).)

On the whole, the frequencies in table 3.5 correspond reasonably well to the findings of Moreau (1986:152), and such differences as there are may result, not only from a divergence in methodology at certain points, but also from some major differences in the type of data. For example, it is hardly surprising that the celebrities studied by Moreau should produce larger quantities of aller + infinitive than the informants in the present study, since such persons are probably expected to be able to speak about their future rather more than the average layperson.

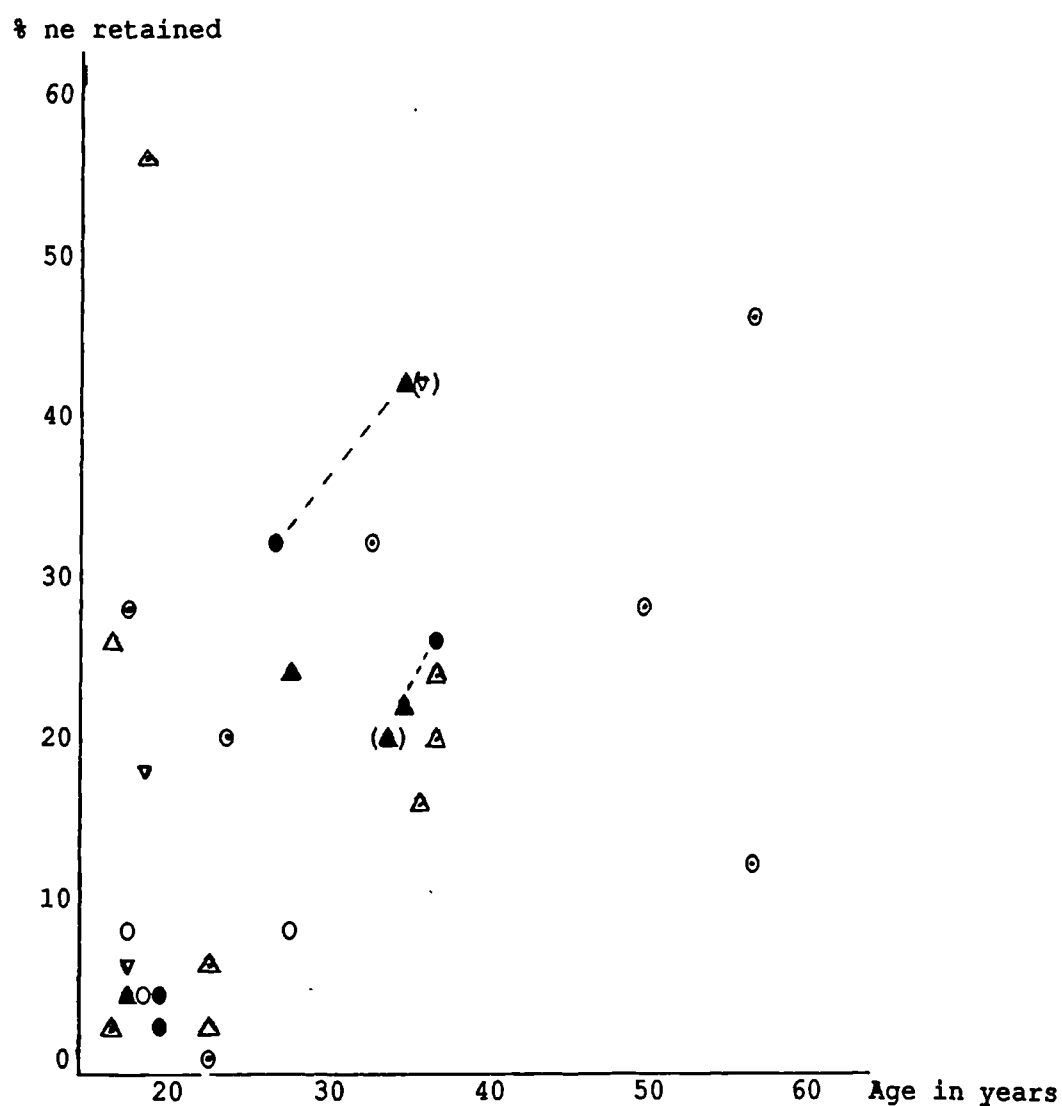
It is interesting to note that, from the viewpoint of generative grammar, which considers a language to be an infinite set of sentences (with a greater potential for NP than pronominal subjects), it might be said that ne omission is very much a minority phenomenon, since it affects less than one third of NP subjects (cf table 3.2). On the other hand, a sociolinguistic approach to language (in common with other functional approaches) is concerned with language in use and with the overall frequency

of variants in speech communities. Since pronominal subjects are far more frequent in spoken language than are NP subjects, from a sociolinguistic point of view, the absence of ne will be considered a majority phenomenon.

3.11 Interpersonal variation

Let us now move on then to examine the evidence for interpersonal variation in the corpus. Figure 3.1 shows the relative frequency of ne retention for each of the thirty informants in this study. The three inverted triangles represent the three male informants who could not be categorised for social class, and the two triangles in parentheses represent the scores of two male informants who produced so few tokens of (ne) that the relative frequencies cannot be considered reliable: 13 in one case, 8 in the other. (This was not due to any mysterious aversion to negation, but rather simply due to the fact that, in both cases, the interviews were cut short by interruptions.) Despite this, it is noteworthy that the relative frequencies produced from such small numbers of tokens are very much in line with those of similar informants.

Figure 3.1 Retention of ne by individual speakers



	Female	Male
Upper class	●	▲
Intermediate class	⊙	△
Working class	○	△

▽ = males from unspecified social class

() = frequency calculated from less than 15 tokens

--- = married couples

The speaker with the highest frequency is a 19 year-old intermediate-class male. His score is so out of line with the other young informants that it requires some explanation. This speaker was already an assistant camp director, and it may be that he felt a responsibility to present, in his interview, a favourable image of the camp, and that this led him to adopt a more formal style of speech. In addition, he had just spent a year as an unqualified primary school teacher, and this work may well have eroded his vernacular style somewhat: much sociolinguistic research has pointed to the influence of occupation on linguistic variability (cf the notion of the linguistic market, as widely used in research on Montreal French: Sankoff & Laberge, 1978). However, these factors do not seem to provide the entire explanation, since there were other young adults who were either assistant directors or trainee teachers (though not both, in fact), who have rates of ne retention a good deal lower than this informant. The crucial factor would seem to be simply that he felt that it was necessary to speak in a more formal style during the interview. Indeed, at the beginning of the recording, he asked whether he should speak slowly so that English listeners could understand more easily, and despite the efforts of the fieldworker to persuade him that this was not necessary, he did, in fact, speak in a more careful style. (It may also be significant that this was the first interview conducted by the fieldworker, and that he failed to put the informant at his ease.) In view of such indications that this speaker did not, then, speak in a style very close to his vernacular, there seems to be a strong case for considering his score apart from those of the other informants.

Let us now consider more generally whether there are any correlations between the linguistic variability and the three speaker characteristics of age, sex and social class. Of the three, it appears that age is by far the most important parameter. Among the fourteen speakers aged 17-22 years, ten have scores of less than 10%, and the only informant older than 23 who has such a score is the 28 year-old working-class woman. Moreover, she resembled the younger speakers, rather than most of the older speakers, in that she was neither married nor in employment, and since it seems likely that both marriage and employment often have a conservative, or "vernacular-eroding", influence on one's language (because of the breaking down of peer networks, and increased pressure from standard norms), it is not surprising that her behaviour with regard to (ne) should be more like that of the younger informants.

In contrast, it does not seem that the sex of the speaker is a very significant differentiating factor in the variability, with the possible exception of the 24-37 age range, where three women have higher scores than five men from similar social classes. Incidentally, we may note that among the informants were two couples who had been married for several years - their scores are linked by lines in figure 3.1. Although they were recorded separately, it is striking that in both cases the scores of the partners are rather similar. This could be because, after several years of living (and, presumably, talking) together, they had accommodated to each other linguistically. There is considerable evidence from research into the social psychology of language

that individuals make their speech more similar to that of interlocutors whose approval they seek (Giles & Smith, 1979). Speech accommodation theory is based, in part, on similarity-attraction theory, and it is, of course, also possible that in the case of these two couples, their sociolinguistic similarity predated (and perhaps even assisted) their mutual attraction.

As for the social class of the speaker, it is again difficult to discern any very clear differences. It is true that four of the five working-class informants have scores below 10%, but this may simply be due to the fact that they are also young. What is rather striking is the relative homogeneity of the younger speakers, regardless of sex or social class. On the whole, the younger speakers omit ne on a massive scale, not so very far removed from the near zero usage found in Montreal.

Another way to look at the social differentiation of (ne) is to consider the mean frequencies for various sub-groups of speakers. Table 3.6 shows the frequencies of ne retention for sub-groups of the speakers according to age, sex and social class. These scores have been arrived at by averaging those of the individuals belonging to each subgroup, not by aggregating the data from individuals and then calculating the average of that. (This latter method, although unavoidable when many individuals have low numbers of tokens, has the unfortunate consequence of biasing the average scores in favour of those informants who produce more tokens.) The two unclassifiable young males have, of course, been omitted from the social class groupings, but not from those of

age and sex. The two informants who produced very small numbers of tokens have been excluded entirely, as has the young male with the anomalously high ne retention rate, since his inclusion would have distorted the average scores of the sub-groups to which he belongs.

Table 3.6 Frequencies of ne retention for groups of speakers
according to age, sex and social class
(categorical tokens excluded)

Group (no. of speakers)	<u>ne</u> retained	Average N per speaker
17-22 yrs (13)	8.4%	87.3
24-37 yrs (11)	23.9%	131.9
50-60 yrs (females, intermediate) (3)	28.8%	83
<hr/>		
Female (17-37 yrs) (11)	14.8%	85.7
Male (13)	16.1%	126.4
<hr/>		
Working (5)	9.2%	72.4
Intermediate (9) (except older females)	16.4%	117
Upper (8)	19.3%	122.9
<hr/>		
All speakers (27)	17%	105

Table 3.6 generally confirms the points we noted with regard to figure 3.1. Age of speaker emerges as the most important differentiating factor of (ne) in this corpus, whilst sex of speaker seems to play little or no part. The effect of social class is more difficult to interpret: the difference between the working class and the other two classes seems clear enough, but

as already mentioned, this also reflects the youth of that group. The difference between the scores of the other two social classes appears slight, although it is the type of difference we would expect, with the upper class using more of the standard variant than the intermediate class.

The only previous survey in France with which the results in table 3.6 can be meaningfully compared is Ashby's (1981) survey. However, methodological differences and the uneven distribution of informants in the present study mean that such a comparison is less than straightforward. Membership of Ashby's three classes was based as much on education as occupational criteria (1981:676, and personal communication). Secondly, for his investigation of (ne), he took data from 35 informants (from a total of over 100) distributed evenly between two age groups, 14-21 years and 51-64 years. Thirdly, the group scores appear to have been calculated on the basis of speakers' aggregated data, rather by averaging the relative frequencies of individuals (p.682).

Bearing in mind these differences, we can nevertheless say that, in several respects, the findings of the two studies do correspond. A first point is that the substantial age differentiation found in figure 3.1 and table 3.6 replicates, though on a smaller scale, that found by Ashby. Secondly, the division between the working class and the other two classes is also a replication, again to a lesser degree, of the social differentiation in the Tours data. However, in contrast to the present study, Ashby found a slight, but significant, difference

between the sexes, with the male informants having a higher rate of ne retention than the females. In addition, there are some remarkable differences between the sub-groups. The young upper-class males, for example, have the relatively high score of 38.4%, whereas their female counterparts have the much lower relative frequency of 12.9%. The older speakers of the upper and intermediate classes have very high rates of ne use of between 60% and 75%, whereas the older working-class men score 39.6% and their female counterparts just 10.3%. Ashby interprets these findings as evidence of a change in progress ("from below"), principally, it seems, because of the substantial age differences. However, some of the sex differences do not correspond with those which have been discovered in other cases of changes from below in progress. In particular, the relatively conservative scores of the young upper-class males and the older working-class males are unexpected, since changes from below tend to be led by males, rather than females.

Furthermore, in view of what is known of the historical and geographical diffusion of the loss of the negative particle, it might well be anticipated that Tours would, along with Paris, be in the vanguard of this change, and that the overall rate of ne retention would by now be considerably lower than 37%, especially when we consider that a group of speakers from the Somme (a more conservative area, as we saw in 3.5) produced an overall rate of only 18.8%.

3.12 Style-shifting

Before concluding this investigation into the variable (ne), let us consider whether there is any evidence, from the present corpus and from other sources, that it is a sociolinguistic marker rather than merely an indicator, ie that the interpersonal variation we have just seen is reflected in style-shifting by individuals.

In view of the data collection methods employed, it is not, on the whole, possible to compare distinct styles for the speakers in this study, but there is one special case which merits closer examination. This is the 37 year-old intermediate-class male whose score is shown in figure 3.1 as 25%. The interview with this informant, a camp director, took place in two distinct parts. It began at about 9 pm one evening, inside the camp buildings, in the presence of the assistant director, who was also a friend of the director, but who intervened only occasionally in the conversation. After about 20 minutes, the interview had to be interrupted, but was restarted early the following morning, in somewhat different circumstances. In this second part, the fieldworker and the informant strolled around the grounds of the camp, with no third person present, except for the occasional interruption from small children coming up to see the director. In addition to the fact that the situation was perhaps more relaxed than the night before, the topic of conversation was rather less serious also. In the first part of the interview the informant had been carefully explaining the thinking behind the somewhat original organisation of his camp.

The next day, however, the conversation was a good deal more anecdotal, concentrating on the informant's past experience of children's holiday centres. The net result of these various situational changes was a massive style-shift by the informant: from 50% retention of ne in the first part of the interview, to just 11.4% in the second part.

In this connection, it is pertinent to mention the finding by Sankoff & Vincent (1980:302-3) that the very rare instances of ne in the speech of their Montreal informants tended to occur when the topic of conversation was a serious one, such as religion, the education of children and language. However, in contrast with Montrealers, it would appear that the stylistic repertoire of certain French speakers, if not all, allows them to produce high rates of ne for sustained periods in their more formal styles. Indeed, it is clear that highly literate speakers can produce rates of retention of ne of 100%. Pohl (1975:21) gives several examples of "public" speech situations (ie where there is a large audience) where he has observed speakers retaining the negative particle almost consistently: a sermon, meetings, a lecture and the public soutenance of a thesis. In addition, we can again cite Ashby (1981:681), who recorded three informants in two different situations, in an interview at their offices and with their families at home after a meal. Not surprisingly, the overall rate of ne retention fell from 35% in the interview to 16% at home.

Although we clearly need more information about the social and stylistic differentiation of (ne) in various communities, it seems possible that it is what Bell (1984:153) has termed a

"hyper-style" variable. (That is to say, in a given speech community, the degree of style-shifting involving the variable exceeds the degree of social differentiation - in the vernacular style). What is not yet entirely clear is whether there are still, in France, some speakers who have extremely high rates of retention of ne in their vernacular style, for it seems likely that no quantitative study of (ne) (including the present one) has yet succeeded in eliciting the vernacular from a socially diverse group of speakers.

If (ne) is, as we suspect, a hyper-style variable, then it is in very select company, since the only examples of such variables which emerged in Bell's comprehensive survey of the sociolinguistic literature were from Tehran Persian (cf Jahangiri & Hudson, 1982). Contemporary France and Iran might be thought to have little in common, but perhaps one significant sociolinguistic similarity helps to account for the fact that hyper-style variables are to be found in both societies: formal styles, which reflect the conservative written language, seem to involve, in certain respects at least, a quite different type of linguistic behaviour than informal styles in both French and Persian. Bell reports that the Persian variables are losing their interspeaker variation (1984:156). As the loss of the French negative particle nears completion, we might expect that the same process would take place - indeed if the linguistic behaviour of the 17-22 year-olds in the present study is at all typical then that process may already be well under way.

3.13 Conclusion

To conclude, let us consider whether the age differentiation found in the present study and elsewhere (notably in Ashby, 1981) is indeed a case of change in apparent time. Given that the omission of ne is thought to have begun at least 300 years ago, it seems surprising that it should have taken so long to run its course, and that there should have been a dramatic acceleration in the early 1960s, when our 17-22 year-old informants were born. Moreover, as we have seen, certain aspects of the social differentiation found (particularly in the Tours study) do not point straightforwardly to a change in progress. What seems more probable is that there is now a pattern of age-grading, whereby each generation of speakers has virtually a zero rate of ne retention as children and adolescents, but then as they become older modify their speech, under pressure from, and in the direction of, the written language. Clearly, the extent to which individuals do this will vary, according to factors such as their social networks, as shaped by their work, domestic circumstances and leisure-time activities (cf Milroy, 1980), and also to imponderables such as their linguistic adaptability and their motivation (cf Le Page & Tabouret-Keller, 1985:182-6).

According to Lemieux (1985:101), the negative particle ne should still be considered part of the grammar of Montreal French, because speakers are still able to use it in formal styles: even working-class adolescents were found to produce many instances of ne in sketches they performed. The fact that ne is acquired only after children begin school is, according to Lemieux, immaterial.

Yet again, we see here a major divergence between formalist and functionalist views of language. From a sociolinguistic perspective, the evidence seems clear: ne is no longer part of the vernacular linguistic system which is transmitted from generation to generation by the normal processes of acquisition.

CHAPTER 4: INTERROGATIVE STRUCTURES

4.1 Defining the variables and their variant structures

4.1.1 Direct and indirect interrogatives

In this chapter, the range of interrogative structures is first sketched out, and we then set them in their socio-stylistic, historical and dialectal context. Previous quantitative treatments of interrogatives are then reviewed briefly, and their overall results compared with those of the present study. In the last section we consider two problems which have not generally been confronted in quantitative studies of this area: the acceptability and equivalence of the variant structures.

It is important to emphasise at the outset that this and the following chapters are concerned exclusively with direct interrogatives, even though other structures may be used, for example, as requests for information or action. So, indirect interrogatives such as in (1),

- (1) B: je sais pas si vous allez à Paris souvent?/
A: non / (12775)

(which are frequent in the corpus) are not considered in the present study. (The ? in (1) indicates that final rising intonation was used.) This is in line with standard practice in the variationist paradigm, where the aim is to investigate variation involving different forms in a restricted sub-system of the language be it on the phonological, grammatical or lexical levels. This approach can therefore be characterised as

"system-oriented", rather than "speaker-oriented". A "speaker-oriented" approach, within a pragmatic or discourse analytical framework would perhaps take a function, such as a request for information, as the starting point, and investigate speakers' motivations for using different structures, including non-interrogatives, to express this function (cf the work of Lavandera, for example, discussed by Milroy, 1987:165-8).

4.1.2 Interrogative comment clauses

Although there is indeed variation in the form of interrogative comment clauses (eg vous voyez?, voyez-vous?) they will be excluded from the main analysis of interrogatives. This is because certain informants habitually use these forms with very great frequency, and others do so very little indeed. To count them in with full interrogative clauses would produce a skewing effect on the data: in most cases, it would add very large numbers of the [SV] variant to certain speakers' scores. A similar problem has been encountered in studies of /t,d/ deletion in American English, where occurrences of and are not counted. For many speakers, /d/ is almost categorically deleted in and, and since the word occurs so frequently, its inclusion in the quantification would vastly increase the number of instances of /d/ deletion for these speakers (cf Neu, 1980). Interrogative comment clauses will, however, be considered briefly at the end of chapter 6.

4.1.3 Yes/no interrogatives and WH interrogatives

The alternating interrogative structures will be analysed in this and the following chapters in terms of two variables, Yes/No interrogatives and WH interrogatives, henceforth abbreviated to (YNQ) and (WHQ). This decision is motivated both linguistically and sociolinguistically: linguistically, because the two types of questions have distinct syntactic, semantic, and pragmatic properties; and sociolinguistically, because (YNQ) has three variants, whereas (WHQ) has five or more, and these variants have distinct sociolinguistic values attached to them. For example, grammarians accept [SV] as standard (in speech), whereas [QSV] is either considered non-standard (Wagner & Pinchon, 1962:545; Grevisse, 1986:657), or simply not mentioned. The few examples of alternative questions encountered in the data will be treated as occurrences of (YNQ), because of their structural similarity to Yes/No interrogatives.

As in other complex areas of variation (eg consonant cluster simplification in Black English; cf Labov, 1972b), it would probably be possible to analyse the variation in terms of several simpler variables. Sociolinguistic variables do not necessarily come clearly defined by the data: rather, they must be carefully specified by the analyst in each particular case.

4.1.4 Variant structures occurring in the corpus

It will be convenient in subsequent discussion to refer to the variants involved by means of formulae, which represent, albeit

schematically, their structural descriptions. These are similar, though not identical, to those used by Pohl (1965), Price (1971:265-70) and Al (1975:57-63). The principal structures to be examined here are set out in (2-9), together with an example of each:

- (2) [SV] vous voulez la tisane? (1621336)
- (3) [ESV] est-ce que c'est bon aussi? /(21223)
- (4) [V-CL] aurais-tu du feu? (example noted, but not recorded, from a non-informant)
- (5) [QSV] comment elles s'appellent? (2843)
- (6) [QESV] qu'est-ce qu'il faut? (21659)
- (7) [SVQ] on en était où là? (4671)
- (8) [QV-CL] où allons-nous? (13762)
- (9) [QV NP] en quoi consiste votre enquête? (211142)

(key: S = subject clitic or Noun Phrase; CL = subject clitic; NP = subject Noun Phrase; V = verb; E = est-ce que/qui; Q = WH word or phrase). The [SV] structure is often accompanied by a final-rising intonation in order to distinguish it from a corresponding declarative, although Fónagy & Bérard (1973:95) found that 36.8% of their [SV] interrogatives did not have this distinctive intonation, but were nevertheless interpreted as questions because of the conversational context.

4.1.5 Clitic inversion and NP inversion

Grammars of French have often presented the structures involving inversion rather differently from the way they will be dealt with here. Traditionally a distinction is made between 'simple' inversion on the one hand whether with a clitic subject as in (4) and (8), or a NP subject as in (9), and 'complex' inversion in the other, where a NP subject is accompanied by a coreferential subject clitic, as in (10):

- (10) pourquoi les animateurs ne feraient-ils pas le jeu en même

temps? (example noted, but not recorded, from a non-informant)

Despite the presence of the subject NP, however, it can be argued that such cases of 'complex' inversion involve fundamentally the same structure as 'simple' inversion with a clitic subject (cf Kayne, 1972:81). In fact, the informants in this study produced no occurrences of complex inversion. Had they done so, however, small numbers of these could have reasonably been grouped with other instances of clitic inversion. Moreover, the traditional category of 'simple' inversion is unsatisfactory for both syntactic and sociolinguistic reasons, and must be replaced by the distinction between clitic and NP inversion. Kayne (1972) and Kayne and Pollack (1978), among others, have demonstrated the quite different behaviour of, and constraints on, these two structures. Secondly, it will be shown later that the socio-stylistic differentiation of the two structures is quite different.

4.1.6 The [Q=S V] structure

A problem which arises in defining the variants of (WHQ) is how to classify interrogatives such as (11), where the WH word is traditionally seen as the subject:

(11) qui y a été?/ (23982)

Formal syntactic arguments might suggest it should be considered as an instance of [QSV], with an 'empty' subject position (cf Lefebvre, 1982:56-57). However, examples such as (11) are accepted by grammarians as standard, whereas other cases of [QSV] (eg (5) above) are considered to be non-standard. Since one of the principal objectives of variationist work is to reveal

patterns of social and stylistic differentiation, structures with different socio-stylistic values should be classed as distinct variants, despite certain formal similarities. For this reason, examples such as (11) will be classified as a distinct variant, [Q=S V], ie a structure where the WH expression is traditionally considered to be the subject.

On structural grounds again, interrogatives beginning with qu'est-ce qui might likewise be treated as a distinct variant. However, it seems that for grammarians and native-speakers alike they are assimilated to [QESV], and this is how they will be classified in the present study.

4.1.7 Structures involving a dislocated NP coreferential with subject clitic

A further problem in the definition of the variants involves how dislocated structures such as (12) and (13) should be handled:

(12) "vélo c'est quoi?" / (28141)

(13) il est connu Brassens euh au en Angleterre? / (25461)

These involve respectively a left-dislocated and a right-dislocated NP (underlined) coreferential with the subject clitic. There may be a case for treating these as distinct variants. However, it was decided in this study to count them as occurrences of variants of parallel non-dislocated structures (ie of [SVQ] and [SV] in the cases of (12) and (13)). This was done partly for purely practical reasons: the very small numbers of them in the corpus did not seem to warrant distinct variants. But also because dislocation can be considered a more general process (or transformation) which is by no means specific to

interrogatives and which should be analysed separately. In fact, all the interrogative structures except for [QV NP] and [Q=S V] can be combined with left or right dislocation.

4.1.8 Structures not occurring in the corpus

4.1.8.1 The interrogative particle *ti*

From the late 19th century onwards the post-verbal interrogative particle *ti* has attracted considerable attention among linguists (eg Paris, 1877; Foulet, 1921; Guiraud, 1965). However there is little disagreement today that it has become extremely rare, at least in France, although according to Grevisse (1986:641) it still occurs in certain regions, eg Normandy.

In France the structure appears to have been more or less restricted to Yes/No interrogatives (but see below for Canada): the only example of *ti* in a WH interrogative encountered in the various works consulted is the following:

(14) Où j'ai-*ti* vu ce nom-là?

This is given by Foulet (1921:280), who adds, however, that *ti* is very rare in WH interrogatives.

Certain instances of what might, at first glance, appear to be *ti*, such as (15), are probably cases of hypercorrection:

(15) Il y a-t-il selon vous des articles qu'il faudrait ajouter?... (on a printed questionnaire to passengers on North Sea Ferries in 1986: Renchon, 1969:88-9, gives several similar examples.)

Grevisse explains (1986:642) a similar example by Foucault, by

pointing out that the standard inverted form y a-t-il involves an initial y a, which is also found in the non-standard form of il y a, when il is omitted.

One area of the French-speaking world where this structure with ti or tu appears to be very much alive is Canada (cf Patterson, 1971, and St-Pierre, 1977), where it occurs in both Yes/No and WH interrogatives. (Patterson, 1971:29, gives several examples of the latter.)

In contrast, during the two months of fieldwork for this research project, the only instances of ti which were heard were several examples of ça va ti? produced by elderly speakers outside a village church near the coast of the Somme.

4.1.8.2 Structures said to be 'rivals' of est-ce que

A wide range of non-standard structures, mostly variants of (WHQ) and seen as related 'rivals' of est-ce que, are discussed by certain grammarians, eg Grevisse (1986:653), Judge & Healey (1983:423). Some of these, such as (16) and (17) occur quite frequently in working-class speech (cf Behnstedt, 1973):

- (16) [QsekSV] qui c'est qui voulait une poire? / (X2417)
(17) [QkSV] quel age que t as? (example produced by a non-informant)
(In these formulae 'se' = c'est, 'k' = que/qui.)

Other less frequent structures include the following:

- (18) [seQkSV] C'est quoi qui se passe?
(19) [QEsekSV] Qui est-ce que c'est que vous cherchez?

Both examples are from Judge & Healey (1983:423).

Some observers have suggested that the omission of que in est-ce que, as in (20) produces yet another variant structure:

(20) qu'est-ce tu veux? (21284)

In the present study, however, this is considered to be simply an instance of consonant cluster simplification (here resulting in the elision of [k]), triggered by the following plosive [t]. (20) is thus to be classified as an example of [QESV].

4.1.8.3 The structure [QEV NP]

Wagner & Pinchon and Grevisse both point out that est-ce que can be combined with NP inversion to produce yet another (WHQ) variant, as in (21) and (22):

(21) Qu'est-ce qu'a dit ton père? (Wagner & Pinchon, 1962:538)

(22) Avec qui est-ce que travaille Nicole Dupont? (Grevisse, 1986:64)

Though not very frequent, this structure appears to be accepted as standard, and even considered elegant by some.

4.1.8.4 The structure [Q=S V-CL]

Occasionally a WH phrase (particularly one involving quel, lequel or combien), functioning in traditional terms as subject, may be followed by clitic inversion, as in (23):

(23) Combien d'entre nous auraient-ils droit au titre d'hommes?
(Vercors, cited by Grevisse, 1986:644)

This 'hybrid' structure is justified, according to Wagner & Pinchon (1962:538) by the need to avoid confusion with the exclamative, (24):

(24) Combien d'entre nous auraient droit au titre d'hommes!

Not all examples can however be explained the same way, for

example (25):

- (25) De ces fillettes, lesquelles sont-elles les tiennes?
(Mauger, 1968:147)

4.1.8.5 The structure [E NP V-CL]

A fourth variant of (YNQ) may occasionally be heard, involving a combination of est-ce que with 'complex' inversion, as in (26):

- (26) Est-ce que demain les sauveteurs pourront-ils s'approcher
des alpinistes en détresse? (Mauger, 1968:381)

Although Mauger implies that this structure is not uncommon as an emphatic device in spoken French, it does not seem to be mentioned by other grammarians, and it appears in fact to be a very infrequent hypercorrection. Indeed, elsewhere Mauger describes the corresponding WH interrogative structure, as in (27), as incorrect (even though it was written by Ionesco):

- (27) Qu'est-ce que le rédacteur de la rubrique des chats écrasés
entend-il par un pachyderme? (Mauger, 1968:146)

4.2. The socio-stylistic evaluation of the variant structures by grammarians

As a consequence of the great emphasis laid on the teaching of grammar in French schools, traditional grammars have been considerably more influential in France than in, for example, Britain (cf Hintze, 1986). The relationship between these grammars and linguistic behaviour is of course symbiotic, in that they both influence, and are influenced by, usage. In areas of variation, such as direct interrogation, grammarians categorise variants in terms of evaluative socio-stylistic labels such as

distingué, soigné, usuel, familier, relâché, incorrect,
populaire, proscrit, vulgaire. These labels are arrived at partly
on the basis of their predecessors' judgements, and partly on the
basis of the grammarians' perceptions of who uses which variant
and in which circumstances. The distinctions between 'educated'
and 'uneducated' users and between speech and writing are
particularly important. Wagner and Pinchon (1962:539) state: "Le
choix de l'un ou l'autre [ie type of structure] dépend du degré
de culture de la personne qui parle, des circonstances,
c'est-à-dire du ton familier ou soutenu du discours, enfin de
certaines exigences du style (rythme, harmonie)." The stigma
attached by grammarians to certain variants has, via the
education system in particular, very probably played a major part
in restricting the use of non-standard structures such as [QkSV],
and may also have contributed to the virtual disappearance of ti.

Few grammarians refer to the whole range of structures described
in the previous section: among the most comprehensive are
Grevisse, and Wagner & Pinchon. Moreover, the grammarians
sometimes disagree on the socio-stylistic classification of the
variant, eg between 'populaire' and 'familier'. On the whole,
however, there is a large measure of agreement, and Table 4.1
attempts to reflect this consensus. It is probable that
subjective evaluation tests on French native-speakers would
produce judgements coinciding with this table since the
grammarians' evaluations are propagated through the education
system.

Table 4.1 Socio-stylistic evaluation of variants
of (YNQ) and (WHQ)

- careful style, writing [V-CL], [QV-CL], [QV NP], [QEV NP]
- neutral [Q=S V]
- neutral (though sometimes [ESV], [QESV]
 'inélégant' in writing)
- colloquial ('familier'), [SV], [SVQ]
 but generally accepted
 as standard in speech
- 'mistake' or hypercorrection [E NP V-CL], [Q=S V-CL]
 (though evaluation varies)
- colloquial/working-class [QSV], [seQkSV]
 ('populaire'), incorrect
- uneducated ('vulgaire'), [QkSV], [QsekSV], [QEsekSV]
 incorrect
- rural/working-class/uneducated [SV-ti]

4.3. The history of interrogative structures in French

4.3.1 Introduction

In order to provide a historical perspective to the description of contemporary variation in interrogatives, a brief outline will be sketched here of the evolution of these structures since Old French. This outline derives entirely from secondary sources, notably Chevalier et al (1964), Grevisse (1986), Harris (1978), and Price (1971).

It should, of course, be borne in mind that a historical description of this, or any other, aspect of 'the French language' is not, strictly speaking, comparing like with like. Firstly, because the geographical area in which French has been spoken has increased greatly from the Middle Ages, thus

(presumably) increasing the amount of regional variation within the variety. And secondly, because descriptions of the state of the language in earlier periods inevitably reflect the written language more than the spoken language.

4.3.2 Old French: before c. 1300

In Old French, inversion was the commonest interrogative strategy, both for Yes/No and WH questions. NP inversion existed not only in WH interrogatives as in (28) but also in Yes/No interrogatives, as in (29), where today it is totally unacceptable:

(28) Quant fust avenus chis affaires? [QV NP]

(29) Font ensi li autre a Paris? (both from Foulet, 1968, cited by Harris, 1978:31, 34)

The following example of complex inversion (ie [NP V-CL]) from the Chanson de Roland (early 12th century) is quoted by Price (1971:266), but this structure was not very widespread in Old French.

(30) L'aveir Carlun est il apareilliez?

According to Grevisse (1986:658), examples of [QSV] such as (31) were not uncommon in Old French;

(31) Que ce puet estre? (Roman de Renart)

The [SV] structure was only occasionally found in Old French and was restricted to echo questions and exclamative-type questions (Grevisse, 1986:648, 653). The first examples of the use of est-ce que as an interrogative marker came in the 12th century (Grevisse, 1986:653) in WH questions where the WH word was que or qui (Price 1971:267; Harris 1978:32). Only later was est-ce que used with other WH words and in Yes/No questions. Evidence for the grammaticalisation of est-ce (que) can be gleaned from the fact that it was often written in the Middle Ages as esse, even though

the tense of the copular element could vary until about 1600 (Grevisse, 1986:653).

4.3.3 Middle French: from c. 1300 to c. 1600

During this period, [V NP] became ungrammatical: according to Kayne (1972:116) this happened by about the 15th century, but Chevalier et al claim it was only by the 17th century. In any case, the decline of this structure was accompanied by a corresponding growth in the alternatives. [NP V-CL] had become common by the 14th century and by the 16th had overtaken the [V NP] structure, according to Price (1971:266). The corresponding WH structure, [Q NP V-CL], became far more frequent between the 15th and 17th centuries according to Harris (1978:32). [SV] however was still relatively rare in Middle French (Grevisse 1986:563).

By the 14th century, est-ce que was being used quite widely with other WH words, in addition to qui and que, (Price, 1971:267), and the use of [QESV] continued to grow after 1500 according to Gougenheim (1973:236) (cited by Harris, 1978:32). The tense of the copular element of est-ce que was still variable in about 1400 as is shown by (32):

- (32) Quant sera ce que nos i serons? (Froissart c. 1333- c. 1400, quoted by Price, 1971:267)

In the late Middle French period est-ce que came to be widely used for Yes/No questions also - from the 16th century, according to Grevisse (1986:653). Also in the 16th century, there is the first attested use of the post-verbal particle ti with verbs in persons other than the third, according to Harris (1978:33), though according to Price (1971:268), this was only in the 18th

century.

4.3.4 Classical French: from c. 1600 to c. 1800

By the beginning of this period, [V NP] had become quite unacceptable (Chevalier et al, 1964:91), giving way entirely to the alternative structures: [NP V-CL], [ESV] and presumably [SV] also. Malherbe (1555-1628) censured what is perhaps the last attested example of this structure:

- (33) Viendra jamais le jour qui doit finir ma peine? (from the poet Desportes, quoted by Grevisse, 1986:648)

Notice here the heavy NP subject, a factor which today continues to favour NP inversion in WH interrogatives and certain subordinate clauses.

The socio-stylistic evaluation of est-ce que was the subject of some controversy in this period (and still is, for some).

Vaugelas (1647) reported that some condemned the [QESV] structure and preferred [Q V-CL]. He himself, however, considered that [QESV] was 'fort bonne' (Chevalier et al. 1964:91).

In the latter part of this period certain constraints came to operate on [QV NP]. Firstly, the structure became unacceptable with direct objects and other complements closely linked to the verb. (34) is a late case of freedom from this constraint:

- (34) En quoi blesse le ciel une visite honnête? (Molière, quoted by Chevalier et al, 1964:91).

Secondly, it was still acceptable in the 17th century for pourquoi to occur in [QV NP] (Grevisse, 1986:646), whereas today, in principle, it is not, although the occasional example can

still in fact be heard.

4.3.5 Modern French: post-1800

Several observers of the French language have remarked on the decline of clitic inversion in modern French, both in Yes/No and WH interrogatives (eg Harris, 1978:31). Just how far this trend has gone will become clear later in this chapter.

The [QkSV] structure was first attested in 19th century 'popular' speech according to Price (1971:270), and some have suggested that it is gaining ground there (eg Harris, 1978:35). According to Behnstedt (1973:36), the structure is more characteristic of the northern dialects than of the south.

On the progress of ti there has been considerable disagreement: Price (1971:269), for example, has suggested it is making headway in colloquial and working-class speech. Recent surveys, however, tend to point towards the opposite conclusion (cf also Désirat and Hordé, 1976:152-3).

It is rather striking that very little, if any, mention is made in the sources consulted of the history of the [SVQ] structure. Its development seems to have gone largely unnoticed, and even today most grammars have remarkably little to say about it.

4.4 Interrogatives in Picard dialect (as spoken in the Somme)

4.4.1 Yes/No interrogatives

Since the informants in this study were from the Picard dialect area, and in certain cases had parents who spoke Picard (or even did so themselves), it is appropriate here to present a brief outline of interrogative structures in the Picard dialect as spoken in the Somme.

In his description of the Picard spoken north-east of Amiens, Debrie states 'Le patois ignore pratiquement l'inversion dans la phrase interrogative' (1960:113). Whilst other sources generally confirm this, we may note that Tuaillon (1975:91) shows that, in the Atlas Linguistique de la France, the informant at point 266 (south-west Somme) did in fact use inversion in giving the patois equivalent of Où vas-tu?.

According to Debrie, Yes/No interrogatives are formed with the post-verbal particle ti, as in (35) and (36):

(35) j'é ti? (= ai-je?)

(36) t'o ti? (= as-tu?)

In compound tenses ti follows the (first) auxiliary verb:

(37) O z avon ti kanté? (= avons-nous chanté?) (all examples from Debrie, 1983:40)

Confirmation of this structure is provided by Flutre (1955:55) as in (38):

(38) Ț dire ti wēi? (= dirai-je oui?)

However, Flutre also describes another particle, Țu, derived apparently from the Old Picard sai jou? (= sais-je?) and which

occurs both pre-verbally and post-verbally, as in (39) and (40) respectively:

- (39) žu k t i vó? (= est-ce que tu y vas?)
(40) ež se žú? (= est-ce que je sais?)

Debrie makes no mention of this particle. However, he says that the Picard equivalent of the interrogative adverb quand is kanjou (1983:102) which appears to be derived historically from kan plus the particle jou. Indeed in 1960, Debrie had written this as two separate words: kã ju (1960:113).

4.4.2 WH Interrogatives

According to Debrie (1983:68-9, 102) all fronted WH words are followed by èke (= que), reducible to k. This includes kwè (= quoi), and tchèche (= qui, but etymologically, qui est-ce).

Examples of this are (41), (42) and (43):

- (41) Kwè k'tu fro kan t'éro tin brévé?
(= Que feras-tu quand tu auras ton brevet?)
(42) Tchèche k'o z'intère édmin?
(= Qui enterre-t-on demain?)
(43) Kmin k't'o fwé ?
(= Comment as-tu fait?)

Debrie also gives, without comment, an example which could be analysed as a different, though related, structure:

- (44) Tchèche èke ché k'i l fro?
(= Qui le fera?, literally Qui que c'est qui le fera?)

4.4.3 Other Structures

According to Debrie then, the Yes/No interrogative structure in Picard is [SV-ti], and the WH structure is [QkSV]. For (YNQ) Flutre gives [SV-žu] and [žu SV] in addition. It seems very likely

that the [SV] structure also exists in Picard, given its frequency in colloquial French.

Additionally, Flutre (1955:53) provides an example which appears to be the [SVQ] structure:

(45) šə pur tšèš? (= c'est pour qui?)

It is interesting that there is no mention in these sources of the [QSV] structure, which Behnstedt has suggested is characteristic of southern, Occitan, varieties, rather than of northern dialects (Behnstedt, 1973:36).

4.4.4 Picard structures and the informants in the present study

One informant from the west of the Somme, who reported that she spoke Picard, produced (46) and (47), when asked to give the Picard equivalent of Qu'est-ce qu'on fait aujourd'hui?:

(46) [kwe k ʃe] qu'on va faire? (1520)

(47) [kwe k ʃe] qu'on fait? (1522)

These are clearly analogous to the non-standard [QkseSV] structure.

However, the interrogatives produced spontaneously by the informants in the present study seem to show no obvious influence from Picard: neither [SV-ti] nor [QkSV] were produced by the informants. However, observations outside of the recorded interviews suggested that [QkSV] is in fact quite widely used by children and some adults: (48) and (17) above were both produced by children:

(48) d'ou qu'il est l'enregistr_ueur? (d'ou = ou, here)

As already mentioned, ti was not encountered at all in the centres visited, but was heard from some elderly rural speakers in the set phrase ça va-ti?

4.5 Previous quantitative studies

4.5.1 Data, informants and methodology

4.5.1.1 Introduction

Since the 1960s several quantitative studies of interrogatives in spoken French have been carried out, and these provide invaluable comparative material for the present study. Here, we will first describe, for each survey, how and from whom the data were obtained, together with some comments on the methodology where appropriate, before presenting the data and principal findings of these studies.

4.5.1.2 Gougenheim et al (1964)

The principal objective of the Français Fondamental project was to provide a quantitative description of "basic" French for the teaching of French as a foreign language. Most data for the project were obtained from 275 informants, of diverse geographical origins, and of whom just 11 were schoolchildren. Of the adult informants 104 were from professional backgrounds (especially teachers and students), but a wide range of other occupations was also represented. The total number of words in the corpus was 312,135, and the number of direct interrogatives

was over 1,000. Rather surprisingly, in their study of interrogatives, Gougenheim et al did not count the [SV] variant of (YNQ). For this reason, and also because their classification is rather different from that employed in this study, we shall rely instead on the presentation of the Français Fondamental data in Pohl (1965).

4.5.1.3 Pohl (1965)

For part of his study, Pohl made use of some of the data from the Français Élémentaire project (which was later renamed as Français Fondamental). Not all the data were examined, however, since the total number of interrogatives figuring in this part of Pohl's study was just 645. Pohl grouped the informants into working-class and middle-class, according to their level of education (rather than their occupation). Pohl also studied the interrogative structures produced over several weeks by his elderly parents, both Belgian, middle-class and well educated. (His father, aged 80+, was an ingénieur ; his mother was in her seventies.) Unusually, Pohl made a written record of the interrogatives rather than attempting to tape-record all of his informants' speech during the fieldwork. He stresses, however, that his parents were not aware that he was systematically observing and noting down their questions. Pohl conflated the figures for both [QESV] and [QkSV], and also for [QV-CL] (except for 'complex' inversion) and [QV NP], as have also several other researchers. More unusually, he classified the few occurrences of ti and of questions beginning with the formulae N'est-ce pas que? or Hein que? as cases of [ESV].

4.5.1.4 Terry (1970)

Terry's (1970) study was based on a corpus of the scripts of contemporary popular plays ("théâtre du boulevard"), dating from 1957 to 1964, and in which the characters were mainly from middle-class backgrounds. Some 31 categorical interrogatives (mostly echo questions) were excluded from the quantification of the data.

Söll (1982) recategorised Terry's data in order to compare them with his own, and since Söll's categories correspond more closely to ours, we shall use his presentation of Terry's data. However, it seems that a printing error (Söll, 1982:49-50) resulted in the omission of some 54 of Terry's tokens. These have been retrieved from Terry (1970), and are reflected in the relative frequencies for his data in the tables to be presented in 4.5.2.

Clearly the source of Terry's data means that they reflect only writers' perceptions of spoken language. This somewhat diminishes the interest of these data from a sociolinguistic perspective, and may explain discrepancies between Terry's results and those of the other surveys. A further problem is Terry's decision to conflate tokens of [QSV] and [SVQ] (and the very small number of [QkSV]), on the grounds that they all involve "incomplete syntactic transformations".

4.5.1.5 Söll (1982)

Söll's data came from another pedagogically oriented survey, the CREDIF's Enquête sur le langage de l'enfant français, carried out in 1964-5. The 79 children involved were all aged 9, and had diverse social and geographical backgrounds. Söll's data comprised 452 tokens of (YNQ) and 364 of (WHQ). The recordings included both free conversation and a game.

4.5.1.6 Behnstedt (1973)

In what is possibly to date the most complete quantitative study of interrogatives in spoken French, Behnstedt (1973) examined three varieties of spoken French: français soutenu, français familier, and français populaire, which we shall call formal middle-class, colloquial middle-class, and working-class (colloquial), respectively. (Français soutenu is in fact Al's term: Behnstedt called this variety Rundfunksprache, or 'radio language'.) The formal middle-class data (16,436 tokens) were obtained from recordings of 4,000 interviews and conversations on the radio. The colloquial middle-class data were from the speech produced by 18 persons during one week, together with three individuals observed over a three-day period. The working-class data included observations made by Behnstedt while working as a co-driver of a lorry.

The results to be given in tables 4.2 and 4.3 are taken from Al's re-presentation of Behnstedt's data, cf Al (1975:57-64).

4.5.1.7 Ashby (1977)

In his (1977) study, Ashby used the corpus recorded by Malécot (1972). The 35 informants were all middle-class Parisian adults. Echo questions were excluded from the quantification. Unlike most of the other surveys, Ashby's study did attempt to find evidence of social differentiation, despite the small number of tokens (215 variable cases).

4.5.1.8 Lefebvre (1981)

Lefebvre's small-scale (1981) study examined only (WHQ) in the speech of 42 Montreal children and adolescents aged 8-18. There were just 72 tokens, some produced in informal peer-group interaction, others in formal style, including interviews.

Working within a generative framework (as opposed to the ad hoc structuralist descriptive framework of the other studies), Lefebvre conflates our [QSV] and [QkSV] categories on the grounds that, in both cases, the WH expression is in the TOPIC position under S". Similarly, she conflates [QV-CL] and [QV NP] on the grounds that, in these, the WH expression is in the WH (COMP) position. As stated earlier, the view taken here is that structures with different socio-stylistic evaluations should be treated as distinct variants.

Lefebvre excluded WH interrogatives with qui from her quantitative study, and also other WH interrogatives such as the clefted interrogative [seQkSV].

4.5.1.9 Lafontaine & Lardinois (1985)

For their study of interrogatives used by 163 7-12 year old children in Belgium, Lafontaine and Lardinois (1985) devised a guessing game in order to elicit large numbers of tokens, and the children's speech was recorded surreptitiously. The total number of tokens obtained does not seem to be mentioned, though it was presumably at least sufficient for the statistical tests of significance that were carried out. 230 tokens of (WHQ) were also obtained, for comparative purposes, from Belgian TV journalists.

Lafontaine & Lardinois point out that the game did not in fact succeed in eliciting natural, informal speech from the children: "... nous avons souvent eu l'impression que les enfants s'efforçaient de 'bien parler'." (p.66). This no doubt helps to explain the surprising results of this study, and it seems unlikely that these are typical of the everyday speech of Belgian children. The results show that verbal games do not automatically provide a successful means of eliciting large quantities of particular grammatical forms in an informal speech style.

4.5.2 Results of previous quantitative studies

4.5.2.1 Reorganisation of the data

In order to compare the findings of the surveys referred to in 4.5.1, it is necessary, in some cases, to reorganise and reinterpret the figures presented in the original studies. This

is for at least two reasons. Firstly, several of the studies used different categories, especially for the classification of WH interrogatives - eg categories based on the grammatical function of the WH word involved. In such cases the data have been regrouped into the variant structures outlined in 4.1. Secondly, the observed frequencies have been converted to relative frequencies, where this had not been done previously.

4.5.2.2 Yes/No interrogatives

The relative frequencies for the variants of (YNQ) in the varieties examined in these surveys are shown in table 4.2. The varieties have been approximately ordered according to their expected linguistic conservatism, more conservative varieties occupying higher positions in the table.

It is clear that, in most varieties, [SV] is used over 80% of the time. The two exceptions to this are the formal speech of middle-class speakers (on the radio), and, extraordinarily, the speech of Belgian children. With regard to the use of [ESV] and [V-CL], these two varieties again stand out. In the others, it is clear that middle-class adults use rather more inversion than do children and working-class adults. In several groups of speakers, the productive use of inversion (ie apart from formulae and quotations) is minimal.

Table 4.2 Relative frequencies for (YNQ) in previous studies

Variety (Researcher)	[SV]	[ESV]	[V-CL]	N =
formal middle-class (Behnstedt)	41	39	20 a	(12,069)
plays (Terry)	85.5	3.2	11.2 b	(3,016)
middle-class (Ashby)	80	10.8	9.2 c,d	(130)
middle-class (Pohl)	82.2	8.9	8.9	(270)
colloquial middle-class (Behnstedt)	90	— 10 e —		? f
elderly Belgian couple (Pohl)	85.5	14 g	0.5 h	(816)
working-class (Pohl)	90.7	4.6	4.6 i	(151)
working-class (Behnstedt)	95	— 5 e —		? f
9 year-olds (Söll)	90.9	7.7	1.3 j	(452)
Belgian 7-12 year-olds (Lafontaine & Lardinois)	49	31	20	? k

Notes on table 4.2

- a. 4% of the total N being complex inversions.
- b. There being just 15 cases of complex inversion.
- c. None of these being complex inversions.
- d. Most of these being "formulae" such as voulez-vous?, comment dirai-je?, etc.
- e. Al does not give separate figures for [ESV] and [V-CL] for colloquial middle-class or for working-class.
- f. Al does not give the figures for N here, and it has not been possible to retrieve them from Behnstedt (1973). However, since in other studies the ratio of (YNQ) tokens to (WHQ) tokens averages out at almost exactly 2:1, it seems probable that for Behnstedt's colloquial middle-class data, N is about 900, and for his working-class data about 1100.
- g. This includes two (jocular) cases of T'as-ti tout?, one of Hein que ...? and one of N'est-ce pas que ...?
- h. All 4 tokens being from Madame Pohl.
- i. All except one or two were "formulae".
- j. All except one were quotes from a film, the exception being "As-tu regardé Télé-Dimanche?" (Söll, 1982:48).

k. N is not given by Lafontaine & Lardinois.

4.5.2.3 WH interrogatives

The relative frequencies for the variants of (WHQ) are given in table 4.3. The interpretation of this table is rather more difficult than that of table 4.2, due both to the larger number of variants and to the fact that figures for certain variants are sometimes conflated. In particular, it is not always clear whether or not there were any tokens of the variants [QkSV], [QksekSV], [QsekSV] and [seQkSV], as some researchers consider these to be sub-types of [QESV], but without making this explicit.

Table 4.3 Relative frequencies for (WEQ) in previous studies

Variety (researcher)	[SVQ]	[QSV]	[QksV]	[QksksV]	[QksksV]	[QksksV]	[QV-CL]	[QV NP]	N=
formal									
middle-class (Behnstedt)	25	10	0	0	0	-	3 50 a	12	(4,367)
plays (Terry)	12.5			-	-	-	38.9	48.6	(1,515)
Belgian TV journalists (Lafontaine & Lardinois)	14.9	5.3	0	-	-	-	35.5	44.3	(230)
middle-class (Ashby)	38.8	7.1 b	0	0	0	0	38.8 13 c	2.4	(85)
middle-class (Pohl)	10.3	7.7	-	-	-	-	46.5 d	35.5 e	(155)
colloquial middle-class (Behnstedt)	33	46	0	0	4	-	12 3 e	2	(446)
elderly Belgian couple (Pohl)	2	3	-	-	-	-	66 d	28	0.5 (184)
working-class (Pohl)	8.7	27.5	-	-	-	-	53.6d	10.2 e	(69)
working-class (Behnstedt)	12	36	26	6	3	-	8	9 f	(587)
Belgian 7-12 year-olds (Lafontaine & Lardinois)	22	11	9	-	-	-	18	40 g	?
Montreal 8-18 year-olds (Lefebvre)	-	87.5	-	-	-	-	-	12.5 h	(72)
9 year-olds (Söll)	12.9	31.6	6.9	0	3.3	0.3	41.5	3.6	(364)

- indicates that there is no information on the structure concerned.

Notes on table 4.3

- a. 3 of this 50 represents complex inversion.
 - b. All these tokens being with comment.
 - c. Including just one example of complex inversion.
 - d. These figures may include some tokens of [QkSV] and [QsekSV], cf Pohl (1965:504).
 - e. With no cases of complex inversion.
 - f. Al conflates the two types of inversion, but according to Valdman (1982:225), all of these tokens were [QV NP].
 - g. The older children used greater proportions of inverted structures than did the younger children.
 - h. All tokens were produced in formal style, and only one was produced by a speaker younger than 14.
-

Nevertheless, a number of general tendencies can be discerned in these sets of figures. Firstly, the groups of middle-class adult speakers do not use at all the stigmatised variants [QkSV] and [QksekSV], whereas the children and adult working-class speakers do use them, in some cases quite frequently. The non-standard [QSV] structure is used by all the groups, but is favoured more by children and working-class adults than by most middle-class groups. However, in Behnstedt's colloquial middle-class variety, [QSV] is by far the most popular (WHQ) structure. The [SVQ] structure is also used by all groups, but there seems to be no clear pattern of socio-stylistic differentiation.

The [QESV] structure is used to a considerable extent in most of the varieties represented in the table, although the figures in some cases include tokens from other variants, as indicated in note (d). It is striking that [QESV] is used least in the three groups of speakers studied by Behnstedt. In his formal middle-class variety this can be partly explained by the very strong preference for [QV-CL]. Indeed, the relatively high frequencies for inversion again distinguish several (though not all) of the more conservative varieties from those of the

children and of the working-class. In these latter groups, however, it can be noted that inversion is generally more frequent in (WHQ) than in (YNQ): this is partly accounted for by the fact that "inversion" in (WHQ) includes [QV NP]. In several varieties, complex inversion, both in (YNQ) and (WHQ), does not occur at all.

4.6 Interrogatives in the present corpus: obtaining sufficient data and identifying the tokens

4.6.1 Obtaining sufficient data

As we have seen in chapters 1 and 2, obtaining sufficient interrogatives for a quantitative study can be problematic, especially when the data consist basically of interviews. According to Sebastião Votre (personal communication), in the Rio de Janeiro corpus of 64 hours, interviewees produced in all just ten questions. The present corpus, although of only 18 hours, fortunately yielded rather more. This is probably due to a number of different factors. Firstly, the interviews were, in many cases, often rather more like conversations, in which participants have more or less equal rights to ask questions - this was partly due to the fieldwork circumstances, and partly because of the interviewer's deliberate tactic of talking about himself and about Britain at some point in the interview. This quite often resulted in the interviewee asking questions of the interviewer. Secondly, several informants used interrogatives for various rhetorical or conversational purposes quite apart from their typical function as requests for information: there are,

for example, many interrogatives which occur when speakers quote themselves, or another person, or even a hypothetical person. (These various communicative functions will be discussed and exemplified in chapter 5.) Thirdly, the interviews were, of course, interrupted on a number of occasions, especially when the informants were camp directors as they tended to have a steady flow of people calling in to see them. These interruptions produced a quite different type of discourse from that of the main body of the interview, characterised by short turns from both participants, and with the "interrupter" asking at least one question, but the informant also not infrequently asking one or more questions. As the interrupters were, in some cases, also informants who were interviewed on another occasion, the interruptions were a particularly fruitful source of interrogatives. The number of interrogatives produced in these three ways is shown in table 4.4.

Table 4.4 Sources of interrogatives produced by the informants

	(YNQ)	(WHQ)
questions addressed to the fieldworker	108	21
interrogatives used for other rhetorical and conversational purposes	50	92
interrogatives occurring during interruptions	22	9
Totals	<u>180</u>	<u>122</u>

In spite of these various factors which increased the number of interrogatives in the corpus, it is clear that the quantity produced, although large enough for a study of linguistic and pragmatic constraints, is still quite small for thirty

informants, and will consequently be insufficient for a detailed analysis of the behaviour of individual speakers.

4.6.2 Identifying occurrences of the variants: general problems

The identification of occurrences of the variants of (YNQ) and (WHQ) is a good deal less problematic than is the case with variables involving potentially ambiguous or vague forms such as on (cf Laberge, 1977; Boutet, 1986), or être + past participle (cf Sankoff & Thibault, 1977). There are however a certain number of categories of utterances which either resemble, or clearly are, interrogative, but which nevertheless have been excluded from the data to be quantified in the present study. Firstly there are examples of immediate and identical repetition of the interrogative, usually because the first instance was misheard or misunderstood by the addressee, as in (49):

- (49) B: ... [CALLING TO X] Mariline tu les as fait goûter? ////
A: les animatrices fument /
B: hein Mariline - is ont pas goûté les enfants? /
X: hein? /
B: is ont pas goûté les enfants? // [X DOES NOT RESPOND] is
ont pas goûté les enfants? /
X: non non mais il est pas quatre heures / ... (29667)

In such cases it was decided to count only the first utterance, and to exclude from the quantitative study the repetitions, since, from the speaker's point of view, these utterances may be seen as merely mechanically reproducing the original rather than as expressing new meanings. (In practice, of course, it makes no difference whether it is the first or the last of the identical utterances which is counted.) However, in cases where the repetition is not identical, syntactically and/or lexically, both interrogatives are counted, since the second version may have

been motivated by the speaker's wish to modify the meaning of the original. Thus in (50) the speaker's switch from the present tense to the imperfect alters the meaning slightly, and therefore both the underlined utterances have been counted:

- (50) A: ... ta - l'anniversaire c'est bientôt? /
B: janvier /
A: ah d'accord parce que le mien c'est le trente et un août /
B: oui mais tu seras parti. /
A: oui oui - oui /
B: et qu'est-ce que tu veux me faire dire pour ça? /
A: hein? /
B: qu'est-ce que tu voulais me faire dire? /
A: rien /
B: [LAUGH] (491820)

A second category involves a kind of conversational repair in which an incomplete utterance is immediately repeated (with or without modification) and completed. In some of these the initial utterance is nevertheless sufficiently complete to show which interrogative structure has been chosen by the speaker. However, again from the speaker's viewpoint, it seems that this first attempt is, in effect, erased by the follow-up utterance.

Therefore utterances such as that underlined in (51), have not been counted, whereas the interrogative which follows has been:

- (51) i t'ont pas laissé euh / t'ont pas laissé rester là-bas? /
(25189)

Thirdly there are a few cases of what might be termed anacoluthon, where the speaker seems to lose their way syntactically, and, despite the presence of some marker of interrogativity, no clearly interrogative clause can be identified. This is the case in (52), for example, where the speaker presumably intended est-ce que to be followed by tu veux dire que, but inadvertently omitted it after the subordinate clause (ils here refers to British animateurs who had worked in a centre with B):

- (52) A: et est-ce qu'ils ont apporté quelque chose de spécial /
aux enfants pendant la colo? /
B: pour l'un oui - sur - sûrement oui / mm /
A: en quoi exactement? /
B: euh feuh / s - spécial euh - j d e - fin est-ce que quand
tu dis "spécial" c'était euh la la connaissance du du -
d'une culture différente * ou (l'appréhension)? * /
A: * oui - oui * par exemple / (14427)

Incidentally, both false starts and anacolutha seem to be more frequent in the speech of certain individuals rather than others, but, on the whole they are not particularly common in the corpus on which the present study is based, and this lends further support to Labov's (1972a:203) claim that the so-called "ungrammaticality" of everyday speech has been greatly exaggerated.

4.6.3 Problems in identifying tokens of (WHQ)

In addition to these three categories, which involve a temporary or partial communication breakdown, there are a certain number of cases where it is not clear whether the utterance should be interpreted (and was intended by the speaker) as an interrogative or another structure. The [QSV] variant of (WHQ) closely resembles subordinate WH clauses, and in (53) and (54) the underlined utterances have been interpreted, not as direct interrogatives, but rather as indirect interrogatives, even though their syntactic links with the preceding context are not entirely clear:

- (53) ... et ça c'est vrai que c'est un problème d'éducation euh /
c'est euh comment on fait passer des tas de choses - auprès des gamins / y a pas que le football à faire passer. / ... (14640)
(54) ... on avait dit que ça nous poserait d'énormes problèmes /
si on se retrouvait par exemple avec deux musulmans pratiquant le ramadan / euh en période de centre de vacances comment on allait faire pour l'assumer parce que / ha on va pas embaucher un moniteur musulman pour eux. / ... (4888)

On the other hand, the underlined utterance in (55) has been interpreted as an interrogative, since, after very close listening, it was felt to be a quoted question which the teenagers previously mentioned had asked of a British animatrice. Although the interpretation here was problematic (partly, of course, because of the brief inaudible section in the preceding context), the final rising intonation on repose is an objective indication that the utterance was indeed intended as an interrogative:

(55) ... on avait discuté avec les ados parce que (. . . certains) Rosemary "mais / euh à sur quoi ça repose? le fait qu vous faites - fin si un très gros repas l matin midi presque rien puis l soir (bon) / ... (21006)

Occurrences of another (WHQ) variant, [SVQ], are also sometimes difficult to identify due to their similarity to sequences in which a WH word is not linked syntactically to the preceding context. This preceding context may be either incomplete, as in (56), or complete, as in (57):

(56) ... à l'Ecole des Beaux Arts oui. / j'en ai fait euh combien? quatre euh quatre cinq ans ... (2017)

(57) ... c'est une question de moyens / la preuve c'est que les gens - à faible revenu ils ont beaucoup d'enfants. pourquoi? parce qu'ils sont aidés et assistés. / ... (11843)

In (56) the filled pause euh suggests that the speaker could not recall the number of years that he had done, and had to interrupt his declarative utterance, and ask himself how many years (he had done), before continuing. In (57) the speaker was fairly animated and was speaking quite quickly, and so there is no pause between the declarative and the independent WH word. But in this case the falling tone on enfants indicates unambiguously the end of the clause, and that the following pourquoi is intended as a response to the assertion conveyed by that clause. (A single full stop, it

will be recalled, is used in the transcription to represent a falling tone, in places where it seems particularly important to indicate this.) A rather different instance in which a sequence of words appears superficially to be a variant of [SVQ] is that in (58):

(58) ... quand y a un mariage avant les la mairie est en face de de l'église carrément on étend un grand tapis rouge / j te parle euh d ça y a mes grandparents l'ont connu / (bon) y a combien alors? - ça devait être euh vers dix-neuf cent vingt des choses comme ça ... (26409)

The sequence is, of course, not a [SVQ] interrogative clause, because y a here is functioning as a preposition rather than as an equivalent to there is/are.

All the examples of apparent or actual (WHQ) tokens considered so far have involved the two variants [QSV] and [SVQ], which have no syntactic marker of interrogativity other than the WH element. On the other hand, we may suppose that tokens of [QESV] and [QV-CL] are likely to be easier to identify, because of their overt markers of interrogativity: est-ce que and inversion, respectively. However, there is one instance in the corpus of a [QESV] token which could conceivably have been intended as an indirect interrogative:

(59) ... le travail d la journée c'est un peu voir / "bon - où on en est? euh qu'est-ce qui s passe? euh" // avec des interventions euh - ou ponctuelles ou ou même plus longues - ... (28507)

The underlined interrogative (and the immediately preceding one too) has been interpreted, partly on prosodic grounds, as a quotation, and consequently as being a direct, rather than an indirect, interrogative. The use of qu'est-ce que/qui instead of ce que/qui in indirect interrogatives is a fairly widespread non-standard feature, and one which is used by several informants

in the present study. (It seems to be particularly common in Montreal: cf Kemp, 1979.)

4.6.4 Problems in identifying tokens of (YNQ)

With regard to Yes/No interrogatives, [ESV] tokens are quite straightforward to identify, as would be [V-CL] tokens. However, certain utterances which end with an interrogative tag, such as hein? or non?, are clearly Yes/No questions, but their main clauses have generally been interpreted here as declaratives, unless the intonation suggests an interrogative. (60), therefore, has not been counted as a token of [SV]:

(60) c'est moi qui qu'ai la clé non? / (21331)

However, in (61) the rising tone on vus suggests that the underlined clause is interrogative, and the following non? seems to signal the speaker's expectation of a negative answer, possibly because of a facial gesture from A:

(61) B: ... et on a mis en système un carnet de chèques - les
chéquiers tu les as vus? non? /
A: ah non - pas encore // (5382)

4.7 A preliminary comparison of interrogatives in the corpus with previous quantitative studies

Having systematically excluded certain categories of utterances from the data to be quantified, we can now proceed to compare the interrogatives in the present corpus with the previous quantitative studies referred to in 4.5. Even now, however, the comparison will not be entirely valid: firstly because some (but not all) of the previous studies excluded categorical tokens (ie

those where there is no possibility of another variant being used), and this has not yet been done for our data; and secondly because we recognise an additional (WHQ) variant, [Q=S V], which in the previous studies appears generally to have been subsumed under one of the other categories.

The total number of (YNQ) tokens and the relative frequencies of each variant are shown in table 4.5, and the equivalent details for (WHQ) are given in table 4.6. (The observed frequencies for each variant are not provided, but can obviously be deduced from the information in the tables.)

Table 4.5 Relative frequencies of variants of (YNQ) in the corpus

[SV]	[ESV]	[V-CL]	N=
79.4%	20.6%	0%	(180)

Table 4.6 Relative frequencies of variants of (WHQ) in the corpus

[SVQ]	[QSV]	[QESV]	[QV-CL]	[QV NP]	[Q=S V]	N=
15.6%	23.8%	48.4%	6.6%	2.5%	3.3%	(122)

The overall ratio of tokens of (YNQ) to tokens of (WHQ) is almost precisely 3:2, and it is interesting to note that a very similar ratio has emerged from previous surveys. Indeed in all those quantitative studies referred to in 4.5 which provide figures for both types of interrogatives, the number of (YNQ) tokens is consistently greater than the number of tokens of (WHQ). This preponderance may be partly due to the fact that Yes/No interrogatives presuppose less than do WH interrogatives, and are therefore more likely to be used in situations where the "pool"

of presuppositions shared by the speaker and addressee is more limited (cf Brown & Yule, 1983:79-83). It is also no doubt due to the fact that they are used for a rather wider range of communicative functions than are WH interrogatives (as will become clear in 4.12). The higher frequency in discourse of Yes/No interrogatives may also be linked to the fact (reported by Sadock & Zwicky, 1985:178-84) that they seem to be present in all languages, whereas some languages do without WH interrogatives by using corresponding Yes/No interrogatives with indefinite subjects. Yes/No interrogatives can thus be considered to be the more basic syntactic structure, both typologically and in discourse.

Looking now at the relative frequencies of the (YNQ) variants in table 4.5, we see that the speakers in this study resemble the working class and the nine year-olds in table 4.2, in their overwhelming preference for [SV] and their avoidance of [V-CL]. However, [ESV] is considerably more frequent in the present corpus than in all varieties represented in table 4.2, except the formal middle-class (radio) variety.

Comparing now table 4.6 with table 4.3, our data seem to resemble most closely Pohl's working-class variety (though the latter may have included small numbers of tokens of [QkSV] and other structures). Our informants differ from the children in table 4.3, and also from Behnstedt's working-class speakers, in that they did not produce any examples of the stigmatised [QkSV] structure. However, they also differ from most of the more conservative varieties in table 4.3 in their higher rate of use

of the non-standard [QSV] structure.

4.8 The problems of acceptability and equivalence

4.8.1 The need for judgements of acceptability and equivalence

In the analysis of linguistic variation, it is normally assumed that in variable contexts all variants are both possible and equivalent, in the sense that they do not change the meaning of the utterance. However, as we have already seen in chapter 1, although this assumption can be made more or less automatically in the study of phonological variation, this is certainly not the case with variation involving meaningful units. It is argued here that a rigorous analysis of grammatical or syntactic variables requires the resolution of the twin problems of the acceptability and equivalence of alternatives. Behind the straightforward comparison of the relative frequencies of variants lies the tacit assumption that for all tokens there exists a choice of "different ways of saying the same thing" - ie that, for each token, the alternative structures would be acceptable in that context, and that their use would produce utterances which would be semantically and pragmatically equivalent to the utterance actually produced.

The problem of acceptability is particularly acute in the case of (WHQ), since there are many tokens where it is clear that one or more of the alternative structures is unacceptable. For example, occurrences of [Q=S V], such as (62), have acceptable equivalents with the [QESV] variant, eg (63), but there are no possible

equivalents with the [QSV], [SVQ] or [QV-CL] variants: (64), (65)

and (66) are all quite unacceptable:

- (62) qui y a été? (23982) [Q=S V]
- (63) Qui est-ce qui y a été? [QESV]
- (64) *Qui il y a été? [QSV]
- (65) *(Il) y a été qui? [SVQ]
- (66) *Qui y a-t-il été? [QV-CL]

The [QV NP] structure is of course impossible here, since there is no NP subject.

Another example, which is rather more significant numerically than that just discussed, involves (WHQ) tokens where the WH word is que or quoi (which are in complementary distribution). For example, (67), which is a token of [QESV], does not have a possible alternative with the [QSV] variant, since (68) is quite unacceptable to native-speakers, as is (69):

- (67) qu'est-ce que tu faisais? (1418) [QESV]
- (68) *Que tu faisais? [QSV]
- (69) *Quoi tu faisais? [QSV]

(Examples such as (69) are, it seems, acceptable in Canada - cf Lefebvre, 1981:235 - and may also exist, albeit very marginally, in France.) Tokens such as (67) thus involve neither categorical nor completely variable contexts - some, but not all of the other variants are available to the speaker. We shall refer to these as "semi-variable" contexts: ie where at least two, but not all, variants are acceptable and equivalent.

The surveys referred to in 4.5 were mostly carried out before the current debate on syntactic variation (eg Lavandera, 1978), and, not surprisingly therefore, did not take account of the problems of acceptability and equivalence. (Certain researchers did, however, exclude categorical cases, ie tokens where no

alternative structure would give an acceptable utterance equivalent to the one actually produced.) The findings of these earlier studies may therefore have been skewed to a certain extent, as a result of large numbers of tokens in semi-variable, and even categorical, contexts.

4.8.2 The equivalence of variants

An even more thorny problem is that of equivalence: with both (YNQ) and (WHQ) it is essential to consider whether the alternatives to a given token would not only convey the same propositional meaning, but also fulfil the same communicative function. (70), for example, was used as a request for action, but the theoretical alternatives to this utterance, (71) and (72), although acceptable syntactically, are not pragmatically equivalent to (70), since they cannot function as requests for action:

- (70) tu coupes un petit peu? (12103) [SV]
- (71) Est-ce que tu coupes un petit peu? [ESV]
- (72) Coupes-tu un petit peu? [V-CL]

Both (71) and (72) would normally be requests for information.

4.8.3 Solutions to the problems of acceptability and equivalence

The solution to the two problems of the acceptability and the equivalence of variants lies firstly in systematically making such judgements for each token in the corpus, and secondly in taking these judgements into account in the quantification of the variability.

There is a certain irony in the fact that a variationist analysis should require the use of intuitive judgements of this kind, since sociolinguists have, in the past, attacked the uncontrolled use of intuitions as practised by linguists of other persuasions. However, Labov (1972a:191-9) did not in fact reject entirely the use of intuitions; rather he argued that they should be interpreted in the light of data drawn from the spontaneous use of language. In any case, it seems that, if we are to adhere to the principles of the variationist paradigm, we must inevitably have recourse to intuitive judgements of acceptability and equivalence in analysing grammatical or syntactic variation.

No doubt it would be preferable to obtain these judgements from a large group of native-speakers, preferably indeed the informants themselves. This would be a major undertaking in itself, involving, for the present study, 360 judgements of acceptability and equivalence for the (YNQ) tokens, and 610 such judgements for the (WHQ) tokens. For practical reasons, therefore, the vast majority of these judgements had to be made by myself, a non-native speaker. Notwithstanding recent research which apparently shows significant differences in the intuitions of native- and non-native speakers with regard to certain structures (Coppieters, 1987), most of the judgements seemed unproblematic.

In cases of uncertainty, however, and also for all the negative Yes/No interrogatives, the assistance of native-speakers was sought. For the negative (YNQ) tokens, the judges were presented with the recording and transcript of the token with a substantial amount of context, together with an explanation of the topic of

conversation and any relevant background information. The judges were asked, for each token, whether the speaker could have used the two alternative structures to convey the same meaning as the original utterance. In fact the judges frequently found it difficult to give a definite answer. The problem seemed to lie not so much in judging the syntactic acceptability of the alternative utterances, but rather in determining the communicative function of the interrogative, and then deciding whether the alternatives could fulfil the same function. The judgements of the two native-speakers coincided in only a little over half of all the cases considered. The other cases were generally taken to be acceptable and equivalent, on the grounds that for a sequence to be considered completely unacceptable, there should be agreement among the native-speakers consulted.

If, as is assumed here, judgements of acceptability have at least some connection with the judge's estimation of the probability of occurrence of a sentence (cf Al, 1975, and Greenbaum, 1976, for support for this view), then these judgements should ideally be scalar. However, once again for reasons of practicality, the judgements here have been made in rather simplistic "yes or no" terms. Partly because of this, it was felt necessary to err on the side of liberality, and so, for example, many potential instances of inversion with je, such as (73), were judged acceptable, despite their archaic ring:

(73) Quel âge pouvais-je avoir?

The method of quantifying the variability which has been devised in order to take account of semi-variable contexts will be explained in chapter 6. But first it will be necessary to analyse

in some detail the communicative functions of the interrogatives in the corpus, and this will then enable us to judge, for each token, whether the use of the alternative variants would, or would not, produce "different ways of saying the same thing".

CHAPTER 5: THE COMMUNICATIVE FUNCTIONS OF INTERROGATIVES

5.1 Introduction

As was seen in 4.8, a variationist analysis of interrogatives requires that, for each occurrence, the potential utterances involving other variants should convey the same communicative function as the actual utterance. Not only is this an important theoretical consideration, it is also an essential precondition to investigating the influence exerted by communicative function on the choice of interrogative structure. Before this can be done, it is necessary to establish, as far as possible, the communicative function of each interrogative in the corpus. This is a particularly problematic exercise, since there is much disagreement over how (and to what extent) pragmatic aspects of the meaning of utterances can be represented. The classification to be employed in this chapter draws eclectically from several different sources: the categories of illocutionary force developed by Austin (1962) and Searle (1969), and applied to data from spoken Montreal French (with an encouraging degree of success) by St-Pierre (1977); secondly, the categories of conversation analysis elaborated by various researchers, and conveniently summarised by Levinson (1983:284-369); and thirdly, the communicative analysis of interrogation and negation in English by Leech (1983:157-69).

5.2 A set of distinctive features for categories of communicative function

Before discussing and exemplifying the various communicative functions^(CFs) of the interrogatives in the corpus, we shall outline in this section the set of features which differentiate these CFs. Each CF will therefore have a unique set of values for the features, which will provide its pragmatic definition.

These features fall into three groups. The first concerns the relationship between the speaker, the utterance and the addressee. The second covers relevant aspects of the knowledge, beliefs and assumptions of the speaker, and, to a lesser extent, of the addressee too. The third set of features deals with the propositional content of the question.

5.2.1 Features concerning the relationship between speaker, utterance and addressee

There is a substantial number of interrogatives in the corpus which are quotations and/or self-addressed utterances, and a smaller number which are echo utterances. Three features which serve to differentiate these utterances from each other, as shown in table 5.1, are:

- A: the speaker of the utterance is simultaneously the original "author" of the words of the utterance
- B: the author is/was also the addressee (ie the person addressed by the author)
- C: the utterance is an identical or near-identical repetition (or echo) of a recent utterance by another speaker

Table 5.1

	A	B	C
self-addressed question	+	+	-
quoted question	-	-	-
quoted self-addressed question	-	+	-
echo question	-	-	+
"ordinary" question	+	-	-

Feature A specifies that the speaker is simultaneously the author, so those quoted interrogatives in which the speaker is quoting their own words from a previous occasion will be [-A]. Feature B caters for self-addressed questions, whether or not they are also quoted. It should be mentioned that, in quoted questions, the addressee (ie the person addressed by the author) is usually not the the same person as the hearer (ie the person addressed by the speaker). Feature C means that, included among echoes, will be those utterances which involve a slight modification in the words used, such as a change in the subject from tu to je. It also specifies that the utterance being echoed need not immediately precede the echo - in some instances, the echoed utterance occurs two or more conversational turns prior to the echo itself. In table 5.1 the semantic category "question" is used as a cover term for the large number of communicative functions (such as request for information, pre-announcement etc.), which will be discussed shortly. It should be added that a CF which has a negative value for a particular feature has, as one of its properties, the negated definition of that feature. For example, one of the properties of quoted questions is [-B], ie that the author is not also the addressee.

5.2.2 Features concerning the author's knowledge, beliefs and assumptions

The second group of features concerns the author's knowledge, beliefs and assumptions, and serves on the one hand to distinguish between "biased" (or "conducive") questions of various types and neutral questions, and on the other to set apart those questions which normally require a response (requests, offers) from those which do not, and which, pragmatically, are assertions of various kinds (rhetorical questions, pre-announcements etc.). Table 5.2 shows how various categories of question are differentiated by the following features:

- D: the author does not know (for certain) q / whether p is true
 - E: the author believes the addressee knows (or may know) q / whether p is true
 - F: the author wants to know q / whether p is true
 - G: the author assumes (or believes the addressee assumes) that p is true
 - G': the author assumes (or believes the addressee assumes) that p is not true
 - H: the author has had (or believes the addressee has had) a disposition to believe that p is true
 - H': the author has had (or believes the addressee has had) a disposition to believe that p is not true
 - I: the author believes q / p is particularly significant
 - I': the author has known q / whether p is true (but has forgotten)
- (where p is a proposition and q is the missing variable in a WH question)

Table 5.2

	D	E	F	G	G'	H	H'	I	I'
unbiased (neutral) question	+	+	+	-	-	-	-	-	-
affirmative question expecting "oui"	+	+	+	+	-	-	-	-	-
affirmative question expecting "non"	+	+	+	-	+	+	-	-	-
negative question expecting "non"	+	+	+	-	+	-	-	-	-
negative question expecting "si"	+	+	+	+	-	-	+	-	-
tentative assertion	+	-	-	-	-	-	-	-	-
affirmative emphatic assertion expecting "non"	-	+	-	-	+	+	-	+	-
negative emphatic assertion expecting "si"	-	+	-	+	-	-	+	+	-
sub-topic-introducing question	-	-	-	-	-	-	-	+	-
pre-announcement	-	-	-	-	-	-	-	+	-
post-announcement	-	+	-	-	-	-	-	+	-

A number of points call for comment or explanation here. Features D-H are derived, with only minor modifications, from the implicatures (with the same letters) in Leech's analysis of neutral and biased questions in English (1983: 165-168). As Leech points out (pp 164-5), some of his implicatures correspond closely to Searle's felicity conditions, although Leech's analysis is based rather on Grice's conversational maxims. D-F have, however, been altered so as to handle WH as well as Yes/No questions : q represents the missing variable in a WH question, and p the propositional content of Yes/No questions. A second point is that the qualifier "for certain" has been added in parentheses to D, in order to accommodate even "strongly biased" questions. Similarly in E the parenthetical qualifier "or may know" has been inserted, since it is quite possible to ask a question whilst believing that there is only a slight probability that the addressee knows the answer. G and G' are the author's actual (or new) assumptions or expectations, whereas H and H' are their cancelled (or old) assumptions or expectations (cf also Quirk et al, 1985:808). The parenthetical qualifier in G-H', "or believes the addressee assumes/has had", has been carried over from Leech's implicatures, although most, if not all, instances

of biased questions in the present corpus derive their bias from the author's assumptions rather than those of the addressee. In table 5.2 the fuller descriptions of the different types of biased questions have been abbreviated, so that "positively-conducive" questions are indicated in the table as expecting "oui/si", and "negatively-conducive" questions as expecting "non".

Those interrogatives whose CF is that of an emphatic assertion are traditionally known as rhetorical questions, although some writers employ the latter term in a very broad sense to refer to any question where the speaker either already knows the answer or where no answer is expected. We follow St-Pierre (1977), Borillo (1981), Quirk et al (1985) and others in restricting the term to those interrogatives which serve to convey an assertion of the opposite polarity to that of their surface polarity. Consequently all such interrogative emphatic assertions ("rhetorical questions") are necessarily conducive. In these utterances the author does already know the answer and believes that the addressee knows too, hence these are [-D] and [+E] in table 5.2. The author does not normally go on to provide an answer to the question, since this is unnecessary. In contrast, tentative assertions are [+D] and [-E], since in such utterances the author does not know, or is not certain of, the answer, and may indeed go on to say so explicitly. These interrogative assertions could generally be paraphrased by a corresponding declarative which included an adverbial such as peut-être - hence the term "tentative assertion". (Notice, however, that St-Pierre uses the term "assertion tentée ou mitigée" to refer to rhetorical

questions, apparently because they convey opinions rather than objective facts.) The last three CFs in table 5.2 bear names which suggest that their essential functions are conversational, but in so far as they have illocutionary force, it is that of assertions rather than of requests of any kind. They all serve to highlight, or emphasise, certain information, hence they are [+I] in table 5.2. In all three cases the author cannot be said to want to know the answer, since it is already known - hence they are all [-D] and [-F]. Post-announcements highlight some information already known to the addressee and so are [+E], whereas pre-announcements and sub-topic-introducing questions highlight information about to be given by the author and which they believe the addressee does not already possess, hence [-E]. (It should be added that, in both pre-announcements and post-announcements, the relevant missing variables - q or the truth value of p - are those of the indirect interrogatives which are preceded by tu sais, tu vois or tu as vu.) All three of these CF categories involve neutral questions and consequently have negative values for G-H'. In table 5.2 the sub-topic-introducing question and the pre-announcement have the same values for all features, but they will be differentiated by one of the features to be discussed presently. Finally it may appear from table 5.2 that I' is redundant, since all the CFs have a negative value for this feature. However it will be needed for the specification of other CFs, and has been presented here since it is a feature concerning the author's knowledge.

5.2.3 Features concerning propositional content

The third set of features concerns the propositional content of the question (or, in certain cases, notably rhetorical questions and so-called "indirect speech acts", that of an implicated statement). These features cover similar ground to some of the preparatory conditions, as well as the propositional content, which form part of the definition of Searle's speech acts. They are, of necessity, a less tightly knit set of features than the first two sets that we have discussed, and it will be apparent that there is considerable redundancy, since each CF will have not more than one positive value for a feature specifying its propositional content. In table 5.3 a large group of communicative functions are distinguished from each other by this third set of features, which are as follows:

J: p is presented as objective fact
L: p = 'you have cognizance of X'
M: p = 'you mean X'
N: p = 'I should do X'
P: p = 'you (with or without others, but excluding me) could do X'
Q: p = 'I and you (with or without others) will do X'
R: p = 'you will do X'
S: p = 'you want X (which I am able and willing to help
provide/fulfil)'
T: p = 'I have your consent to do X'
(where X is any variable)

Table 5.3

	J	L	M	N	P	Q	R	S	T
request for information	+	/	/	/	/	/	/	/	/
request for opinion	-	/	/	/	/	/	/	/	/
check (on knowledge of addressee)	+	+	-	-	-	-	-	-	-
request for clarification	+	-	+	-	-	-	-	-	-
request for advice	-	-	-	+	-	-	-	-	-
request for reminder	+	/	/	/	/	/	/	/	/
suggestion (for action without author)	-	-	-	-	+	-	-	-	-
suggestion (for action with author)	+	-	-	-	-	+	-	-	-
request for action	+	-	-	-	-	-	+	-	-
offer	+	-	-	-	-	-	-	+	-
request for permission	+	-	-	-	-	-	-	-	+
tentative assertion	-	/	/	/	/	/	/	/	/
emphatic assertion	-	/	/	/	/	/	/	/	/
sub-topic-introducing question	+	/	/	/	/	/	/	/	/
pre-announcement	+	+	-	-	-	-	-	-	-
post-announcement	+	+	-	-	-	-	-	-	-

Again it must of course be borne in mind that table 5.3 does not provide full feature specifications for the CFs, and consequently certain CFs have the same set of values as each other in this table, but are differentiated by their values for other features. This is the case, for example, with request for information and sub-topic-introducing question, and it is no coincidence that instances of the latter, taken out of context, appear to be straightforward requests for information. A negative value for a given feature continues to mean that the negated definition of the feature forms part of the specification of the CF in question. Thus one of the properties of an offer is [-R], ie that the propositional content is not 'you will do X'. Where some realisations of a CF are +, and others -, a given feature, this is indicated in the table by /. It can be seen from the table that it is not possible to specify the propositional content of CFs such as request for information, request for opinion, request for reminder, tentative assertion, emphatic assertion or sub-topic-introducing question, since realisations of these CFs have a vast, if not infinite, number of different propositional

contents. The feature J serves to distinguish, on the one hand, CFs such as requests for information and for action (which are [+J]), and, on the other, requests for opinion and for advice (which are [-J]). There is no need for a feature "p is presented as subjective opinion", since the values for this would obviously be the exact mirror image of those for J. Realisations of CFs which are [-J] can always include an overt expression of subjectivity, such as à ton avis, tu trouves que?, whereas these would be quite inappropriate in utterances which are [+J].

In feature L the expression "have cognizance" is intended to cover the lexical items savoir, connaître and voir (the latter in both the present and perfect tenses): positive values for L are found in the CFs check (on the knowledge of the addressee), pre-announcement and post-announcement.

The features M to T each yield only one positive value in table 5.3, since they all deal with the propositional content of utterances with one specific type of illocutionary force. Following Martins-Baltar (1976), for example, we distinguish here between suggestions for action which exclude the author and those which include the author. This distinction is motivated not only by the respective linguistic properties of the two CFs (notably the different subjects involved), but also by the fact that "inclusive" suggestions amount to the first half of a joint resolution (the second half being the addressee's response), which is not the case with "exclusive" suggestions.

There is perhaps a case for setting up "request for object" as a

separate category from "request for action" on the grounds that the linguistic realisations of the two CFs do not overlap entirely. In particular the formula est-ce que je peux avoir is conventionally used to (politely) request objects, but not to request actions. Additionally, requests for action and requests for objects might be seen as differing in the extent to which they threaten the face of the addressee, and hence typically requiring different degrees of politeness. However, it is also true that the two types of request have many linguistic realisations in common, and the ranking of a request, whether for an object or action, will vary considerably according to the cultural value attached to the object or action requested. Moreover, if we were to make this distinction in our taxonomy, it would be necessary, for the sake of consistency, to set up two types of offer also: "offer for action" and "offer for object". On balance, it seems that the case is not sufficiently strong to justify two further categories.

The CFs figuring in tables 5.1, 5.2 and 5.3 do not account for all those occurring in the corpus - there are in addition several "complex" CFs, such as quoted offers and so forth. A matrix providing full feature specifications is to be found in Appendix 1. In the next section we proceed to discuss and give specific examples of the CFs from the corpus.

5.3 Exemplification and discussion of categories of communicative function

5.3.1 Introduction

In the rest of this chapter we will consider examples of each CF occurring in the corpus, and discuss, in the context of previous studies where appropriate, their linguistic and pragmatic properties. On the whole we will follow the order of presentation of the CFs established in section 5.2, except that conducive questions will be discussed as a group at the end, partly because conduciveness is a property which runs in parallel with the various other categories, and partly because, by grouping together different types of biased questions, we shall be able to examine a larger number of examples in this complex area. Additionally, quoted questions and self-addressed questions will be discussed immediately after the unquoted, other-addressed questions to which they are related. It is of course the case that quoted questions form part of larger utterances (generally, assertions) made by the speaker, but it is perhaps more helpful at this stage to classify them according to the CF apparently intended by the original author. Where possible, examples of both WH and Yes/No realisations of each CF will be given.

Interpreting the CF of utterances is, of course, rather problematic at times, as has often been reported (eg by van Kleeck et al, 1985). Nevertheless, the fact that, in the present study, the analyst was present in the speech situation, had substantial background knowledge and, in most cases, was also the addressee, facilitated this task considerably. Of relevance here

is Kreckel's (1981) finding that participants in a conversation were significantly more successful at interpreting the meanings intended by speakers than were outside observers.

A particular problem which is encountered in any attempt to specify the CF of utterances is the phenomenon of indirect speech acts, which have been discussed extensively in the literature (cf the summary in Levinson, 1983:263-76). The position which will be adopted here is the Standardization Thesis of Bach & Harnish (1979:192-5), according to which a strategy is standardly used to convey a particular illocutionary force, if it is mutually believed (within the speech community) that when someone utters this strategy, their intent is to indirectly convey that illocutionary force, and "when the utterance of this strategy is in a context where its literal direct interpretation would run counter to one's conversational expectations" (Fraser, 1983:50).

5.3.2 Request for information

As might well be expected, straightforward non-conducive requests for information account for a large number of the interrogatives in the corpus, both Yes/No and WH. In extract (1) the informant made two successive requests for information, the first a WH and the second a Yes/No question:

- (1) A: ...une année moi j'ai fait deux sessions de suite en colo /
c'est un peu trop hein je trouve/ et je sais pas comme * euh *
B: * et tu * avais - qu'est-ce que tu faisais? t étais
animateur? /
A: oui oui / (1418)

Di Sciullo & St-Pierre (1982:168) state that, in Montreal French,

the formula s'il vous plaît can only occur in requests for action, and not in requests for information. This appears not to be the case in metropolitan French, witness (2), which was produced by a third person who interrupted an interview to ask this question of the informant:

(2) X: est-ce que vous savez où sont les épuisettes s'il vous plaît? / (X211015)

We consider the initial est-ce que vous savez to be a formulaic politeness marker here and not to form part of the propositional content of the question, hence our classification of this utterance as a request for information rather than a check on the knowledge of the addressee. Moreover, in France, s'il vous plaît can occur freely even in those requests for information which do not have this initial politeness marker. Although there are in fact no examples of this in the corpus, (3) seems quite unexceptional as a polite request for information:

(3) Où sont les épuisettes s'il vous plaît?

Di Sciullo & St-Pierre suggest (p.165) that, in Montreal, s'il vous plaît serves in requests for action to mark insistence, whereas in these examples it seems rather to be a politeness marker.

There are a smaller number of quoted requests for information in the corpus, such as (4) and (5):

(4) ... (hein et) / "les gens veulent travailler?" "oui monsieur" on leur envoie une feuille / (28600)

(5) ... j sais pas si c'est toi qui demandais "qu'est-ce qu'ils mangent ce soir?"... (21685)

and just one example each of a self-addressed request for information, (6), and of a quoted self-addressed request for information, (7):

(6) B: [LOOKING THROUGH LIST OF EACH CHILD'S POCKET MONEY IN NOTEBOOK] ...i m semble avoir vu des trois cents francs voilà deux cents deux cents / là bon ben - cinquante là / c'est qui celui-là? /

A: Noël ... (3495)

(7) ...je vérifie les fiches puisque j'ai pour chaque enfant une fiche sanitaire / je vérifie - la santé disons durant l'année ou la santé euh de l'enfant / "est-ce qu'y a des problèmes cardiaques?" / (190)

In (6), despite the fact that it was the addressee who answered (he was also looking at the book), the speaker addressed the question not to him but to herself, since she would not have expected him to know the answer better than herself (he was a temporary visitor to the centre). Nor was her question a request to her memory to remind her of something she had previously known but had then forgotten. Rather, she was expecting to find the answer externally, ie from the book. In (7), there is no overt indication (such as je me dis) that this is in fact a quoted rather than a "live" question, however this does seem to be the most reasonable interpretation. Similarly, there is no overt indication that the question is self-addressed rather than addressed to another person, but the fact that the speaker has implied that she does this work alone (je vérifie and not on vérifie) again makes this the most plausible interpretation. It will be clear from this that interpreting the communicative function of quoted questions is particularly problematic, and that this is mainly due to the lack of context in such cases.

5.3.3 Request for opinion

Somewhat surprisingly, given the quite large number of requests for information made by the informants, there is just one occurrence each of a request for opinion and a quoted request for

opinion in the corpus:

(8) ...tu trouves qu c'est beau en tant qu pays la France? -
(25620)

(9) ...et puis ils me disent "bon ben je voudrais manger ci je
voudrais manger ça euh / est-ce que c'est faisable?" et puis on
essaie d'équilibrer malgré tout... (3426)

In (8) the request for opinion is made explicitly by means of the
formula tu trouves que. Some overt marker of subjectivity seems
to be more or less obligatory here, since, if the speaker had
asked (10) instead,

(10) C'est beau en tant que pays, la France?

he would have been implying that he himself did not have an
answer to the question, which would have been odd given that he
was a native who had travelled quite widely in France and the
addressee was a non-native visitor.

5.3.4 Check (on the knowledge of the addressee)

In contrast to the paucity of requests for opinion, checks on the
knowledge of the addressee are fairly numerous in the data. In
terms of conversational function, they could be divided into two
sub-categories. In the first the speaker has just introduced a
new entity into the discourse, and immediately checks whether the
addressee has cognizance of the entity mentioned, eg (11):

(11) ...bon y a eu un jeu de l'oie / tu connais le système du jeu
de l'oie? / (4495)

The frequency of such instances can be explained in part by the
fact that in most cases the addressee was not a native of the
same speech community as the speaker and thus could not be
assumed to possess all the speaker's background knowledge. The
second sub-category comprises cases where the speaker wishes to

introduce a new topic or sub-topic, and does this by asking whether the addressee knows the proposed topic entity (or in some cases a related one), as in (12):

- (12) B: ...- euh vous avez déjà entendu parler du traiteur (là Jacques Bord qui fait tant de réclame?) /
A: non /
B: non / eh bien lui il a beaucoup - il fait - il fait beaucoup de cantines comme ça / il apporte à manger / euh tout préparé / (21526)

Conversation analysts have sometimes called utterances such as these "pre-announcements" (cf Levinson, 1983:349-56), but we prefer to reserve this term for certain questions where the addressee clearly could not know the answer, which is not the case in (12) and other utterances of the same type. There are no examples of quoted checks or self-addressed checks, and whilst the former are not difficult to imagine, the latter seem to be inconceivable in normal speech, since a normal speaker presumably never has to ask themselves whether they have cognizance of something (except for cases where the speaker temporarily forgets, and these are classed as self-addressed requests for reminder; cf 5.3.7).

5.3.5 Request for clarification

There are a small number of requests for clarification, both unquoted and quoted, in the corpus and these are exemplified in (13-15):

- (13) A: ...des trucs comme tomate hein /
B: la tomate oui /
A: oui - tous ces jeux de ballon ou bien des jeux d'intérieur qui sont connus aussi / enfin / à peu près tous ces jeux-là sont pas connus chez nous hein /
B: "la tomate" tu parles de (le) ballon au milieu son - on se l'envoie et on doit pas le passer entre les jambes? / (9789)
(14) A: est-ce que tu crois qu'il est important que les animateurs gardent leurs distances vis-à-vis des gamins? /

- B: qu'est-ce que tu appelles "les distances"? /
 A: ben * (si tu veux . .) *
 B: * jusqu'à quel point les * distances euh? / (4401)
 (15) ...les gamins / y en avait un certain nombre qui voulaient faire du vélo. // euh - bon - dans le système euh fonctionnement habituel // tous les jours - on redémarrait pour savoir qui veut faire du vélo / et là / euh on avait euh à partir de / "bon qu'est-ce que vous voulez faire?" / "vélo" // on avait dit "mais - 'vélo' c'est quoi? / vous voulez - apprendre à faire du velo? / ... (28141)

Not surprisingly, these requests for clarification tend to occur at the beginning of the author's turn when they are not quite clear what the addressee meant by a particular expression. The conversational function of such utterances has been termed "next turn repair initiator" (cf Levinson, 1983:339), as they invite the addressee, in their next turn, to repair some aspect of their preceding turn which has caused a communication breakdown, eg a misunderstanding.

Certain types of echo questions have the CF of request for clarification, although there is just one such example in the corpus:

- (16) A: [INAUDIBLE]
 B: qu'est-ce que je pense (.)? /
 A: [INAUDIBLE] (9370)

Despite the inaudibility of A's preceding and following turns (inaudible, that is to the transcriber, but not, it seems, to B), B's utterance was certainly an echo of a question from A. Levinson (1983:351) seems to imply that all echo questions serve conversationally as next turn repair initiators. Quirk et al (1985:835-7), on the other hand, not only distinguish between echo questions and echo exclamations, but also suggest a distinction between explicatory and recapitulatory echo questions. The former correspond to what we are calling echo requests for clarification, whereas the latter call for

repetition rather than explanation. Most echo questions in the present corpus are not in fact requests for clarification, nor indeed, next turn repair initiators, but simply fillers, as we shall see later.

5.3.6 Request for advice

Given the nature of the data, and the relationship between the principal interlocutors (the fieldworker and the informants) it is perhaps unsurprising that there are no requests for advice as such in the corpus. However there is one quoted request for advice, (17), and a few quoted self-addressed requests for advice, including (18) and (19):

(17) ... lors ils discutaient avec Christine "comment tu veux qu ce soit?" / bon puis "qu'est-ce qu'i faut?" "mais oui mais i fait chaud dans cette pièce-là faut faire l'aération comment on va le faire?" - etcetera ... (21660)

(18) ... alors je me posais pas (d'autres) questions par rapport / fin si je m'en posais (je me suis dit) "bon euh qu'est-ce que je fais maintenant? j'ai vingt-quatre ans est-ce que je continue?" ... (141834)

(19) ... ben j pa j commençais à paniquer pas mal "(non) mais qu'est-ce que j vais faire moi?" parce qu'i fallait bien / (2541)

As already mentioned, the relative lack of context in quoted questions can pose problems of interpretation, and in (19) it is possible that the quoted request for advice was not in fact intended to be self-addressed, though there is no indication of an interlocutor, and indeed the low volume at which the utterance was produced would suggest that the author may not even have actually vocalised the original utterance.

5.3.7 Request for reminder

We distinguish requests for reminder from requests for information on the grounds that in the former the speaker has previously known the answer, but has forgotten it. In (20) the speaker asks to be reminded of the question which had been put prior to an interruption in the interview:

- (20) A: ... et quand ils partent en randonnée // qu'est-ce que tu leur donnes comme enfin / tu leur prépares déjà des menus tout ça? / * (.) *
B: * non non non non * c'est (eux) - pardon? / [INTERRUPTION]
bon qu'est que t - * qu'est-ce que tu m'avais demandé? * /
A: * alors pour les randonnées * / euh /
B: oui? / (3416)

Far more numerous in the corpus, however, are self-addressed requests for reminder. In most of these the forgotten information is a matter of content, as in (21) and (22), but in others it is one of form (a proper name or single lexical item), as in (23):

- (21) ... et on est dans une région // qui euh est-ce que c'est encore le Morbihan? oui / oui le Morbihan va jusqu'au / jusqu'aux Côtes-du-Nord au-dessus hein / (1457)
(22) A: et des activités manuelles ça te plaît * des? * /
B: * ah oui j'adore ça * /
A: * quel genre de chose? * /
B: * qu'est-ce que j'ai * fait? / [LAUGH] l'année dernière j'avais fait la vannerie. / (7180)
(23) ... euh je suis resté dans une colonie - en tant que comment (on) pourrait dire? - que visiteur quoi ... (694)

These self-addressed requests for reminder of form are restricted to WH interrogatives, and indeed in many cases are realised by more or less formulaic utterances, such as that in (23).

5.3.8 Suggestion

In contrast with the fairly large number of interrogatives in the corpus which have what could broadly be called

information-eliciting CFs, the number of action-oriented tokens is quite small. It seems highly probable that this is principally because the data are drawn from what were basically interviews, and that action-oriented CFs would occur more frequently in settings where the participants are involved in some activity in addition to conversation, such as eating, playing, buying/selling, making decisions etc.

(24) and (25) are the only examples of non-conducive suggestions, exclusive (ie excluding the author), and inclusive, respectively. The one in (25) was in fact produced, not by an informant, but by a third person:

- (24) ... vous passez par ici? / [GESTURING FOR ADDRESSEE TO GO THROUGH ONE DOOR RATHER THAN ANOTHER] (211052)
(25) B: d'accord - non mais là on peut bosser / dix minutes un quart d'heure après euh / y a toujours un moment donc (des) animateurs peuvent descendre alors - on continue si tu veux (hein) /
X: (après - * c'est bien) * /
A: * oui oui * /
X: on fait comme ça? /
A: oui oui / (X286)

(26) is the sole example of a quoted non-conducive suggestion (inclusive) in the data:

- (26) ... mais c'est vrai que i disent pas i diront jamais / "tiens euh demain est-ce qu'on peut faire ça?" c'est "maintenant je veux faire ça." / (18733)

(As mentioned earlier, conducive questions of various CFs will be discussed together in sections 5.4 and 5.5.)

5.3.9 Offer

Although there are several quoted offers in the corpus, such as (27) and (28), there are just two "live" offers, (29) and (30). In the latter, one informant interrupted an interview with

another informant in order to offer a herbal tea to the fieldworker and the interviewee (the vous is for plurality, not politeness in this case):

- (27) ... enfin Bertrand a téléphoné (à Paris) en disant à Jacques euh - un - un moniteur un animateur / "est-ce que tu veux aller au mariage de Christine?" "d'accord bon on arrive" et puis ils sont allés comme ça / (3674)
- (28) ... et ils ont dit aux ados "voilà y a tel et tel film / à aller voir euh qu'est-ce que vous voulez voir?" ... (4478)
- (29) ... et qui qu'est-ce que tu veux qu je dise? que j t raconte comme on a fait? / (29303)
- (30) X: vous voulez la tisane? /
B: oui si tu veux - j veux bien en tout cas. / (1521336)

5.3.10 Request for permission

The only occurrences of requests for permission in the data were produced by third persons addressing informants who were being interviewed, and who were in a position of authority over them. In (31) an animateur asks his director for permission to take a group of children out of the camp:

- (31) X: est-ce qu'on peut sortir avec les trente-cinq gosses? - *
(on est deux) * /
B: * vous êtes * combien? / (1279)

In (32) it is one of the children who asks a director/animateur for permission, and in (33) it is the informant's own small child who asks for permission. (This last example would suggest that even very young children are aware that their parents have authority or power over them.)

- (32) X: François - j peux le peindre? //
B: oui si tu veux. / (8225)
- (33) B: ... * (. . . .) * -
X: * maman je peux mettre * un pansement? /
B: oui mets-le / (30498)

5.3.11 Request for action

Some speech act taxonomies have distinguished between three types of "impositive" or "directive" acts, commands, orders and requests (ie for action). This is not done here, however, since these categories do not differ from each other in their values for any of the features outlined in 5.2. Commands and orders seem in any case to be practically identical - commands possibly being restricted to institutional settings - and the fact that English and French have two nearly synonymous performative verbs is, in itself, insufficient grounds for setting up two parallel speech act categories. Moreover, any difference between requests and the other two would have to be in terms of their relative politeness, which is deliberately excluded from the CF features used in this study.

The small number of requests for action in the data are all Yes/No interrogatives, as one would expect, and they all occurred as a result of an intervention from a third person. In (34) the request for action was made by the third person (who was also, however, interviewed as an informant on another occasion) to the interviewee, whereas in (35) the informant (jokingly) made a request to the third person, who was a passive (and occasionally active) participant in the conversation. Finally, in (36) the informant addressed his request for action to the fieldworker - to switch off the tape recorder, because of the imminent arrival of a local official, with whom he needed to have a fairly lengthy discussion:

(34) X: elle est là Odile? /

- B: Odile? /
 X: oui /
 B: elle est elle doit faire la sieste Odile. /
 X: tu pourrais bien lui donner son tee-shirt? - (elle l'a trouvé.) / (15249)
- (35) A: alors euh est-ce que je peux te poser deux questions euh * délicates avant de terminer? * /
 B: * oui à - à cents francs ou euh * / [TO X] bon écoute tu peux * partir deux minutes? il a des questions délicates là. * /
 A: * euh [LAUGH] * / tu as quel âge en fait? / (4908)
- (36) B: ... moi ça m paraît important. // [ENTER X]
 X: c'est monsieur (. .) /
 B: oui / tu coupes un petit * peu? parce * que /
 A: * oui oui * / [RECORDING STOPPED] (12103)

In (37) the speaker quotes a request for action made to her by a very young child (a rare quoted request for action in the corpus):

- (37) ... par exemple y a une petite fille qui me dit "ah j'ai / je me frotte les - noe_ils euh tu me mets de de la couleur sur les noe_ils?" bon ben c'est - c'est les yeux quoi - (17317)

In fact the speaker here appears almost certainly to have invented this quotation, in order to give a context to the word noe_ils (the plural of oe_il in child language, with the initial [n] presumably deriving from the child's reanalysis of the resyllabified liaison [n] in un oe_il, mon oe_il etc as the initial consonant of the word for "eye"). This serves to emphasise that the form of quoted questions is determined ultimately by the speaker, who can change the precise words used by the original author of the question, or even invent the quoted question themselves, as in the case above. However it also seems that speakers make the structure and words of the question appropriate to the real or supposed author of the original utterance.

5.3.12 Tentative assertion

All of the CFs discussed so far have had a positive value for the feature F, ie the author of the utterance wants to know the answer to the question. However, the next group of CFs to be discussed are [-F], and are all, in a broad sense at least, assertions.

A category which seems not to have figured at all in previous taxonomies of speech acts is the tentative assertion. (St-Pierre's "assertion tentée", it will be recalled, corresponded to the "rhetorical question", which in this study we call the emphatic assertion.) Unlike emphatic assertions, interrogatives with this CF are not necessarily conducive: those which are will be discussed later. All the neutral tentative assertions are in alternative questions, and are therefore classified here as Yes/No rather than WH interrogatives. Indeed it is not at present clear whether the latter are possible for conveying neutral tentative assertions.

An example of a non-conductive tentative assertion is in (38), which implicates (or could be paraphrased as) (39):

(38) ... culturellement on a eu une influence là-dessus. / mais est-ce que c'est bien ou mal? j sais pas. de toute façon ils étaient obligés / ... (14468)

(39) Je ne sais pas si c'est bien ou mal.

The only other occurrences of non-conductive tentative assertions are the affirmative members of a pair of juxtaposed interrogative clauses (together forming an alternative question) as in (40) and (41), which implicate (42) and (43) respectively:

- (40) ... (j'avais) envie de le dire à son père. / comment son fils se comportait. / et puis euh bon ben (ah - non) / alors est-ce que j'ai bien fait? est-ce que j'ai pas bien fait? j'en sais rien. / et faites-moi confiance que c'est vraiment un cas problème. / (11642)
- (41) ... parce que bon euh / on - on no - on fonctionne toujours dans les mêmes colos / ors on prend peut-être des habitudes euh / euh et puis euh est-ce que c'est bon? est-ce que c'est pas bon? euh / donc là on s confronte un petit peu aux autres pour voir euh - euh - ce que c'est / hein c que / si y a une évolution à donner ou pas. / (22160)
- (42) Peut-être que j'ai bien fait.
- (43) Peut-être que c'est bon.

(Incidentally, it seems that the order of such clauses is necessarily affirmative + negative, rather than vice versa.)

Quirk et al (1985:826) use the term "ratiocinative questions" for some similar examples, but the use of this term for all questions where the speaker does not expect the hearer to reply seems too broad.

5.3.13 Emphatic assertion

It will be recalled that emphatic assertions ("rhetorical questions") are distinguished from tentative assertions by their being [-D] and [+E], ie the author does know for certain the answer to the question, and the author also believes that the addressee knows the answer. Indeed these utterances are generally used as a rhetorically marked (hence emphatic) way of expressing what the speaker considers to be self-evident truths. A second difference between tentative and emphatic assertions is that the latter are necessarily conducive, whereas the former are not. Despite their conduciveness, emphatic assertions will nevertheless be discussed at this point in order to contrast them with tentative assertions and for the sake of the completeness of the taxonomy.

The only Yes/No emphatic assertion in the data is that in (44).

(Here it seems necessary to provide a substantial amount of context, including the addressee's question in the preceding turn.)

(44) A: et le système de fonctionnement que tu as élaboré là ...
est-ce que ça correspond à une idéologie particulière? - à ton
avis / est-ce qu'i y a derrière / une philosophie particulière
derrière ça? /////

B: m - crois pas / ... en fait bon ça ça part de // fin du
respect euh - des adultes et euh et et des gamins et / enfin
c'est - hein - le choix puis - j veux dire j sais pas si c'est
euh vraiment une philosophie / euh / euh de dire aux gamins euh /
"ors là vous jouez euh un quart d'heure et si vous avez envie
d'aller ailleurs vous / vous allez jouer ailleurs" enfin bon
c'est c que font tous les gamins euh / hein / bon / euh - est-ce
que c'est vraiment est-ce qu'y a une philosophie derrière le fait
de dire euh "les gamins ne remontent pas dans leurs chambres
parce que bon euh i faut un coup d sifflet à c moment-là pour
euh "?/ et puis euh de / de les mettre / en situation de jouer -
quatre heures c'est-à-dire de huit heures du matin jusqu'à midi /
euh plutôt qu'une heure et demie dans un fonctionnement un peu
traditionnel / ... (28438)

The underlined interrogative in (44) implicates (45):

(45) Il n'y a pas de philosophie derrière le fait de dire ...

We can notice here the presence of both vraiment (in the "false
start" to the rhetorical question) and il y a, which are among
the linguistic elements frequently occurring in rhetorical
questions according to Borillo (1981). The fact that there are no
other Yes/No rhetorical questions in the data seems to be
fortuitous - Borillo (1981) gives numerous such examples, albeit
apparently confected (there is no indication of the source of her
data, nor is any context provided, and clitic inversion is far
more frequent than would be expected in conversational data).

WH rhetorical questions have been discussed somewhat less in the
literature. Quirk et al (1985:826) state that they are equivalent
to emphatically asserted declaratives in which the WH phrase is

replaced by a negative element. In French the following correspondences will operate: qui -> personne, que/quoi -> rien, où -> nulle part, quand -> jamais, combien de/ (le)quel -> aucun etc. However as we shall see below, there are certain cases of WH rhetorical questions which are not quite so straightforward. Like their Yes/No counterparts they are necessarily conducive, and their polarity is the opposite of the implicated declarative, as in (46) and (47):

- (46) ... on a conservé l'activité tennis. / bon // elle a fonctionné mais m mal / tu vois ça n c'est une activité - qui n s'adressait pas à cette tranche d'âge-là. / et on a fait une erreur. / bon - euh - qui ne fait pas d'erreurs? [LAUGH] hein / mais c'est (bien de faire l'erreur parce qu'on s'aperçoit .) // ... (22306)
- (47) Personne ne fait pas d'erreurs. (= Tout le monde fait des erreurs.)

In the implicated declarative (47), the WH element is replaced by a corresponding negative element. However this produces a somewhat awkward double negative, and the paraphrase given in parentheses is a far more natural version. Comparing this with the original interrogative, we can summarise the relationship thus: WH [negative] -> V [affirmative]. If we now consider affirmative examples in addition, we see that this formula needs to be modified to: WH [α polarity] -> V [$-\alpha$ polarity], though again the actual wordings of the implicatures thus produced are somewhat infelicitous:

- (48) ... c'est vrai que c'est difficile de faire autre chose que / déjà la maison les enfants le travail dehors / faire encore autre chose après euh / i faut vraiment du courage. / oui mais alors où allons-nous? hein / c'est là ou c'est la fin hein - des centres. / (13762)
- (49) Nous n'allons nulle part. (= Nous allons vers la fin.)
- (50) ... y a certains professeurs qui euh / qui acceptent / euh que / qu les enfants à l'éc à l'école euh les tutoient euh / fin / j dis où est l'inconvénient si? mais moi je n'accepte pas qu les enfants m tutoient. / tu vois / (22565)
- (51) L'inconvénient n'est nulle part. (= Il n'y a pas d'inconvénient.)

(52) ... moi j'ai fait un système simple - avec quatre piquets. /
je prends quatre piquets / que j'ai plantés dans - du moins là
non j'ai pas pris quatre piquets - qu'est-ce que je dis? / j'ai
pris trois (.) - qu'étaient en triangle disposés en triangle. /
(23511)

(53) Je ne dis rien. (= Je dis des bêtises./Ce que je dis n'est
pas vrai.)

(46) reveals quite clearly one of the reasons why speakers use
these rhetorical questions to express an emphatic assertion
instead of a more direct means. The speaker in (46) has just
admitted making a mistake, which, in politeness theory, is
classed as an act which threatens the positive face of the
speaker (Brown & Levinson, 1987:68). In order to protect his
face, the speaker then excuses the mistake with the rhetorical
question, which has the added advantage of implicating, rather
than expressing directly, the excuse. He thus leads the hearer to
join him in constructing mentally the implicated proposition,
thereby, he hopes, securing the hearer's agreement that this
proposition is true and is a valid excuse.

In the next three examples the implicated declarative involves
not only the replacement of the WH phrase by a corresponding
negative element, but also the addition of an appropriate form of
the modal verb pouvoir. These examples show in addition the use
of a verb of volition (vouloir here) in the expression of an
absurd or unrealisable idea (as pointed out for Yes/No rhetorical
questions by Borillo, 1981:10).

(54) on trouve des constantes comme ça / alors c'est à la fois un
petit handicap mais qu'est-ce tu veux? ça fait partie des règles
de vie on / en tout cas on n se donne pas pour but de les casser.
// (2798)

(55) Tu ne peux vouloir rien d'autre. (= C'est exactement ce à
quoi il faut s'attendre.)

(56) A: oui / est-ce que les enfants en général - enfin à partir
de l'âge de dix ans est-ce qu'ils participent un petit peu au
travail dans les fermes tout ça? /

B: ah non. / ah non. / i faut compter quinze ans. / ah oui - oui

- alors - ah oui / ah non non c'est parce qu'y a pas de / y a pas d'ouvrages pour eux / qu'est-ce que vous voulez qu'ils fassent? / et s tout est mécanisé euh / i faut des hommes euh / hein /

(13346)

(57) Vous ne pouvez pas vouloir qu'ils fassent quoi que ce soit.
(= Ils ne peuvent rien faire.)

(58) ... mais / ils vont trois jours en Angleterre pour trois cents francs. / avec le professeur d'anglais. / puis trois jours qu'est-ce que vous voulez qu'ils fassent pour trois jours en Angleterre? ils n'apprennent rien. / (13446)

(59) Vous ne pouvez pas vouloir qu'ils fassent quoi que ce soit pour (ie pendant) trois jours en Angleterre. (= Ils ne peuvent rien faire ...)

Finally we come to the most complex example of an emphatic assertion in the data (in fact, it is a quoted utterance, though this in itself does not seem to add to the complexity):

(60) ... mais par contre - on fixe des interdits - dans la salle à manger tu n'as vu - personne fumer. / parce que - c'est interdit et c'est - innégociable / mais c'est innégociable on explique pourquoi / "la raison c'est qu regardez autour de vous tout est en bois / euh si jamais ça crame là-dedans qu'est-ce qui va se passer?" / bon - donc c'est interdit à c niveau-là / ...

(2235)

In this case the simple replacement of the WH element by a negative, and the conversion of the interrogative into a declarative would produce (61), which fails to correspond even approximately to the author's intended meaning:

(61) Rien ne va se passer.

If, however, "negative element" can be taken to include not only the negative complements, but also the phrase quelque chose de négatif, then we could take (62) to be an implicature of (60):

(62) Quelque chose de négatif va se passer. (= Une catastrophe va se passer.)

However perhaps, instead of following this rather tortuous reasoning, we should simply recognise that some rhetorical questions implicate not a declarative of the opposite polarity, but rather one which says that the answer is obvious. (cf also the "questions tautologiques" discussed by Fauconnier, 1981:45, such as Le Pape est-il catholique?, which are an emphatic way of

saying C'est évident!)

Before leaving rhetorical questions, we should mention a number of points made by St-Pierre (1977) with which we have to disagree in the light of the data considered here. Firstly, she states (p.88) that the propositional attitude for these illocutionary acts is that the speaker's certainty regarding the truth of the proposition is not total, whereas in 5.2 it was stated that one of the features of emphatic assertions was [-D], ie that the author does know for certain whether the proposition is true. This divergence is partly explained by the fact that St-Pierre's category assertion tentée maps on to the two categories of emphatic and tentative assertions in the present study. Secondly we do not agree with St-Pierre (p.97) that vraiment cannot occur in rhetorical questions. On the contrary, along with Borillo (1981:10) we see this as one of their characteristic linguistic properties (cf also Quirk et al, 1985:808, on really). Thirdly St-Pierre (p.106) suggests that it is inappropriate, following a rhetorical question, to say "That is a lie ", since one cannot accuse someone of lying when they are expressing their opinion. However, it seems entirely appropriate to respond to certain rhetorical questions by saying "That is not true", ie to disagree with the assertion made.

5.3.14 Exclamatory assertion

Although no interrogative exclamatory assertions occur in the data, it is not inappropriate to discuss them briefly at this point, especially since, as Huddleston (1984:379) points out with

reference to examples such as (63) and (64), they have some affinity with rhetorical questions:

- (63) Isn't it a lovely day!
- (64) Was I furious!

Negative exclamatory assertions such as (63) implicate an affirmative assertion, as do the corresponding rhetorical questions. But when we consider affirmative exclamatory assertions, such as (64), we see that the implicated declarative is also affirmative, as in (65), rather than negative, as is the case with rhetorical questions:

- (65) I was furious.

It is clear, then, as several writers have pointed out, that in exclamatory assertions the proposition to be conveyed must be positive, and linked to this is the fact that they do not allow non-assertive items such as yet. One of the formal properties of exclamatory assertions in English is that the uncontracted form not is excluded:

- (66) *Is it not a lovely day!

This, and other factors, led Hudson (1975:9) to go so far as to suggest that the n't in exclamatory assertions such as (63) is not really a marker of negation at all.

Turning now to French, several grammars state that the [V-CL] structure can be used to make exclamatory assertions, and give the following examples, among others:

- (67) Est-ce beau! (Wagner & Pinchon, 1962:533)
- (68) Est-elle jolie! (Chevalier et al, 1964:96)
- (69) Est-il intelligent! (Mauger, 1968:382)

Diller (1984:87) also gives examples with the [ESV] structure:

- (70) Est-ce que tu es fou!

However the use of [V-CL] in this way seems to be on the verge of extinction: native-speakers are rejecting cases like (69) as unacceptable, and the more recent Judge & Healey (1983:428) grammar comments that examples such as (71) now seem archaic:

(71) Combien ce pauvre homme a-t-il fait d'efforts!

There are no occurrences in the present corpus of interrogatives being used to make exclamatory assertions. In the vernacular French of Montreal, however, it seems that such use is not at all uncommon: St-Pierre (1977:110) gives the following examples:

(72) Il est-tu beau avec son habit neuf!

(73) C'est-tu fou d'affliger un enfant d'un nom comme Rogatien!

(It will be recalled that -tu is the Québécois form of the interrogative enclitic -ti.)

5.3.15 Sub-topic-introducing question

Another CF for which interrogatives are used in the data is similar to what Comrie (1984:40-1) has called "topic-introducing questions" in Russian. He reports that the latter are used to establish the topic, especially in journalistic, political and popular scientific discourse (presumably written). In our data these questions do not introduce totally new topics, but serve rather to allow the speaker to develop the topic along the lines they wish. Although the majority of the tokens with this CF are WH interrogatives, there is the following example of a Yes/No interrogative:

(74) A: est-ce que vous faites un bilan ensemble après la colo? /
fin y a des gens qui font un bilan j sais pas au mois d'octobre
comme ça. / vous faites ça ou pas? /

B: ben j s je sais pas parce que / je sais pas c que m - c que
prévoit euh la caisse des écoles de Bégard / et - ça veut dire -
est-ce que l y a des moyens // euh financiers pour ce genre de
chose? - bon je - je pense qu'en demandant euh - i y aurait pas

de problème / ... (28525)

In (74) the speaker says that the answer to A's question depends on the answer to another question, which he then puts in the form of a direct, rather than an indirect question, for expressive or rhetorical reasons which will be discussed shortly.

In contrast to the solitary example of a Yes/No sub-topic-introducing (S.T.I.) question, there are some twelve instances of WH interrogatives with this CF, among them (75) and (76):

- (75) ... on verra à la fin quand on fera le bilan s'i y en a qui veulent revenir et puis on discutera si / si nous aussi on a envie (qu'i reviennent) avec nous ou pas / hein // mais - comment d'habitude se fait se faisait notre groupe? - eh bien - bon je te dis je prends Marie-Françoise / qui est une monitrice moi je l'ai eue quand j'étais moniteur. / sa mère était économe dans la colonie où on était avec Christine à c moment-là / ... (21173)
- (76) ... hein parce que / l'économat ça (me) prenait quand même pas tout mon temps. / alors ça consiste en quoi? eh ben - ça consiste euh à animer / à animer un groupe d'animateurs c'est le cas de le dire. / (comme) je vous l'expliquais tout à l'heure euh i faut quelquefois les pousser ... (1124)

It is interesting to note the use of discourse markers immediately before and after these questions. In (75) mais precedes the question, perhaps signalling a more substantial shift in topic, whereas the lesser shift in (76) is marked by alors. Similarly these questions are frequently followed by bien (or its elided form ben) - optionally preceded by eh - or in some cases by bon. These markers are characteristically used at the beginning of an answering turn.

What is the function of this type of question? Intuitively, the WH S.T.I. questions seem to fulfil a similar function to that of certain pseudo-cleft structures, in that the WH phrase can be seen as heralding a corresponding element later on in the

utterance which the speaker wishes to emphasise. Thus the pseudo-cleft in (77) seems to present the information in a broadly similar fashion to the S.T.I. question in (76):

(77) Ce en quoi ça consiste c'est à animer un groupe d'animateurs.

However the S.T.I. question has the advantage (for both speaker and addressee) of generally being syntactically simpler than the pseudo-cleft, and also rhetorically more expressive. The speaker converts a part of their turn into a simulated dialogue, involving a question-answer sequence (the S.T.I. question is formally identical to a request for information), often accompanied by the appropriate discourse markers.

Intuitively, the effect seems to be that the addressee's curiosity to learn the answer to the question is increased. The main communicative function of these questions lies, therefore, not in their illocutionary force (they could be said to form assertions with the following utterances), but rather in the conversational and rhetorical purposes they serve.

5.3.16 Pre-announcement

Another CF in which the conversational rather than the strictly illocutionary purpose predominates is the pre-announcement. This term is used here in a somewhat narrower sense than that found in, for example, Levinson (1983). We reserve the term for cases such as (78), (79) and the quoted example in (80):

(78) non y a un jeu qu'is aiment bien les les grands c'est /
c'est la ronde. / t sais euh - (. . .) - t sais t sais pourquoi?
/ (parce) qu'on fait des jeux comme ça et puis euh / le principe
du jeu c'est que y a un garçon / qui va au milieu du rond / qui
va s mettre à genoux devant une fille - et qui l'embrasse /
(25392)

(79) ... il a passé son permis / il a très bien conduit // et puis l bonhomme euh / il était sûr de l'avoir son permis hein / le bonhomme lui a dit "vous (n) l'avez pas" / (alors il comprenait pas) / tu sais pourquoi il (n l) a pas eu? parce qu'il avait oublié d mettre la ceinture. // il avait oublié d mettre la ceinture ... (25743)

(80) ... l'information - commençait à / bon à se - euh à se dire / et il aurait suffi qu'une personne un peu malveillante du village dise "mais dites donc vous savez c qui s passe là-haut?" / un petit coup de telephone à la brigade volante / et après c'est la galère quoi euh bon / je sais pas si ça serait allé jusque là ... (28725)

According to Levinson (1983:354), pre-announcements serve to check the newsworthiness of the subsequent announcement. A number of the tokens in the present corpus which have been categorised as "checks on the knowledge of the hearer" serve this conversational purpose (although many others serve simply to check whether the addressee knows, or knows of, an entity referred to by the speaker). And it is clear that formally the interrogatives in the extracts above are identical to what we have called checks, and so, in a different context they could indeed be checking the newsworthiness of the subsequent announcement. Indeed, it has been pointed out by St-Pierre, among others, that misunderstandings sometimes arise due to the ambiguity of these utterances, as in (81):

- (81) - Sais-tu qui a gagné les élections?
 - Non. Qui?
 - Ah! je l'sais pas, j'pensais que tu l'savais. (St-Pierre, 1977:163)

However, in (78) and (79), it cannot be said that the interrogatives were intended to check the newsworthiness of the subsequent announcement, since the speaker does not wait for the addressee to respond. Moreover, in (79) the addressee could not possibly have known the answer to the embedded question, since the person being referred to was totally unknown to him. And in (78) it certainly appears that the speaker assumed that the

addressee did not know the answer, since he proceeds to describe the game as if it were new to the addressee. (80) is rather more problematic, since, as with many quoted questions, there is very little context. In particular, it is not clear whether the author (who is hypothetical in this case) was indeed checking whether the addressee (also hypothetical) had heard the news.

The function of the interrogatives in (78) and (79) seems to be to herald, or emphasise, what follows, in a way similar to S.T.I. questions. (We can note that in (79) the speaker repeats the answer, which would support the suggestion that he wishes to emphasise it.) However, in terms of the distinctive features outlined in 5.2 they are [+I], ie the author believes that the answer to the embedded question is particularly significant. This is typically one (often surprising) piece of information, in contrast to the generally more complex (but unsurprising) information which constitutes the sub-topic introduced by a S.T.I. question. This interpretation is very similar to that proposed by Rehbein (1981:242), with regard to examples such as (82):

(82) Do you know what I dreamt of last night?

He comments that, with the formula do you know (as well as guess what, you just can't imagine etc.), "the speaker refers to something the hearer does not know, but which the speaker characterises as worth knowing." In terms of distinctive features, pre-announcements differ from S.T.I. questions also by virtue of their being [+I], ie that their propositional content is basically "you have cognizance of X", whereas the propositional content of S.T.I. questions is as varied as that of

requests for information.

Pre-announcements also occur characteristically as the first utterance in riddle-type jokes, as in (83-85):

- (83) Vous savez pourquoi les Belges se bouchent les oreilles dans les églises? Ben parce que Jesus crie. (Walker et al., 1986:1.52)
(84) Tu sais pourquoi y a pas de bancs dans les églises à Cuba? Parce que les fidèles cassent trop. (Mark Rae, personal communication)
(85) Vous savez pourquoi il y a tant de pharmacies à Villeurbanne? Parce que le maire éternue. (France Inter, 1988, when Charles Hernu was mayor of Villeurbanne.)

The interrogative here is again a means of highlighting, as particularly significant, the answer: in these cases, the punchlines.

5.3.17 Post-announcement

By analogy with pre-announcements, interrogatives which highlight some preceding information which the speaker considers particularly significant are categorised here as post-announcements, eg (86) and (87):

- (86) ... sa mère était économe dans la colonie où on était avec Christine à c moment-là / j la connais elle avait neuf ans. / maintenant elle a vingt-quatre ans. / tu vois un peu c qui s'est passé? elle est prof de gym et tout maintenant / ... (21178)
(87) ... tu peux apprendre à conduire avec euh / avec un ami / et puis euh / tu tu t présentes pour passer ton permis. / mais en général euh / les inspecteurs i te / i te donnent pas ton permis parce que t as pas donné d'argent à la boîte. // t vois c qu i veux dire? / alors c'est pour ça // ... (25729)

Whereas in pre-announcements the speaker seems literally to be asking the addressee whether they know something which they normally cannot possibly be expected to know, in post-announcements the speaker seems to be asking whether the addressee has understood something which the speaker has just said. Since in both instances the answer is obvious, it is not

surprising that the answer literally goes without saying: the listener says nothing, and the speaker continues. In addition, both pre-announcements and post-announcements occur only as Yes/No interrogatives - in the present corpus, at least. However, the parallel between the two CFs is not total, since post-announcements highlight a larger chunk of discourse than do pre-announcements.

The interrogative in (88) has been classified as a quoted post-announcement, even though it is somewhat different from cases such as (86) and (87):

(88) ... alors dans des domaines différents i sont arrivés à êt - à pouvoir parler d'égal à égal aux autres si tu veux. / "toi t as peut-être une grosse tête mais moi t as vu c que j'ai fait avec mes bras? (et puis) - puis quelque chose j sais faire" / tu vois? alors y a le côté euh se revaloriser / ... (265)

As with other quoted utterances, the interpretation of (88) is somewhat problematic. On the face of it, it appears that here the author is highlighting, not something which he has just said, but something recently done. In this particular instance, the action was done by the speaker himself, and since he is clearly proud of it, the utterance has an element of boasting, or at least requesting praise, about it. However, where the action referred to has not been performed by the speaker, the effect can be something like a reproach:

(89) T'as vu ce que t'as fait?

Utterances similar to (86) and (87) have been placed by some writers into rather broader categories. Crystal & Davy (1975:95-7) include you see among what they call "softening connectives" or "softeners", and they state that it functions, medially and finally, to summarise what has been said so far, and

also to ask permission to continue, by providing the listener with an opportunity to interrupt. Quirk et al (1985:1115) include you see among "comment clauses", which, they claim, serve to call for the hearer's attention and also to "express the speaker's informality and warmth toward the hearer". Both these interpretations seem, perhaps inevitably, highly subjective, and neither provides a really satisfactory explanation of the function of the interrogatives in (86) and (87). Biggs & Dalwood (1976b:130-1) say that certain "fillers" (including voyez? and voyez-vous?) serve to maintain a rapport with the hearer, or to elicit a reaction. The former, and perhaps the latter, could be said of examples such as (86) and (87), but we must perhaps acknowledge that it is sometimes impossible for the analyst (even when they are also the addressee) to be certain of the communicative function of a speaker's utterance. And indeed in some cases it may be that the speaker had no very specific intention at all.

5.3.18 Echo question

Echo questions are those interrogatives which repeat, in part or in entirety, an utterance from an interlocutor's preceding turn (though not necessarily the immediately preceding one). According to Levinson (1983:339), echo questions can be classified, in conversation analytical terms, as a type of "next turn repair initiator". This is indeed the function of some echo questions, but not all, as will be shown below. Quirk et al (1985:835-7), as we have already seen, distinguish two types of echo question: explicatory, which request clarification rather than mere

repetition; and recapitulatory, which, whilst overtly requesting repetition or confirmation, may have other functions, "such as to express irony, incredulity, or merely to fill a conversational gap." Comrie (1984:37-8), on the other hand, makes a primary distinction between echo questions which follow another question and those which do not. He states that the function of both types is to request confirmation, though the first type is often used in practice by the speaker "as a stalling device to gain time to work out their answer." He adds that "In French and German an echo question of this type is constructed as a question in indirect speech (in French introduced by si, in German by ob, 'if, whether') but with the rising intonation characteristic of a general question in direct speech" (p.39). A (confected) French example of this would be B's echo in (140):

- (90) A: Vous allez au cinéma?
 B: Si je vais au cinéma?

However, no such examples occur in the present corpus, and they seem to be characteristic of formal, rather than informal, speech styles.

It is clear that echo questions do not in fact form a homogeneous class in terms of their CF. Of the echo questions occurring in the corpus, several are conducive requests for either information or clarification, and will be discussed later. Others, however, can best be described as "fillers" in that their function does indeed seem limited to filling a gap. The one Yes/No example of this is the echo in (91):

- (91) B: ... [ENTER X] euh * quand * -
 X: * Bertrand * est-ce que je peux avoir des agrafes et de la colle? /
 B: est-ce que tu peux avoir des agrafes et de la colle? / [B

STARTS LOOKING IN DRAWER OF DESK AT WHICH HE IS SITTING]
X: s'il te plaît. / (2174)

Since X had just entered the room and interrupted B in mid-sentence to make the request, it is quite possible that the echo was intended to be a slight rebuke, although the prosody does not convey this. X's subsequent s'il te plaît, at any rate, expresses extra politeness, but whether this was prompted by B's echo is by no means certain. (The relevance of the speakers' identities in this exchange will be discussed in chapter 6.)

The other echo fillers in the corpus are WH interrogatives, (92) being a fairly typical example:

- (92) A: et sinon qu'est-ce que tu fais d'autre? /
B: sinon qu'est-ce que je fais d'autre? // à vrai dire j fais pas grand-chose [LAUGH] / (231019)

It seems clear that the echo here has no other function than to fill a gap and to give the speaker time to think. To leave such a gap unfilled could be interpreted by an interlocutor as uncooperative, or impolite, behaviour.

5.3.19 Indeterminate communicative function

Despite a degree of uncertainty in some instances, it has proved possible to classify the vast majority of interrogatives in the data according to their communicative function. Nevertheless, at the end of this process there remain some ten tokens whose CFs cannot be determined precisely, although even in these cases we can generally say that the utterance had one of two CFs. The ambiguity can, for the most part, be attributed to one of the following three factors. Firstly, in quoted interrogatives there

is often only very limited contextual information to assist with the interpretation of the CF, especially in the cases where the quotation is hypothetical rather than real. Secondly, in some instances it is not even clear whether the question is quoted or not, since it has not been preceded by an overt marker such as j'ai dit. Thirdly, in quoted utterances the functional ambiguity of the interrogative is aggravated when on is present, since even in context it can be difficult to determine who on is referring to, and indeed whether it has definite or indefinite reference.

In (93) the interrogative is clearly marked as a quotation (albeit hypothetical), but due to the lack of context it is not clear whether it is intended to be self-addressed or not:

(93) si ç avait été un accident j'aurais dit "mon dieu / qu'est-ce qui s'est passé?" euh - c'est vrai à la suite euh des fois / d'une malveillance de quelqu'un ... (21870)

Assuming that the author is asking about the cause of the accident (and not simply making an exclamation), this utterance's CF is either a quoted request for information or a self-addressed request for information. However, given that the quoted utterance was hypothetical rather than real, it is entirely possible that the speaker, without intending it to be ambiguous, did not in fact intend it to be unambiguously self- or other- addressed either.

In extract (94) the four interrogatives are not overtly marked as being quotations, but at the same time they are clearly not requests for advice addressed to the hearer. If we assume that they are in fact quotations, or something similar, it is still by no means clear whether they are self- or other- addressed:

(94) ... les gens sont quand même euh à l'écoute des propositions des enfants. / euh - euh - les enfants proposent des activités bidon / bon - y en a qui ont été faites - parce que les enfants les ont proposées / et que les animateurs se sont dit "s'ils le proposent c'est qu ça (corre mande corre) quand même ça correspond à un besoin donc on va en faire / mais euh / on va mettre en parallèle - une activité comme ça - et puis une autre activité / et puis les enfants vont se déterminer / si" / c'est c'est tout - tout le vieux problème de "est-ce qu'on écoute? est-ce qu'on écoute pas les gamins? est-ce qu'on manipule? ou est-ce qu'on manipule pas?" même nous par rapport aux animateurs. / si on écoute les gamins ben ils passeraient les gamins passeraient leur temps à jouer au football. // ... (146334)

The five other interrogatives with indeterminate CF were all produced by an informant whose style of speaking was characterised by particularly long turns and a considerable amount of quotation. He was clearly at pains to explain his ideas on the activities in children's holiday centres and the sometimes complex issues involved. In extract (95) he was discussing the organisation of a (potentially chaotic) water-play activity for very young children:

(95) ... c qui s'est passé ce soir c'est que // ils se déshabillaient i venaient jouer - on les séchait - i remontaient. / donc c'est relativement clair. / s'is étaient venus habillés / y avait des shorts euh des / des maillots d corps des chaussettes dans tous les coins / on les retrouvait plus / euh certaines étaient mouillées f euh qu'est-ce qu'on fait? on s rhabille sur place et on retourne jouer ailleurs euh / euh - bon là y a une espèce de flottement et / et la journée s terminait en - comme ça enfin on / on sait pas trop comment quoi. / donc y a // une espèce de m // de de liberté pour les gamins de venir de n pas venir de / mais euh / et ça c j crois qu c'est une chose qu'on retrouve très rarement / c'est euh / c'est l'anticipation - c'est-à-dire / quand on fait une activité on sait comment on la maîtrise / hein - et la maîtriser ça veut pas dire euh coup d sifflet euh coup d gueulante ou etcetera enfin c'est euh / "bon / ça y est on arrête euh tu t rinces euh / y en a encore un qui traîne dix minutes de plus bon ben ça y est / i traîne on en parle plus quoi." / tandis que là bon si j'avais pas arrêté l'histoire euh / euh à trois heures et demie y avait tous les gamins ici quoi. / alors après y a - bon - plus personne contrôle rien euh / on sait plus c qu'on fait euh "est-ce qu'i remonte?" "ben non i retrouve plus ses fringues i sont mouillés" bon etcetera - tandis que là c'était / et ça c'est quand même tres difficile à expliquer aux gens. / ... (28298311)

In the first interrogative in this extract, the switch into the present tense in fait (from the imperfect tense of the preceding étaient, retrouvait and y avait) is the only indication that the interrogative was possibly intended to be a quotation (again, a hypothetical one), with perhaps the answer being given immediately by the hypothetical addressee: "on s rhabille ... ailleurs". The interpretation is once again made more difficult by the presence of on. In a question-answer adjacency pair (such as this), on most likely has first person plural reference. If so, then the (hypothetical) author of the question was probably a child, and the (hypothetical) addressee giving the answer was probably an animateur. If this interpretation is basically what the speaker intended, the CF would be a quoted request for advice. On the other hand, if the tense switch was simply a stylistic switch to the "conversational historical present", then there is nothing to indicate that the interrogative was intended as a quotation, and the only reasonable interpretation of it would then be as a sub-topic-introducing question. There does not seem to be any non-arbitrary way to choose between these two interpretations. The second interrogative poses similar problems: the ben non marks that utterance as an answer to the preceding question, but whether this quoted question was intended as a request for information or as a request for advice is by no means clear.

The same informant produced the two interrogatives in extract

(96):

(96) ... le travail d la journée c'est un peu voir / bon - où on en est? euh qu'est-ce qui s passe? euh // avec des interventions euh - ou ponctuelles ou ou même plus longues - ... (285067)

Once again there is no overt indication that these are quotations, except for bon and the pauses which set the interrogatives off from the preceding context. If they were indeed intended as quoted requests for information, it is still not clear whether they are self- or other-addressed. (97) includes the final two cases of indeterminate CF (once again from the same speaker):

(97) ... parce qu'on en avait discuté / et - par rapport à euh // euh /// euh - à - à c qui aurait pu s passer éventuellement / si cette information continuait à - à - à circuler hein / c'est-à-dire éventuellement petite enquête mais euh bon euh alors on interroge les gens mais euh pendant c temps-là les gamins qu'est-ce qu'on en fait? et puis euh / bon et les gens bossent plus parce que "oui qu'est-ce qui va arriver?" etcetera / c'est uniquement par rapport à la / euh à ça que je qu'i sont partis enfin / ... (287367)

In the first of the two, the present tense of the verb cannot be taken as indicating a quotation, since the shift to the historical present has already occurred in the preceding verb interroge. The left-dislocated indirect object NP (les gamins) seems, however, to suggest that the interrogative is intended to be quoted, but whether as a self- or other-addressed request for advice is again unclear. If it was not a quotation, then the only possible interpretation of its CF seems to be as a S.T.I. question, though a somewhat curious one, since the speaker does not continue with the sub-topic thereby introduced.

The second interrogative is marked as a quotation both prosodically and by the oui (which would otherwise be entirely inappropriate here). However, the total absence of quoted context again makes it impossible to say whether the intended addressee was the author or another person. As with several of the other tokens discussed in this section, the hypothetical nature of the

quotation makes the interpretation of its communicative function especially problematic.

5.4 Conducive questions

The parameter of conduciveness or bias intersects with the various CFs discussed in 5.3 and so has been left until now. The discussion of negative questions in 5.5 forms a bridge between the two areas of grammar which are the focus of this study: interrogation and negation.

5.4.1 Terminology

There is an astonishing number of terms, both in English and in French, to refer to questions which are biased or oriented to expect a particular response. In English these questions have been termed "biased" (Moravcsik, 1971; Sadock & Zwicky, 1985), "conducive" (Hudson, 1975; Quirk et al, 1985) and "loaded" (Leech, 1983); and in French, "dirigées" (Wagner & Pinchon, 1962), "incitatives" (Price, 1978), "implicatives" (Fónagy & Bérard, 1973) and "orientées" (Borillo, 1979). Wagner & Pinchon (1962:546) also use the term "insistantes" to refer to questions such as (98):

(98) Il est venu aujourd'hui, n'est-ce pas?

However, since these involve a declarative clause with an interrogative tag, they are not of direct concern to us here.

In his taxonomy of speech acts, Martins-Baltar (1976:132) has just one category to cover all types of conducive questions -

"demander accord sur la vérité d'un fait" - but this seems to ignore several important distinctions. St-Pierre (1977), on the other hand, distinguishes between "demande de confirmation" and "demande de reconfirmation", the former corresponding to "questions insistantes" such as (98). In the latter, the speaker presumes that the proposition is true, but is not completely certain, whereas in the former, they know that the proposition is true. Whilst this distinction is a valid one, it still only handles one dimension of "bias", ie the relative strength of the speaker's doubt regarding the proposition. The polarity of the interrogative, the orientation (positive or negative) of the expected response and the interaction between these two are important factors in the pragmatic meaning of conducive questions, and they may constrain the speaker's choice of interrogative structure. Moreover, in taxonomies such as Martins-Baltar (1976) and St-Pierre (1977), no allowance is made for the fact that speech acts other than requests for information can be conducive.

5.4.2 Four types of conducive questions

Following Leech's extremely lucid exposition (1983:166-9), we recognise here four types of conducive interrogatives, which are distinguished from each other both by their logical structure (an abbreviated formula of which, adapted from Leech, is given on the second line in each case, with an ordinary language paraphrase below that), and also by their implicatures concerning the speaker's beliefs regarding the truth of the proposition (as discussed in 5.2 above):

- (i) affirmative question expecting a positive answer ("aff oui")
[? TRUE [pos X]]
'Is it true that p (is the case)?'
- (ii) affirmative question expecting a negative answer ("aff non")
[? TRUE [pos TRUE [pos X]]]
'Is it true that it is true that p (is the case)?', ie 'Is it really true that p (is the case)?'
- (iii) negative question expecting a negative answer ("neg non")
[? TRUE [neg X]]
'Is it true that not p (is the case)?'
- (iv) negative question expecting a positive answer ("neg si")
[? TRUE [neg TRUE [pos X]]]
'Is it true that it is not true that p (is the case)?', ie 'Isn't it true that p (is the case)?'

(Key: [] enclose "propositionals" (ie a type of sentence sense covering propositions, questions and mands); X is the predication (or propositional content) shared by each member of a set of sentence types; neg is the negative operator, and pos the positive; TRUE and FALSE are predicates; ? is the interrogative operator.)

The logical structures given above are, of course, those of biased Yes/No questions, not of WH questions. Type (ii) is not in fact discussed by Leech, but the formula of its logical structure, and the paraphrase of this, have been constructed following the same principles as for the three types which he does deal with.

Sadock & Zwicky (1985:181) state that, although negative Yes/No questions are often positively-biased, in some languages (eg Japanese) they are neutral questions about a negative proposition. This would presumably imply that type (iii) above (and perhaps, by extension, also type (i)) are not biased at all. However, Leech (1983:166-7) treats (i)-(iv) as 'loaded' questions on the grounds that they are logically equivalent to the simple question [? X], and also therefore logically more complex and pragmatically more oblique than is strictly necessary. In her discussion of negative questions, Borillo (1979) calls (iii) negatively-oriented (or Q.CONFIRM-NON), and (iv)

positively-oriented (or Q.CONFIRM-OUI).

Types (i) and (iii) are realised exclusively as Yes/No interrogatives in this corpus, whereas types (ii) and (iv) involve certain WH interrogatives additionally. Although the vast majority of conducive questions in the corpus are Yes/No, there are a few WH interrogatives also, and therefore the phrases "expecting a positive/negative answer" must be interpreted more broadly than simply "expecting oui/non/si", despite the use of these to give abbreviated labels to various conducive CFs. Moreover, since some conducive questions involve CFs in which the speaker does not in fact expect (or receive) an answer - such as "rhetorical questions" - the phrase "expecting a positive/negative answer" should in these cases strictly be read as "implicating a corresponding positive/negative statement".

The number of tokens produced by the thirty informants for these four types of biased question is shown in table 5.4 (together with neutral questions for comparison).

Table 5.4 Frequency of biased and neutral questions in the corpus

	Yes/No	WH	BOTH
neutral questions [? X]	96	114	210
(i) "aff oui" [? TRUE [pos X]]	43	0	43
(iii) "neg non" [? TRUE [neg X]]	29	0	29
(ii) "aff non" [? TRUE [pos TRUE [pos X]]]	7	7	14
(iv) "neg si" [? TRUE [neg TRUE [pos X]]]	5	1	6
Totals	<u>180</u>	<u>122</u>	<u>302</u>

It is interesting to note that the frequency of these types of

question correlates precisely with their degree of markedness (assuming the latter to be increased by negation and greater logical complexity).

These four types of conduciveness can combine with many of the various communicative functions discussed in 5.3: in the next two sections we shall examine cases of affirmative biased questions, ie types (i) and (ii).

5.4.3 Affirmative questions expecting a positive answer

Since it is assumed here that all negative questions are conducive, as in English (cf Quirk et al, 1985:808), questions of types (iii) and (iv) are easily identifiable. However, with affirmative questions, it is necessary to rely to some extent on various linguistic and contextual indicators of bias. The first category of such indicators is the presence in the interrogative of a lexical (or discoursal) item which suggests that the speaker, although not knowing for certain whether p is true (hence [+D]), simultaneously assumes that p is true (hence [+G]). These items include alors and donc, which indicate explicitly that the speaker is making an inference from the preceding context, or the situation, (or indeed both), as in (99) and (100):

- (99) A: ... je fais une petite interview pour pour mon reportage
sur les colos. /
B: ah bon? /
A: oui /
B: oui alors vous êtes en route d'enregistrer là? / (1914)
(100) A: ... toutes les formalités tout ça tous les règlements de
la loi française - tout * ça ça m'a intimidé. * /
B: * ah c'é / c'était en France * donc? / (1426)

Other such markers include quand même (as in (101)), by which the

speaker signals that they assume the proposition to be true in spite of the preceding context, and ah (as in (102)), whereby the speaker expresses surprise at the proposition which they now

assume to be true:

- (101) B: ... pour toi c'est (.) j suppose qu c'est quand même plus pratique en anglais. /
A: [LAUGH] j sais pas / j commence à le perdre hein - un petit peu. /
B: ça fait combien d temps qu t es arrivé là? /
A: depuis le six juillet / quand même / non mais euh -
B: tu rêves encore en anglais quand même? / ... (210)
(102) B: ... oui je je crois / [ENTER X WHO HANDS B A MESSAGE] encore? /
A: encore un message. /
X: tu l'as lu? /
B: ah ben non /
X: eh ben lis-le /
B: ah i fallait le lire tout d suite? / ... (8300)

((99-102) are all biased requests for information.) Three other markers of positive bias in affirmative questions which all indicate explicitly an attenuation of the speaker's doubt vis-à-vis the truth of the proposition are bien, as in (103) (which is a quoted biased request for information), peut-être, as in (104) (a biased suggestion excluding the speaker) and (105) (a self-addressed biased request for a reminder), and also certainement, as in (106) (a check on the knowledge of the addressee):

- (103) ... mais j pense que les parents ont dû s faire des soucis hein / je pense - ils ont dû se dire euh "vivement qu'ils arrivent" ils ont dû aller à l'usine et dire "est-ce qu'ils sont bien arrivés?" et tout ça euh / ... (30646)
(104) ... c'est une école primaire / lorsqu'i sortent de cette école ils vont à Montélimar c'est-à-dire qu'ils rentrent euh / dans les cours élémentaires / (oui) - vous pourriez peut-être visiter cette école? / ... (21553)
(105) ... le jour où t es arrivé là j sais pas (si c'était) c jour-là que j suis parti on est partis manger avec Christine c'était la veille peut-être? / oui c'était la veille - ... (21449)
(106) A: j connais un petit peu les Vosges parce que j'ai fait des études à Nancy /
B: euh ben tu connais Gerardmer certainement? / ... (26163)

(Although the detailed study of the discourse functions of such

items is relatively recent, over fifty years ago Fromaigeat (1938:12) pointed out that the expectation of a positive response to a question could be expressed explicitly by the use of alors, donc, bien, sans doute etc.)

Another type of indicator of bias is the presence of an utterance-final tag, which usually consists of a single word, such as là or a disjunctive pronoun such as toi or ça (as in (107), (108) and (109) respectively):

- (107) B: ... alors euh l - l - l'animation si on peut dire / ici dans la campagne c'est c'est commencé là? on peut parler? /
 A: oui oui oui - ... (1348a)
- (108) A: alors étant étranger c'est pas tellement facile de trouver une organisation qui veut bien te prendre. / alors euh si tu es * (en France) * /
 B: * tu vis en France toi? * / ... (9562a)
- (109) B: c'est ton nom ça? [SEEING NAME ON BAG] /
 A: oui c'est ça / ... (25512)

Rather less usually, the tag may consist of a phrase, as in (110) and (111), although the latter example strikes some (non-Picard) native-speakers as odd:

- (110) ... mais toi c'est ta tu as déjà fait quatre colos tu m'as dit? / (9536)
- (111) X: euh - ah oui faut que je descends au Tréport pour payer le boucher aussi. /
 B: oui ben à ce moment-là on va descendre ensemble tout à l'heure. /
 X: quand? /
 B: ben il est vers onze heures là il est? /
 A: ah - * moins vingt-cinq oui * /
 B: * on descend vers onze heures * il nous payera l'apéro comme ça. / (4781)

Unfortunately it has not been possible to deal with intonation in any systematic way in this study, but there is no doubt that the intonation of these positively-biased questions often differs from that of neutral questions. Impressionistically, the final syllable in these biased questions generally seems to be said not with rising intonation but with a low tone, which is also

characteristic of utterance-final tags and right-dislocated constituents (such as those in (107-111) above). When, however, the speaker wishes to express considerable surprise at the proposition, which they are now assuming to be true, there is a very high rising final intonation, as in (112) (which is a biased check on the knowledge of the addressee):

- (112) A: [BOTH SPEAKERS ARE LOOKING AT LIST OF NAMES IN BOOK]
 alors * i y a monsieur Brown qui * -
 B: * j suis même en correspondance * / vous le connaissez? ↑ /
 A: non non non - ... (21125)

The final linguistic indicator of positive bias in these affirmative questions is the simultaneous use of a cleft structure, which results in contrastive stress being placed on the final NP of the main clause, as in (113) and (114) which are both biased requests for clarification:

- (113) B: quand tu parles comme ça d'animateur c'est en (. .)
colonie d vacances que tu parles? /
 A: oui / (29284)
 (114) ... mais tu sais tu me demandais au niveau des animateurs
 (c qui était le) / au niveau animateur sur les qualités là -
 d'animateur d colonie d vacances c'est ça qu tu voulais savoir? /
 ... (29762)

Sadock & Zwicky (1985:182) state that focused Yes/No questions "resemble biased questions to some extent, in that they display a belief by a speaker that a proposition is likely to be true". In the approach adopted here, these questions do not merely resemble biased questions - they are biased questions, due to their logical structure and their implicatures. Sadock & Zwicky appear to have the additional requirement that biased questions must have some distinctive formal property.

In addition to the above linguistic indicators, there are certain

factors in the discoursal context, which signal positive bias. Firstly, there may be a strong indication in a previous utterance that the proposition is true, as in (115) (a biased request for information): the addressee has just said that he did not in fact live in France, and so the speaker now has strong grounds for believing that he lives in England, since these were the only two countries which had been mentioned in the conversation:

- (115) B: ... tu vis en France? - à longueur d'année? /
 A: non non non - non - non * je vis *
 B: * tu vis en Angleterre? * /
 A: en Angleterre - oui oui / (9564)

In some cases it may even be the case that a previous speaker has stated that the proposition is true, as in (116), an echo biased request for information. In this instance, the relevant piece of evidence is in the immediately preceding conversational turn, but it is not difficult to imagine it being located further back in the conversation.

- (116) A: ...est-ce que tu as des fois des difficultés à
 comprendre les tout petits là? / les gamins? - parce que /
 moi j'en ai - hein * souvent * /
 B: * toi tu en as? /
 A: oui / (17309)

The following discourse context can also provide evidence that the speaker was assuming the proposition in their question to be true. In (117) the speaker does not even pause for her addressee to answer her biased request for clarification:

- (117) A: ... est-ce que les gens sont pratiquants en général dans
 la région - je veux dire la région tout ça - * ils vont à
 l'église? /
 B: * chez - où? - ici? euh au camp ou? * /
 A: non euh * je veux dire / chez toi plutôt. /
 B: ah n - moi j - non - je pense que ça se perd beaucoup la la
 religion / fin ///, oui enfin y a toujours les gens euh /
 plutôt l - tu veux^{dire} dans les jeunes? dans les jeunes y a pas
 tellement ... (16502)

In many cases, of course, two or more of the factors mentioned in

this section are present simultaneously. But in *two* instances there seems to be no overt indication of positive bias. *Both* of these, however, are affirmative interrogatives with the CF of tentative assertions - a rhetorical device in which the speaker regularly continues speaking instead of pausing for the addressee to provide the answer, and often proceeds to say that they themselves do not know the answer.

The two examples are in (118) and (120), and they implicate (119) and (121) respectively:

- (118) B: ... quand on voit les filles maintenant sont vraiment pas (des . .) hein //
- A: aha - [LAUGH]
- B: si c'est vrai j'enfin c'est l'impression j'en ai hein je crois que c'est de vraiment de pire en pire. // non c'est s - j sais pas enfin. // font euh font de petites manières euh - ou est-ce que c'est l'âge? enfin je n sais pas (moi) // (3326)
- (119) Peut-être que c'est l'âge.
- (120) B: ... et la réunion voyez-vous? j'organise une réunion. / ils viennent pas ils viennent à trois /
- A: * des parents? * /
- B: * non * - "oui faites-le." / alors est-ce que je fais trop? / peut-être que je / j'ai des amies qui - c'est partout pareil. / j'ai des amies qui sont découragées ... (13664)
- (121) Peut-être que je fais trop.

(The alors in (120) does not seem to have the same function as the cases of alors discussed above which showed that the speaker was drawing an inference from something said by a previous speaker.)

5.4.4 Affirmative questions expecting a negative answer

The number of negatively-biased affirmative questions in the corpus is far lower than that of the positively-biased ones, and

they fall into a quite small number of CF categories. However, in contrast with the former type, they include some WH questions.

The only example of a negatively-biased affirmative request for information is in (122):

- (122) B: ... mais est-ce que vous mangez vraiment de la viande en Angleterre? /
A: beaucoup moins qu'en France. /
B: oui - beaucoup moins / oui / (21482)

Of particular significance here is the presence of vraiment, and the fact that the intonation of the final seven syllables was low rather than high rising: both of these properties indicate the speaker's expectation of a negative answer. When A's answer (more or less) confirms this, the speaker echoes A and says "oui" to show that A's answer was in accordance with her expectation.

All the other instances of this type of biased question are either tentative assertions or emphatic assertions, and therefore there is naturally no answer to the question from the addressee. Additionally, there is in all cases a low, or falling, final intonation.

In one of the Yes/No affirmative tentative assertions with negative bias, (123), vraiment is also present, and in (124) uniquement seems to fulfil a similar function to vraiment, ie to express the speaker's augmented doubt regarding the truth of the proposition and hence the implication of a negative statement:

- (123) ... l'enfant quand il sort de la colonie bon on a eu beau faire / euh le maximum / bon il en a retenu quelque chose. / il en aura un souvenir bien sûr. / mais est-ce qu'on l'aura vraiment aidé pour qu'il reparte euh sur d'autres bases? euh - pt - c'est [NOISE] - c'est pas évident du tout. / ... (11389)
(124) A: et le stage franco-allemand binational que tu as fait? ...

B: ...en fait ce que j'en retire ... c'est peut-être quand même des des - plus de points négatifs que de points positifs / maintenant ah - est-ce que c'est - dû euh / est-ce que c'est dû uniquement à la présence euh d'Allemands ou / ou un mixage qui n'a pas été réussi ou / ou une différence culturelle? parce que je crois que / ça ça intervient hein ... (14217)

(In another example, given above as (44) and therefore not repeated here, the interrogative is an emphatic assertion - and so biased by definition - and vraiment is present, not in the interrogative itself, but in the "false start".)

In three other examples of affirmative tentative assertions with negative bias there are no apparent linguistic or prosodic markers of this bias:

(125) A: ... ça garantit aussi euh une bonne équipe d'animateurs chaque année plus ou moins hein /
B: ah ben c'est sûr. / ça ça évite des problèmes bon maintenant / est-ce que c'est bon aussi? j veux dire est-ce qu'on vit pas un peu sur nous-mêmes? (et) ça j'en sais rien. / ... (21223)

(126) A: et tu comptes continuer à faire des colos / pendant longtemps ou / * est-ce qu'il * va venir un moment ou tu dis / "bon j'en ai marre" ou? //
B: * [LAUGH] * // j sais pas / ... ou j peux très bien euh // de travailler dans une boîte ou (. . . .) - c'est pas très // bon / (peux) supposer que // que j'en j'en ferai encore quoi - mais euh est-ce que ça s ça sera dix ans? est-ce que / alors est-ce que ça sera toujours des maternels? euh (j'y suis pour rien quoi)/ (28416)

The underlined interrogatives here implicate the following statements respectively:

(127) Peut-être que ce n'est pas bon.

(128) Peut-être que ça ne sera pas dix ans.

(129) Peut-être que ça ne sera pas toujours des maternels.

With regard to WH interrogatives, all the affirmative negatively-biased questions are emphatic assertions (one being a quotation) and have already been discussed in 5.3.13. They do not appear to have any distinctive linguistic or contextual properties which can be attributed to their conduciveness as

such, rather than to their being rhetorical questions.

5.5 Negative questions

5.5.1 Logical structure

Negative questions are a particularly complex area, especially pragmatically, and their history in French has been the subject of considerable controversy. We shall first of all examine a little more closely their logical structure, as an understanding of this is essential to the issues to be discussed later. In 5.4.2, it was established that all negative interrogatives (ie those in which the finite verb is negated) are biased or conducive, and that there are two types, (iii) and (iv), whose abbreviated logical formulae (after Leech, 1983: 163) are repeated here, for convenience, together with English and French illustrative sentences:

- (iii) [? TRUE [neg X]] 'Didn't she buy any flowers?'
'Elle n'a pas acheté de fleurs?'
- (iv) [? TRUE [neg TRUE [pos X]]] 'Didn't she buy some flowers?'
'Elle n'a pas acheté des fleurs?'

The difference between these two types is similar to the distinction in modal logic between internal and external negation (cf Lyons, 1977:769), in that the negation is internal to the innermost propositional in (iii), but external to it in (iv). (In passing, we can note also the non-assertive/assertive contrast between de and des, parallel to that of any and some in English.)

5.5.2 History

In Classical French the type (iv) ("neg si") negative interrogative was frequently used without the negative particle

ne, as in (130):

(130) Les yeux peuvent-ils pas aisément se méprendre? (Racine,
cited by Wagner & Pinchon, 1962:547)

Vaugelas (1647:210) discussed this alternation between two

different ways of saying the same thing, as in (131) and (132):

(131) N'ont-ils pas fait?

(132) Ont-ils pas fait?

He stated that both forms were acceptable, but that he considered

the second to be the more elegant. However, in 1704 the

Académie ruled that the ne was obligatory, but in spite of this,

archaic examples of this structure without the ne are still

occasionally found in some contemporary writing (Price, 1978:599,

gives examples from Etienne and De Gaulle).

Price has argued convincingly that this structure did not involve

the "omission" of ne, as is quite widely supposed, but rather

that it originated in a structure which was not in fact negative

at all. In Old French the word point was widely used with a

positive meaning, as were the other "words of minimal quantity"

which were also used to reinforce the negative particle ne. Thus,

following Schulze (1888), Price (p.603) suggests that, in Old

French examples such as (133), the word point did not have a

negative value, but meant rather "a little bit":

(133) Congnois le tu point? (= Le connais-tu un point, si peu que
ce soit?)

(cf also the quite widespread use, in contemporary spoken French,

of un peu in questions, especially as an attenuator, as in (86)

above and (154) below). According to Price, in the mid-fifteenth century the use of a reinforcing particle such as point (and by then pas also) became usual, and consequently a negative value came to be attributed to structures such as (134):

(134) Voyez-vous point?

However, Price adds (p.606) that there was presumably a period of variation, during which (134) meant both (135) and (136):

(135) Voyez-vous peut-être?

(136) Ne voyez-vous pas?

(The non-negative origins of this interro-negative structure perhaps suggests that Hudson's (1975:9) suggestion that, in inversion exclamatives, n't is not a marker of negation is less fanciful than might be supposed at first glance.)

As already pointed out, in contemporary French the absence of ne in [V-CL] interrogatives is archaic, and it would appear from the grammars that both "neg non" and "neg si" questions can occur with any interrogative structure. (However, we shall consider later to what extent, if any, the choice of interrogative structure is constrained by communicative function.)

5.5.3 Negative questions expecting a negative answer

In the negatively-biased negative questions in the corpus, it is clear that in most cases the expectation of a negative answer derives from contextual or situational evidence. Most obviously, a turn from the interlocutor overtly expressing a negative proposition, may lead the speaker to ask the question with a strong expectation of a negative answer, as in (137) (a biased echo request for information):

- (137) B: ...is ont pas goûté les enfants? /
 X: non non mais il est pas quatre heures /
 B: il est pas quatre heures? ↑ /
 X: non (il est seulement . . .) / (29673)

(Les enfants in B's first question is, of course, not a direct object, but a right-dislocated NP coreferential with the subject pronoun.) The speaker's surprise at the apparent truth of this negative proposition was expressed by a very high rising tone on the final syllable of his echo question. Secondly, the speaker may take part of an interlocutor's preceding turn to implicate the negative proposition, as in (138) (a biased request for information, in which the presence of alors signals the inferencing process), and the two examples in (139) (the first, a biased check, the second a biased request for information):

- (138) B: ... à quelle heure est-ce que vous mangez l'assiette anglaise? / * (ou) * le jambon charcuterie euh * froid (quoi) * /
 A: * euh * / * enfin y a beau on man * on mange beaucoup moins de charcuterie chez nous ... [DIGRESSION ON ENGLISH SAUSAGES AND CHEESE]
 B: ... alors / y a pas c'est pas à onze heures du matin ou à? / * (pour le) repas hein. * /
 A: ah non * quand même - fin * / non non ... (13969)
 (139) A: ... ça veut dire quoi conseiller prud'homme? /
 B: conseiller prud'homme? - tu sais pas c que c'est que les conseils de prud'hommes? / ça existe pas en Angleterre? /
 A: je sais pas enfin je sais pas exactement ce que ça veut dire. / (4668ab)

In (140) a negative facial expression from her interlocutor leads the speaker to make her request for clarification expecting a negative response:

- (140) A: est-ce que vous pourriez me décrire un petit peu une journée moyenne - pour vous - pendant une colonie d vacances? /
 B: c'est-à-dire euh le prix de de le revient d'une journée? / non
 * c'est pas ça que vous voulez * dire? /
 A: * euh - fin * - non - ce que vous faites pendant la journée - ... (21244)

A rather more complex case is involved in (141), where A's negatively-biased negative question arose from A's drawing a

wrong inference from B's previous turn, and B then requests clarification, by echoing A's question, with a rise-fall intonation on the last two syllables. Her meaning is clearly: "What do you mean by asking 'Ils sont pas arrivés?'?".

- (141) A: ... est-ce que tu as jamais travaillé / dans une colo un camp d'ados avec des animateurs étrangers? /
 B: non /
 A: non? /
 B: non j crois pas non /
 A: (c'était pas arrivé?) /
 B: un stage oui on devait avoir des Anglais une fois. /
 A: mais * ils sont pas arrivés? * /
 B: * (mais) j'ai pas travaillé * avec euh des non /
 A: ils sont pas arrivés les Anglais - alors - pour le stage? /
 B: ils sont pas arrivés? /
 A: ben - tu disais "on devait avoir des Anglais" // * pour le stage * /
 B: non je * je dis "on * devait" parce que non - "on devait" ça veut dire je me souviens plus s'i y avait des Anglais (mais) / ch - s - on é - y avait des Anglais mais je sais plus si c'était dans mon stage ou un stage qu'y avait à côté ... (16622)

Several other examples of negatively-biased negative questions in the corpus occurring as the second part of alternative questions have been given above in extracts (40), (41) (both biased tentative assertions) and (94) (which included two instances with indeterminate CF). In (142) the CF of the underlined interrogative is a quoted self-addressed negative request for advice expecting a negative answer:

- (142) ... bon ben j'ai vu son père tout à l'heure qui me l'a ramené sur la plage. / c'est peut-être un tort j'en sais rien. / mais après je me dis "je le fais? je l fais pas?" / (j'avais) envie de le dire à son père. / comment son fils se comportait. / ... (11639b)

(Of course taking the alternative question as a whole, the author was not expecting a negative answer.)

Borillo (1979) has provided a detailed catalogue of items which,

in negative questions, force on the one hand a "neg non" interpretation, and on the other a "neg si" interpretation. Following Quirk et al (1985:782-5) on corresponding English items, we shall use the term "non-assertive" for those items which may occur in "neg non" questions, and "assertive" for those that can occur in "neg si" questions. Given the limited number of negative interrogatives in the corpus, it is unsurprising that only a few of the items listed by Borillo actually occur. Of the non-assertive items, however, there are examples in the following extracts of de (as the plural indefinite article), the adverbs toujours and trop, and the negative intensifying expression du tout:

- (143) B: ... et vous n'avez toujours pas d - n'avez pas de colos qui marchent chez vous? /
 A: non non / (3694)
 (144) X: William /
 B: oui /
 X: (. .) - Gerard i m'a donné une claque /
 B: mm / alors ça / où il est Gerard? / on verra ça tout à l'heure. là t'as pas trop mal? / non? / (28560)
 (145) B: ... vous n'êtes pas aidés du tout par le gouvernement (vous)? / ... (211044)

The two negative questions in extract (146) provide further examples of de and du tout:

- (146) B: mais comment ça marche exactement chez vous les colos? /
 A: les colos? - * y en a * pas - justement hein . /
 B: * mm * / y a y a rien du tout comme to colos? ↑ y a y a pas d centres d vacances organisés pour les (gosses)? ↑ /
 A: fin quelques-uns peut-être ... (2618 9)

The speaker's astonishment at A's statement was conveyed not only by his quasi-repetition of the question, but also by very high rising final intonation for both questions. It is interesting to note that the two native-speakers who gave acceptability/equivalence judgements on the non-occurring interrogative structures in negative questions seemed uncertain

about the interpretation of these two questions - ie as "neg non" or "neg si" - despite the presence of the non-assertive items du tout and de. The explanation for this is possibly that the marked intonation suggests that the speaker's doubt regarding the truth of the negative proposition is so great that it approaches a belief in the corresponding positive proposition. (Additionally, Borillo, 1979:30, suggests that de can in fact occasionally be found in "neg si" questions.)

The first question in (146) also includes rien, which, although generally non-assertive, is not unambiguously so (as Borillo, 1979:33, points out), since, like aucun and personne, it can lose its negative value in interrogative clauses. Thus (147), a confected example, is ambiguous between a "neg non" and a "neg si" interpretation:

(147) Tu ne vois rien?

5.5.4 Negative questions expecting a positive answer

There are just six occurrences in the data of positively-biased negative questions, and therefore our comments on these will necessarily be somewhat limited. Of the three which are biased requests for information (given in (148) and (149)), two are the quasi-formulaic c'est pas ça?, which can occur as a tag following a declarative, in the same way as n'est-ce pas - but it seems not to occur in as wide a range of contexts, or as frequently, as hein or n'est-ce pas itself:

- (148) A: ... c'était une soirée de poésie bretonne je crois /
B: ah oui y avait un genre de fête dans la ville c'est pas ça? /
A: je me rappelle plus exactement ... (6543)

- (149) A: chez nous y a pas de service militaire. /
 B: non? / * (. .) * /
 A: oui / * (j sais pas) * depuis pas mal de temps d'ailleurs
 depuis une bonne vingtaine d'années /
 B: mm /
 A: mais ce gouvernement-là justement j crois qu'ils seraient
 plutôt pour le réintroduire . / mais ils auraient du mal hein /
 B: mais l'Angleterre ça s tend euh c'est pas un système euh / de
deux grands partis quoi? c'est pas ça? * c'est (. .) * /
 A: * oui - du point d vue politique * tu veux dire? (23966ab)

In the first of B's two interrogatives in (149), we can note the use of the assertive item un (as is expected in a "neg si" question) rather than the non-assertive de. Indeed in this context, de could not occur even if the "neg non" meaning were intended:

- (150) *Ce n'est pas de système de deux grands partis?

This is probably because of the coreferentiality of the subject pronoun ce and the complement NP. If, however, c'est pas is replaced by il n'y a pas, then both de and un can occur, the former suggesting a "neg non" interpretation, and the latter a "neg si" reading:

- (151) Il n'y a pas de système de deux grands partis?
 (152) Il n'y a pas un système de deux grands partis?

The negative interrogative in (153) is a quoted suggestion (inclusive of the speaker) and is a good example of the use of negative questions as a negative politeness strategy:

- (153) ... y a toujours des exceptions t as quelques gamins qui sont bien mûrs et qui eux viennent discuter avec toi / "on pourrait pas faire ceci?" "moi (j'ai) j voudrais bien aller à l / faire une rando comme ça" euh ou bien - tu vois? - ça arrive / ... (2397)

(The speaker, a camp director, was quoting teenagers, and since the relationship between the staff and the teenagers is more solidary and less hierarchical than in a children's camp, it seems likely that the author of the question was including the

addressee in the reference of on, and therefore also intended the question as a suggestion rather than a request for permission.)

The final example of a Yes/No "neg si" question, in (154), has the CF of a positively-biased negative tentative assertion, implicating (155), and includes the assertive item un peu:

- (154) B: ... mais c'est toujours comme ça c'est-à-dire que / c'est l copain d quelqu'un qu est déjà venu ici // et en fait l'autre se porte un peu garant d la personne qu'il a introduit au groupe tu vois? // ... ça ça évite des problèmes. bon maintenant / est-ce que c'est bon aussi? j veux dire est-ce qu'on vit pas un peu sur nous-mêmes? (et) ça j'en sais rien. / (21224)
(155) Peut-être qu'on vit un peu sur nous-mêmes.

The only occurrence in the corpus of a negative WH interrogative is that in (156) (already given above as (46), but repeated here for convenience). Unlike the tentative assertion in (154), the speaker here does not provide any answer to the question, clearly assuming that the truth of the implicated affirmative statement (157) is self-evident: hence its CF is a positively-biased negative emphatic assertion.

- (156) ... et on a fait une erreur. / bon - euh - qui ne fait pas d'erreurs? [LAUGH] hein / mais c'est (bien de faire l'erreur parce qu'on s'aperçoit .) // ... (22306)
(157) Tout le monde fait des erreurs.

Looked at from another angle, it could be said that the speaker is implying that (158), which is a presupposition of (156), is absurd:

- (158) Quelqu'un ne fait pas d'erreurs.

In the question in (156), the non-assertive item de is present, but native-speakers find des equally acceptable in the same context, and with the same CF:

- (159) Qui ne fait pas des erreurs?

As we have seen, in Yes/No "neg si" interrogatives an assertive form (eg des) is usual, and a non-assertive form (eg de) is generally unacceptable. This may be due to the fact that, in the logical structure of Yes/No "neg si" questions, the negative operator is located outside the innermost propositional, where the assertive form is situated - [? TRUE pos [TRUE neg [pos X]]]. For Yes/No questions, therefore, we could propose the following rule:

(160) If, in the logical structure, the negative operator is situated within the same propositional as an item having both assertive and non-assertive forms (eg the French indefinite article), then the non-assertive form (eg de) must be used; if not, then the assertive form (eg des) is used.

However, the logical structure of a WH question of this type, such as that in (156), is [neg ('person x' fait des erreurs)], (after Leech, 1983:116), where the negative operator is located within the same propositional as the assertive item des. ('person x' here is the unknown variable realised as the WH word qui, and the parentheses enclose the predication). Here the negative operator is indeed located in the same propositional as the indefinite article, and yet the assertive form des is just as acceptable as the non-assertive form. We must conclude either that rule (160) does not apply to WH questions, or that the distribution of des and de is not as strictly complementary as is that of other assertive/non-assertive pairs.

Although Quirk et al (1985:820) give several examples of perfectly normal negative WH interrogatives being used as requests for information, it does seem that such forms are generally rather restricted. Not only is there just one

occurrence in the present corpus, but Givón (1978:95) suggests that a large number of negative WH interrogatives, such as (161-5) are more or less unacceptable:

- (161) ?Where didn't you leave the keys?
- (162) ?When did John not arrive?
- (163) ?How fast did John not run?
- (164) ?How did he not do it?
- (165) ?With what didn't he cut the meat?

To interpret such questions as requests for information, it would indeed seem necessary to construct a rather elaborate pragmatic context. Thus, (162) would be acceptable (with contrastive stress on not) if John had been expected to arrive on not just one but several occasions. But in addition, it seems that all except (163) would be acceptable as rhetorical questions - again, provided that there was contrastive stress on the negative element (cf Liberman & Sag, 1975, on the intonational differences between requests for information and rhetorical questions in American English). Most of (161-5) would also, of course, be acceptable as echoes.

Quirk et al (1985:821) also point out that questions beginning with Why don't you or Why not are often used as invitations or suggestions. Similar forms exist in French - (166) was noted (but not recorded) during the fieldwork:

- (166) Pourquoi les animateurs ne feraient-ils pas le jeu en même temps?

Since pourquoi does not have a simple negative equivalent, the corresponding declarative implicated by (166) requires an additional clause:

- (167) Il n'y a pas de raison pour laquelle les animateurs ne feraient pas le jeu en même temps. (= Les animateurs pourraient faire le jeu en même temps.)

As with (156), we could say that the speaker was implicating that the presupposition of this interrogative, (168), is laughable, if not absurd:

(168) Les animateurs ne feraient pas le jeu en même temps.

And since the addressees were some of the animateurs being referred to, the CF of (166) seems to have been a positively-biased negative suggestion.

Several writers have noted that positively-biased negative interrogatives are often used for politeness (eg Gaatone, 1971:211-2). According to Brown & Levinson (1987:122-3), this is because they indicate that the speaker knows the hearer's wants and attitudes and thus partially redress the threat to the hearer's positive face involved in acts such as offers.

We shall return to negative interrogatives at various points in chapter 6, to investigate, in particular, whether the variant structures are constrained by the pragmatic factors considered here.

5.6 Conclusion

In this chapter we have firstly sketched out a taxonomy of communicative functions, distinguished from each other by a set of pragmatic and semantic features, and then proceeded to illustrate these functions in some detail with interrogative utterances drawn from the corpus. It proved possible to account reasonably satisfactorily for the vast majority of the interrogatives in this framework, but a small number of

utterances remained at the end, whose communicative function could not be determined, principally because of the lack of context with quoted utterances. Finally it was seen how the pragmatic parameter of bias (of various types) can act as a kind of modifier of a wide range of communicative functions. This analysis of the pragmatic meaning of the interrogatives in the corpus enables us, in the next chapter, to consider the constraints that this places on the speaker's choice of structure.

CHAPTER 6: CONSTRAINTS ON INTERROGATIVES

6.1 Introduction

In this chapter we examine the various factors which influence a French speaker's choice of structure in direct interrogative clauses. In accordance with the standard methodology of the variationist paradigm, we first isolate the categorical tokens, where, for pragmatic or linguistic reasons, the speaker has no choice in the structure to be used. We then move on to the analysis of the variable tokens, looking in turn at the linguistic and pragmatic constraints on the variation. Finally we consider whether there is any evidence of social and stylistic differentiation in the use of (YNQ) and (WHQ).

6.2 Coding

In order to facilitate the identification of constraints, the codings for all tokens were entered into two computer files, one for (YNQ) tokens and the other for (WHQ). A sample line from the (YNQ) file appears in (1):

(1) 210 [inf aff oui]a SV 1 ESV - VCL - tu reves !

The reference number at the start of the line serves to locate the token within the corpus. Next, within square brackets, and in abbreviated form, is the communicative function of the utterance - in this case it is an affirmative request for information expecting a positive answer. Thirdly the lower-case a which immediately follows the right-hand bracket indicates that the question was addressed to "A", the fieldworker. Other letters

in this position are used to show that a token was addressed to a third person, or from another informant to the interviewee, or from another informant to the fieldworker. Next follows the coding for the syntactic structures: SV 1 means that [SV] was the structure actually used, ESV - and VCL - mean that these two structures either would be unacceptable, or would be acceptable but would produce utterances non-equivalent to the utterance actually produced. The subject of the interrogative clause, tu, is given next, followed by the verb form, rêves, and finally the exclamation mark indicates that this token is categorical, since the speaker had no choice regarding the structure to be used.

(2) is a sample line from the file for (WHQ) tokens:

(2) 2179 [inf]x QSV 0 QESV 0 SVQ 1 QVCL 0 QVNP - Q=SV - tu cbn

The pattern here is similar to that in (1), except that the WH word is given as the final element, and there are six variant structures rather than three. Zero after the variant formula indicates that the structure would be acceptable and would produce an utterance equivalent to the one actually produced. The CF in this instance is a request for information, and the x following the right-hand bracket shows that the utterance was addressed to a third person. (The files of the codings for all (YNQ) and (WHQ) tokens in the corpus are given in Appendices 2 and 3, respectively: for reasons of space, however, the initial reference number of each token is not given there.)

Since each token was coded onto a separate line, a global command (within the Context Editor, or similar), such as (3), (4), (5) or (6), then makes it possible to view all lines on which certain

strings occur:

- (3) scan@all /file ; ESV 1 ;
- (4) s@a /f ; cbn ;
- (5) s@a /f ; aff oui];
- (6) s@a /f ;]a ;

(In (4), (5) and (6) the elements of the command appear in their conventional abbreviated form.) The output from this operation enables us to visually inspect all the lines where the specified sequences occur: ie with (3), all the tokens of [ESV]; with (4), all the (WHQ) tokens where combien is the WH word; with (5), all the tokens with a positively-biased affirmative CF; and with (6), all the tokens which involved questions addressed to the fieldworker. In this way sub-sections of the data can be examined to check for the effect on the variation of different factors. Much of the discussion in this chapter will be based on such visual inspection of a large number of scans of these two files.

6.3 Categorical tokens

6.3.1 The exclusion of categorical tokens in variationist methodology

The identification of categorical tokens, and their subsequent exclusion from the quantitative analysis, is a crucial stage in the variationist methodology, and constitutes an important difference between the latter and earlier quantitative approaches. We shall see later the extent to which the exclusion of categorical tokens alters the relative frequencies of the variants of (YNQ) and (WHQ).

6.3.2 Categorical tokens of (YNQ)

There are several categories of (YNQ) tokens which are categorical on account of pragmatic factors. Perhaps the least surprising of these is echo questions, where, by definition, the speaker is obliged to re-use the structure produced by another speaker, as in (7):

- (7) X: Bertrand est-ce que j peux avoir des agrafes et d la colle? /
B: est-ce que tu peux avoir des agrafes et d la colle? /
(2174)

Another case where the structure is constrained by a previous utterance, though in this instance one by the same speaker, is that of alternative questions (which we treat, on syntactic grounds, as tokens of Yes/No interrogatives). The second interrogative of the pair (which, in our data and perhaps universally, is negative) has, necessarily, the same structure as the first. Thus, in (8) the first two interrogatives are both [SV], and the second two are both [ESV]:

- (8) ... après je me dis "je le fais? je l fais pas?" / ... alors est-ce que j'ai bien fait? est-ce que j'ai pas bien fait? j'en sais rien. / ... (11639642)

What we do not find is cases such as (9), where the two clauses in an alternative question have different structures:

- (9) #Est-ce que je le fais? Je ne le fais pas?

(Pragmatic or discoursal inappropriateness is indicated by the symbol #.)

Thirdly there are certain CFs which constrain categorically the variant used, in some, if not all, cases. Requests for action

which take the form tu/vous + non-modal verb (in present tense)

are categorically [SV], as is shown by (10-12):

(10) ... y a une petite fille qui me dit "... tu me mets de de la couleur sur les noëils? " ... (17317)

(11) #Est-ce que tu me mets de la couleur sur les noëils?

(12) #Me mets-tu de la couleur sur les noëils?

Although (11) and (12) are quite acceptable syntactically, they cannot be used as requests for action. (10) shows that this constraint holds also when the request for action is quoted rather than "live". In contrast with (10), if a modal verb is present, as in (13), or if the verb is in the future tense, as in (16), then all three variants of (YNQ) are possible, as is shown by (14-15) and (17-18), respectively:

(13) ... bon écoute tu peux partir deux minutes? ... (4908)

(14) Est-ce que tu peux partir deux minutes?

(15) Peux-tu partir deux minutes?

(16) B: ...tu le laisses sur le bureau et * puis * /

X: * tu lui * diras? - ... (X26482)

(17) Est-ce que tu lui diras?

(18) Lui diras-tu?

(According to St-Pierre, 1977:189-99, in Montreal the conditional form of the verb is also commonly used in this way.) The constraint does not hold either in requests for action when a variant of the formula je + peux + avoir + NP is used:

(19) ... est-ce que j peux avoir des agrafes et d la colle? / ... (242173)

(20) maman - j peux avoir du fil? / (X30399)

(21) Puis-je avoir le cahier?

Moreover the constraint seems to be lexically-sensitive, since not all verbs are equally possible in the present tense with the [SV] structure, as requests for action - partir, for example, seems far less acceptable than the nearly synonymous s'éloigner:

(22) ?*Tu pars deux minutes (s'il te plaît)?

(23) Tu t'éloignes deux minutes (s'il te plaît)?

However the precise nature and extent of this constraint is not at present clear.

A second, and related, CF which imposes a categorical constraint on some tokens, is the suggestion. Our data here are very limited, but it seems that both inclusive and exclusive suggestions (ie those which include, and exclude, respectively, the speaker as agent of the proposed action) oblige the speaker to use the [SV] structure, if there is no modal verb:

- (24) X: on fait comme ça? /
 A: oui oui /
 B: mm - d'accord / (X286)
- (25) #Est-ce qu'on fait comme ça?
- (26) #Fait-on comme ça?
- (27) ... vous passez par ici? / [GESTURING FOR ADDRESSEE TO GO THROUGH ONE DOOR RATHER THAN ANOTHER] (211052)
- (28) #Est-ce que vous passez par ici?
- (29) #Passez-vous par ici?

However if a modal verb is present, the constraint does not apply, and any of the three (YNQ) variants can be used:

- (30) ... i diront jamais / "tiens euh demain est-ce qu'on peut faire ça?" / ... (18733)
- (31) On peut faire ça?
- (32) Peut-on faire ça?

A third CF which categorically constrains the choice of structure is what we have termed the post-announcement, and again the constraint applies whether or not the utterance is quoted:

- (33) ... i te donnent pas ton permis parce que t as pas donné d'argent à la boîte. // t vois c qu j veux dire? / alors c'est pour ça // ... (25729)
- (34) #Est-ce que tu vois ce que je veux dire?
- (35) #Vois-tu ce que je veux dire?
- (36) ... i sont arrivés à êt - à pouvoir parler d'égal à égal aux autres si tu veux. / "toi t as peut-être une grosse tête mais moi t as vu c que j'ai fait avec mes bras? ... (265)
- (37) #Est-ce que tu as vu ce que j'ai fait avec mes bras?
- (38) #As-tu vu ce que j'ai fait avec mes bras?

There is one (YNQ) token in the corpus which is categorically [SV] because it involves what is a fixed locution, when used as a

greeting at least:

- (39) X: ça va? (t as) fait la sieste? /
A: oui j'ai fait la sieste ... (162413)

Finally there are two types of biased questions which, with some exceptions, categorically require the use of [SV]:
positively-biased affirmative and negatively-biased negative questions. Numerically these are the most significant sources of categorical Yes/No interrogatives.

Almost all the interrogatives with positively-biased affirmative CFs (ie which, in terms of the features outlined in chapter 5 are [+G] and [-H]) are categorical tokens of [SV]. Among these, there is a large number of biased requests for information, such as in (40) - where the option of using [ESV] or [V-CL] is not open, hence the inappropriateness of (41) and (42):

- (40) A: ... tous les règlements de la loi française tout * ça ça m'a intimidé. * /
B: * ah c'é / c'était en France donc? /
A: ah oui - oui oui. / (1426)
(41) #Ah, est-ce que c'était en France, donc?
(42) #Ah, était-ce en France, donc?

In addition to biased requests for information, smaller numbers of similar biased CFs oblige the speaker to use [SV]: check, request for clarification, exclusive suggestion and self-addressed request for reminder of content, as in (43-46) respectively:

- (43) A: ... et puis j'avais écrit aussi à la Jeunesse et Sports /
B: mm /
A: de Nancy pour avoir * des adresses tout ça * /
B: * mais tu connais bien Nancy alors (comme ville?) * /
A: * oui * / ... (16736)
(44) A: ... est-ce que les gens sont pratiquants en général dans la région? ...
B: ... tu veux dire dans les jeunes? dans les jeunes y a pas tellement ... (16502)

- (45) B: ... vous pourriez peut-être visiter cette école? / non? /
 A: peut-être./ (21553)
 (46) ... le jour où t es arrivé là j sais pas (si c'était) c
 jour-là que j suis parti on est partis manger avec Christine
c'était la veille peut-être? / oui c'était la veille - ...
 (21449)

In none of these was it possible for the speaker to use [ESV] or [V-CL], hence the inappropriateness of (47-54):

- (47) #Mais est-ce que tu connais bien Nancy alors comme ville?
 (48) #Mais connais-tu bien Nancy alors comme ville?
 (49) #Est-ce que tu veux dire dans les jeunes?
 (50) #Veux-tu dire dans les jeunes?
 (51) #Est-ce que vous pourriez peut-être visiter cette école?
 (52) #Pourriez-vous peut-être visiter cette école?
 (53) #Est-ce que c'était la veille peut-être?
 (54) #Etait-ce la veille peut-être?

(There are also several positively-biased affirmative echo questions, but we assume that the categorical nature of these is due principally to the fact that they are echoes.)

Although the overwhelming majority of interrogatives with affirmative positively-biased CFs are categorically [SV], there are three exceptions to this in the data. Two of these are biased tentative assertions with the [ESV] structure, as in (55), but which could equally have been expressed with [V-CL], as in (56):

- (55) ... c'est de vraiment de pire en pire. // non c'est s - j
 sais pas enfin. // font euh font de petites manières euh - ou
est-ce que c'est l'âge? enfin je n sais pas (moi) // ... (3326)
 (56) Ou est-ce l'âge?

The third exceptional case is a quoted positively-biased affirmative request for information (the only one in the data), in which the speaker used [ESV], but could have used [V-CL]:

- (57) ... ils ont dû aller à l'usine et dire "est-ce qu'ils sont bien arrivés?" et tout ça euh / ... (30646)
 (58) ... et dire "Sont-ils bien arrivés?"

In this instance the fact that the question was quoted rather

than live seems to rule out the possibility of using [SV].

Turning now to negatively-biased negative questions, the situation is rather less straightforward, due to a larger number of exceptions. As with the majority of positively-biased affirmative questions, the categorical tokens of this type are of the [SV] structure. They include requests for information, checks and requests for clarification, as in (59), (60) and (61):

- (59) B: ... [TO X] Mariline tu les as fait goûter? ///
A: les animatrices fument /
B: hein Mariline - is ont pas goûté les enfants? /
X: hein? / (29669)
- (60) A: ... elle habite où? /
B: à (. .) /
A: connais pas - c'est quel coin? /
B: j sais pas j'y suis jamais allée. / mais / elle va venir euh j sais pas à Noël. /
A: ah oui? /
B: mm /
X: tu la connais pas? /
B: non j / c'est-à-dire ... (324208)
- (61) A: est-ce que vous pourriez me décrire un petit peu une journée moyenne - pour vous - pendant une colonie d vacances? /
B: c'est-à-dire euh le prix de de le revient d'une journée? /
non * c'est pas ça que vous voulez * dire? /
A: * euh - fin * - non - ce que vous faites pendant la journée - ... (21244)

In each of the underlined utterances in (59-61) it is clear that the speaker believed (or at least suspected) that the proposition in their utterance was not true, and (62-67) would be inappropriate in these contexts, in the sense that they would produce utterances non-equivalent to the originals:

- (62) #Est-ce qu'ils n'ont pas goûté, les enfants?
(63) #N'ont-ils pas goûté, les enfants?
(64) #Est-ce que tu ne la connais pas?
(65) #Ne la connais-tu pas?
(66) #Non est-ce que ce n'est pas ça que vous voulez dire?
(67) #Non n'est-ce pas ça que vous voulez dire?

However there are some eight examples of negatively-biased negative questions which have not been coded as categorical tokens. In all cases this is because one or both of the native-speaker judges (who made the judgements of acceptability and equivalence on the negative questions) felt that one or other of the alternative structures would in fact have given an utterance equivalent to the one actually produced. This is almost certainly because their interpretation of the communicative function of the utterance differed from my own: specifically, they seemed to interpret some of these utterances as being positively-biased, rather than negatively-biased. We have, then, a mismatch for these tokens between their negatively-biased CF (as determined by myself, a participant in the conversation, and frequently the person to whom the question was addressed) and the acceptability/equivalence judgements of the native-speakers, when these were more "liberal" than my own judgements. This mismatch, although unsatisfying, seems unavoidable, given that the analyst is not a native-speaker, and that the native-speaker judges were not participants in the conversations.

Perhaps of relevance here is Borillo's (1979:27) conclusion, arrived at on the basis of rather different data from ours, that it is not entirely clear whether [ESV] can be used in negatively-biased negative questions.

Before moving on to look at categorical tokens of WH interrogatives, it is worth noting that all the constraints mentioned in this section have been pragmatic rather than linguistic (unless the alternative question constraint is to be

considered linguistic). The number of categorical tokens of (YNQ) in the various categories is shown in table 6.1.

Table 6.1 Categorical tokens of (YNQ)

Categorically the same variant as a preceding utterance:	
- echo question	7
- second interrogative in alternative question	5
Categorically [SV]:	
- request for action of the form: tu/vous + non-modal verb (in present tense)	3
- suggestion (neutral) of the form tu/vous/on + non-modal verb (in present tense)	1
- post-announcement	4
- formulaic greeting	1
- positively-biased affirmative question (excluding quotations and tentative assertions)	37
- negatively-biased negative questions	13
Total	71

6.3.3 Categorical tokens of (WHQ)

The number of categorical tokens of (WHQ) in the corpus is far smaller, and they are of just four types. Firstly, and as with (YNQ), the echo questions, such as in (68), normally allow the speaker no choice of structure:

- (68) A: ... un bon animateur - pour toi qu'est-ce que c'est? /
 B: un bon animateur qu'est-ce que c'est? /
 A: oui / (22387)

However there is one WH echo question in the data in which, exceptionally, the speaker did not use the same structure as the interlocutor:

- (69) B: ... mais enfin je suis plus proche des plus vieux que des plus jeunes. /
 A: mais ça pose quelle sorte de problèmes? /
 B: quel genre de problème ça pose? des f - des euh âges? //
 oh f //// ça dépend ... (9109)

The explanation for this unusual change of structure may lie in

the fact that the original utterance, produced by the non-native fieldworker, was an inappropriate use of the structure [SVQ]. Moreover the [SVQ] structure would have produced a very odd-sounding echo, and so the speaker was obliged to use a different variant. (He also changed sorte to genre, since, in collocation with problèmes, it is more normal to use the latter.) The variant he chose was [QSV], but he could equally have chosen [QESV] (or even [QV-CL], although this would have entailed "complex inversion" which, as we have seen in chapter 4, is extremely rare in speech):

- (70) Quel genre de problème est-ce que ça pose?
 (71) Quel genre de problème cela pose-t-il?

A second pragmatic/discoursal factor which produced two categorical tokens of (WHQ), was the quoting of words from a children's singing game:

- (72) B: ... le Loup dans la Forêt aussi "promenons-nous dans les bois pendant que le loup n'y est pas si le loup y était il nous mangerait - mais comme il n'y est pas il nous mangera pas. / * où es-tu? - que fais-tu? * / ...
 A: * oui oui oui * ... (15288)

Clearly here the speaker had no choice but to use the structures which occur in the words of the game. (Naturally this would also be the case with (YNQ) tokens, and whenever the source of the quotation is some set form of words - and particularly poetry, songs and ceremonial speech events.)

In the previous section we found that all the categorical (YNQ) tokens owed their categorical nature to pragmatic constraints. Amongst the categorical (WHQ) tokens, on the other hand, there are two linguistic constraints operating. Firstly there are the interrogatives in which the WH word is que and where it is also the logical subject (the grammatical subject being qui):

- (73) ... qu'est-ce qui s passe? ... (28507)
 (74) ... "oui qu'est-ce qui va arriver?" ... (28737)
 (75) ... "mon dieu / qu'est-ce qui s'est passé?" ... (21870)
 (76) ... "qu'est-ce qui va s passer?" ... (2235)

(It will be recalled that examples such as these are classed as [QESV].) The [Q=S V] structure is not acceptable here since the WH word is que/quoi, and neither que nor quoi can function as a grammatical subject:

- (77) *Que/Quoi se passe?

(cf however Grevisse, 1986:1108, for some archaic and marginal examples similar to (77).) Nor are any of the other structures possible, as they all require the WH element and the logical subject to be realised as distinct constituents.

It is assumed here that these utterances do not in fact involve the corresponding impersonal verbs, although that would indeed be a plausible interpretation, given that il is frequently pronounced as [i] (cf Grevisse, 1986:1084). If that were the case, they would still be instances of [QESV], with il as the grammatical subject, as in (78), but they would not then be categorical, as is shown by (79-80):

- (78) Qu'est-ce qu'il se passe? [ESV]
 (79) Il se passe quoi? [SVQ]
 (80) Que se passe-t-il? [QV-CL]

(It is, however, a moot point whether (79) is pragmatically equivalent to (78) and (80).) Qu'est-ce qui can of course be used with many other verbs, including those where there is no possibility of an impersonal equivalent, for example:

- (81) Qu'est-ce qui te fait peur?
 (82) *Qu'est-ce qu'il te fait peur?

The second linguistic constraint which results in categorical

tokens of (WHQ) is illustrated in (83):

- (83) A: moi la dernière colo qu j'ai visitée c'était une colo
maternelle / ... le personnel d service occupait tout l
troisième étage du château. /
B: ah oui alors là vraiment c'était une colonie pour quoi
alors * ça? * /
A: * c'était * une colonie * de * -
B: * pour * une entreprise alors? /
A: non ... (1947)

This interrogative has been classified as [SVQ], and it is clear
that no other (WHQ) variant would be acceptable:

- (84) *Pour quoi c'était une colonie, alors, ça? [QSV]
(85) *Pour quoi est-ce que c'était une colonie, alors, ça? [QESV]
(86) *Pour quoi était-ce une colonie, alors, ça? [QV-CL]

One explanation for this may be the need to avoid a "homonymic
clash" with pourquoi: if, in (84) for example, pour quoi (= "for
what (organisation)") were replaced by pourquoi (= "why/for what
reason"), then this would produce (87), which is acceptable, but
obviously has a quite different meaning from (83-6):

- (87) Pourquoi c'était une colonie, alors, ça?

This explanation is ultimately, then, based on communicative
need: if pour quoi did occur in the fronted position, then it
would normally be interpreted by the hearer as pourquoi, firstly
because the sets of sentences in which pourquoi and pour quoi
occur probably overlap considerably, and secondly because
pourquoi is (we assume) far more frequent in discourse than is
pour quoi.

Moreover, it seems that the constraint against the occurrence of
pour quoi in the fronted position of tensed clauses is an
absolute one. (Its occurrence before infinitives as in pour quoi
faire, is, of course a quite different structure, in which pour
and quoi do not form a prepositional phrase.)

This explanation receives some support from the fact that prepositions other than pour, can occur quite freely with quoi in the clause-initial position, presumably because they have no homophones with which they could be confused:

- (88) ... par quoi ça passe? ... (2343)
- (89) ... sur quoi ça repose? ... (21006)
- (90) ... en quoi consiste votre enquête? ... (211042)
- (91) Avec quoi c'était comme organisme?

We might reasonably expect this principle to hold for the [SVQ] structure also, and we would then suppose that pourquoi would be unacceptable in the non-fronted position. However, this is not the case:

- (92) Il fait ça pourquoi? (Byrne & Churchill, 1986:454)

(But, as we shall see later, there is a strong tendency for pourquoi to occur in the fronted position.)

Furthermore it seems that "homonymic clash" does not provide the whole explanation for the unacceptability of (84-86), since, if we consider the same sentence with different prepositions, we see that they fare no better than pour in the fronted position:

- (93) *Avec quoi c'était une colonie, alors, ça?
- (94) *De qui c'était une colonie, alors, ça?

We shall return to the problems posed by these examples when we look at the variable pragmatic constraints on [SVQ]. In the meantime, let us note that the principal difference between (91) and (93) seems to lie in the presence in (93) of an indefinite NP as the complement of était.

Two further contexts where the speaker has no choice as to which

WH structure to use may be mentioned briefly here, even though there are no examples of them in the present corpus. Firstly if the WH word is the subject, and it involves combien, quel or lequel (or their morphological variants), then the structure is obligatorily [Q=S V]:

(95) Combien de pommes/Quelles pommes/Lesquelles sont tombées?

However, according to several grammarians, a hypercorrected structure, [Q=S V-CL], is found occasionally, especially when the WH phrase is "heavy":

(96) Combien d'autres ennemis du pouvoir ne font-ils pas partie des commissions et comités de lecture [...]? (Ionesco, dans le Figaro, 3 août 1974, cited by Grevisse, 1986:644)

Secondly if the subject is a NP, the verb is the copula and the WH element involves qui, quel or lequel, then [QV NP] must be used:

(97) Quelle est cette histoire d'Hélène? (Giraudoux, cited by Chevalier et al, 1964:92)

Finally in this section let us recap on the types of categorical (WHQ) tokens we have discussed by showing, in table 6.2, the number of tokens in the corpus of each type.

Table 6.2 Categorical tokens of (WHQ)

Categorically the same variant as a previous utterance:

- echo questions	4
- words (of a singing game)	2

Categorically:

- [QESV], where Q= <u>que</u> and S=Q	4
- [SVQ], where Q= <u>pour quoi</u>	1

Total 11

6.4 Quantification of non-categorical tokens

We are now in a position to exclude these categorical tokens from the data, and to focus on the variability which remains. Tables 6.3 and 6.4 show the relative frequencies of non-categorical tokens of the variants of (YNQ) and (WHQ) respectively, and, for comparison, the frequencies for all tokens (non-categorical plus categorical), which we have previously seen in Tables 4.4 and 4.5.

Table 6.3 Relative frequencies of non-categorical tokens of (YNQ)

	[SV]	[ESV]	[V-CL]	N=
non-categorical tokens	70.6%	29.4%	0%	109
all tokens (categorical + non-categorical)	(79.4%)	(20.6%)	(0%)	(180)

Table 6.4 Relative frequencies of non-categorical tokens of (WHQ)

	[SVQ]	[QSV]	[QESV]	[QV-CL]	[QV NP]	[Q=S V]	N=
non-categorical tokens	16.2%	26.1%	46.0%	5.4%	2.7%	3.6%	111
all tokens (categorical + non-categorical)	(15.6%)	(23.8%)	(48.4%)	(6.6%)	(2.5%)	(3.3%)	(122)

In table 6.3 we see that the exclusion of categorical tokens makes a substantial difference in the relative frequencies of the [SV] and [ESV] variants, and since the majority of the categorical tokens were variants of [SV], the frequency of [ESV] has of course risen. If we assume that the proportions of categorical (YNQ) tokens in previous quantitative studies were

similar to those of the present study, then we may conclude that the relative frequencies of those studies would have been modified in a similar fashion, had the categorical tokens been excluded. In contrast, we see from table 6.4 that the elimination of the eleven categorical tokens of (WHQ) makes very little difference to the relative frequencies of the different variants.

6.5 Categorical linguistic constraints

We now move on to look at categorical constraints, in groups, according to which structures they render unacceptable. We shall be considering only those case where a potential utterance would meet the structural description of a variant, but where it is ruled out due to some specific linguistic factor. There would be little point, it seems, in stating certain rather obvious "constraints" which simply involve a failure to meet the structural description: for example, that, [QV NP] is unacceptable when the subject is a clitic.

6.5.1 [V-CL] and [QV-CL]

As we saw in chapter 4, there is strong evidence that subject clitic inversion in interrogatives has been declining in French, and, in informal spoken varieties, is now found only marginally. In addition the structure seems to be becoming unacceptable in an increasing number of contexts. Perhaps the most notable of these is where the subject is je and the verb is in the present tense. Verbs of the first conjugation (ie regular -ER verbs) have been increasingly unacceptable in this context since Classical French.

Thus (98-101), which are theoretical alternatives to four interrogatives occurring in the corpus, have been judged unacceptable:

- (98) *Continué-je?
- (99) *Arrêté-je?
- (100) *Qu'oublié-je?
- (101) *Comment m'appliqué-je là-dedans?

(Examples similar to these are still occasionally found in twentieth-century written French, although they are invariably labelled as archaic by grammarians.)

Indeed the number of verbs which, in the present tense, can be inverted with je is now highly restricted. Table 6.5 shows the forms most commonly listed in grammars as being still in use, together with the judgements of Grevisse (1986), Mauger (1968), Byrne & Churchill (1986) and Price (1971).

Table 6.5 Grammarians' judgements on inversion of je
with the present tense forms of ten verbs

	Grevisse	Mauger	Byrne/Churchill	Price
suis	+	+	+	+
ai	+	+	+	+
puis	+	+	+	+
dis	+	+	+	+
dois	+	+	+	0
sais	+	+	+	0
vais	+	+	+/-	0
fais	+	+	(+)	0
vois	+	+	(+)	0
veux	+	0	0	-

(key: + = currently in use, - = has gone out of use, 0 = not mentioned, (+) = less common, +/- = only auxiliary vais in use)

From table 6.5 we see that these grammarians are agreed on just four of the verbs: suis, ai, puis and dis, and that for all the others, at least one grammarian suggests that they are no longer

in everyday use. Assuming that these judgements are based on accurate and wide-ranging observation of contemporary usage, it appears that the inversion of je with verbs in the present tense is on the brink of extinction, at least in spoken French, and that the change has been progressing by a process that might be termed "lexical diffusion in reverse".

The written language, of course, being more conservative, still provides examples of inversion not only with the verbs in table 6.5, but also with lis, mens, viens, prends and even polysyllabic verbs such as entends (Grevisse, 1986:1204-6). On the other hand the most recent examples of veux-je in any of the sources consulted are (102) and (103), both from nineteenth-century literature (the latter being an inversion in a non-interrogative context):

(102) Que veux-je conclure de là? (E.Quinet)

(103) Aussi ne veux-je plus que vous ayez d'appointements. (H. de Balzac) (both examples cited by Wagner & Pinchon, 1962:173)

This fact lends support to the claim by Price (1971:266) that inversion with veux has gone completely out of use, and so (104), a theoretical alternative to (105), has been judged unacceptable in the present study: (104) *Que veux-je dire?

(105) ... qu'est-ce qu j veux dire? ... (22206)

However several potential instances of inversion with fais and vais (as an auxiliary) have been judged to be still possible today:

(106) Que fais-je? (cf 15771)

(107) Comment vais-je vous expliquer ça? (cf 21715911)

(In the data there are three potential occurrences of vais-je and

five of fais-je, but no actual occurrences of either.)

As far as the contemporary language is concerned, the "explanation" for this constraint is simply that inversion of je with a verb in the present tense now seems archaic. However, we may wish to look back in time and seek the original functional motivation which led to this linguistic change. It might be supposed that je would be the least frequent of the subject pronouns in interrogatives, since one is less likely to ask questions about oneself. In the present corpus, however, this supposition is not borne out: je is the subject in 37 interrogative clauses. But one significant difference between je and the other subject clitics (apart from ce, which occurs only with être), is that its vowel is the schwa, which, in modern northern French pronunciation, does not occur in final position. It is, of course, true that, in one case at least, a final schwa is converted into [ø]: in the direct object pronoun le, when following an imperative verb, eg mangez-le! For whatever reason (avoidance of homonymic clash with jeu?), this does not happen with je. Consequently, an inverted verb with je will necessarily end in the consonant [ʒ]. This means firstly that the communicatively important enclitic subject is conveyed by a bare consonant, which a listener might well misinterpret as being part of the previous word. Secondly, it means the creation of a virtually unpronounceable consonant cluster for regular -ER verbs, whose stems all end in a consonant (given the loss of their final schwa also) - eg [tʒ] for chante-je, [mʒ] for aime-je. A temporary solution in Classical French was to convert the final schwa of the verb into [ɛ] (spelt, unusually, as -é),

as in (98-101) above. However, this still failed to make the je more prominent, and hence easier to perceive, and moreover it created two further difficulties. Firstly, it meant that there were now two different verb forms for the first person singular in the present tense - an uneconomical, and, for French certainly, highly unusual state of affairs. Secondly, these -é forms were homophonous with the established imperfect forms, such as aimais-je? For verbs other than the first conjugation, it appears that the demise of inversion with je came about even earlier. Again, the need to avoid homonymic clashes has sometimes been invoked for forms such as cours-je, vends-je and mens-je. We might add that the pronunciation of veux-je, which seems to have died out only recently, would have involved either the unusual presence of [ø] in a closed syllable, or the conversion of [ø] into its mid-low counterpart [œ], with the consequence that the same form of the verb would have two distinct pronunciations. It seems, then, that the progressive disappearance of inversion of present tense verbs with je has indeed been motivated by communicative factors (cf Foulet, 1921:260-1).

Another categorical constraint on subject clitic inversion (albeit one which does not affect any tokens in the present corpus) is when there is a NP subject and the WH word is que:

(108) *Que ton père a-t-il pensé? (Dubois & Dubois-Charlier, 1970:223)

(But cf Grevisse, 1986:645, for some archaic examples of this.)

If, however, the NP subject slot is occupied by cela, then clitic inversion is acceptable, and occasionally actually used:

(109) Que cela veut-il dire? (heard on France Inter, 1.10.88)

This is one of the very few contexts in which the interrogative pronoun que can occur without immediately preceding a verb: generally, it can be considered a clitic (cf Kayne, 1972:113-4). (And perhaps cela can occur here only because it is phonologically and semantically "light", and serves as a reinforcer of que similar to donc and diable.) We shall consider in 6.7.3 certain variable linguistic constraints which further restrict the use of subject clitic inversion.

6.5.2 [QSV]

The clitic status of interrogative que also means that the [QSV] structure is excluded, whether the subject is a NP, as in (110), or a clitic, as in (111):

- (110) *Que ton père a pensé? (Dubois & Dubois-Charlier, 1970:223)
 (111) *Que je veux dire?

(But Grevisse, 1986:657, gives a rare example of que in [QSV], the subject being ceci.) Given the high frequency of WH interrogatives with que, this constraint weighs very heavily in the present data.

In contrast, another categorical constraint on [QSV] affects just one token, but is somewhat more difficult to specify. (112), an example of [QV NP], has (113) as its theoretical [QSV] alternative:

- (112) ... où est l'inconvénient? ... (22565)
 (113) *Où l'inconvénient est?

The unacceptability of (113) seems to be due to the imbalance between the relatively "heavy" NP, and a verb which is very "lightweight" (both phonologically and semantically). Two other

tokens with NP subjects have [QSV] equivalents which are quite acceptable - perhaps because their verbs are weightier than the copula of (113):

(114) En quoi votre enquête consiste? (cf 211042)

(115) Comment notre groupe se faisait? (cf 21173)

The acceptability of (114) suggests that the constraint does not require that the verb be longer, in syllables, than the subject.

If we now consider (116) and (117),

(116) Où il est?

(117) Où l'inconvénient se trouve?

then we see firstly that [QSV] can be perfectly acceptable with the copula (provided the subject is a clitic), and secondly that a verb which is longer than the copula (but with little more meaning) improves (113) considerably. However, it is not only the copula which is affected by this constraint: other monosyllabic verbs appear to fare no better than est with a NP subject:

(118) *Où Françoise va?

(119) *Lesquels ton frère a?

We may tentatively conclude, then, that the constraint can be formulated as follows: [QSV] is unacceptable if the subject is a NP and the verb is monosyllabic (cf Fromaigeat, 1938:42-3). If, however, a monosyllabic verb is reinforced by a discourse particle such as déjà, then the result is acceptable, as we see by comparing (120) with (118):

(120) Où Françoise va déjà?

6.5.3 [QESV]

The categorical constraint just described for [QSV] also excludes [QESV]:

(121) *Où est-ce que l'inconvénient est?

and this may suggest that the principles involved are of a more general nature and not simply variant-specific constraints. On the whole, however, it is noticeable that [QESV] is freer of linguistic constraints than the other variants of (WHQ).

6.5.4 [QV NP]

The structure involving NP inversion (or "stylistic inversion", according to Kayne & Pollock, 1978) is subject to two interesting and quite subtle categorical constraints, which we shall consider in some detail, even though they do not affect any of the WH interrogatives in the corpus.

The first constraint is that [QV NP] is unacceptable if the WH word is pourquoi:

(122) *Pourquoi vient Pierre? (Kleiber, 1986:601)

Renchon (1969:48) suggests that it is the juxtaposition of two stressed elements, the verb and pourquoi, which accounts for the unacceptability. This suggestion, however, does not seem to be supported by any empirical evidence. Pourquoi is not generally stressed any more than any other WH element, and certainly not more than lengthy WH phrases with combien, quel or lequel. As for the verb, it has often been remarked, on the contrary, that the verb in [QV NP] interrogatives tends to be lighter syntactically, and perhaps also semantically, than elsewhere, and we would thus expect it to be unstressed.

A more promising line of enquiry has been opened up by the observation that pourquoi is not directly "linked" to the verb,

but rather bears on the sentence as a whole. Blinkenberg (1928:147) pointed out that this is also the case in causal clauses (introduced by parce que, puisque etc.), where inversion is similarly disfavoured. More recently Korzen (1983, 1985) has commented that pourquoi represents an argument which is peripheral to the situation described in the sentence. We might add that the answer to a question with one of the other WH words can consist of one word, whereas the answer to a pourquoi question is typically an entire sentence expressing a new proposition. Korzen argues further (as have others) that for [QV NP] to be acceptable, there must be a high degree of "cohesion" between the WH element and the verb ("cohesion" here presumably being related to the concept of valency). Whilst this is undoubtedly correct, it seems, in itself, to be only the beginning of an explanation.

Two other facts relating to this constraint are worth noting at this point. Firstly it was still possible in the seventeenth century to use pourquoi with [QV NP]:

(123) Pourquoi n'ont pas péri ces tristes monuments? (La Fontaine)

(124) Et pourquoi commandent les hommes, si ce n'est pour faire que Dieu soit obéi? (Bossuet; both examples cited by Renchon, 1969:49)

Secondly (a fact perhaps related to the first), there is by no means universal agreement that the constraint is in fact categorical. Foulet (1921:294-5), for example, felt that it was not impossible to say:

(125) Pourquoi crie cet enfant?

And Kayne (1972:113) comments merely that stylistic inversion with pourquoi is "dubious", both in direct interrogatives such

as:

(126) ?Pourquoi part ce garçon? (p.70)

and in indirect interrogatives, where the constraint also operates:

(127) ??Il tient à savoir pourquoi est partie la fille. (p.74)

Moreover, the structure is even actually used on occasion:

(128) Il ne saura jamais pourquoi est morte sa femme. (read by a reporter from a script, France Inter, 1988)

(129) Pourquoi donc ne devraient pas en faire autant les vénérables Pères du futur Concile oecuménique? (Libre Belgique, 19.3.1962; cited by Renchon, 1969:276)

The second constraint on [QV NP] concerns the presence of a post-verbal complement or modifier, "closely linked" to the verb (again), and which would be separated from the latter by the inverted subject NP:

(130) *Quand deviendra ce comédien célèbre?

Somewhat less unacceptable are examples in which the NP follows the complement:

(131) ?Quand deviendra célèbre ce comédien?

And a very "heavy" NP produces a quite acceptable sentence:

(132) Quand deviendra célèbre le comédien que nous avons vu si bien jouer l'autre jour à la télévision? (all three examples, and judgements from Kayne, 1972:70-2)

(131) shows that this constraint in fact involves rather more than simply the separation of the verb and its complement. This is illustrated further by the unacceptability of (133):

(133) *Où gare sa voiture votre frère?

despite the fact that the verb and complement are adjacent to each other. Unlike Old French, Modern French does not permit such NP + NP sequences (unless one of the NPs is a time adverbial), because the lack of overt markers of syntactic function means

that, at least occasionally, it would not be possible to distinguish the subject from the direct object (cf Foulet, 1921:262).

As with the pourquoi constraint, it is interesting to note that examples of freedom from the constraint can be found in literature from as late as the seventeenth century:

(134) Où prend mon esprit toutes ces gentilleses? (Molière, cited by Grevisse, 1986:648)

We can see, therefore, that the last attested examples (in literature) of the following three structures date from about the same time:

- interrogative [V NP] (cf 4.3.4 above)
- [QV NP], where Q=pourquoi
- [QV NP], where NP separates V from its direct object:

It is surely not surprising that, following the complete disappearance of the Yes/No interrogative [V NP], the corresponding WH interrogative structure should become subject to quite severe constraints on its distribution.

Let us now attempt to develop the explanation for these constraints on [QV NP] along communicative or functional lines. It will be recalled that, according to Blinkenberg (1928), Korzen (1983, 1986) and others, the unacceptability of pourquoi in [QV NP] is due to the fact that pourquoi is not closely linked to the verb. More specifically, pourquoi corresponds to an argument which:

- (i) is never an obligatory part of the valency (or "sub-categorisation") of any verb;
- (ii) is often expressed by a distinct clause (in contrast with

the phrase or single word corresponding to other WH words);
(iii) normally follows all the other arguments attached to a verb.

We are now in a position tentatively to suggest a functional explanation for the greater unacceptability of pourquoi in [QV NP] (and indeed in other contexts where NP inversion occurs). Instances of [QV NP], since they involve a deviation from the normal, unmarked word order of contemporary French, pose a potential problem of comprehension for the listener, who has to identify the subject and direct object (when there is one) without the help of the case-marking inflections of Old French. If the direct object is represented by a fronted WH phrase, then the subject can be identified more easily, as there will then be only one NP after the verb (apart from any time adverbials, and NPs within PPs), and therefore only one plausible candidate as the subject. Indeed most fronted WH phrases will aid comprehension as they will reduce by one the number of post-verbal arguments from which the subject has to be identified. A fronted pourquoi, on the other hand, does not help the listener to identify the subject, since it does not reduce the number of post-verbal constituents at all - a causal clause (or a causal PP) cannot occur in a WH interrogative, in any case:

- (135) *Où est parti ton frère parce qu'il avait faim?
(136) *A quelle heure arrive le car à cause de la grève?

Since the decoding task presumably involves reconstructing the logical structure underlying a WH interrogative utterance, the listener must be able to locate the point from which the WH element may be said to have moved. We might reasonably suppose

that this would be easier for WH elements representing arguments which are obligatory for a given verb, and which typically occur soon after the verb. Similarly, it would be harder for WH elements representing arguments which are optional and which are located some distance after the verb. Since pourquoi represents a causal complement, which is both optional and follows all other complements, it is, of all WH words, the one which makes the listener's task most difficult in decoding [QV NP].

In the same vein, we may reasonably suppose that, since most intra-clause sequences of V + NP are correctly interpreted as predicates, hearers may fail to identify the subject in a [QV NP] interrogative if the NP could plausibly be the complement of the verb. A verb which can only be intransitive would thus pose fewer problems than one which is sometimes, or always, transitive. More generally, we could assume that any case of [QV NP] which includes, as a result of the inversion, a sequence which could be incorrectly parsed and thus misinterpreted by the hearer, will be judged unacceptable, to a greater or lesser degree. If this assumption is correct, then it would help to explain, on the one hand, the relative acceptability of some instances of [QV NP] with pourquoi, and on the other, the unacceptability of certain examples of [QV NP] with other WH words. Thus the relative acceptability of (137) and the actual occurrence of pourquoi with NP inversion in (138) (albeit not in a direct interrogative) would be partly due to the fact they do not include any easily misinterpretable sequences of V + NP:

(137) ?Pourquoi part ce garçon?

(138) Il ne saura jamais pourquoi est morte sa femme.

In (137) although part is homophonous with the preposition par,

the brevity of the question makes a misunderstanding unlikely, and in (138) the fact that mourir can never take a direct object means that there is no possibility of misinterpretation.

Similarly with (139) and (140):

(139) *Quand deviendra ce comédien célèbre?

(140) ?Quand vote Pierre? (Kleiber, 1986:603)

the unacceptability of the former is partly due to the fact that it separates the verb from its complement and simultaneously produces a sequence of Det + N + Adj which a hearer might well misinterpret as a NP. Similarly the doubtful status of (140) is perhaps partly attributable to the possibility of the V + N sequence being misinterpreted as a singular imperative followed by a direct object. No doubt there are other factors at play here also, such as the relative lengths of the verb (plus any complements) and the subject NP (and we shall return to this factor later).

Implicit in this line of explanation is an assumption that French listeners employ a set of perceptual strategies (for the parsing and comprehension of utterances) similar to those which have been proposed for English (cf the discussion of these in Clark & Clark, 1977:58-71). The relevance of such strategies to explaining syntactic change (and, by implication, syntactic variation) was demonstrated by Bever & Langendoen (1971). However, until psycholinguistic experiments or corpuses of misunderstandings provide us with firm evidence for the existence of similar perceptual strategies for French, we must acknowledge that the communicative explanations outlined above are provisional and speculative.

6.6 Categorical pragmatic constraints

In addition to the pragmatic constraints which result in categorical tokens, and which we have discussed in 6.3, there are others which rule out one of the variant structures and result in what we have called "semi-variable" tokens.

6.6.1 Yes/No interrogatives

The first and most significant of these is that [SV] is unavailable when the communicative function is a tentative assertion, as we see from comparing (141), (143) and (145), with, respectively, (142), (144) and (146):

- (141) B: ... quand on voit les filles maintenant sont vraiment pas (des . .) hein //
A: aha - [LAUGH]
B: si c'est vrai j'enfin c'est l'impression j'en ai hein je crois que c'est de vraiment de pire en pire. // non c'est s - j sais pas enfin. // font euh font de petites manières euh - ou est-ce que c'est l'âge? enfin je n sais pas (moi) // (3326)
- (142) #Ou c'est l'âge?
- (143) ... l'enfant quand il sort de la colonie bon on a eu beau faire / euh le maximum / bon il en a retenu quelque chose. / il en aura un souvenir bien sûr. / mais est-ce qu'on l'aura vraiment aidé pour qu'il reparte euh sur d'autres bases? euh - pt - c'est [NOISE] - c'est pas évident du tout. / ... (11389)
- (144) #Mais on l'aura vraiment aidé pour qu'il reparte sur d'autres bases?
- (145) ... culturellement on a eu une influence là-dessus. / mais est-ce que c'est bien ou mal? j sais pas. de toute façon ils étaient obligés / ... (14468)
- (146) #Mais c'est bien ou mal?

It is worth noting that this constraint operates, whether the tentative assertion is biased (eg (141) and (143)), or not (as in (145)). It seems probable that emphatic assertions and sub-topic introducing questions are also subject to this constraint, although we cannot be certain of this, as there is just one

occurrence of each in the corpus - (147) and (149) respectively:

(147) A: ... est-ce qu'i y a derrière / une philosophie particulière derrière ça? /////

B: m - crois pas / ... bon / euh - est-ce que c'est vraiment est-ce qu y a une philosophie derrière le fait de dire euh "les gamins ne remontent pas dans leurs chambres parce que bon euh i faut un coup d sifflet à c moment-là pour euh?" / et puis euh de / de les mettre / en situation de jouer - quatre heures ... (28438)

(148) #Il y a une philosophie derrière le fait de dire ...?

(149) A: est-ce que vous faites un bilan ensemble après la colo? / ...

B: ben j s je sais pas parce que / je sais pas c que m - c que prévoit euh la caisse des écoles de Bégard / et - ça veut dire - est-ce que l y a des moyens // euh financiers pour ce genre de chose? - bon je - je pense qu'en demandant euh - i y aurait pas de problème / ... (28525)

(150) #Et ça veut dire - il y a des moyens financiers pour ce genre de chose?

(Fromaigat, 1938:13, made the observation some fifty years ago that [SV] could not be used for rhetorical questions.) It is perhaps not surprising that the constraint should apply to these three CFs, as for all of them it can be said that the speaker does not want to know the answer to the question - in terms of the distinctive features outlined in chapter 5, they are all [+A] and [-F].

However there are two other CFs which also share this property, but which are not concerned by this constraint.

Post-announcements, as we saw in 6.3.2, are categorically [SV], whereas with pre-announcements ("live" or quoted), only [ESV] is excluded:

(151) ... tu sais pourquoi il (n l) a pas eu? parce qu'il avait oublié d mettre la ceinture. // ... (25743)

(152) Sais-tu pourquoi il ne l'a pas eu?

(153) #Est-ce que tu sais pourquoi il ne l'a pas eu?

(154) ... et il aurait suffi qu'une personne un peu malveillante du village dise "mais dites donc vous savez c qui s passe là-haut?" / ... (28725)

(155) "Mais dites donc! Savez-vous ce qui se passe là-haut?"

(156) # "Mais dites donc! Est-ce que vous savez ce qui se passe là-haut?"

There is one other semi-variable Yes/No interrogative where [SV] is not available, apparently due to a pragmatic constraint. This is the quoted positively-biased affirmative request for information in (156):

(157) ... ils ont dû aller à l'usine et dire "est-ce qu'ils sont bien arrivés?" et tout ça euh / ... (30646)

(158) # "Ils sont bien arrivés?"

The non-equivalence of (158) to the interrogative in (157) seems to be due to the fact that the latter was both quoted and a positively-biased affirmative question.

6.6.2 WH interrogatives

There are just two categorical pragmatic constraints on (WHQ) (apart from those discussed in 6.3.3). The first is relatively trivial, since it concerns what could almost be considered a fixed locution. [QESV] is excluded when the CF is a self-addressed request for a reminder (of form), the WH word is comment, the verb is dire (in the future tense) and the subject is je:

(159) #Comment est-ce que je dirai?

(Naturally, (159) would be acceptable if it were, for example a request for advice addressed to an interlocutor.) In addition to four instances in the corpus of [QV CL], as in (160), there is also an occurrence of [QSV], (161):

(160) ... c'est des mots qui sont / comment dirai-je? / qui sont changés - ... (16479)

(161) ... c'était une semaine où tu faisais euh m - comment je te dirai? / c'était la semaine où tu allais dans un dans une deuxième île ... (461)

This would suggest that the precise [QSV] equivalent of (160) could occur with this CF, and therefore tokens such as (160) are not classed as categorical.

The second categorical pragmatic constraint on (WHQ) is, in contrast, anything but trivial. In no less than 42 of the 111 non-categorical tokens, the [SVQ] structure is excluded as a result of this constraint. The structure seems particularly inappropriate for emphatic assertions (rhetorical questions) - (163) could not be used in the context of (162), nor (165) in the context of (164):

- (162) ... alors c'est à la fois un petit handicap mais qu'est-ce tu veux? ça fait partie des règles de vie ... (2798)
(163) #Tu veux quoi?
(164) ... y a certains professeurs qui euh / qui acceptent / euh que / qu les enfants à l'éc à l'école euh les tutoient euh / fin / j dis où est l'inconvénient si? mais moi je n'accepte pas qu les enfants m tutoient. / ... (22565)
(165) #L'inconvénient est où?

But apart from this, the precise nature of this constraint against the use of [SVQ] is obscure: it does not appear to be solely due to the communicative function, since a very broad range of different CFs are represented among the 42 interrogatives which it affects. Nor is it due to any obvious linguistic factor. The following examples include diverse CFs and linguistic elements:

- (166) A: ... tu leur prépares déjà des menus tout ça? / * (.) *
B: * non non non non * c'est (eux) - [TO X] pardon? /
[INTERRUPTION] bon qu'est-ce que t - * qu'est-ce que tu m'avais demandé? * /
A: * alors pour les randonnées * / ... (3416)
(167) #Tu m'avais demandé quoi?
(168) ... dans le même style tu as par exemple comment elle s'appelle? euh m euh Marie-Claire / ... (4351)
(169) #Elle s'appelle comment?
(170) ... on téléphone on écrit en disant "mais quand est-ce que vous venez nous inspecter? on veut vous voir euh" / ... (14780)
(171) #Vous venez nous inspecter quand?

- (172) ... qu'est-ce qu'il y a Charlotte? / ... (28855)
(173) #Il y a quoi, Charlotte?

Certain self-addressed requests for reminder, such as the interrogative in (168), are rather like set phrases, and perhaps for this reason are frequently incompatible with the [SVQ] structure - (169), on the other hand, would be perfectly appropriate as a request for information addressed to an interlocutor. More generally, it is clear that there is something very incongruent about the *odd*-numbered examples above being used in the context of their *even*-numbered counterparts. The reasons for this will perhaps become clear when we examine those interrogatives where the [SVQ] structure is actually used, and the discourse contexts in which they occur.

6.7 Variable linguistic and pragmatic constraints

6.7.1 Preliminary comments

Once the categorical tokens and categorical contexts have been set aside, the next step in the analysis is to attempt to explain the variation which remains. One approach which has been widely adopted, especially when the number of tokens is large and when a program such as VARBRUL is available, is to examine, in turn, several factor groups (eg the nature of the subject, the verb) in order to check for any significant effects they have on the choice of variant.

However, this may not always be the most illuminating approach, since it can fail to reveal constraints which involve the interaction of two or more factors from different factor groups.

We have already seen in 6.5 that there are a number of categorical linguistic constraints of this kind which operate on interrogatives - eg *[QSV] and *[QESV], if S = a NP, and V is monosyllabic; *[V-CL] and *[QV-CL], if CL = je, and V is in the present tense (except for a few common verbs). It is important that the methodology we adopt should assist rather than hinder us in the search for variable constraints, especially as these are rather more difficult to identify than categorical constraints.

VARBRUL is not at present available to us, and, in any case, the number of variable tokens to be analysed is possibly too small for it to be used. Nevertheless, we can turn this relative paucity of tokens to our advantage by adopting a rather different approach. By scrutinizing each variable token, we shall attempt to reveal the remaining linguistic and pragmatic factors which favour or disfavour the use of each interrogative structure. One advantage of this approach is that it perhaps makes it easier to identify constraints which combine two or more factors from different factor groups. In looking for pragmatic constraints, we will sometimes need to refer to groups of related communicative functions - we have already seen in 6.6.1 that there is at least one categorical pragmatic constraint which operates on such a group of CFs.

At the end of this process, any variation which remains unexplained by linguistic or pragmatic constraints will be examined further in the light of the identities of the speaker and addressee(s). We will then be in a position to see whether there is evidence of social and stylistic differentiation in the

data.

Naturally, there is a danger that such an approach may degenerate into a series of ad hoc pseudo-explanations. We shall try to avoid this by the use of supporting quantitative evidence, and, where possible, by referring to the findings of previous studies. Where quantitative evidence is presented, the observed frequency for each variant will be calculated as in (174):

$$(174) \text{ total of actual occurrences of variant} \\ \frac{\text{total of actual + potential occurrences of variant}}{\text{total of actual + potential occurrences of variant}}$$

There is, of course, nothing unusual about this formula. However, the decision to exclude from the "potential occurrences" the categorical contexts of the semi-variable tokens does represent a departure from standard methodology. Furthermore, since the "total of actual + potential occurrences" will rarely be identical for any two variants, the frequencies for the different variants will not always be easy to compare. Most strikingly, the relative frequencies for all the variants often will not add up to 100%. Table 6.6 shows the observed frequencies for the variants of (YNQ) calculated according to this method, and table 6.7 shows the same for variants of (WHQ).

Table 6.6 Observed and relative frequencies of variants of (YNQ)
(variable contexts only)

[SV]	[ESV]	[V-CL]	
77/95	32/106	0/106	(Total N=109)
(81.1%)	(30.2%)	(0%)	

Table 6.7 Observed and relative frequencies of variants of (WHQ)
(variable contexts only)

[SVQ]	[QSV]	[QESV]	[QV-CL]	[QV NP]	[Q=S V]
18/65	29/58	51/105	6/104	3/3	4/4
(27.7%)	(50%)	(48.6%)	(5.8%)		

(Total N=111)

The relative frequencies of [SV] and [ESV] in table 6.6 are basically similar to those in table 6.3 - the only significant difference being that [SV] is shown to occur in a rather higher proportion of the contexts where it is possible. In contrast, table 6.7 is in several respects radically different from table 6.4. Firstly, [QV NP] and [Q=S V] are shown to realise all their potential occurrences. Secondly, the non-standard [QSV] variant is now seen as being on a par with [QESV], in the sense that it too is chosen by speakers in half the contexts where it is an option. Thirdly, [SVQ], whilst still a minor variant in relation to [QSV] and [QESV], is now shown to be selected in a rather higher proportion of contexts than might have been supposed from table 6.4. The relative frequencies of [QV-CL] and [QESV] are both quite close to those in table 6.4 - the exclusion of the categorical contexts has little effect on them, simply because there are few contexts in which they are not possible.

(An alternative approach for (WHQ), adopted by some of the quantitative studies referred to in chapter 4, would be to calculate the scores for the WH pronouns separately from the WH adverbs. This would deal effectively with the fact that [QSV] is

unacceptable if the WH word is que, but it would not handle the other categorical contexts which result in semi-variable tokens. Moreover, by, in effect, dividing (WHQ) into two sub-variables, it would exacerbate the problems arising from small numbers of tokens.)

6.7.2 [SV] and [ESV]

As was apparent from tables 6.3 and 6.6, [V-CL] does not occur at all in the data. Of the 109 variable and semi-variable tokens of (YNQ), 70.6% are of [SV], which we can reasonably consider to be the unmarked variant. Our principal objective in this section, in a sense, is to explain why speakers sometimes use the marked variant, [ESV].

There do not appear to be any very obvious linguistic constraints affecting the choice between [SV] and [ESV], but several writers have suggested that certain pragmatic factors favour the use of [ESV]. Söll (1982:47), in his study of children's interrogatives, found a tendency to use [ESV] at the beginning of a "conversational unit" or to introduce a new topic. Perhaps not unconnected to this is Grundström's (1973:25) suggestion that [ESV] is used for more lengthy questions, so that the speaker can send an early signal to the addressee of the interrogative nature of the utterance. Finally, Vigner (1978:88), among others, comments that [ESV] is used to convey politeness.

In view of the total absence of [V-CL], it follows that the categorical pragmatic constraints which exclude either [SV] or

[ESV] (and which we discussed in 6.6.1) result in the virtual elimination of choice for the speaker. We could say that tokens affected in this way are "near-categorical". Table 6.8 recapitulates (from 6.6.1) the number of such Yes/No interrogatives.

Table 6.8 "Near-categorical" tokens of (YNQ)

- pre-announcements, quoted or "live" ([ESV] is excluded):	3 [SV]
- quoted positively-biased affirmative request for information ([SV] is excluded):	1 [ESV]
- tentative or emphatic assertions, S.T.I questions ([SV] is excluded):	13 [ESV]

Total 17

Of the remaining 18 tokens of [ESV], 12 are quoted utterances. The fact that an interrogative is quoted does not, of course, oblige the speaker to employ [ESV], but it nevertheless substantially increases the likelihood of them doing so. Table 6.9 shows the frequencies of [SV] and [ESV], in variable and semi-variable tokens (but excluding the "near-categorical" tokens), for three groups of communicative functions. The first group of CFs are all unquoted, non-echo, non-self-addressed and their speakers want to know the answer to the question - in terms of the features outlined in chapter 5 they are [+A, -B, +F]. The group includes requests for information (neutral and biased), requests for opinion, requests for clarification, checks on the knowledge of the addressee, offers and requests for action. The second group is that of quoted CFs (which can be characterised as

[-A, -C, -I]), and the third is the small group of self-addressed CFs (which are [+A, +B]).

Table 6.9 Frequencies of [SV] and [ESV] for three groups of communicative function

	[SV]	[ESV]
- CFs where speaker wants answer from addressee	66/71 (93%)	5/71 (7%)
- quoted CFs	7/19 (36.8%)	12/19 (63.2%)
- CFs where speaker wants answer from self	1/2	1/2

(categorical and "near-categorical" tokens excluded)

From table 6.9, it is clear that, for CFs where the speaker wants an answer from their interlocutor, [SV] is overwhelmingly the structure preferred. (The five cases where [ESV] is used will be carefully scrutinized later.) For quoted CFs, in contrast, there is a tendency for the speaker to prefer [ESV]. We can now propose a unified explanation for this and for the use of [ESV] in tentative assertions, emphatic assertions and S.T.I questions:

(175) In Yes/No interrogatives, a speaker wishing to elicit the answer from an interlocutor tends to use [SV], whereas a speaker not wishing to elicit an answer tends to use [ESV].

One type of question which, for obvious reasons, does not occur at all in the present corpus, is what has been called the "teacher's" or "display" question. Fromaigeat (1938:14) noted that in such questions [SV] is virtually excluded: a teacher who asked "Il y a des images dans le morceau que nous avons lu?" would be implying that they did not know the answer themselves (or else the question would be positively-biased). In view of this, (175) should perhaps be amended as follows:

(176) In Yes/No interrogatives, a speaker who does not know the answer to the question and who wishes to elicit the answer from an interlocutor tends to use [SV], whereas a speaker not wishing to elicit an answer, or who already knows the answer, tends to use [ESV].

There remain 6 tokens where we have yet to account for the use of [ESV]. To do so will require a consideration of certain characteristics of the speaker and addressee, notably their age, sex and social class. We shall postpone discussion of these 6 interrogatives until 6.9.2

Givón (1978:89) has claimed that external negation tends to be marked in natural languages by special marked patterns, and so we might expect to find a tendency for [ESV] rather than [SV] in positively-biased negative questions. But in fact this is not borne out by our data: although there is one example of [ESV], in (177), (where again this can be attributed to the fact that the utterance is a tentative assertion), there are four examples of [SV] in questions of this type, including that in (178) and the two in (179):

(177) ... ça ça évite des problèmes bon maintenant / est-ce que c'est bon aussi? j veux dire est-ce qu'on vit pas un peu sur nous-mêmes? (et) ça j'en sais rien. / ... (21224)

(178) ... t as quelques gamins qui sont bien mûrs et qui eux viennent discuter avec toi / "on pourrait pas faire ceci? moi (j'ai) j voudrais bien aller à l / faire une rando comme ça" ... (2397)

(179) B: mais l'Angleterre ça s tend euh c'est pas un système euh / de deux grands partis quoi? c'est pas ça? * c'est (. .) * /

A: * oui - du point d vue politique * tu veux dire? ... (23966ab)

Borillo (1979:27) similarly points out that [SV] can be used in positively-biased negative questions. (We shall consider in the next section whether Givón's claim finds any support with regard

to inversion in French.)

6.7.3 [V-CL] and [QV-CL]

Although [V-CL] and [QV-CL] are treated in the present framework as variants of different variables, they are obviously closely related, and since they appear to be subject to several of the same linguistic constraints, it will be convenient to discuss them together.

Assuming that forms such as n'est-ce pas? and est-ce que involve fossilized forms of clitic inversion, there are no productive occurrences of [V-CL] in the 106 contexts in the data where the speaker could have used this structure. In contrast, there are 8 tokens of [QV-CL]: 2 of these are categorical tokens, since they were quoted from the words of a singing game (as discussed in 6.3.3). 4 others were examples of comment dirai-je?, which could probably be regarded as a more or less fixed locution, even though we have argued (in 6.6.2) that these should not be considered categorical occurrences. The same might be said of one other case, that in (180):

- (180) B: ... c'est dire à c moment-là m on - on l'avait très
jeune hein / [ENTER X] je vous présente la jeune
gestionnaire. / [X INAUDIBLE]
A: oui oui on s'est déjà rencontrés /
B: ah bon. //// euh / où en étais-je? /
A: oui on parlait de du certificat d'études. / (21608)

A number of other more or less fixed locutions involving clitic inversion in interrogatives are generally reckoned to be still fairly current in informal spoken French. These include:

- (181) Quelle heure est-il?

- (182) Comment veux-tu que ...?
(183) Comment se fait-il que ...?

((182) is quite commonly used in rhetorical questions.) But even the number of these locutions seems to be dwindling. From the list which Fromaigeat (1938:20) suggested were still widespread, plaît-il? and qu'est-ce à dire? now seem archaic.

The final example of [QV-CL], from our data, and possibly the only one to involve the genuinely productive use of this structure, is that in (184):

- (184) ... c'est vrai que c'est difficile de faire autre chose que / déjà la maison les enfants le travail dehors / faire encore autre chose après euh / i faut vraiment du courage. / oui mais alors où allons-nous? hein / c'est là ou c'est la fin hein - des centres. / ... (13762)

The communicative function of this utterance is an emphatic assertion, or "rhetorical question", and it is surely not too surprising to find a marked interrogative structure being used for a highly marked type of question.

With such a small number of variable tokens, the data clearly cannot offer us any positive evidence for the existence of linguistic constraints on clitic inversion. However, we shall briefly mention some important constraints which have been suggested in previous studies and illustrate these with negative evidence, ie instances where an informant could have used inversion, but chose not to, and was perhaps influenced by the constraint in question.

Behnstedt (1973:279) found that in formal middle-class speech clitic inversion was much less likely to occur when cela/ça was the subject. Cela/ça itself, of course, cannot be inverted at

all, as in (184), and can only occur with an inverted "reprise"
il, as in (186) and (187), and even then sounds rather awkward:

- (185) *Va-ça/cela servir en Angleterre?
(186) Cela va-t-il servir en Angleterre?
(187) En quoi cela consiste-t-il?

Ce, which occurs only with forms of être (though even there it is
rivalled by ça when the verb begins with a consonant, eg ça
sera), seems only barely acceptable today with anything but est:

- (188) ?Était-ce ce matin ou hier?
(189) ?Sera-ce dix ans?
(190) ?Sont-ce des garçons intelligents? (Kayne, 1972:121)

(However such examples were still being produced in speech some
sixty years ago, according to Damourette & Pichon,
1911-1934:598-9.) As with je, of course, the final vowel of ce is
schwa, which would normally be elided in group-final position,
leaving the [s] to be attached to the preceding syllable and thus
giving pronunciations such as [sɔ̃s] for (190). We have already
argued in 6.5.1, with reference to je, that this process may have
caused perceptual difficulties which contributed to the decline
of inversion in such cases.

With regard to inversion with il(s), elle(s) and on, it is worth
noting that these entail the pronunciation of the liaison
consonant from the preceding verb (or the insertion of -t-, where
there is none), and that this consonant is resyllabified. This
seems to have the effect of "camouflaging" the clitic, and there
is anecdotal evidence that this can lead to misunderstandings.

(191) occurred (as a rhetorical question) at the end of a T.V
news item, and at least one native-speaker reported that, even
after listening to the recording several times, she was unable to
make sense of the utterance, because she interpreted the first

two words as ces styles:

(191) Cessent-ils pour autant d'être des chômeurs? (Antenne 2, 5 February 1988)

In Montreal, according to St-Pierre (1977:3), clitic inversion occurs almost exclusively when the subject is tu or vous, but it is apparently quite frequent in such cases, judging from the numerous examples given by St-Pierre from her recordings of informal speech. Impressionistically, it seems the same could be said for metropolitan French, at least for those speakers who still use the structure at all productively: forms such as pouvez-vous, voulez-vous, veux-tu and croyez-vous have a relatively familiar sound to them. The existence of such forms helps to account for Behnstedt's (1973:286-9) finding that certain verbs, such as vouloir and pouvoir, have a rather stronger tendency than other verbs to be inverted.

Another factor which has been found to disfavour inversion, in those varieties where it is still used, is when the verb is of three or more syllables as in the following:

(192) Manipule-t-on?

Similarly, it has been found that verbs in the imperfect tense tend to be inverted less frequently than those in other tenses:

(193) Marchaient-ils bien?

(194) Dans quoi faisiez-vous ça?

(Both these findings are attributable to Behnstedt, 1973:291-2.)

In the present corpus, all 8 occurrences of [QV-CL] involve mono- or disyllabic forms of very frequent verbs: es, étais, fais, allons and dirai.

A further factor which seems to inhibit clitic inversion (but for which we again lack positive quantitative evidence) is the presence of one or more non-subject proclitics, such as me, les, lui and also the negative particle ne. With one or more of these, even verbs which invert relatively frequently begin to sound somewhat awkward:

(195) Les as-tu vus?

(196) T'en souviens-tu?

(197) Que m'avais-tu demandé?

(Notice, that this last example, however, is nevertheless acceptable, in contrast with the quite unacceptable cases where the clitic que is separated from the verb by a subject clitic, as we saw in 6.5.2.)

When several of these inhibiting factors are simultaneously present, the awkwardness of the inversion seems greater:

(198) Où en était-on?

(199) Comment leur enseignais-je ça?

All of these factors have the effect of lengthening the verbal group, and hence of delaying the crucially important information of the subject itself. As French has, over the centuries, become an increasingly analytic language, we may suppose that cases of clitic inversion which involve one or more of these factors became increasingly difficult for children to master, and so gradually disappeared from the vernacular. Today, as we have seen, even the comprehension of such cases can pose problems.

However, it would be misleading to suggest that all the quantitative evidence of constraints can be accounted for in terms of this general explanation. Among the factors which favour

the use of clitic inversion, Behnstedt found that one of the strongest was when the verb was in the future tense - though not the conditional (1973:292). Since the future, like the imperfect, is an inflected tense, it increases the morphological complexity of the verbal group, and we might therefore expect speakers to avoid the added complexity of inversion. At present, we can offer no explanation for the fact that they do not.

Before leaving clitic inversion, let us briefly consider its use in negative interrogatives. Borillo (1979:27) has stated that, [V-CL] is used only for positively-biased negative questions, and not for negatively-biased ones. In seeking to explain this, she points out that in [V-CL] the negative particle ne is located at the front, or, in a sense, "outside" the sentence, whereas this is obviously not the case in negative [SV] or [ESV] interrogatives. She claims that consequently this structure is the most natural one for the expression of external negation, ie [? TRUE [neg TRUE [pos X]]], where the negative operator is located outside the central proposition.

Attractive as it may at first appear, this explanation, nevertheless, runs into several difficulties. Firstly, in cases of "complex" inversion, it is the subject NP which has the first position in the sentence, not the ne. Secondly, surely in contemporary French pas has as much a claim as ne to be considered the principal marker of negation (though it is generally true today that ne is obligatory in [V-CL] interrogatives). Thirdly, in Classical French (and still in rare archaic examples today) ne was omitted precisely in

positively-biased questions, but not in negatively-biased ones. Finally, if negative [V-CL] is indeed the most natural structure for the expression of positively-biased negative questions, then we would surely expect it to occur more frequently than the other structures. But in the present corpus it is in fact avoided in the 5 questions where it could have been used, and we may reasonably suppose that it was similarly avoided by speakers who contributed to the larger corpora of certain previous quantitative studies (cf especially the elderly Belgian couple and the 9-year-old children in table 4.2 in chapter 4).

As an alternative explanation, we would suggest that, as [V-CL] has gradually disappeared from the vernacular, it has become the most marked variant of (YNQ). Consequently (or perhaps just simultaneously), its pragmatic functions have come to be restricted to the more marked cases, such as rhetorical questions, and positively-biased negative questions, which, as we saw in 5.2.3, are considerably less frequent in the present corpus than are negatively-biased negative questions. In Classical French, the marked pragmatic function of positive bias in negative questions was the domain of the marked negative interrogative structure, ie \emptyset + [V-CL] + pas, whereas for the less marked negatively-biased negative questions, the normal ne + [V-CL] + pas structure was used. This development is summed up in table 6.10.

Table 6.10 The use of [V-CL] in negative interrogatives
in Classical and Contemporary French

	Classical French	Contemporary French
negatively-biased (internal negation)	<u>ne</u> + [V-CL] + <u>pas</u>	-
positively-biased (external negation)	Ø + [V-CL] + <u>pas</u>	<u>ne</u> + [V-CL] + <u>pas</u> (formal styles only)

Kontra (1980) describes an interesting parallel to this in modern English. The negative interrogative not + Subject structure, as in:

(200) Did not Robert go out before us?

has been losing ground since the seventeenth century, is often now judged to be appropriate only in printed language and is often unacceptable, especially with pronoun subjects:

(201) ?*Could not he have learned that?

(202) ?Is not linguistics really a branch of psychology?

As with negative [V-CL] interrogatives in French, this increasingly marked structure is, according to Kontra, now restricted to the expression of positive bias. Of course, in both French and English, positively-biased negative questions are expressed additionally, and indeed rather more frequently, by other, less marked structures. On the basis of the evidence from these two languages, then, the generalisation by Givón (1978) referred to in 6.7.2 should perhaps be modified as follows:

(203) Marked negative interrogative structures (eg near-archaic ones) tend to be restricted to the expression of the marked case of negation, ie external rather than internal.

6.7.4 [Q=S V]

Turning our attention now exclusively to WH interrogatives, let us first consider what is perhaps the most straightforward variant, [Q=S V]. This structure occurs in all four contexts in the data where it is possible, in preference to the only alternative [QESV]. It is interesting to note that the [QESV] variant seems less appropriate as an alternative for (204), and particularly for (206), than it does for (208) and (210):

- (204) ... j'avais calmé tout le monde / elle était repartie faire sa tournée de tentes alors les gosses ils l'ont chahutée. / et paf c'est redémarré encore un tour / alors qui prend les sanctions apres derrière? c'est moi qui prends les sanctions / euh vis-à-vis des gosses ... (4395)
- (205) ?Alors qui est-ce qui prend les sanctions apres derrière?
- (206) ... c'est une activité - qui n s'adressait pas à cette tranche d'âge-là. / et on a fait une erreur. / bon - euh - qui ne fait pas d'erreurs? [LAUGH] hein / (mais c'est bien de faire l'erreur parce qu'on s'aperçoit .) // ... (22306)
- (207) ??Qui est-ce qui ne fait pas d'erreurs?
- (208) B: * mais ça s * - comment ça s passe en Angleterre par exemple là en cas d - pendant en cas de guerre? - qui y a été? / s'i y a pas de d'armée / * s'i y a pas de (con) service militaire * /
- A: * euh c'est * une armée de metier ... (23982)
- (209) Qui est-ce qui y a été?
- (210) ... alors y avait effectivement un choix d'activité / ça se faisait tous les matins - les quatre-vingt-six / dans la même pièce. / bon // et (enfin) seulement y avait - y avait aussi les vélos. //alors "qui veut faire du vélo?" / ça veut dire qu'y en avait vingt-cinq - qui levaient l doigt / et qui à un moment se s'avançaient "moi moi moi moi moi moi moi" / hein ... (28124)
- (211) Qui est-ce qui veut faire du vélo?

The lesser acceptability of [QESV] in the contexts of (204) and (206) seems to be due to the fact that the interrogatives are, in the former, a S.T.I question, and, in the latter, an emphatic assertion (rhetorical question). Whilst it is certainly not the case that [QESV] is excluded from expressing these CFs - there are a number of occurrences in the data which do so - it does seem that, where there is a choice, [Q=S V] is preferred. It may

be that [Q=S V], the older structure, is still preferred to [QESV] in literature (cf Grevisse, 1986:650), and that these "rhetorical" functions tend to be expressed by the more marked variants associated with conservative varieties.

As we have seen, the only WH word to occur in this structure in the corpus is qui. Combien, quel and lequel are also possible, but there is perhaps a tendency for them to be avoided in contexts where there is a risk of ambiguity with [QV NP], ie where there is a NP complement which could be misinterpreted as the subject, as in (212):

(212) Qui aime Jean? (Grevisse, 1986:647)

6.7.5 [QV NP]

In contrast to the massive avoidance of clitic inversion by our informants, the [QV NP] structure was used in all three contexts where it was possible. The very low frequency of potential occurrences is at least partly attributable to the more general rarity of subject NPs in spontaneous, relatively uninterrupted spoken discourse. NPs representing entities new to the conversation tend to be introduced in positions other than subject, and, in interviews, of course, they are often introduced, not by the interviewee at all, but by the interviewer, who generally exerts greater control over the topic of discourse.

At first glance, it might be supposed that WH interrogatives with a right-dislocated NP (coreferential with a pronominal subject)

should also be counted as potential instances of [QV NP], as in
(213):

- (213) X (a child): William /
B: oui /
X: (. . .) - Gerard il m'a donné une claque /
B: mm / alors ça / où il est Gérard? / (28559)

However, if we compare this with the [QV NP] interrogative (214),
(214) #Où est Gérard?

we see quite clearly that the two are not pragmatically equivalent, with (214) being inappropriate in the context of (213). This seems to be due to the fact that right-dislocated NPs, such as Gérard in the underlined utterance in (213), tend to be used for entities which are "given" (or at least inferrable) in the discourse (cf Lambrecht, 1981:84-98). Such NPs are generally less prominent prosodically than the NP in [QV NP], and are traditionally considered to be external to the clause - in writing they are usually set apart from it by a comma.

If we now consider the three occurrences of [QV NP] in the data, we see that the subject NP is "new" in the discourse, absolutely so in the case of (215), and relatively so in (216) and (217):

- (215) ... y a certains professeurs qui euh / qui acceptent / euh que / qu les enfants à l'éc à l'école euh les tutoient euh / fin / j dis où est l'inconvénient si? mais moi je n'accepte pas qu les enfants m tutoient. / ... (22565)

- (216) ... bon - Aline par contre euh qui est là / Aline / elle ne voulait plus travailler avec l'ancienne équipe. / elle ne s'accordait pas avec / et elle ne voulait plus travailler avec donc - ça l'intéressait de travailler avec la nouvelle équipe / bon ben déjà donc ça (en) fait déjà deux j'ai l'impression qu si jamais on fait encore deux mois l'année prochaine c'est sûr qu'ils voudraient revenir (au moins on en a) prévu pendant l'année. / David m'a déjà dit qu'il voudrait lui aussi revenir / bon les autres (on n'en a) pas encore discuté on verra à la fin quand on fera le bilan s'i y en a qui veulent revenir et puis on discutera si / si nous aussi on a envie (qu'i reviennent) avec nous ou pas / hein // mais - comment d'habitude se fait se faisait notre groupe? - eh bien - bon je te dis ... (21173)

- (217) A: ... fin la fin d'août je compte aller en Lorraine un petit peu parce que j'ai des amis (qui habitent là-bas

. . . .) /
 B: mais en quoi consiste votre enquête? /
 A: alors euh ... (211042)

Indeed in (216) and (217) the fact that a new topic or sub-topic is being introduced is highlighted by the presence of mais before the question. The need for the [QV NP] structure in these contexts can also be explained by the maxims of "end-weight" and "end-focus" (cf Leech, 1983:64-6; Quirk et al, 1985:1356-64), according to which longer and more informative constituents tend to occupy the end-position in the clause. We concluded in 6.5.2 and 6.5.3 that the unacceptability of the [QSV] and [QESV] versions of the interrogative in (215) was attributable to the "imbalance" between the monosegmental copula est and the subject NP l'inconvénient. We can now say, rather more specifically, that they were unacceptable because the heavier constituent, the NP, was not in the end-position, and that they therefore violated the maxim of end-weight. In the interrogatives in (216) and (217), the verbs are more substantial than the copula of (215), and so the use of other (WHQ) structures does not entail any serious violation of the end-weight maxim:

- (218) ?Comment d'habitude notre groupe se faisait? [QSV]
- (219) ?Comment d'habitude est-ce que notre groupe se faisait?
[QESV]
- (220) Comment d'habitude notre groupe se faisait-il? [QV-CL]
- (221) ?En quoi votre enquête consiste? [QSV]
- (222) ?En quoi est-ce que votre enquête consiste? [QESV]
- (223) En quoi votre enquête consiste-t-elle? [QV-CL]

The two [QV-CL] versions, (220) and (223), are quite acceptable, as the verbs are reinforced by the enclitics and thus become weightier than the subject NPs themselves. (However, such sentences are, as we have seen, very rare in informal spoken French.)

With just three occurrences of [QV NP], it is clearly not possible to conclude from the present corpus whether the various WH words favour or disfavour the use of this structure. However, we can briefly summarise some of the evidence from secondary sources. In addition to the very nearly categorical constraint against the use of pourquoi, grammarians mention that there is a tendency to avoid lequel, quel (and their morphological variants), qui and combien, where there would be a risk of ambiguity with the [Q=S V] structure (as already mentioned in 6.7.4). This is, however, only a tendency - such more or less ambiguous utterances do sometimes occur (cf (212) above and Grevisse, 1986:647, for several other quite recent examples).

One final linguistic constraint on [QV NP] involves the collocation of que, as the WH element, and a form of the copula. Grevisse (1986:645-6) cites several examples from modern literature, some of which seem more unacceptable than others:

- (224) Que sont ces petits des petits? (Michelet)
- (225) Mais qu'était leur condamnation misérable [...]? (Malraux)
- (226) Qu'est la vieille langue? (Littre)
- (227) Qu'est cela? (Jouhandeau)

Commenting on the rarity of such instances, Grevisse suggests that this is because of their lack of clarity. This indeed seems to be the case, and if we consider in addition a confected example,

- (228) Que sera notre avenir?

then we can suggest that the more phonologically substantial the form of the copula, the more acceptable is the utterance - two syllables being more substantial than one, and a CV syllable being more substantial than one comprising just a vowel.

It is interesting to note that nearly seventy years ago Foulet wrote that the rival construction, [Q NP V-CL] or ("complex inversion"), was more natural than the literary-sounding [QV NP] structure (1921:295). Judging from the present corpus, and the previous studies referred to in chapter 4, it is the former, rather than [QV NP], which has virtually disappeared from spontaneous spoken French. Nevertheless, NP inversion is both highly constrained linguistically and quite infrequent in informal varieties.

6.7.6 [SVQ]

We move on now to consider what is perhaps the most enigmatic interrogative structure in contemporary French. Most grammars have relatively little to say on it, other than that it is more commonly encountered in colloquial styles of speech than in writing. Successive quantitative studies have shown that it can be found in all varieties of spoken French, with an overall frequency which varies, seemingly at random, from 2% (the elderly Belgian couple of Pohl's 1965 study) to 38.8% (the middle-class Parisians studied by Ashby, 1977). Impressionistically, it appears to be considerably more frequent and less marked than its English equivalent. Let us attempt to discover what factors motivate speakers to choose this [SVQ] structure, and what constraints inhibit its use.

As with the analogous English structure, [SVQ] can be used for echo questions, when the speaker is requesting repetition or clarification of the constituent replaced by the WH element, and

simply repeating the rest of the sentence from an interlocutor's preceding turn. Several writers have claimed that the echo and non-echo occurrences actually involve different structures (eg, for French, Grevisse, 1986:628-9, and, for English, Huddleston, 1984:370), although it is difficult to see precisely how they differ syntactically. In French it is certainly the case that the [SVQ] structure is far more frequent in non-echo contexts, and also that socio-stylistically it is evaluated as less standard in such contexts, and even condemned by purists (cf for example, Fromaigeat, 1938:16). According to Huddleston, the English structure, Subject + Verb + WH element, is most commonly found "in contexts involving sustained interrogation, by an investigator, quizmaster or the like." (1984:370).

In 6.6.2, we saw that [SVQ] is excluded for emphatic assertions, and, on closer inspection of the actual and potential occurrences of the structure, there does appear to be a very general variable constraint deriving from the communicative function. Table 6.11 gives the observed and relative frequencies of [SVQ] for four groups of CF.

Table 6.11 Frequencies of [SVQ] for four groups
of communicative function

- CFs where speaker wants answer from addressee	10/19	(52.6%)
- CFs where speaker wants answer from self	6/22	(27.3%)
- quoted CFs	1/15	(6.7%)
- S.T.I. questions	1/9	

(categorical tokens and categorical contexts excluded)

Although the number of tokens per group is not large, we may tentatively conclude from table 6.11, and from our previous observation on emphatic assertions, that [SVQ] is more likely to be used when the speaker is asking a question to which they want an answer.

The most apparent feature which sets [SVQ] apart from the other variants of (WHQ) is, of course, the fact that the WH phrase is not fronted, but remains in what is generally the end-position in the clause. Since, as we have seen, this position tends to be occupied by a longer and more informative constituent (in accordance with the maxims of end-weight and end-focus respectively), let us investigate whether this is true of the WH element in examples of [SVQ].

With regard to the end-weight maxim, we may hypothesise that the WH element in occurrences of [SVQ] would be longer than the rest of the clause, whereas in non-occurrences of [SVQ], the converse would be true. In order to test this, we have calculated the average length (in syllables) of both the WH element and of

the SVC (Subject + Verb + Complement) for actual occurrences of [SVQ] and also for those occurrences of other (WHQ) structures where [SVQ] would have been an acceptable and equivalent alternative. For interrogatives with a fronted WH element, the SVC was deemed to go up to the point from which the WH element could be said to have moved. Thus, in (229) the SVC includes anglaise, but in (230) it does not include pour trois jours en Angleterre:

- (229) à quelle heure est-ce que vous mangez l'assiette anglaise? (13939)
 (230) qu'est-ce que vous voulez qu'ils fassent pour trois jours en Angleterre? (13446)

Est-ce que has been excluded from the calculations, as it cannot, of course, occur in [SVQ]. As for the WH element, the entire WH phrase has been measured, since in WH-fronted structures it is this constituent that is fronted - except for occasional instances, in formal styles, where combien is fronted without its complement. Thus, in (231) and (232) we have counted the number of syllables in avec quoi and combien d temps, respectively, but, in (233), d'agrafes has not been included since it is right-dislocated, and not a part of the WH phrase:

- (231) et c'était avec quoi comme organisme? (30444)
 (232) ça fait combien d temps qu t es arrivé là? (28)
 (233) t en veux combien d'agrafes? (2179)

Table 6.12 Mean length (in syllables) of WH element and of SVC
in occurrences of [SVQ] and of other (WHQ) variants
 (where [SVQ] would be an acceptable and equivalent alternative)

	WH element	SVC	difference between WH & SVC
tokens of [SVQ] (N=18)	1.9	2.6	0.7
tokens of other (WHQ) variants (N=45)	1.3	3.1	1.8

(Note: 2 [QESV] tokens have been excluded here because they are defective, ie the verb is inaudible or omitted.)

From table 6.12, we see that, in occurrences of [SVQ], the WH element is not in fact longer, on average, than the Subject + Verb (+ Complement). (Indeed in only 4 out of the 18 occurrences is that the case.) However, the WH constituent is generally heavier in [SVQ] than it is in the other structures, and the difference in length between the WH element and the S + V (+ C) is rather less in tokens of [SVQ] than in those of the fronted WH structures. We may conclude, then, that there is some evidence here of the effect of the end-weight maxim, though not to the extent that was initially hypothesised.

A number of previous studies have found the identity of the WH word to be a significant influence on the choice of structure in (WHQ). In particular, Behnstedt (1973:93) found that quand and especially combien have a clear and consistent tendency to favour the use of [SVQ], whereas comment, pourquoi and, to a lesser extent, où, disfavour that structure. In the present corpus, the number of occurrences of each of these WH words is very low, but there are nevertheless some similarities to the tendencies found

by Behnstedt: table 6.13 shows the observed frequencies for the use of [SVQ] with the various WH words (with relative frequencies only where the denominator is at least 10).

Table 6.13 Observed frequencies of the use of [SVQ] with different WH words
(excluding categorical tokens and categorical contexts)

qui	quel	que/quoi	combien	quand	comment	où	pourquoi
2/2	1/2	6/33 (18.2%)	7/7	0/1	1/9	1/10 (10%)	0/1

It is perhaps worth noting from table 6.13 the low rates of use of [SVQ] with où and comment, and the high rate of use with combien - especially as the latter confirms a trend found in previous studies: eg, in addition to Behnstedt, Gougenheim et al (1964:228-9).

The principle of end-weight does not account entirely satisfactorily for these differences among the WH words, for, whilst it is true that the relative heaviness of combien (especially when followed by a complement) would lead us to anticipate the tendency to place it in the end-position, it is two other heavy WH words, comment and pourquoi, which are the least likely to occur there.

An alternative explanation is perhaps to be found in comments made by Bolinger (1978:137-8) in connection with multiple WH questions in English. Observing that how and, especially, why are far less acceptable than other WH words in the end-position, he suggests that the looser the connection between the WH word and

the verb, the easier, or even the more necessary, it is for the WH word to be fronted. And to clarify what he means by "connection", he points out that the types of complements most frequently required by verbs are, in descending order, as follows:

(234) nominal - locative - temporal - manner/means -
reason/purpose

Moreover, this is also the normal order in which complements follow the verb in declarative sentences. In addition to accounting for the tendency to avoid final comment and pourquoi, this also points to an explanation for the high frequency with which combien is found in end-position. Where it occurs, combien tends to be the essential, if not the only, complement of the verb, as can be seen from the seven WH interrogatives with combien from the present corpus:

- (235) ça fait combien d temps qu t es arrivé là? (28)
- (236) t en veux combien d'agrafes? (2179)
- (237) pour euh une leçon c'est combien à peu près? (10378)
- (238) vous êtes combien? (1280)
- (239) c'est à combien? (20175)
- (240) y avait combien? (20567)
- (241) mais vous êtes en France pour combien de temps? (211040)

Notice, however, that the order of complements in (234) would predict, incorrectly, a low frequency of occurrence of quand in end-position. Secondly, the constraint which excludes pourquoi in the [QV NP] structure is also an exception to Bolinger's general principle.

So far we have been concentrating very much on the WH element in the [SVQ] structure: let us now shift our attention to the elements which make up the rest of the clause. We have already seen in table 6.12 that the average length of the Subject + Verb

(+ Complement) is shorter in [SVQ] than it is in the WH-fronted structures. One of the principal factors contributing to this is that there are fewer complements in the occurrences of [SVQ] in the data than in the occurrences of the other (WHQ) variants. More generally, it seems that the presence of certain types of complement disfavors the use of [SVQ]. Table 6.14 shows the numbers of several types of complement which are found in the occurrences of [SVQ] and in those occurrences of the WH-fronted structures where [SVQ] would be an acceptable and equivalent alternative. Only those complements which are located after the verb but before the WH element (or the point from which it may be said to have been moved) are taken into account here.

Table 6.14 Complements in [SVQ] and other (WHQ) variants
(where [SVQ] would be acceptable and equivalent)

	[SVQ]	other (WHQ) variants
- infinitival	0	14
- clausal	0	2
- ça	1	1
- prepositional phrase	2	0
- none	15	28
Total	<u>18</u>	<u>45</u>

(Note: 2 defective [QESV] tokens have again been omitted here.)

Table 6.14 would suggest that the presence of an infinitival or clausal complement, as in (242) and (244) respectively, makes it less likely that a speaker will use the [SVQ] structure, such as (243) and (245):

- (242) et qu'est-ce que tu veux me faire dire pour ça? (4918)
 (243) Tu veux me faire dire quoi pour ça?
 (244) qu'est-ce tu veux qu je dise? (29303)
 (245) Tu veux que je dise quoi?

Having considered the end-weight maxim as a motivating factor for the use of the [SVQ] structure, let us now investigate whether the parallel maxim of end-focus can shed any further light on the matter. At first glance, it may seem rather unpromising, since it surely cannot be said that a WH element - that which is unknown by the speaker - is more informative than any other part of an utterance.

Although all WH interrogatives presuppose the SV(C) part of the clause, Bolinger (1978:131) points out that the WH-final structure in English presupposes more than does the WH-fronted structure. Later he says: "The speaker not only assumes a fact but assumes his interlocutor's agreement. More often than not the speaker incorporates something that the interlocutor has already said [...] or something that the speaker infers from what the interlocutor has said or appears to have meant." (p.132) This interpretation seems equally valid for the corresponding structure in French.

Looking at the 18 non-categorical occurrences of [SVQ] in our data, we find that, in various ways, the S + V (+ C) parts are of minimal informational value. In addition to the fact that the subjects are all pronouns (as is the case with the vast majority of the interrogatives), 15 of the verbs are semantically empty - être, il y a etc. In one instance the subject, verb and complement are all represented by pro-forms, referring back to

information given by the addressee in the preceding context (and hence "textually evoked", to use the terms of Prince's (1981) framework):

- (246) A: on a essayé notre association a essayé de faire des colos binationales (. . . .) enfin pendant quinze ans quand même / moi j'en avais pas fait mais il paraît que ça marchait * relativement * bien /
B: * (ils) marchaient bien? * /
A: et ils en ont fait en Angleterre en France un petit peu mais maintenant avec toutes les affaires la crise chez nous / ils ont commencé à fermer des centres qu'on utilisait * (. . . .) * -
B: * euh vous * faisiez ça dans quoi? dans des dans des écoles dans des? / ... (3706)

In (247), the SVC part of the underlined interrogative is inferrable information, that is, it can be inferred from the previous context, since B interprets X's first interrogative as a request for B to give X some staples and glue:

- (247) X: Bertrand est-ce que j peux avoir des agrafes et d la colle? /
B: est-ce que tu peux avoir des agrafes et d la colle? /
X: s'il te plaît /
B: j crois que là y a une agrafeuse. / (combien) pas beaucoup y en a encore. /
X: merci /
B: t en veux combien d'agrafes? / ... (2179)

In the following two extracts, the SVC parts of the interrogatives have not actually been mentioned recently in the discourse, but, in neither case is the information in any sense new. In (248), it could be said to be "textually evoked", albeit "displaced" rather than "current" (cf Brown & Yule, 1983:183), since the previous mention of it was some twenty minutes earlier in the conversation:

- (248) B: ... mais toi c'est ta tu as déjà fait quatre colos tu m'as dit? /
A: cinq / ... [TWENTY MINUTES LATER]
B: ... mais tu es arrivé à l'animation comment? / (91098)

And in (249) the information in the SVC part of the interrogative is also given, in that it is "situationally" evoked: from A's

physical presence in front of her, B could fairly safely assume that he was indeed in France at the time:

(249) [LONG SILENCE]

B: mais vous êtes en France pour combien de temps? / (211040)

We can conclude, then, that, in the present data, in occurrences of [SVQ], the informational value of the SVC part is minimal. Let us now consider whether the converse holds in the cases where the [SVQ] variant was judged not to be a possible alternative, on discoursal/pragmatic grounds.

In contrast to the predominance of more or less empty verbs in occurrences of [SVQ], in the contexts where that structure is not possible, there are no instances of the copula, and instead a wide range of lexically rich verbs: for example, dire, aller, venir, manger, se déceler, s'appliquer. In addition, there are a number of very clear cases where the SVC section of the interrogative includes new information. In (250), for example, the speaker presupposes that the English (vous) eat what is known in France as l'assiette anglaise - this was indeed news to her interlocutor:

(250) A: ... mais * on euh on les (manie) un petit peu (.) * -

B: * oui - et je voulais savoir / euh quand / à quelle heure est-ce que * vous mangez l'assiette anglaise? /

A: * euh * -

B: * (ou) * le jambon charcuterie euh froid (quoi) / (13939)

(It is worth pointing, out, however, that the NP l'assiette anglaise was not new, since it had been mentioned a few seconds earlier: what was new was the proposition as a whole.) The use of the [SVQ] structure here would produce a very odd utterance, since the speaker would be presupposing very strongly that the SVC part of the interrogative was true and that her interlocutor

had confirmed this:

(251) #Vous mangez l'assiette anglaise à quelle heure?

Similarly, in the context of (252), the use of [SVQ] seems incongruous because the SVC part cannot be taken for granted - again, it may well have been news to the interlocutor:

(252) ... on téléphone on écrit en disant "mais quand est-ce que vous venez nous inspecter? on veut vous voir euh" / alors qu'auparavant on [LAUGH] / on voulait pas ... (14780)
(253) #Vous venez nous inspecter quand?

It may be the case that the exclusion of [SVQ] from emphatic assertions (rhetorical questions) can be explained in similar terms. In 5.1.12, it was suggested that the interrogative (254) has the communicative function of an emphatic assertion, since it implicates (255):

(254) ... où est l'inconvénient? ... (22565)
(255) L'inconvénient n'est nulle part. (= Il n'y a pas d'inconvénient.)

Now if the speaker had used the [SVQ] structure, as in (256), he would have been strongly presupposing (257), at the same time as implicating (255), which contradicts (257):

(256) #L'inconvénient est où?
(257) L'inconvénient est quelque part.

A further point is that, in the [SVQ] structure, the final stress falls on the WH word (or at least a part of the WH phrase), and this would be particularly inappropriate when the interrogative is not intended to elicit an answer, but rather to convey an emphatic assertion. Notice, however, that this structure can be used for sub-topic-introducing questions:

(258) ... j'aidais ma femme donc dans ses activités. / hein parce que / l'économat ça (me) prenait quand même pas tout mon temps. / alors ça consiste en quoi? eh ben - ça consiste euh à animer un groupe d'animateurs c'est le cas de le dire ... (1124)

With regard to quoted interrogatives, the tendency to avoid [SVQ] may be due to the relative absence of discourse context - hence there is little chance of information being presupposed. More generally - but also more speculatively - one reason for [SVQ] being considered more appropriate in colloquial styles, may be that in such contexts, where interlocutors tend to be socially intimate and therefore to share a good deal of background knowledge, the "presupposition pool" is larger - ie more can be taken for granted.

It should be stressed, then, that this discoursal constraint on [SVQ] does not, on the whole, apply automatically. Rather, it is essentially a pragmatic constraint, since it leaves the speaker a degree of choice, in some, if not all, cases, as to how strongly they wish to presuppose certain information. This principle is demonstrated in the following pairs of examples, where an occurrence of [SVQ] precedes a token of another variant, used for the same communicative function and in a similar context:

- (259) ... on a fait ça jusque / j'avais quel âge? dix-neuf vingt ans quand j'ai commencé à faire des colos / ... (321)
 (260) ... moi j'étais déjà mariée à l'époque quand j'ai fait le stage - j'avais vingt euh - quel âge je pouvais avoir? / j'étais vieille déjà [LAUGH] / ... (3135)
 (261) ... c'était la taille au-dessus c'était jusqu' oh y avait quoi? - six sept des choses comme ça // ... (3335)
 (262) ... y avait des fringues aussi - on faisait des vêtements. / euh qu'est-ce qu'i y avait encore? / euh - ... (15231)
 (263) [INTERRUPTION]
 B: oui on en était où là? /
 A: oui je voulais te demander ... (4671)
 (264) [INTERRUPTION]
 B: oui oui - où j'en etais là? / [LAUGH] //
 A: La Rochelle tout ça. / (28569)

Finally, let us return to the categorical occurrence of [SVQ],

(265), which we briefly discussed in 6.3.3, to see whether we can now explain more adequately why it is impossible to front the WH phrase:

(265) ... c'était une colonie pour quoi alors ça? / ... (1947)

Perhaps the only discoursally appropriate structure here is [SVQ], because the SVC part of the interrogative has to be strongly presupposed by the speaker, since, in this context, it is a tautology:

(266) C' (=the colonie you are talking about) était une colonie.
and like almost all tautologies, it is totally uninformative (cf Lyons, 1977:417). In fact this explanation is not far removed from the notion of "l'interrogation incidente", where, it is said, the WH element is incidental to the sentence as a whole (cf Grevisse, 1986:628), and so cannot be fronted. In this example, it could be said that the SVC part of the sentence (which, being a tautology, is not "in question") is not in fact part of the question.

6.7.7 [QSV]

The non-standard structure [QSV], seems to be affected by few variable linguistic or pragmatic constraints of significance. It may be that, as with [SVQ], the structure tends to be avoided for emphatic assertions (rhetorical questions). (268) seems somewhat odd in the context of (267) (which is the only example in the corpus of an emphatic assertion where [QSV] would be a syntactic possibility):

(267) ... faire encore autre chose après euh / i faut vraiment du courage. / oui mais alors où allons-nous? hein / c'est là ou c'est la fin hein - des centres. / ... (13762)
(268) ?Où nous allons?

Otherwise, there seem to be no pragmatic restrictions on the use of [QSV], as is apparent from table 6.15.

Table 6.15 Frequencies of [QSV] for four groups
of communicative function
(categorical tokens and categorical contexts excluded)

- CFs where speaker wants answer from addressee	3/16 (18.8%)
- CFs where speaker wants answer from self	14/25 (56%)
- quoted CFs	6/8
- S.T.I. questions	5/7

The relatively high rate of use of [QSV] in self-addressed questions is to some extent due to the presence of several more or less formulaic utterances - we shall discuss these further in connection with the use of comment as the WH word.

In contrast with the very slight influence of the communicative function, the identity of the WH word has, in some cases, a very clear effect on the likelihood of this structure being selected. Once again, the scarcity of occurrences of certain WH words in our data means that we must rely to an extent on secondary sources for evidence. First, nevertheless, let us examine what evidence there is from the present corpus. Table 6.16 shows the observed frequencies of the use of the [QSV] structure with the various WH words. As before, categorical tokens and contexts are excluded, and relative frequencies are given only where the denominator is 10 or above.

Table 6.16 Frequencies of the use of [QSV]
with different WH words
(categorical tokens and contexts excluded)

qui	quel	quoi	combien	quand	comment	où	pourquoi
0/2	2/4	2/6	0/7	0/2	19/25 (76%)	5/11 (45.5%)	1/1

The only conclusion which can safely be drawn from this table is the strong tendency for [QSV] to be used when the WH word is comment. This is partly accounted for by 13 instances of certain more or less formulaic interrogatives which include comment: comment X s'appelle?, comment on appelle X?, comment ça marche? etc (but none of comment ça va?). It seems reasonable to suppose that this was equally the case in previous studies, which also reported a preference for comment to occur in [QSV] (eg Behnstedt, 1973:166-7).

The absence of interrogatives with combien can be seen as merely a consequence of the constraint on the [SVQ] structure which favours such tokens, rather than as a constraint disfavouring combien with [QSV].

One important constraint which our data fail to reveal is the strong tendency to avoid [QSV] when the WH word is quand. Indeed for some native-speakers, this is close to being a categorical constraint, with instances such as (269) and (270) being judged unacceptable (except as echoes, with contrastive stress on quand and final rising intonation):

- (269) *Quand vous venez nous inspecter? (cf 14780)
(270) *Quand vous pouvez venir? (cf 28119)

However, there is evidence that for other native-speakers this constraint is not categorical. In a large-scale acceptability test, Behnstedt (1973:253) found that slightly over 50% of the judges considered the following two sentences to be acceptable:

(271) Quand ça commence?

(272) Quand il arrive?

In view of this, it was decided in the present study to code the two previous examples, (269) and (270), as acceptable.

Although this constraint is not categorical, then, it is not very far from being so. Behnstedt (1973:166-7) reported that quand was used with [QSV] significantly less than other WH words, and that *this was particularly striking in working-class speech (where, it will be recalled, the use of [QSV] was the most widespread)*. Moreover, it is possible that Behnstedt's figures for occurrences of [QSV] with quand included cases of prepositional WH phrases, as in:

(273) Depuis quand vous êtes là? (Foulet, 1921:324)

which are certainly more frequent than instances such as (271) and (272). (There is perhaps a case for counting prepositional WH phrases separately from other instances of the same WH words.)

We may again look to perceptual strategies for a functional explanation for this constraint. Quand is used very frequently as a subordinating conjunction, and in these clauses, the order is normally quand + S + V. Moreover, quand clauses often precede the main clause. Listeners then are accustomed to interpret an utterance-initial sequence of quand + S + V as a subordinate clause. With a [QSV] interrogative introduced by quand, the first

indication that the clause is an interrogative may come at the end, with the (usually) falling intonation, followed typically by a pause (as opposed to the slight rise at the end of a subordinate clause). Perhaps, then, speakers generally avoid the [QSV] structure with quand in order not to conflict with their listeners' expectations that quand + S + V will tend to be a subordinate clause.

There are two pieces of evidence which lend support to this explanation. Firstly, the frequency of quand in contexts other than direct interrogative clauses is much greater than for several other WH words. In the present corpus, for example, the informants produced 799 occurrences of quand in such contexts, in comparison with 128 for comment, 34 for pourquoi and 16 for combien. (These figures were obtained from the OCP wordlists.) Secondly, the fact that there is a real possibility of "homonymic clash" between interrogative and subordinate clauses is suggested by the difficulties experienced in identifying tokens of [QSV] which were reported in 4.6.3.

6.7.8 [QESV]

The final variant of (WHQ) to be considered is [QESV]. Like the corresponding Yes/No interrogative structure, it can be used freely to convey any of the communicative functions we have discussed. Table 6.17 shows the frequency of occurrence of [QESV] for four groups of CFs.

Table 6.17 Frequencies of [QESV] for four groups
of communicative function

(categorical tokens and categorical contexts excluded)

- CFs where speaker wants answer from addressee	13/28 (46.4%)
- CFs where speaker wants answer from self	17/37 (46%)
- quoted CFs	11/19 (58%)
- CFs where speaker knows answer already	8/18 (44.4%)

The fourth CF group includes both sub-topic-introducing questions and emphatic assertions - all 4 instances of the latter with this structure happen to be negatively-biased and affirmative, but positively-biased negative examples can also occur:

(274) Qu'est-ce qu'elle n'a pas vu? (= Elle a tout vu.)

The slightly higher rate of use of [QESV] for quoted CFs echoes the similar tendency with [ESV] which we saw in table 6.9.

The effect of the WH word involved in the interrogative is rather more striking, however, and can be seen in table 6.18.

Table 6.18 Frequencies of the use of [QESV]
with different WH words

(excluding categorical tokens and categorical contexts)

qui	quel	que/quoi	combien	quand	comment	où	pourquoi
0/6	1/4	45/54	0/7	2/2	0/20	3/11	0/1
		(83.3%)				(27.3%)	

The very high frequency for que/quoi in table 6.18 can be attributed to the fact that [QSV] is unacceptable where the WH

word is que or quoi (unless the latter is part of a prepositional phrase), and that, of the alternative structures, [QV-CL] is on the brink of extinction, and [SVQ] is normally used in very specific pragmatic circumstances, as we have seen. Other studies have found a tendency for où and quand also to favour [QESV] (Behnstedt, 1973:169; and, to a lesser extent, Gougenheim, et al., 1964:228). Où, like que, is perhaps sometimes felt by speakers to need the phonological reinforcement provided by est-ce que (or, by que, for speakers who use the [QkSV] structure), and quand is, as we have seen, excluded from [QSV] by many speakers. The absence of tokens with comment and combien in table 6.18 can be seen as an automatic consequence of the high frequencies of these in [QSV] and [SVQ], respectively.

[QESV] is able to accommodate any WH word, but it is, in a sense, a redundant marker of interrogation, except where it serves to reinforce the WH word. In informal speech, at least, speakers tend to avoid redundancy. In the terms of the four "charges to language" identified by Slobin (1977), speakers make their language as quick and easy to produce as possible within the constraints of two of the other charges, clarity and processibility. Redundancy may be tolerated, however, where it serves the fourth charge, that of being semantically and rhetorically expressive.

In the data there are 6 occurrences of [QESV] which are with a WH word other than que, and which call for further comment. In two of them, those in (275) and (276), the presence of quand helps to explain the use of [QESV], but this is also partly due to the

quoted nature of the utterances. In addition, they are quotations from the telephone or letters (where more formal language is usual: cf Coupland, 1980):

- (275) ... on téléphone on écrit en disant "mais quand est-ce que vous venez nous inspecter? on veut vous voir euh" / ... (14780)
(276) ... j'avais remplacé un directeur qui venait d'être licencié par l'organisme. / ... / i m'ont téléphoné "quand est-ce qu vous pouvez venir?" "ben écoutez l temps d faire les valises euh j'arrive." / ... (28119)

Moreover, the quoted CFs involved are, respectively, a request for action and a request for information, and the interlocutors are clearly not intimates. Therefore it is appropriate that, in addition to using the polite vous form, they employ a polite interrogative structure, [QESV].

The 4 other instances of this structure being used without que are "live" rather than quoted utterances, and to explain the speaker's decision to use [QESV] in them, we must refer to aspects of the identity of the speaker and of the addressee, and of the relationship between the two.

6.8 The effect of speaker identity and audience design

6.8.1 A quantitative analysis

The conventional approach to investigating the effect of the speaker's identity on linguistic variability is to compare the frequencies of use of variants for individual speakers and/or for groups of speakers. We shall proceed along these lines initially, before we go on to consider a second, more interpretative, approach.

Since the number of tokens per speaker, for both (YNQ) and (WHQ), is very low, there is clearly no possibility of obtaining any meaningful patterns from individual scores. Nor is it possible to obtain scores for groups of speakers by calculating the mean of the individuals' relative frequencies. We are obliged, instead, to take all the tokens for a group of speakers and to then calculate the relative frequency for the group as a whole, as if it were a single speaker. This is, of course, something of a distortion of the data. The groups we shall consider are those established in chapter 2.

6.8.2 Yes/No interrogatives

Table 6.19 shows the observed and relative frequencies for the two variants of (YNQ) which occur in the data, [SV] and [ESV]. In addition, the total number of variable and semi-variable tokens is given, together with the average number of interrogatives produced per individual for that group of speakers.

Table 6.19 Observed and relative frequencies of the use of
variants of (YNQ) for groups of speakers according
to age, sex and social class
(categorical tokens and categorical contexts excluded)

Group	[SV]	[ESV]	N=	Mean no. of tokens per speaker
17-22 yrs	41/45 91.1%	4/43 9.3%	45	3.2
24-37 yrs	29/50 58%	24/52 46.2%	53	4.1
50-60 yrs (females, intermediate)	7/10 70%	4/11 36.4%	11	3.7
Female (17-37 yrs)	19/27 70.4%	10/29 34.5%	29	2.6
Male	51/58 87.9%	18/66 27.3%	69	4.3
Working	4/7	3/7	7	1.4
Intermediate (except older females)	33/37 89.2%	10/42 23.8%	43	4.3
Upper	15/23 65.2%	15/30 50%	30	3.3
All speakers	77/95 81.1%	32/106 30.2%	109	3.6

(Note: Tokens for 3 unclassifiable informants are omitted from the social class groupings.)

Although the number of tokens is small, and the practice of grouping speakers somewhat contentious, there are a number of points in table 6.19 which are worthy of comment. Firstly, there is a very clear difference between the middle and younger age groups, with the latter having a much stronger preference for [SV], and a very low rate of use of [ESV]. The 24-37 age group,

on the other hand, shows a lower rate of use of [SV] than any other group, and a high rate of use of [ESV]. Given the very low number of tokens for the group of older females, we can only note, with considerable caution, that their usage seems to be closer to that of the middle age group, and virtually identical to that of the other women. The male speakers seem to show a rather stronger tendency than the females to use [SV], but use [ESV] to about the same extent. Looking now at the speakers in terms of social class, we can, of course, say nothing of the working class, in view of the paucity of tokens. The upper class, on the other hand, have the highest frequency for [ESV] of all the groups in table 6.19, whereas the intermediate group shows a very high rate of use for [SV]. If we assume that [ESV] has a higher socio-stylistic value than [SV] (as most grammarians would seem to suggest), then the results in table 6.19 are quite unsurprising, with older speakers, speakers from higher social classes and (perhaps) women tending to use [ESV] rather more than the average, and younger speakers, males and the intermediate social class having a stronger than average preference for [SV].

6.8.3 WH interrogatives

Let us now move on to consider what evidence there is of social differentiation in the use of WH interrogatives. Table 6.20 gives the observed and relative frequencies for four variants of the (WHQ) variable. The figures for [Q=S V] and [QV NP] are omitted here because, as we saw in 6.5.4 and 6.5.5, the number of potential occurrences of these variants is just 4 and 3, respectively, and there is therefore no prospect of us finding

evidence of interpersonal or inter-group variation. However, these tokens are represented in the table in so far as they are potential occurrences of other variants. Consequently, N means the number of tokens of all variants of (WHQ), including [Q=S V] and [QV NP]. As before, these figures exclude both categorical tokens and the categorical contexts of semi-variable tokens.

Table 6.20 Observed and relative frequencies of the use
of variants of (WHQ) for groups of speakers
according to age, sex and social class
(categorical tokens and categorical contexts excluded)

Group	[QSV]	[SVQ]	[QESV]	[QV-CL]	N=	Mean no. of tokens/speaker
17-22 yrs	9/16 56.3%	5/19 26.3%	13/28 46.4%	0/27 0%	28	2
24-37 yrs	18/35 51.4%	12/43 27.9%	32/65 49.2%	4/65 6.2%	71	5.5
50-60 yrs (female, intermediate)	2/7	1/4	6/12 50%	2/12 16.7%	12	4
Female (17-37 yrs)	1/10 10%	6/17 35.3%	16/23 69.6%	1/23 4.4%	24	2.2
Male	26/41 63.4%	11/45 24.4%	29/70 41.4%	3/69 4.4%	75	4.7
Working	0/5	3/7	4/7	0/7	7	1.4
Intermediate (excluding older females)	13/19 68.4%	4/26 15.4%	25/44 56.8%	3/44 6.8%	48	4.8
Upper	11/23 47.8%	9/26 34.6%	16/38 42.1%	1/37 2.7%	40	4.4
All speakers	29/58 50%	18/65 27.7%	51/105 48.6%	6/104 5.8%	111	3.7

If caution was in order for the interpretation of table 6.19, then this is even more so for table 6.20, since here a similar total number of tokens is distributed among four variants, rather than the two of (YNQ). This said, let us nevertheless consider what differences there appear to be among the various groups represented. Firstly, we can note that the relative frequencies of use of the three variants [QSV], [SVQ] and [QESV] are very

close indeed for the two main age groups, but with the younger age group having a slightly higher score for the non-standard [QSV] structure. With regard to [QV-CL], it is noticeable that the younger speakers avoid this variant entirely. In addition, it is perhaps worth mentioning that all four occurrences produced by the 24-37 age group were of the expression Comment dirai-je?, and that therefore the only productive use of clitic inversion in the corpus was the two tokens produced by two of the older female speakers. Looking next at differences between the sexes, the females appear to make considerably more use of [QESV] than do the males, and, on the other hand, much less use of the non-standard [QSV] structure - though for the latter, in particular, the evidence is not at all robust, given that the potential occurrences of [QSV] for the females is only 10. There is a slight difference between females and males for use of [SVQ], but it is not clear how this should be interpreted. The same holds for the rather larger difference in rates of use of this variant by the intermediate and upper social classes. (Again, we are unable to comment at all on the speakers of the working class, due to the scarcity of tokens.) Finally, we may note that the difference between the two main social class groups in the use of [QSV] is in the direction we would expect, ie that the intermediate class uses this non-standard variant more than does the upper class, but that the same is not true for the more formal [QESV] structure, where it is again the intermediate class which has the slightly higher score. Overall, then, the picture is rather less clear for (WHQ) than it was for (YNQ) - which is to be expected, given that, for (WHQ), a similar number of variable and semi-variable tokens is spread over a larger number

of variants.

6.8.4 Further considerations

In tables 6.19 and 6.20, the mean number of tokens per speaker was given for each group, and it would appear from this information that, for both variables, informants who are working class, and/or young and/or female, produced on average fewer interrogatives than did other groups of speakers. It might be supposed that this is due to the generally subordinate social status of these groups - the right to ask questions being a reflection of power more generally.

However, these figures do not take account of the fact that individuals varied in the total amount of speech that they produced, and that certain speakers used interrogatives for a wider range of communicative functions than did others. In fact, the younger speakers produced just 3 of the 40 quoted interrogatives, and 2 of the 32 tokens which are emphatic/tentative assertions or S.T.I. questions. (These figures exclude categorical occurrences.) We interpret this as being due to the tendency of such interrogatives to occur in longer conversational turns and, generally, in the speech of informants who felt more in control of the speech event, ie the 24-37 age group and the older women. These observations raise a rather important question: do the quoted interrogatives tend to have the [ESV] and [QESV] structures because they are quoted, or because they are mostly produced by speakers who use more [ESV] and [QESV] anyway? We incline towards the former explanation, since

we have argued that there is a communicative motivation for using these structures in quotations, ie to signal to the interlocutor that no answer is expected.

In order to to discover to what extent the different groups asked questions, let us consider only the questions addressed to the fieldworker. Taking (YNQ) and (WHQ) together, Table 6.21 gives the number of such interrogatives, per 10,000 words, for each group of speakers. The figures have been calculated for the group as a whole rather than by averaging the scores of individuals, and categorical tokens have been excluded.

Table 6.21 Number of questions addressed to the fieldworker
per 10,000 words by different groups of speakers
(categorical tokens excluded)

Group	Questions to fieldworker	No. of words produced	Questions per 10,000 words
17-22 yrs	43	84,254	5.1
24-37 yrs	25	91,383	2.7
50-60 yrs (female, intermediate)	7	15,574	4.5
<hr/>			
Female (17-37 yrs)	20	57,108	3.5
Male	48	118,529	4.1
<hr/>			
Working	4	23,024	1.7
Intermediate (excluding (the older females)	31	70,853	4.4
Upper	16	66,536	2.4
<hr/>			
All speakers	75	191,211	3.9

(Note: 5 questions produced by an older female informant have been excluded, since they occurred in a post-interview conversation which was not transcribed in entirety and hence does not figure in the total of words produced.)

From table 6.21 we can see that the picture is rather different from that conveyed by tables 6.19 and 6.20. It is confirmed that working-class speakers addressed the fewest questions to the fieldworker, but we also find the upper-class speakers with a low score. Contrary to expectations, it is the younger speakers who have the highest scores, followed by the older females. Whilst the other females asked rather fewer questions than the males, the difference is not as great as might have been anticipated. It

seems likely that other factors were also at play here, including imponderables such as the rapport between informant and fieldworker, the personality of the former, and the informants' perception of the speech event and its norms.

A criticism which has sometimes been levelled at sociolinguistic research based on interviews with a fieldworker is that speakers may shift away from their most informal style and adopt a more formal style which they perceive as more appropriate to a relatively unfamiliar interlocutor, who may also be somewhat distant from them in social terms. We are, unfortunately, unable to compare the informants' speech with large quantities of their speech with more familiar interlocutors. However, we can contrast the questions which they addressed to the fieldworker with those they addressed to third persons (in all cases these were adult native-speakers of French), to see whether they tended to use different variants in each case. We shall, of course, consider only variable and semi-variable tokens in which the speaker expects an answer from the addressee. Table 6.22 shows the observed and (where possible) relative frequencies of the variants of both (YNQ) and (WHQ) addressed to the fieldworker and to third persons. Tokens for all speakers are taken together.

Table 6.22 Tokens of (YNQ) and (WHQ) addressed to fieldworker
and to adult native-speaking third persons

	[SV]	[ESV]	[QSV]	[SVQ]	[QESV]	[QV-CL]
addressed to fieldworker	57/61 93.4%	4/61 6.6%	2/10 20%	6/12 50%	9/19 47.4%	0/18 0%
addressed to third person	9/10 90%	1/10 10%	1/6	4/7	4/9	0/9

Unfortunately, the very low numbers of tokens addressed to third persons prevents any meaningful comparison with those addressed to the fieldworker. However, looking just at the latter we can observe that there is an overwhelming predominance of the less formal [SV] structure in the Yes/No interrogatives. For (WHQ) the picture is less clear: there are no instances of [QV-CL], and the rate of use of [QESV] is very close to the average for all the informants (cf table 6.20). Overall, then, it would appear that there is no evidence to suggest that informants tended to use more formal interrogative structures when addressing questions to the fieldworker.

6.9 A socio-pragmatic interpretation of the variability

6.9.1 Problems with a quantitative treatment of the data

Despite the exclusion of categorical tokens and categorical contexts from the data which we analysed in tables 6.19 and 6.20, it is clear that the variation which we interpreted there as correlating with, and (by implication) somehow resulting from the age, sex and social class of the speakers, is also constrained,

in no small measure, by pragmatic and linguistic factors. The effect of these is intertwined with that of the three speaker characteristics (and no doubt of others too). Without running a program such as VARBRUL on the data, there is no way of knowing the relative influence of the speaker characteristics and the linguistic and pragmatic constraints.

In addition, we should perhaps be a little sceptical of the apparent social differentiation in the use of (YNQ) and (WHQ), not only because of the procedure of grouping the observed frequencies from different speakers to obtain a relative frequency, and the doubtful validity of some of these even at the end of that process, but also because the patterns which seemed to emerge do not correspond very closely to the much more reliable ones which we found in the analysis of the, rather less complex, (ne) variable.

In view of these doubts about the validity of a quantitative approach when there is an insufficient quantity of data per speaker to sustain an analysis of the effect of several speaker characteristics, let us now adopt instead an interpretative approach to the variability which remained unaccounted for by linguistic and pragmatic constraints.

6.9.2 Yes/No interrogatives

It will be recalled that 6 tokens of [ESV] seemed to be unmotivated by linguistic and pragmatic factors. One of these is a self-addressed request for a reminder, but as there are only

two self-addressed variable tokens of (YNQ) in the corpus we cannot tell whether there is a preference to use [ESV] in such instances. The speaker of this utterance belonged to two of the groups which, from table 6.19, appeared to favour this variant - the 24-37 years age band and the females - but also to the intermediate social class, which used [ESV] slightly less than the average.

In the other 5 remaining occurrences of [ESV], the speaker apparently expected an answer from their interlocutor. One of these was used to convey a request for action, and is shown underlined in (277):

- (277) B: ... [ENTER X] euh * quand * -
X: * Bertrand * est-ce que je peux avoir des agrafes et de la colle? /
B: est-ce que tu peux avoir des agrafes et de la colle? /
[B STARTS LOOKING IN DRAWER OF DESK AT WHICH HE IS SITTING]
X: s'il te plaît. / (242174)

In attempting to explain the use of [ESV] here, we can note both that the interrogative is located at the beginning of a new conversation, and that the utterance is fairly lengthy. But the most important fact is surely that X recognises that the request for action (an act threatening the negative face of the addressee), which interrupts B in mid-sentence (an act which threatens the addressee's positive and negative faces: cf Brown & Levinson, 1987:67), requires a substantial expression of politeness. Moreover, it is significant that the speaker, X, was an 18 year-old working-class female and the addressee, B, was a 35 year-old upper-class male. In addition to the social distance between them, there was also, even within the seemingly egalitarian setting of a holiday centre, a substantial power differential: X was a kitchen assistant, whilst B was the camp

director. All these factors served to increase the weightiness of X's request for action as a face-threatening act, and contributed to the use of [ESV] and other markers of politeness - the indirectness of the utterance (ie its resemblance to a request for permission) and the addition of s'il te plaît, apparently as an afterthought.

The four other examples of [ESV], all addressed to the fieldworker, are: two requests for information, (278) and (279), a check on the knowledge of the addressee, (280), and a negatively-biased affirmative request for information, (281):

- (278) B: ... par exemple euh / euh demain j vais préparer
d la goulache est-ce que vous êtes encore là? /
A: (. . . si je . .) / (21747)
- (279) B: ... (vous avez fait des jeux) - est-ce que vous avez
fait aussi euh / euh - (des) feux? /
A: oui on a fait un / repas trappeur / (211058)
- (280) ... mais est-ce que vous connaissez les les / les (bus - on
en) parle souvent - (. .) / des des braderies? pas vraiment les
braderies c'est quand i y a des / des stocks - (qu'ils) donnent
euh la la vaisselle ... (10851)
- (281) B: ... / mais est-ce que vous mangez vraiment de la viande
en Angleterre? /
A: beaucoup moins qu'en France. / (21482)

It is not possible to tell from our data whether negatively-biased affirmative questions, such as (281), tend to favour [ESV]: although all seven such questions indeed have this structure, in all cases except (281) this is independently motivated by the fact that they are also tentative or emphatic assertions. In (280) and (281) the use of [ESV] can be partly attributed to the fact that the question marked a new topic or sub-topic in the conversation - hence also the presence of mais. Additionally, it is significant that in (278) and (280) the polite pronoun of address vous was used (in the other two extracts the vous was for plural reference). However, it would be

incorrect to see the presence of singular vous as a linguistic constraint favouring the use of [ESV]. Rather, both vous and [ESV] are linguistic markers of politeness, motivated by the social distance between the addressee (male, 27 years old and upper class - in French terms - though this last characteristic may not have been apparent) and the speakers: in (280) this was an 18 year-old working-class female, and, in the other three, an intermediate-class female in her fifties. Although in these cases the face-threatening acts (requests for information about the addressee or his community) are lesser impositions than that in (277) (a request to provide goods), the speakers clearly felt the need to use [ESV] to convey these acts more politely.

6.9.3 WH interrogatives

Let us now turn our attention to the 4 instances of [QESV] which we left unexplained at the end of 6.7.8. Two of these are self-addressed requests for a reminder:

(282) ... mais j'allais dans d'autres colonies Concarneau La Rochelle où est-ce que j'ai été encore? / dans de toujours avec la S.N.C.F. ... (308)

(283) ... où est-ce que je? - ah oui - je reviens / [EXITS]
(2418)

(282) was produced by an upper-class female of the 24-37 age group, (283) by a young working-class female. Since we have assumed that these utterances were addressed to the speakers themselves, it would seem difficult to argue that the use of [QESV] was motivated by politeness - there is, presumably, no need to be polite to oneself. However, speakers design their speech, not only for their addressee, but also bearing in mind the presence of "auditors" (cf Bell, 1984) - and it is possible

that in the two extracts above, the speakers selected [QESV] in order to express politeness to the auditor (the fieldworker). This is perhaps more plausible in the case of the young working-class female than for the other speaker, who, although female, was of a very similar age and social class to the fieldworker.

The same young working-class female produced one of the two other instances of [QESV] to be discussed:

- (284) X: elle est là Odile? /
B: Odile? /
X: oui /
B: elle est elle doit faire la sieste Odile. /
X: tu pourrais bien lui donner son tee-shirt? - (elle l'a trouvé.) /
B: où est-ce qu'il était? /
X: ben par là - ... (2410)

Her addressee in this case was a woman a few years older than herself, but who was an animatrice in the camp and from the intermediate social class. This, in itself, would seem to provide little motivation for using [QESV] instead of one of the other variants. Again, the presence of the fieldworker (who, it will be recalled, she addressed as vous) as an auditor may have had some effect.

The final example of [QESV] was the following request for information:

- (285) A: ... mais * on euh on les (manie) un petit peu (.) * -
B: * oui - et je voulais savoir / euh quand / à quelle heure est-ce que * vous mangez l'assiette anglaise? /
A: * euh * -
B: * (ou) * le jambon charcuterie euh froid (quoi) / (13939)

Here we can note , once again, the considerable social distance between the fieldworker addressee and the speaker, who

was one of the other older women from the intermediate social class. Although the information she is requesting is not of a personal nature (the vous is for plural reference), it may be that, in France, questions about food are regarded as a weightier face-threatening act than they are in England. Finally, it is worth noting that the 6 [ESV] and 4 [QESV] interrogatives that we have discussed in this section and the last were all produced by female informants, and that in most cases their use seemed to be motivated, to some extent at least, by a need for linguistic expression of politeness.

6.10 Variability in interrogative comment clauses

It was explained in 4.1.2 that interrogative comment clauses would be excluded from the analysis of (YNQ). In this section, however, we shall briefly consider the variation in these forms since they are a further indication of the decline in the use of clitic inversion.

As we shall not attempt to investigate the communicative functions of these tags, we cannot claim that the alternating forms are strictly equivalent. However, it is probably uncontroversial to assume that n'est-ce pas is in alternation with hein (as is suggested by, for example, the Dictionnaire du français contemporain), and that voyez-vous/vois-tu alternate with vous voyez/tu vois - we shall also assume that voyez is an ellipted version of vous voyez.

N'est-ce pas, which is often said to be the normal equivalent of

English tag questions, occurred just once in the corpus. The topic of conversation in the extract given below was sweets:

(286) X: ... cinq francs y en a assez pour moi toute seule. /
[LAUGH] mais grand-mère (qui suce) une sucette là. /
[LAUGHTER]

B: grand-mère elle a une bouche comme toi hein? //
n'est-ce pas? / (19999)

At this point in the recording, the "interview" had effectively ceased and been replaced by a multi-participant conversation consisting largely of light-hearted anecdotes and gentle banter: B was not in fact the grandmother of anyone present, but had been given that nickname. She was one of the three female informants in their fifties. There was no indication, however, that she used n'est-ce pas? regularly, and the occurrence in the extract above was intended humorously and said emphatically, with the schwa of ce pronounced.

In contrast with this lone occurrence of n'est-ce pas?, hein? was produced 1,261 times by the informants. The rate of use per speaker varied from 30.7 per 1,000 words to 0.2, with the mean being 7.2.

Several observers have noted the decline of n'est-ce pas, especially in the speech of younger people: Doppagne (1966), Morin (1973), Maury (1973). Biggs & Dalwood (1976a:98) have also pointed out that, when it is still used, n'est-ce pas often undergoes massive elision, to the point of being pronounced as [spa], [pa] or even [a]. According to Maury (1973:149), in the Français Fondamental corpus, there were 447 occurrences of hein, and 206 of n'est-ce pas. Although the present corpus is clearly not directly comparable with that of the Français Fondamental

project, all the signs are that n'est-ce pas has undergone a recent and massive decline in use.

The second set of comment clauses, which all translate as "(do) you see", are not infrequently accompanied by other discourse markers, as in the following extract:

(287) ... la tente on l'a remontée aussi à la fête du village
l'autre jour enfin tu vois? des choses comme ça quoi. / (21498)

In comparison with hein, these forms are relatively infrequent in the corpus: thirteen informants produced no examples of them at all, and nine others less than 10 each. There were, nevertheless, 236 instances of the [SV] forms (ie tu vois, vous voyez, voyez) and 38 of voyez-vous. All the latter forms were produced by two of the women in their fifties.

Although no account has been taken here of any functional differences among these forms, it seems reasonable to conclude that this is further evidence of the decline of forms involving clitic inversion (other than est-ce que, of course) in the informal speech of French people born since 1945.

6.11 Concluding remarks

In this chapter, we have seen how a very wide range of factors - linguistic, discoursal and pragmatic - constrain the variation in both Yes/No and WH interrogatives. Moreover, with only a modest number of tokens per speaker, some evidence has emerged of social differentiation, especially with regard to (WHQ).

SUMMARY AND CONCLUSION

Drawing on data from the corpus of spoken French described in chapter 2, this study has attempted to apply to negation and interrogation the principles concerning the analysis of grammatical variability which were discussed in chapter 1. The variationist analysis of these two areas of grammar has posed contrasting methodological problems, and revealed very different sets of constraining factors. Various problems and theoretical issues have been discussed as they emerged at different points in this study, and this material will not be repeated here.

The most important linguistic constraint on the (ne) variable was found to be the sequence of items in which the variant is embedded. Generally, the more frequent a given sequence, the greater the tendency to omit the negative particle. Thus, a sequence involving a pronominal subject and a high-frequency verb (eg ce (n') est, je (ne) sais) favours the absence of ne, whereas a sequence which begins with a NP subject disfavours this quite strongly. The effect of a phonological constraint, by which the negative particle tends to be retained intervocalically, seems to account for the relatively high rate of ne retention found in sequences such as qui (n') ont and ça (n') a.

With regard to interpersonal variation, it was found that most speakers aged between 17 and 23 years had extremely low rates of use of the negative particle, whilst speakers aged over 23 had higher ne retention rates, but nevertheless omitted ne more frequently than they used it. A small amount of evidence was

found of speakers' ability to modify substantially their use of this variable according to the speech situation. The findings of this study, taken together with previous research, would suggest that the loss of the negative particle is complete, or very nearly so, in the vernacular of northern France, but that many speakers (especially those from the intermediate and upper classes) increase their use of ne once they are in their mid-twenties.

As a necessary preliminary to the variationist analysis of interrogatives, a taxonomy of communicative functions was set out in chapter 5, and it proved possible to classify nearly all the data from the corpus in terms of the categories established there. The Yes/No interrogatives in the corpus were used for a much wider range of functions than were the WH interrogatives.

A significant proportion of the tokens of the (YNQ) variable were found to be categorically constrained by various pragmatic factors. Of the variable tokens, the marked [ESV] structure was used only rarely when the speaker expected an answer from the addressee. On the few occasions that it was used, it seemed to be motivated by socio-pragmatic considerations, especially politeness. Clitic inversion did not occur at all as a variant of (YNQ). There was only marginal evidence of social differentiation in the use of this variable.

The more complex (WHQ) variable was analysed in terms of six variants, not all of which were acceptable or equivalent in all contexts. In contrast with (YNQ), very few tokens of this

variable offered the speaker no choice of structure. The two least frequent structures, [Q=S V] and [QV NP], were constrained principally by linguistic factors, though in the case of [QV NP], discursual considerations also loomed large. The choice between [SVQ] and a WH-fronted variant proved to be conditioned by a strong discursual constraint: the less informative the subject and the verb, the higher the probability of [SVQ] being selected. Whereas clitic inversion was completely absent from the (YNQ) data, a small number of instances of [QV-CL] were produced by the informants, but only one or two of them could be said to involve the productive use of the structure. [QESV] was the variant least constrained by linguistic factors. Whereas some previous surveys, carried out on much larger and socially more diverse corpora, have revealed very significant social differentiation with regard to (WHQ), the amount found in this study was modest, and was located primarily in the choice between [QSV] and [QESV]. In contrast with the (ne) variable, it was sex that proved to be the most important extra-linguistic variable affecting (WHQ).

The shortcomings of the present study are clear. Five of them may be highlighted here, in the hope that future research may fill some of the gaps which have been left. Firstly, the findings of this research with regard to interpersonal and stylistic variation are limited, due to the nature of the corpus. A corpus sampling a wider range of speakers and styles (but especially of informal styles) would enable us to discover whether (ne) is a case of age-grading, as is suggested here, or whether it is still in the process of change. Secondly, the phonological constraint on (ne), referred to above, was treated only schematically in

chapter 3, and clearly merits further investigation. Thirdly, the taxonomy of communicative functions which was sketched out in chapter 5 could usefully be tested against a richer body of data. Fourthly, although the judgements of acceptability and equivalence on examples given in chapter 6 have been checked with native-speakers, it has not been possible to do this with most of the others (which are given in appendices 2 and 3). It would clearly be desirable to verify all the judgements with a large number of native-speakers. Finally, many of the conclusions arrived at with regard to interrogatives have been tentative, and would be more clear-cut if data were available in sufficient quantities for valid and reliable statistical testing. It would be particularly interesting to analyse the variation in interrogatives with D Sankoff's GOLDVARB program (which has recently become available with documentation for public use), as this is designed to identify the relative strengths of arrays of constraints ^{such as those} discussed in chapter 6. As we have seen, however, it is difficult to obtain such large quantities, given various methodological and data collection problems, and the fact that certain grammatical variants are often poorly represented in even large corpora.

It was argued in chapter 1 that the sociolinguistic variable is an important concept within sociolinguistic theory, notably because it serves to delimit and differentiate speech communities. In this study (as in many others), the variable has also proved to be a powerful analytical procedure with which to discover the numerous linguistic, discoursal and pragmatic factors which constrain and motivate speakers' choices in

variable areas of grammar. In view of the consistent patterns which have been revealed in many studies, there can surely be no doubt that these constraints, most of which can only be discovered by variationist analysis, form part of speakers' knowledge of their language, and are therefore, by any definition, a central concern of linguistic research.

APPENDIX 1 Features matrix for the CFs realised as interrogatives in the corpus (followed by specification of the features, and key to abbreviations).

	A	B	C	D	E	F	G	G'	H	H'	I	I'	J	L	M	N	P	Q	R	S	T
inf	+	-	-	+	+	+	-	-	-	-	-	+	/	/	/	/	/	/	/	/	/
" inf	-	-	-	+	+	+	-	-	-	-	-	+	/	/	/	/	/	/	/	/	/
s inf	+	+	-	+	+	+	-	-	-	-	-	+	/	/	/	/	/	/	/	/	/
" s inf	-	+	-	+	+	+	-	-	-	-	-	+	/	/	/	/	/	/	/	/	/
e inf	-	-	+	+	+	+	-	-	-	-	-	+	/	/	/	/	/	/	/	/	/
inf aff oui	+	-	-	+	+	+	+	-	-	-	-	+	/	/	/	/	/	/	/	/	/
" inf aff oui	-	-	-	+	+	+	+	-	-	-	-	+	/	/	/	/	/	/	/	/	/
e inf aff oui	-	-	+	+	+	+	+	-	-	-	-	+	/	/	/	/	/	/	/	/	/
inf aff non	+	-	-	+	+	+	-	+	+	-	-	+	/	/	/	/	/	/	/	/	/
inf neg non	+	-	-	+	+	+	-	+	-	-	-	+	/	/	/	/	/	/	/	/	/
e inf neg non	-	-	+	+	+	+	-	+	-	-	-	+	/	/	/	/	/	/	/	/	/
inf neg si	+	-	-	+	+	+	+	-	-	+	-	+	/	/	/	/	/	/	/	/	/
op	+	-	-	+	+	+	-	-	-	-	-	-	/	/	/	/	/	/	/	/	/
" op	-	-	-	+	+	+	-	-	-	-	-	-	/	/	/	/	/	/	/	/	/
e op aff oui	-	-	+	+	+	+	+	-	-	-	-	-	/	/	/	/	/	/	/	/	/
check	+	-	-	+	+	+	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-
check aff oui	+	-	-	+	+	+	+	-	-	-	-	+	+	-	-	-	-	-	-	-	-
check neg non	+	-	-	+	+	+	-	+	-	-	-	+	+	-	-	-	-	-	-	-	-
clar	+	-	-	+	+	+	-	-	-	-	-	+	-	+	-	-	-	-	-	-	-
" clar	-	-	-	+	+	+	-	-	-	-	-	+	-	+	-	-	-	-	-	-	-
e clar	-	-	+	+	+	+	-	-	-	-	-	+	-	+	-	-	-	-	-	-	-
clar aff oui	+	-	-	+	+	+	+	-	-	-	-	+	-	+	-	-	-	-	-	-	-
clar neg non	+	-	-	+	+	+	-	+	-	-	-	+	-	+	-	-	-	-	-	-	-
e clar neg non	-	-	+	+	+	+	-	+	-	-	-	+	-	+	-	-	-	-	-	-	-
" adv	-	-	-	+	+	+	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-
" s adv	-	+	-	+	+	+	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-
" s adv neg non	-	+	-	+	+	+	-	+	-	-	-	-	-	-	+	-	-	-	-	-	-
rem	+	-	-	+	+	+	-	-	-	-	-	+	+	/	/	/	/	/	/	/	/
s remc	+	+	-	+	+	+	-	-	-	-	-	+	+	/	/	/	/	/	/	/	/
s remf	+	+	-	+	+	+	-	-	-	-	-	+	+	/	/	/	/	/	/	/	/
s remc aff oui	+	+	-	+	+	+	+	-	-	-	-	+	+	/	/	/	/	/	/	/	/
sug ex	+	-	-	+	+	+	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-
sug ex aff oui	+	-	-	+	+	+	+	-	-	-	-	-	-	-	-	+	-	-	-	-	-
sug inc	+	-	-	+	+	+	-	-	-	-	-	+	-	-	-	-	-	+	-	-	-
" sug inc	-	-	-	+	+	+	-	-	-	-	-	+	-	-	-	-	-	+	-	-	-
" sug inc neg si	-	-	-	+	+	+	+	-	-	+	-	+	-	-	-	-	-	+	-	-	-
off	+	-	-	+	+	+	-	-	-	-	-	+	-	-	-	-	-	-	+	-	-
" off	-	-	-	+	+	+	-	-	-	-	-	+	-	-	-	-	-	-	+	-	-
perm	+	-	-	+	+	+	-	-	-	-	-	+	-	-	-	-	-	-	-	+	-
act	+	-	-	+	+	+	-	-	-	-	-	+	-	-	-	-	-	-	+	-	-
" act	-	-	-	+	+	+	-	-	-	-	-	+	-	-	-	-	-	-	+	-	-
tentass	+	-	-	+	-	-	-	-	-	-	-	-	-	/	/	/	/	/	/	/	/
tentass aff oui	+	-	-	+	-	-	+	-	-	-	-	-	-	/	/	/	/	/	/	/	/
tentass aff non	+	-	-	+	-	-	-	+	+	-	-	-	-	/	/	/	/	/	/	/	/
tentass neg non	+	-	-	+	-	-	-	+	-	-	-	-	-	/	/	/	/	/	/	/	/
tentass neg si	+	-	-	+	-	-	+	-	-	+	-	-	-	/	/	/	/	/	/	/	/
emphass aff non	+	-	-	-	+	-	-	+	+	-	-	-	-	/	/	/	/	/	/	/	/
" emphass aff non	-	-	-	-	+	-	-	+	+	-	-	-	-	/	/	/	/	/	/	/	/
emphass neg si	+	-	-	-	+	-	+	-	-	+	-	-	-	/	/	/	/	/	/	/	/
stiq	+	-	-	-	-	-	-	-	-	-	+	-	+	/	/	/	/	/	/	/	/
prea	+	-	-	-	-	-	-	-	-	-	+	-	+	+	-	-	-	-	-	-	-
" prea	-	-	-	-	-	-	-	-	-	-	+	-	+	+	-	-	-	-	-	-	-
posta	+	-	-	-	+	-	-	-	-	-	+	-	+	+	-	-	-	-	-	-	-
" posta	-	-	-	-	+	-	-	-	-	-	+	-	+	+	-	-	-	-	-	-	-
e filler	-	-	+	+	+	+	-	-	-	-	-	+	/	/	/	/	/	/	/	/	/

A : speaker is simultaneously original author of the words of the utterance
 B : author is also addressee of the utterance
 C : utterance is a partial or complete echo of a recent utterance by another speaker
 D : author does not know (for certain) q / whether p is true
 E : au believes ad knows (or may know) q / whether p is true
 F : au wants to know q / whether p is true
 G : au assumes (or believes ad assumes) that p is true
 G' : au assumes (or believes ad assumes) that p is not true
 H : au has had (or believes ad has had) a disposition to believe that p is true
 H' : au has had (or believes ad has had) a disposition to believe that p is not true
 I : au believes q / p is particularly significant
 I' : au has known q / p (but has forgotten)
 J : p is presented as an objective fact
 L : p = 'you have cognizance of X'
 M : p = 'you mean X'
 N : p = 'I should do X'
 P : p = 'you (+/- others, but excluding me) could do X'
 Q : p = 'I and you (+/- others) will do X'
 R : p = 'you will do X'
 S : p = 'you want X (which I am able and willing to help provide or fulfil)'
 T : p = 'I have your consent to do X'
 (where p is a proposition, q is the missing variable in a WH question, and X is any variable)

/ : + or - (ie some realisations are +, some are -)
 " : quoted question
 s : self-addressed question
 e : echo question
 aff : affirmative (conducive/loaded question)
 neg : negative (conducive question)
 oui / si : (conducive question) expecting affirmative answer
 non : (conducive question) expecting negative answer
 inf : request for information
 op : request for opinion
 check : check on knowledge of addressee
 clar : request for clarification
 adv : request for advice
 rem : request for a reminder
 remc : request for a reminder of content
 remf : request for a reminder of form
 sug ex : suggestion for action (excluding the author)
 sug inc : suggestion for action (including the author)
 off : offer
 perm : request for permission
 act : request for action
 tentass : tentative assertion
 emphass : emphatic assertion
 stiq : sub-topic-introducing question
 prea : pre-announcement
 posta : post-announcement
 filler : filler question
 au : author
 ad : addressee

APPENDIX 2 Coding of tokens of (YNQ)

[" s inf]	SV 0	ESV 1	VCL 0	il ya
[inf]a	SV 1	ESV 0	VCL 0	tu etais
[inf aff oui]a	SV 1	ESV -	VCL -	ce etait !
[s rem c]	SV 0	ESV 1	VCL 0	ce est
[inf aff oui]a	SV 1	ESV -	VCL -	tu reves !
[" posta]	SV 1	ESV -	VCL -	tu as vu !
[e fill]x	SV -	ESV 1	VCL -	tu peux !
[inf]x	SV 1	ESV 0	VCL 0	ce est
[" sug i neg si]	SV 1	ESV 0	VCL 0	on pourrait
[posta]	SV 1	ESV -	VCL -	tu vois !
[tentass aff non]	SV -	ESV 1	VCL 0	ce est
[tentass neg si]	SV -	ESV 1	VCL 0	on vit
[s rem c]	SV 1	ESV 0	VCL 0	ce etait
[s rem c aff oui]	SV 1	ESV -	VCL -	ce etait !
[" off]	SV 0	ESV 1	VCL 0	tu veux
[inf]a	SV 1	ESV 0	VCL 0	ca se fait
[tentass aff oui]	SV -	ESV 1	VCL 0	ce est
[" op]	SV 0	ESV 1	VCL 0	ce est
[inf aff oui]a	SV 1	ESV -	VCL -	tu trouves !
[inf aff oui]a	SV 1	ESV -	VCL -	ils sont !
[" off]	SV 0	ESV 1	VCL 0	tu veux
[inf neg non]a	SV 1	ESV -	VCL -	vous avez !
[inf]a	SV 1	ESV 0	VCL 0	(ils marchaient
[check neg non]b	SV 1	ESV -	VCL -	tu la connais !
[check]a	SV 1	ESV 0	VCL 0	tu connais
[check neg non]a	SV 1	ESV 0?VCL 0?		tu sais
[inf neg non]a	SV 1	ESV 0?VCL 0?		ca existe
[inf aff oui]x	SV 1	ESV -	VCL -	il est !
[check]a	SV 1	ESV 0	VCL 0	tu as vu
[check]a	SV 1	ESV 0	VCL 0	tu as vu
[act]x	SV 1	ESV 0	VCL 0	tu peux
[check]a	SV 1	ESV 0	VCL 0	tu sais
[check]a	SV 1	ESV 0	VCL 0	tu les as vus
[inf]a	SV 1	ESV 0	VCL 0	ca va
[inf]a	SV 1	ESV 0	VCL 0	ca va servir
[inf neg si]a	SV 1	ESV 0	VCL 0	ce est
[inf aff oui]a	SV 1	ESV -	VCL -	tu enregistres !
[inf]a	SV 1	ESV 0	VCL 0	tu veux
[inf aff oui]x	SV 1	ESV -	VCL -	il fallait !
[e op aff oui]x	! SV 1	ESV -	VCL -	np ont ete !
[inf neg non]x	SV 1	ESV 0?VCL 0?		ils ont demande
[inf aff oui]a	SV 1	ESV -	VCL -	tu as deja fait !
[inf aff oui]a	SV 1	ESV -	VCL -	tu vis !
[inf aff oui]a	SV 1	ESV -	VCL -	tu vis !
[inf aff oui]a	SV 1	ESV -	VCL -	tu vis !
[clar]a	SV 1	ESV 0	VCL 0	tu parles
[check]a	SV 1	ESV 0	VCL 0	tu as vu
[inf]a	SV 1	ESV 0	VCL 0	tu as ete
[inf]a	SV 1	ESV 0	VCL 0	tu as fait
[check]a	SV 0	ESV 1	VCL 0	vous connaissez
[inf]a	SV 1	ESV 0	VCL 0	vous avez deja ete
[tentass aff non]	SV -	ESV 1	VCL 0	on l' aura vrmnt aide
[" s adv] alt	SV 1	ESV 0	VCL 0	je le fais
[" s adv neg non] alt	! SV 1	ESV -	VCL -	je le fais !
[tentass] alt	SV -	ESV 1	VCL 0	je ai bien fait
[tentass neg non] alt	! SV -	ESV 1	VCL -	je ai bien fait !

[e inf aff oui]a	!	SV 1 ESV - VCL -	ca marche	!
[act]a	!	SV 1 ESV - VCL -	tu coupes	!
[inf aff oui]a	!	SV 1 ESV - VCL -	ce est	!
[inf aff oui]a	!	SV 1 ESV - VCL -	on peut	!
[tentass aff oui]		SV - ESV 1 VCL 0	je fais	
[inf aff oui]a	!	SV 1 ESV - VCL -	ils se sont sentis	!
[inf]a		SV 1 ESV 0 VCL 0	vous dites	
[inf neg non]a	!	SV 1 ESV - VCL -	ce est	!
[" s adv]		SV 0 ESV 1 VCL -	je continue	
[" s adv]		SV 0 ESV 1 VCL -	je arrete	
[tentass aff non]		SV - ESV 1 VCL 0	ce est	
[tentass]		SV - ESV 1 VCL 0	ce est	
?[" (s) adv] alt		SV 0 ESV 1 VCL 0	on ecoute	alt
?[" (s) adv neg non] !		SV - ESV 1 VCL -	on ecoute	! alt
?[" (s) adv] alt		SV 0 ESV 1 VCL 0	on manipule	alt
?[" (s) adv neg non] !		SV - ESV 1 VCL -	on manipule	! alt
[e inf neg non]a	!	SV 1 ESV - VCL -	tu as remarque	!
[clar aff oui]a	!	SV 1 ESV - VCL -	tu veux dire	!
[inf]a		SV 1 ESV 0 VCL 0	ca se vend	
[e clar neg non]a		SV 1 ESV - VCL -	ils sont arrives	!
[inf aff oui]a		SV 1 ESV - VCL -	tu animes	!
[inf neg non]a		SV 1 ESV - VCL -	tu as pu	!
[inf]a		SV 1 ESV 0 VCL 0	ce est	
[inf neg non]a		SV 1 ESV 0?VCL -	tu peux	
[check aff oui]a		SV 1 ESV - VCL -	tu connais	!
[check]a		SV 1 ESV 0 VCL 0	tu connais	
[off]b		SV 1 ESV 0 VCL 0	vous voulez	
[inf]b		SV 1 ESV 0 VCL 0	elle est	
[act]b		SV 1 ESV 0 VCL 0	tu pourrais	
[inf]ab (greet)		SV 1 ESV - VCL -	ca va	!
[e inf aff oui]a		SV 1 ESV - VCL -	tu en as	!
[" act]		SV 1 ESV - VCL -	tu me mets	!
[check]a		SV 1 ESV 0 VCL 0	tu vois	
[check]a		SV 1 ESV 0 VCL 0	tu les connais	
[check aff oui]a		SV 1 ESV - VCL -	tu connais	!
[" sug i]		SV 0 ESV 1 VCL 0	on peut	
[check]a		SV 1 ESV 0 VCL 0	tu vois	
[inf aff oui]ab		SV 1 ESV - VCL -	vous etes	!
[act]ab		SV 1 ESV - VCL -	vous coupez	!
[check neg non]x		SV 1 ESV - VCL -	(tu sais	!
[check neg non]a		SV 1 ESV - VCL -	vous connaissez	!
[check]b		SV 1 ESV 0 VCL 0	(tu t' en rappelles	
[check]a		SV 1 ESV 0 VCL 0	tu as entendu	
[check aff oui]a		SV 1 ESV - VCL -	vous connaissez	!
[inf]a		SV 1 ESV 0 VCL 0	vous avez visite	
[check aff oui]a		SV 1 ESV - VCL -	vous le connaissez	!
[check]a		SV 1 ESV 0 VCL 0	vous connaissez	
[inf aff oui]a		SV 1 ESV - VCL -	vous aviez	!
[clar aff oui]a		SV 1 ESV - VCL -	vous voulez dire	!
[clar neg non]a		SV 1 ESV - VCL -	ce est	!
[check aff oui]a		SV 1 ESV - VCL -	vous connaissez	!
[inf aff oui]a		SV 1 ESV - VCL -	ce est	!
[inf aff non]a		SV 0 ESV 1 VCL 0	vous mangez	
[check aff oui]a		SV 1 ESV - VCL -	vous connaissez	!
[check]a		SV 1 ESV 0 VCL 0	vous avez entendu	
[sug x aff oui]a		SV 1 ESV - VCL -	vous pourriez	!
[inf]a		SV 0 ESV 1 VCL 0	vous etes	
[inf neg non]a		SV 1 ESV - VCL -	vous etes	!

[inf]a
 [inf aff oui]a
 [inf]a
 [sug x]a
 [inf neg non]a
 [check aff oui]a
 [inf]a
 [s rem c aff oui]
 [tentass]
 [tentass neg non]
 [check]a
 [check]a
 [check neg non]a
 [check]a
 [inf neg si]a
 [inf neg si]a
 [check]a
 [inf]b
 [act]b
 [check]a
 [inf aff oui]a
 [inf neg non]a
 [inf aff oui]a
 [prea]
 [inf]a
 [inf]a
 [inf]a
 [inf]a
 [inf]a
 [inf aff oui]a
 [inf]a
 [op]a
 [posta]
 [inf]a
 [inf]a
 [posta]
 [prea]
 [inf neg non]a
 [inf neg non]a
 [check]a
 [check aff oui]a
 [check]a
 [check]a
 [" off]
 [" off]
 [" off]
 [" off]
 [" off]
 ?[" inf][" adv]
 [tentass aff non]
 [tentass aff non]
 [emphass aff non]
 [stiq]
 [inf neg non]x
 [" inf]
 [" prea]
 [" off]
 [clar aff oui]a

SV 1 ESV 0 VCL 0 vous avez
 SV 1 ESV - VCL - ce est !
 SV 1 ESV 0 VCL 0 vous avez ete
 SV 1 ESV - VCL - vous passez !
 SV 1 ESV - VCL - vous etes alle !
 SV 1 ESV - VCL - vous connaissez !
 SV 0 ESV 1 VCL 0 vous avez fait
 SV 1 ESV - VCL - ce est !
 SV - ESV 1 VCL 0 ce est alt
 SV - ESV 1 VCL - ce est ! alt
 SV 1 ESV 0 VCL 0 tu connais
 SV 1 ESV 0 VCL 0 tu vois
 SV 1 ESV - VCL - tu vois !
 SV 1 ESV 0?VCL 0 tu vois
 SV 1 ESV 0 VCL 0 ce est
 SV 1 ESV 0 VCL 0 ce est
 SV 1 ESV 0 VCL 0 vous connaissez
 SV 1 ESV 0 VCL 0 ce est
 SV 0 ESV 1 VCL 0 je peux
 SV 1 ESV 0 VCL 0 tu vois
 SV 1 ESV - VCL - ce est !
 SV 1 ESV - VCL - ils t' ont laisse !
 SV 1 ESV - VCL - tu veux !
 SV 1 ESV - VCL 0 tu sais
 SV 1 ESV 0 VCL 0 il est
 SV 1 ESV 0 VCL 0 ca a fait
 SV 1 ESV 0 VCL 0 tu as connu
 SV 1 ESV 0 VCL 0 tu etais
 SV 1 ESV 0 VCL 0 tu as connu
 SV 1 ESV - VCL - ce est +ca !
 SV 1 ESV 0 VCL 0 il ya
 SV 1 ESV 0 VCL 0 tu trouves
 SV 1 ESV - VCL - tu vois !
 SV 1 ESV 0 VCL 0 (il ya
 SV 1 ESV 0 VCL 0 (il yavait
 SV 1 ESV - VCL - tu vois !
 SV 1 ESV - VCL 0 tu sais
 SV 1 ESV 0?VCL 0?(il ya
 SV 1 ESV 0?VCL 0?(il ya
 SV 1 ESV 0 VCL 0 tu vois
 SV 1 ESV - VCL - tu connais !
 SV 1 ESV 0 VCL 0 tu vois
 SV 1 ESV 0 VCL 0 tu vois
 SV 0 ESV 1 VCL 0 vous voulez
 SV 1 ESV 0 VCL 0 vous voulez
 SV 1 ESV 0 VCL 0 vous voulez
 SV 1 ESV 0 VCL 0 vous voulez
 SV 1 ESV 0 VCL 0 vous voulez
 SV 0 ESV 1 VCL 0 il remonte
 SV - ESV 1 VCL 0 ca sera
 SV - ESV 1 VCL 0 ca sera
 SV - ESV 1 VCL 0 (il ya
 SV - ESV 1 VCL 0 il ya
 SV 1 ESV - VCL - tu as !
 SV 1 ESV 0 VCL 0 np veulent
 SV 1 ESV - VCL 0 vous savez
 SV 0 ESV 1 VCL 0 tu viens
 SV 1 ESV - VCL - ce est !

[check neg non]a	SV 1	ESV 0?VCL 0?	tu connais
[check]a	SV 1	ESV 0 VCL 0	tu vois
[check neg non]a	SV 1	ESV 0?VCL 0?	tu vois
[inf]x	SV 1	ESV 0 VCL 0	tu les as fait
[inf neg non]x	SV 1	ESV - VCL -	ils ont goute !
[e inf neg non]x	SV 1	ESV - VCL -	il est !
[clar aff oui]a	SV 1	ESV - VCL -	ce est !
[" inf aff oui]	SV -	ESV 1 VCL 0	ils sont arrives

Selective key:

! : categorical tokens

alt : part of an alternative question

]a,x,b,ab : interrogatives addressed to, respectively, fieldworker, 3rd person, interviewee (by 3rd person), fieldworker plus interviewee (by 3rd person).

1: structure used

0: structure acceptable and equivalent to that used

-: structure unacceptable or non-equivalent to that used

0? judgements by native-speaker informants which are inconsistent with analyst's interpretation of the communicative function.

(Corpus reference numbers have been omitted.)

APPENDIX 3 Coding of tokens of (WHQ)

[s rem c]	QSV- QESV1 SVQ0 QVC- QVN- Q=SV- je oublie que
[s rem c]	QSV- QESV1 SVQ0 QVC0 QVN- Q=SV- il ya que
[inf]a	QSV- QESV1 SVQ0 QVC0 QVN- Q=SV- tu faisais que
[inf]a	QSV0 QESV0 SVQ1 QVC0 QVN- Q=SV- ca fait cbn detps
[inf]x	QSV0 QESV0 SVQ1 QVC0 QVN- Q=SV- tu en veux cbn
[stiq]	QSV- QESV1 SVQ0 QVC0 QVN- Q=SV- l on fait que
[" emph aff non]	!QSV- QESV1 SVQ- QVC- QVN- Q=SV- S=Q que va sepasser
[stiq]	QSV1 QESV0 SVQ0 QVC0 QVN- Q=SV- elle fonctionne cmt
[stiq]	QSV1 QESV0 SVQ0 QVC0 QVN- Q=SV- ca passe quoi (par)
[stiq]	QSV- QESV1 SVQ0 QVC0 QVN- Q=SV- on fait que
[" s adv]	QSV- QESV1 SVQ0 QVC0 QVN- Q=SV- je vais faire que
[emph aff non]	QSV- QESV1 SVQ- QVC0 QVN- Q=SV- tu veux que
[s rem f]	QSV1 QESV0 SVQ- QVC0 QVN- Q=SV- elles s'appellt cmt
[" inf]	QSV1 QESV0 SVQ0 QVC0 QVN- Q=SV- ca repose quoi (sur)
[stiq]	QSV0 QESV0 SVQ0 QVC0 QVN1 Q=SV- np se faisait cmt
[s rem f]	QSV0 QESV- SVQ- QVC1 QVN- Q=SV- je dirai cmt
[inf]x	QSV- QESV1 SVQ- QVC0 QVN- Q=SV- tu veux que
[stiq]	QSV1 QESV0 SVQ0 QVC0 QVN- Q=SV- ils vont coucher ou
[" off]	QSV1 QESV0 SVQ0 QVC0 QVN- Q=SV- tu veux cmt
[" inf]	QSV- QESV1 SVQ0 QVC0 QVN- Q=SV- il faut que
[" adv]	QSV1 QESV0 SVQ0 QVC0 QVN- Q=SV- on va le faire cmt
[" inf]	QSV- QESV1 SVQ0 QVC0 QVN- Q=SV- ils mangent que
[s rem c]	QSV0 QESV0 SVQ1 QVC0 QVN- Q=SV- je avais quel age
[s rem c]	QSV1 QESV0 SVQ0 QVC0 QVN- Q=SV- je pouvais quel age
[s rem c]	QSV- QESV0 SVQ1 QVC0 QVN- Q=SV- il yavait quoi
[rem]a	QSV- QESV1 SVQ- QVC0 QVN- Q=SV- tu m'avais dmde que
[s inf]	QSV0 QESV0 SVQ1 QVC0 QVN- Q=SV- ce est qui
[inf]a	QSV0 QESV0 SVQ1 QVC0 QVN- Q=SV- vous faisiez quoi
[s rem f]	QSV1 QESV- SVQ- QVC0 QVN- Q=SV- je te dirai cmt
[s rem f]	QSV1 QESV0 SVQ- QVC0 QVN- Q=SV- on appelle cmt
[s rem f]	QSV1 QESV0 SVQ- QVC0 QVN- Q=SV- elle s'appelle cmt
[stiq]	QSV- QESV0 SVQ- QVC- QVN- Q=SV1 S=Q qui prend
[clar]a	QSV- QESV1 SVQ- QVC0 QVN- Q=SV- tu appelles que
[clar]a	QSV- QESV1 SVQ- QVC0 QVN- Q=SV- tu appelles que
[" off]	QSV- QESV1 SVQ0 QVC0 QVN- Q=SV- vous voulez que
[rem]a	QSV0 QESV0 SVQ1 QVC0 QVN- Q=SV- on en etait ou
[s rem f]	QSV1 QESV0 SVQ- QVC0 QVN- Q=SV- on appelle cmt
[inf]a	QSV- QESV1 SVQ0 QVC0 QVN- Q=SV- tu veux que
[inf]a	QSV- QESV1 SVQ0 QVC0 QVN- Q=SV- tu voulais que
[s rem f]	QSV1 QESV0 SVQ- QVC0 QVN- Q=SV- (on pourrait cmt
[s rem c]	QSV- QESV1 SVQ- QVC0 QVN- Q=SV- on trouvait que
[s rem c]	QSV- QESV1 SVQ- QVC0 QVN- Q=SV- je ai fait que
[e fill]	QSV1 QESV0 SVQ- QVC0 QVN- Q=SV- ca pose quel
[e clar]a	! QSV- QESV1 SVQ- QVC- QVN- Q=SV- je pense que
[inf]x	QSV- QESV1 SVQ0 QVC0 QVN- Q=SV- ce est que
[inf]a	QSV0 QESV0 SVQ1 QVC0 QVN- Q=SV- tu es arrive cmt
[inf]x	QSV0 QESV0 SVQ1 QVC0 QVN- Q=SV- ce est cbn
[s rem c]	QSV- QESV1 SVQ0 QVC0 QVN- Q=SV- 0 noverb que
[stiq]	QSV0 QESV0 SVQ1 QVC0 QVN- Q=SV- ca consiste quoi
[s rem f]	QSV0 QESV- SVQ- QVC1 QVN- Q=SV- je dirai cmt
[s rem c]	QSV- QESV0 SVQ1 QVC0 QVN- Q=SV- il yavait quoi
[s rem f]	QSV0 QESV- SVQ- QVC1 QVN- Q=SV- je dirai cmt
[inf]x	QSV0 QESV0 SVQ1 QVC0 QVN- Q=SV- vous etes cbn
[emph aff non]	QSV- QESV1 SVQ- QVC0 QVN- Q=SV- vous voulez que
[emph aff non]	QSV- QESV1 SVQ- QVC0 QVN- Q=SV- vous voulez que
[" inf]	QSV- QESV1 SVQ- QVC0 QVN- Q=SV- il ya que

[emph aff non]	QSV0 QESV0 SVQ-	QVC1 QVN-	Q=SV-	nous allons ou
[inf]a	QSV0 QESV1 SVQ-	QVC0 QVN-	Q=SV-	vous mangez quelle
[" s adv]	QSV- QESV1 SVQ-	QVC0 QVN-	Q=SV-	je fais que
[" s adv]	QSV1 QESV0 SVQ-	QVC- QVN-	Q=SV-	je m'applique cmt
[" act]	QSV0 QESV1 SVQ-	QVC0 QVN-	Q=SV-	vous venez quand
[s rem c]	QSV- QESV1 SVQ0	QVC0 QVN-	Q=SV-	il yavait que
[" inf] words !	QSV- QESV- SVQ-	QVC1 QVN-	Q=SV-	tu es ou
[" inf] words !	QSV- QESV- SVQ-	QVC1 QVN-	Q=SV-	tu fais que
[s rem c]	QSV- QESV1 SVQ0	QVC0 QVN-	Q=SV-	il ya que
[s rem c]	QSV- QESV1 SVQ-	QVC0 QVN-	Q=SV-	je fais que
[s rem c]	QSV- QESV1 SVQ-	QVC0 QVN-	Q=SV-	je ai fait que
[s rem c]	QSV- QESV1 SVQ-	QVC0 QVN-	Q=SV-	je fais que
[s rem f]	QSV0 QESV- SVQ-	QVC1 QVN-	Q=SV-	je dirai cmt
[inf]a !	QSV- QESV- SVQ1	QVC- QVN-	Q=SV-	ce etait quoi .
[s rem c]	QSV0 QESV0 SVQ1	QVC0 QVN-	Q=SV-	ce est cbn (a)
[s rem c]	QSV0 QESV0 SVQ1	QVC0 QVN-	Q=SV-	il yavait cbn
[s rem c]	QSV- QESV1 SVQ0	QVC0 QVN-	Q=SV-	il ya que
[s rem c]	QSV- QESV1 SVQ0	QVC0 QVN-	Q=SV-	je pourrais que
[s rem c]	QSV0 QESV0 SVQ0	QVC1 QVN-	Q=SV-	je en etais ou
[s rem f]	QSV1 QESV0 SVQ-	QVC0 QVN-	Q=SV-	je vais explqer cmt
[stiq]	QSV- QESV1 SVQ-	QVC0 QVN-	Q=SV-	je ai fait que
?[" (s inf) !	QSV- QESV1 SVQ-	QVC- QVN-	Q=SV-	S=Q que s'est passe
[s rem f]	QSV1 QESV0 SVQ-	QVC0 QVN-	Q=SV-	je vais explqer cmt
[inf]a	QSV0 QESV0 SVQ1	QVC0 QVN-	Q=SV-	vous etes cbn (pour)
[inf]a	QSV0 QESV0 SVQ-	QVC0 QVN1	Q=SV-	np consiste quoi
[s rem f]	QSV- QESV1 SVQ0	QVC- QVN-	Q=SV-	je veux dire que
[emph neg si]	QSV- QESV0 SVQ-	QVC- QVN-	Q=SV1	S=Q qui fait
[e fill] !	QSV- QESV1 SVQ-	QVC- QVN-	Q=SV-	ce est que
[stiq]	QSV1 QESV0 SVQ-	QVC0 QVN-	Q=SV-	ca se decele cmt
[emph aff non]	QSV- QESV- SVQ-	QVC0 QVN1	Q=SV-	np est ou l'inconvt
[stiq]	QSV1 QESV0 SVQ0	QVC0 QVN-	Q=SV-	je leur ensngnais cmt
[clar]a	QSV- QESV1 SVQ-	QVC0 QVN-	Q=SV-	tu noverb que
[e fill] !	QSV- QESV1 SVQ-	QVC- QVN-	Q=SV-	je vais choisir que
[emph aff non]	QSV- QESV1 SVQ-	QVC0 QVN-	Q=SV-	je dis que
[" inf]	QSV1 QESV0 SVQ0	QVC0 QVN-	Q=SV-	tu veux pourquoi
[inf]a	QSV1 QESV0 SVQ0	QVC0 QVN-	Q=SV-	ca se passe cmt
[inf]a	QSV- QESV0 SVQ-	QVC- QVN-	Q=SV1	S=Q qui y a ete
[e fill] !	QSV- QESV1 SVQ-	QVC- QVN-	Q=SV-	je fais que
[inf]x	QSV0 QESV1 SVQ0	QVC0 QVN-	Q=SV-	il etait ou
[s rem c]	QSV0 QESV1 SVQ0	QVC0 QVN-	Q=SV-	je noverb ou
[s rem f]	QSV1 QESV0 SVQ-	QVC0 QVN-	Q=SV-	on peut cmt
[inf]a	QSV1 QESV0 SVQ0	QVC0 QVN-	Q=SV-	ca marche cmt
[s rem c]	QSV1 QESV0 SVQ0	QVC0 QVN-	Q=SV-	elle etait ou (d')
[inf]x	QSV0 QESV0 SVQ1	QVC0 QVN-	Q=SV-	ce est qui
[" inf]	QSV0 QESV1 SVQ0	QVC0 QVN-	Q=SV-	vous pouvez quand
[" off]	QSV- QESV0 SVQ-	QVC- QVN-	Q=SV1	S=Q qui veut
[" off]	QSV- QESV1 SVQ0	QVC0 QVN-	Q=SV-	vous voulez que
[" clar]	QSV- QESV0 SVQ1	QVC0 QVN-	Q=SV-	ce est quoi
[" clar]	QSV- QESV1 SVQ0	QVC0 QVN-	Q=SV-	ca veut dire que
[" off]	QSV- QESV1 SVQ0	QVC0 QVN-	Q=SV-	vous voulez que
[stiq]	QSV- QESV1 SVQ0	QVC0 QVN-	Q=SV-	ils font que
?[stiq][" adv]	QSV- QESV1 SVQ0	QVC0 QVN-	Q=SV-	on fait que
?[" (s inf]	QSV1 QESV0 SVQ0	QVC0 QVN-	Q=SV-	on en est ou
?[" (s inf) !	QSV- QESV1 SVQ-	QVC- QVN-	Q=SV-	S=Q que se passe
[s rem f]	QSV1 QESV0 SVQ-	QVC0 QVN-	Q=SV-	ca s'appelle cmt
[inf]x	QSV1 QESV0 SVQ0	QVC0 QVN-	Q=SV-	il est ou
[s rem c]	QSV1 QESV0 SVQ0	QVC0 QVN-	Q=SV-	je en etais ou
?[stiq][" (s adv]	QSV- QESV1 SVQ-	QVC0 QVN-	Q=SV-	on en fait que

[inf]x	QSV- QESV1 SVQ- QVC0 QVN- Q=SV- il ya que
?[" (s op] !	QSV- QESV1 SVQ- QVC- QVN- Q=SV- S=Q que va arriver
[s rem f]	QSV1 QESV0 SVQ0 QVC0 QVN- Q=SV- (ca s'appelle cmt
[off]a	QSV- QESV1 SVQ0 QVC0 QVN- Q=SV- tu veux que
[s rem c]	QSV0 QESV1 SVQ0 QVC0 QVN- Q=SV- je ai ete ou
[s rem c]	QSV- QESV1 SVQ0 QVC0 QVN- Q=SV- il ya que
[s rem c]	QSV- QESV1 SVQ- QVC0 QVN- Q=SV- elle a eu que
[inf]a	QSV0 QESV0 SVQ1 QVC0 QVN- Q=SV- ce etait quoi (avec)

Selective key:

! : categorical tokens

]a,x,b : interrogatives addressed to, respectively, fieldworker,
3rd person, and interviewee (by 3rd person).

1 : structure used

0 : structure acceptable and equivalent to that used

- : structure unacceptable or non-equivalent to that used

S=Q : subject is the WH word.

(Some linguistic items have been abbreviated, and corpus reference
numbers have been omitted.)

REFERENCES

- Al, B 1975 La notion de grammaticalité en grammaire générative-transformationnelle. Etude générale et application à la syntaxe de l'interrogation directe en français parlé, Leiden: Universitaire Pers Leiden.
- Ashby, W 1976 'The loss of the negative morpheme ne in Parisian French', Lingua, 39:119-137.
- Ashby, W 1977 'Interrogative forms in Parisian French', Semasia, 4:35-52.
- Ashby, W 1981 'The loss of the negative particle ne in French', Language, 57:674-687.
- Ashby, W 1988 'Français du Canada/français de France: divergence et convergence' French Review, 61:693-702.
- Ashton, E 1947 Swahili grammar, London: Longman.
- Austin, J 1962 How to do things with words, Oxford: Clarendon Press.
- Bach, K & Harnish, R 1979 Linguistic communication and speech acts, Cambridge, Mass.: MIT Press.
- Bauche, H 1929 Le langage populaire, Paris: Payot.
- Behnstedt, P 1978 Viens-tu? Est-ce que tu viens? Tu viens? Formen und strukturen des direkten Fragesatzes im Französischen, Tübingen: Narr.
- Bell, A 1984 'Language style as audience design', Language in Society, 13:145-204.
- Bever, T G & Langendoen, D T 1971 'A dynamic model of the evolution of language', Linguistic Inquiry, 2:433-63.
- Biber, D 1988 Variation across speech and writing, Cambridge: Cambridge University Press.
- Biggs, P & Dalwood, M 1976a Les Orléanais ont la parole (Student's book), London: Longman.
- Biggs, P & Dalwood, M 1976b Les Orléanais ont la parole (Teaching guide and transcript), London: Longman.
- Blanc, M & Biggs, P 1971 'L'enquête sociolinguistique sur le français parlé à Orléans', Le Français dans le Monde, 85:16-25.
- Blinkenberg, A 1928 L'ordre des mots en français moderne (1ère partie), Copenhagen: Det Kongelige Danske Videnskabernes Selskab.
- Bolinger, D 1977 Meaning and form, Longman: London.
- Bolinger, D 1978 'Asking more than one thing at a time', in Hiz,

H (ed) Questions, Dordrecht: Reidel.

Borillo, A 1979 'La négation et l'orientation de la demande de confirmation', Langue Française, 44:27-41.

Borillo, A 1981 'Quelques aspects de la question rhétorique en français', DRLAV, 25:1-33.

Boutet, J 1986 'La référence à la personne en français parlé: le cas de "on"', Langage et Société, 38:19-50.

Brown, G & Yule, G 1983 Discourse analysis, Cambridge: Cambridge University Press.

Brown, P & Levinson, S 1978 'Universals in language usage: politeness phenomena', in Goody, E N (ed) Questions and politeness, Cambridge: Cambridge University Press.

Brown, P & Levinson, S 1987 Politeness: some universals in language usage, Cambridge: Cambridge University Press.

Byrne, L S R & Churchill, E L 1986 A comprehensive French grammar (3rd edition, revised by G Price), Oxford: Blackwell.

Carton, F 1981 'Les parlers ruraux de la région Nord-Picardie: situation sociolinguistique', International Journal of the Sociology of Language, 29:15-28.

Casagrande, J & Saciuk, B (eds) 1972 Generative studies in Romance languages, Rowley, MASS.: Newbury House.

Centre Régional de Documentation Pédagogique d'Amiens 1981 Approche de la Picardie, Amiens: Centre Régional de Documentation Pédagogique.

Chambers, J & Trudgill, P 1980 Dialectology, Cambridge: Cambridge University Press.

Cheshire, J 1982 Variation in an English dialect, Cambridge: Cambridge University Press.

Chevalier, J-Cl, Blanche-Benveniste, C, Arrivé, M & Peytard, J 1964 Grammaire Larousse du français contemporain, Paris: Larousse.

Clark, H H & Clark, E V 1977 Psychology and language: an introduction to psycholinguistics, New York: Harcourt Brace Jovanovich.

Comrie, B 1984 'Russian', in Chisholm, W, Milic, L & Greppin, J (eds) 1984 Interrogativity: a colloquium on the grammar, typology and pragmatics of questions in seven diverse languages, Benjamins: Amsterdam.

Coppieters, R 1987 'Competence differences between native and

near-native speakers', Language, 63:544-573.

Coupland, N 1980 'Style-shifting in a Cardiff work-setting',
Language in Society, 9:1-12.

Crampon, M 1978 'Une expérience d'enseignement du dialecte picard
à l'Ecole Normale d'instituteurs d'Amiens', in Picoche, J (ed)
Archives sonores et dialectologie, Amiens: Centre d'Etudes
Picardes, Université de Picardie.

Crampon, M 1981 'La langue picarde', in Debrie, R (ed).

Crystal, D & Davy, D 1975 Advanced conversational English,
London: Longman.

Damourette, J & Pichon, E 1911-1934 Des mots à la pensée: essai
de grammaire de la langue française, t.4, Paris: Editions
d'Artrey.

de Certeau, M, Julia, D & Revel, J 1975 Une politique de la
langue. La Révolution française et les patois: l'enquête de
Grégoire, Paris: Gallimard.

Debrie, R 1960 Etude linguistique du patois de l'Amiénois, Thèse
de Doctorat d'Etat, Université de Paris.

Debrie, R (ed) La Picardie, Paris: Organisation.

Debrie, R 1983 Eche pikar bèl é rade, Paris: Omnivox.

Desirat, C & Horde, T 1976 La langue française au 20e siècle,
Paris: Bordas.

Desrosières, A 1984 'La nouvelle nomenclature des professions et
catégories socioprofessionnelles', in Fouquet, A (ed) Données
sociales (5ème édition), Paris: Institut National de la
Statistique et des Etudes Economiques.

Di Sciullo, A-M & St-Pierre, M 1982 'Les actes de requête dans
les énoncés de forme impérative et interrogative en français de
Montréal', in Lefebvre, C (ed), t.2.

Diller, A-M, 1983 'Subject NP structure and variable constraints:
the case of ne deletion', in Fasold, R (ed) (originally published
in Shuy, R & Schnukal, A (eds) 1980).

Diller, A-M, 1984 La pragmatique des questions et des réponses,
Tübingen: Narr.

Doppagne, A 1966 Trois aspects du français contemporain, Paris:
Larousse.

Dubois, J et al 1971 Dictionnaire du français contemporain,
Paris: Larousse.

Dubois, J & Dubois-Charlier, F 1970 Elements de linguistique
française: syntaxe, Paris: Larousse.

- Eckert, P 1980 'The structure of a long-term phonological process: the back vowel chain shift in Soulatan Gascon', in Labov, W (ed).
- Edwards, V K 1986 Language in a Black community, Cleveland, Avon: Multilingual Matters.
- Eloy, J-M 1987 'Le répertoire linguistique picard-français à Amiens', Langage et Société, 41:22-3.
- Emirikian, L & Sankoff, D 1979 'Coordination et effacement', in Thibault, P (ed).
- Emirikian, L & Sankoff, D 1985 'Le futur simple et le futur périphrastique dans le français parlé', in Lemieux, M & Cedergren, H (eds), t.1.
- Emrik, R 1978 Le parler picard, Amiens: Centre Régional de Documentation Pédagogique.
- Ervin-Tripp, 1979 'Children's sociolinguistic competence and dialect diversity', in Pride, J (ed) Sociolinguistic aspects of language learning and teaching, Oxford: Oxford University Press (first published in Dil, A (ed) 1973 Language acquisition and communicative choice, Stanford: Stanford University Press).
- Fasold, R 1983 (ed) Variation in the form and use of language, Washington: Georgetown University Press.
- Fauconnier, G 1981 'Questions et actes indirects', Langue Française, 52:44-55.
- Flutre, L F 1955 Le parler picard de Mesnil-Martinsart (Somme), Genève: Droz and Lille: Giard.
- Fónagy, I & Bérard, E 1973 'Questions totales simples et implicatives en français parisien', in Grundstrom, A & Léon, P (eds) Interrogation et intonation en français standard et en français canadien (Studia Phonetica 8), Montréal/Paris/Bruxelles: Didier.
- Foulet, L 1921 'Comment ont évolué les formes de l'interrogation', Romania, 47:243-348.
- Foulet, L 1968 Petite syntaxe de l'ancien français, Paris: Champion.
- Fraser, B 1983 'The domain of pragmatics', in Richards, J & Schmidt, R (eds) Language & communication, London: Longman.
- Fromaigeat, E 1938 'Les formes de l'interrogation en français moderne: leur emploi, leurs significations et leur valeur stylistique', Vox Romania, 3:1-47.
- Gaetone, D 1971 Etude descriptive du système de la négation en français contemporain Genève: Droz.
- Gal, S 1979 Language shift: social determinants of linguistic change in bilingual Austria, New York: Academic Press.
- García, E 1985 'Shifting variation", Lingua, 67:189-224.

Gardner-Chloros, P 1985a 'Language selection and switching among Strasbourg shoppers', International Journal of the Sociology of Language, 54:115-35.

Gardner-Chloros, P 1985b Choix et alternance des langues à Strasbourg, thèse de doctorat, Université Louis Pasteur Strasbourg I.

Gaudard, F 1982 'Les moniteurs de colo, ces inconnus familiers', Le Monde de l'Education, juillet/août, 84-86.

Giles, H & Smith, P 1979 'Accommodation theory: optimal levels of convergence', in Giles, H & St Clair, R (eds) 1979 Language and social psychology, Oxford: Blackwell.

Gilliéron, J & Edmont, E 1902-10 Atlas linguistique de la France, Paris: Champion.

Givón, T 1978 'Negation in language: pragmatics, function, ontology', in Cole, P (ed) Pragmatics (Syntax and Semantics 9), New York: Academic Press.

Gougenheim, G 1973 Grammaire de la langue française du seizième siècle (nouvelle édition), Paris: Picard.

Gougenheim, G, Rivenc, P, Michéa, R & Sauvageot, A 1964 L'élaboration du français fondamental (1er degré), Paris: Didier.

Greenbaum, S 1976 'Syntactic frequency and acceptability', Lingua, 40:99-113.

Grevisse, M 1986 Le bon usage (12ème édition, refondue par A Goosse), Gembloux: Duculot.

Groupe Aixois de Recherches en Syntaxe 1980 'Etudes de texte de français parlé: syntaxe et variations', in Gardin, B, Marcellesi, J-B, et le GRECO (eds) Sociolinguistique: approches, théories, pratiques, Paris: Presses Universitaires de France.

*

Guiraud, P 1965 Le français populaire, Paris: Presses Universitaires de France.

Guy, G 1988 'Language and social class', in Newmeyer, F (ed).

Hajičová, E 1983 'On some aspects of presuppositions in questions', in Kiefer, F (ed) Questions and answers, Dordrecht: Reidel.

Harris, J 1984 'Syntactic variation and dialect divergence', Journal of Linguistics, 20:303-327.

Harris, M B 1978 The evolution of French syntax: a comparative approach, London: Longman.

Hintze, M-A 1986 'Standardization: awareness and preservation of a norm in French', in Le Page, R B (ed) Abstract of the proceedings of the workshop of the International Group for the Study of Language Standardization and the Vernacularization of

- 365 -
* Grundström, A 1973 'L'intonation des questions en français standard', in Grundström, A & Léon, P (eds) Interrogation et intonation (Studia Phonetica 8), Montréal: Didier.

Literacy, York: Department of Language, University of York.

Hockey, S & Marriott, I 1982 Oxford Concordance Program users' manual, Version 1, Oxford: Oxford University Computing Service.

Hockey, S & Martin, J 1988 Oxford Concordance Program users' manual, Version 2, Oxford: Oxford University Computing Service.

Horn, L 1978 'Some aspects of negation', Greenberg, J (ed) Universals of human language, vol. 4, Syntax, Stanford, Ca.: Stanford University Press.

Huddleston, R 1984 Introduction to the grammar of English, Cambridge: Cambridge University Press.

Hudson, R 1975 'The meaning of questions', Language, 51:1-31.

Hudson, R 1980 Sociolinguistics, Cambridge: Cambridge University Press.

Huspek, M R 1986 'Linguistic variation, context and meaning: a case of -ing/in' variation in North American workers' speech', Language in Society, 15:149-164.

Institut National de la Statistique et des Etudes Economiques 1981 Picardie: des chiffres et des hommes, Amiens: INSEE, Observatoire Economique de Picardie.

Jacobson, S 1980 'Issues in the study of syntactic variation', in Jacobson, S (ed).

Jacobson, S (ed) 1980 Papers from the Scandinavian symposium on syntactic variation, Stockholm, May 18-19, 1979, Stockholm: Almqvist & Wiksell International.

Jahangiri, N & Hudson, R 1982 'Patterns of variation in Tehrani Persian', in Romaine, S (ed) Sociolinguistic variation in speech communities, London: Arnold.

Judge, A & Healey, F G 1983 A reference grammar of modern French, London: Arnold.

Kayne, R 1972 'Subject inversion in French interrogatives', in Casagrande, J & Saciuk, B (eds).

Kayne, R & Pollock, J-Y 1978 'Stylistic inversion, successive cyclicity and move NP in French', Linguistic Inquiry, 9:595-621.

Keenan, E 1985 'Passive in the world's languages', in Shopen, T (ed).

Kemp, W 1979 'L'histoire récente de ce que, qu'est-ce que et qu'osque à Montréal: trois variantes en interaction', in Thibault, P (ed).

Kemp, W 1981 'Major sociolinguistic patterns in Montreal French', in Sankoff, D & Cedergren, H (eds).

Kleiber, G 1986 Compte rendu de Korzen, H 1985, *Revue de Linguistique Romane*, 50:601-3.

Klippel, F 1984 Keep talking, Cambridge: Cambridge University Press.

Kontra, M 1980 'On English negative interrogatives', in Copeland, J E & Davis, P W (eds) The 7th LACUS forum, Columbia, S.C.: Hornbeam Press.

Korzen, H 1983 'Réflexions sur l'inversion dans les propositions interrogatives en français', Revue Romane (édition spéciale), 24:50-85.

Korzen, H 1985 POURQUOI et l'inversion finale en français: étude sur le statut de l'adverbial de cause et l'anatomie de la construction tripartite, Copenhagen: Institut d'Etudes Romanes, University of Copenhagen/Munksgaards.

Kreckel, M 1981 Communicative acts and shared knowledge in natural discourse, London: Academic Press.

Kroch, A & Small, C 1978 'Grammatical ideology and its effect on speech', in Sankoff, D (ed).

Labelle, G 1976 La performance syntaxique d'enfants de 5 ans de Montréal et de Paris, thèse de doctorat inédite, U.E.R. de linguistique, Université de Paris VII (cited in Lemieux, 1985).

Laberge, S 1977 Etude de la variation des pronoms sujets définis et indéfinis dans le français parlé à Montréal, thèse de doctorat, Université de Montreal.

Laberge, S 1983 'The changing distribution of indeterminate pronouns in discourse', in Fasold, R (ed) (originally published in Shuy, R & Schnukal, A (eds) 1980).

Labov, W ¹⁹⁶⁶ The social stratification of English in New York City, Washington, DC: Center for Applied Linguistics.

Labov, W 1972a Sociolinguistic patterns, Philadelphia: University of Pennsylvania Press.

Labov, W 1972b Language in the inner city, Philadelphia: University of Pennsylvania Press.

Labov, W 1975 What is a linguistic fact?, Lisse: Peter de Ridder Press.

Labov, W 1978 'Where does the linguistic variable stop? A response to Beatriz Lavandera', Working Papers in Sociolinguistics, 44, Austin, Texas: South-west Educational Development Laboratory, University of Texas, Austin.

Labov, W (ed) 1980 Locating language in time and space, New York: Academic Press.

Labov, W, Cohen, P, Robins, C, & Lewis, J 1968 A study of the

non-standard English of Negro and Puerto Rican speakers in New York City. Final Report, U.S. Office of Education Cooperative Research Project 3288, vol. 1, Washington, DC: Office of Education.

Lafontaine, D & Lardinois, B 1985 'Les questions: quelles structures les enfants de 7 à 12 ans utilisent-ils?', Bulletin de Psychologie, 38:63-70.

Laks, B 1977 'Contribution empirique à l'analyse socio-différentielle de la chute des /r/ dans les groupes consonantiques finals', Langue Française, 34:109-25.

Laks, B 1983 'Langage et pratiques sociales. Etude sociolinguistique d'un groupe d'adolescents', Actes de la Recherche en Sciences Sociales, 46:73-97.

Lambrecht, K 1981 Topic, antitopic and verb agreement in non-standard French, Amsterdam: Benjamins.

Lavandera, B 1975 Linguistic structure and sociolinguistic conditioning in the use of verbal endings in SI-clauses (Buenos Aires Spanish), unpublished doctoral dissertation, University of Pennsylvania.

Lavandera, B 1978 'Where does the sociolinguistic variable stop?' Language in Society, 7:171-182.

Leech, G 1981 Semantics, 2nd edition, Harmondsworth: Penguin.

Leech, G 1983 Principles of pragmatics, London: Longman.

Lefebvre, C 1981 'The double structure of questions in French: a case of syntactic variation', in Sankoff, D & Cedergren, H (eds).

Lefebvre, C 1982 'Qui qui vient? ou Qui vient? Voilà la question' in Lefebvre, C (ed)

Lefebvre, C (ed) 1982 La syntaxe comparée du français standard et populaire: approches formelle et fonctionnelle, tomes 1 et 2, Québec: Office de la langue française.

Lefebvre, C (1987) 'Some problems in defining syntactic variables: the case of WH questions in Montreal French', in Fasold, R & Schiffrin, D (eds) (1987) Language change and variation, Amsterdam/Philadelphia: Benjamins (cited in Fasold, R (1986) 'Linguistic analyses of the three kinds', in Sankoff, D (ed) Diversity and diachrony, Amsterdam/Philadelphia: Benjamins.

Lefebvre, C & Maisonneuve, H 1982 'La compétence linguistique des adolescents du Centre-sud: les structures complexes', in Lefebvre, C (ed), t.1.

Lemieux, M 1985 'Pas rien', in Lemieux, M & Cedergren, H (eds).

Lemieux, M & Cedergren, H (eds) Les tendances dynamiques du français parlé à Montréal, tomes 1 et 2, Québec: Office de la langue française.

Le Page, R & Tabouret-Keller, A 1985 Acts of identity, Cambridge: Cambridge University Press.

Lennig, M 1978 Acoustic measurement of linguistic change: the modern Parisian vowel system, PhD dissertation, University of Pennsylvania.

Levinson, S 1983 Pragmatics, Cambridge: Cambridge University Press.

Local, J 1982 'Modelling intonational variability in children's speech', in Romaine, S (ed).

Lüdicke, A 1982 'Zum Ausfall der Verneinungspartikel ne im gesprochenen Französisch', Zeitschrift für Romanische Philologie, 98:43-58.

Lyons, J 1977 Semantics, vol. 2, Cambridge: Cambridge University Press.

Lyons, J 1981 Language, meaning and context, London: Fontana.

McEntegart, D & Le Page, R 1982 'An appraisal of the statistical techniques used in the Sociolinguistic Survey of Multilingual Communities', in Romaine, S (ed).

Malécot, A 1972 'New procedures for descriptive phonetics', in Valdman, A (ed) Papers in linguistics and phonetics to the memory of Pierre Delattre, The Hague/Paris: Mouton.

Marceau, J 1977 Class and status in France, Oxford: Oxford University Press.

Martinet, A 1962 A functional view of language, Oxford: Oxford University Press.

Martins-Baltar, M 1976 'Actes de parole', in Coste, D, Courtillon, J, Ferenczi, V, Martins-Baltar, M, Papo, E & Roulet, E, Un niveau-seuil, Paris: Hatier.

Mauger, G 1968 Grammaire pratique du français d'aujourd'hui: langue parlée, langue écrite, Paris: Hachette.

Maury, N 1973 'Forme et fonction de -hein? d'après un corpus de français ontarien', Canadian Journal of Linguistics/Revue Canadienne de Linguistique, 18:146-156.

Meillet, A 1982 'L'évolution des formes grammaticales', in Meillet, A Linguistique historique et linguistique générale, Slatkine: Geneva and Champion: Paris.

Milroy, J & Milroy, L 1985a 'Linguistic, change, social network and speaker innovation', Journal of Linguistics, 21:339-84.

Milroy, J & Milroy, L 1985b Authority in language, London: Routledge & Kegan Paul.

Milroy, L 1980 Language and social networks, Oxford: Blackwell.

- Milroy, L 1987 Observing and analysing natural language, Oxford: Blackwell.
- Mohren, F 1980 Le renforcement affectif de la négation par l'expression d'une valeur minimale en ancien français, Tübingen: Niemeyer.
- Moravcsik, E A 1971 'Some cross-linguistic generalisations about yes-no questions and their answers', Stanford University working papers on language universals, 7:45-193 (cited in Sadock & Zwicky, 1985).
- Moreau, J-P 1981 'Le milieu et les hommes', in Debie, R (ed).
- Moreau, M-L 1986 'Les séquences préformées: entre les combinaisons libres et les idiomatismes. Le cas de la négation avec ou sans ne', Le Français Moderne, 54:137-160.
- Morin, Y-Ch 1973 'Tag questions in French', Linguistic Inquiry, 4:97-100.
- Mullineaux, L A & Blanc, M H A 1982 'The problems of classifying the population sample in the sociolinguistic survey of Orleans (1969) in terms of socio-economic, social and educational categories', ITL (Tijdschrift van het Instituut voor Toegepaste Linguistiek), Review of Applied Linguistics, 55:3-37.
- Neu, H 1980 'Ranking of constraints on /t,d/ deletion in American English: a statistical analysis', in Labov, W (ed).
- Newmeyer, F (ed) 1988 Linguistics: the Cambridge survey, vol. 4, Language: the socio-cultural context, Cambridge: Cambridge University Press.
- Obenauer, H-G 1981 'Le principe des catégories vides et la syntaxe des interrogatives complexes', Langue Française, 52:100-118.
- Ohkado, M 1989 'INFL and negating particles', Lingua, 77:1-12.
- Paris, G 1877 'Ti, signe d'interrogation', Romania, 6:438-442.
- Patterson, G W 1972 'French interrogatives: a diachronic problem', in Casagrande, J & Saciuk, B (eds)
- Payne, J 1985 'Negation', in Shopen, T (ed).
- Pohl, J 1965 'Observations sur les formes d'interrogation dans la langue parlée et dans la langue écrite non-littéraire', in Actes du Xe Congrès International de Linguistique et de Philologie Romanes, t.2, Paris: Klincksieck.
- Pohl, J 1968 'Ne dans le français parlé contemporain: les modalités de son abandon', in Actes du XIe Congrès International de Linguistique et de Philologie Romanes, t.3, Madrid.
- Pohl, J 1970 'Ne et les enfants', Le Langage et l'Homme, octobre, 41-43.

Pohl, J 1975 'L'omission de NE dans le français parlé contemporain', Le Français dans le Monde, 111:17-23.

Polome, E 1967 Swahili language handbook, Washington, D.C.: Center for Applied Linguistics.

Poplack, S forthcoming 'The care and handling of a mega-corpus: the Ottawa-Hull French project', to appear in Fasold, R & Schifffrin, D (eds) Language variation and change, Amsterdam: Benjamins.

Posner, R 1985 'Post-verbal negation in non-standard French: a historical and comparative view', Romance Philology, 39:170-197.

Price, G 1962 'The negative particles pas, mie and point in French', Archivum Linguisticum, 14:14-34.

Price, G 1971 The French language: past and present, London: Arnold.

Price, G 1978 'L'interrogation négative sans ne', Studii și Cercetări Linguistice, 29:599-606.

Price, R 1987 A social history of nineteenth-century France, London: Hutchinson.

Prince, E 1976 'The syntax and semantics of NEG-raising, with evidence from French', Language, 52:404-26.

Prince, E 1981 'Toward a taxonomy of given - new information', in Cole, P (ed) Radical pragmatics, New York: Academic Press.

Quirk, R, Greenbaum, S, Leech, G & Svartvik, J 1985 A comprehensive grammar of the English language, London: Longman.

Rehbein, J 1981 'Announcing - on formulating plans' in Coulmas, F (ed) Conversational routine, The Hague: Mouton.

Renchon, H 1969 Etudes de syntaxe descriptive II. La syntaxe de l'interrogation (2ème édition), Bruxelles: Académie royale de langue et de littérature françaises.

Rickford, J 1983 'Carrying the new wave into syntax: the case of Black English BIN', in Fasold, R (ed) (originally published in 1975 in Fasold, R & Shuy, R (eds) Analyzing Variation in Language, Washington, DC: Georgetown University Press.

Rickford, J 1987 'The haves and have nots: Sociolinguistic surveys and the assessment of speaker competence', Language in Society, 16:149-78.

Robins, R 1980 General linguistics: an introductory survey, 3rd edition, London: Longman.

Roelandt, A 1975 Etude de l'omission de NE dans les phrases négatives en français populaire: des environs de 1750 à 1870. Mémoire présenté pour l'obtention du grade du licencié en philologie romane, Université libre de Bruxelles (cited in

Sankoff, G & Vincent, 1980).

Romaine, S (ed) 1982 Sociolinguistic variation in speech communities, London: Arnold.

Romaine, S 1984 'On the problem of syntactic variation and pragmatic meaning in sociolinguistic theory' Folia Linguistica, 18:409-437.

Ross, J 1974 'Etude sociolinguistique et description syntaxique', in Verdoodt, A (ed) Proceedings of the 3rd Congress of the International Association of Applied Linguistics, vol. 2, Heidelberg: Groos.

Russell, J C 1981 Communicative competence in a minority group: a sociolinguistic study of the Swahili-speaking community in the Old Town, Mombasa, Leiden: Brill.

Rydén, M 1980 'Syntactic variation in a historical perspective', in Jacobson, S (ed).

Sadock, J & Zwicky, A 1985 'Speech act distinctions in syntax', in Shopen, T (ed).

Sag, I & Liberman, M 1975 'The intonational disambiguation of indirect speech acts', in Papers from the 11th regional meeting, Chicago Linguistics Society (April 18-20, 1975), Chicago, Ill.:Chicago Linguistics Society.

St-Pierre, M 1977 Aspects pragmatiques des interrogatives globales en français de Montréal, thèse de doctorat, Université de Montréal.

Sankoff, D (ed) 1978 Linguistic variation: models and methods, New York: Academic Press.

Sankoff, D 1988 'Sociolinguistics and syntactic variation', in Newmeyer, F (ed).

Sankoff, D & Cedergren, H (eds) 1981 Variation omnibus, Edmonton, Alberta: Linguistic Research Inc.

Sankoff, D & Laberge, S 1978 'The linguistic market and the statistical explanation of variability' in Sankoff, D (ed).

Sankoff, D & Sankoff, G 1973 'Sample survey methods and computer assisted analysis in the study of grammatical variation', in Darnell, R (ed) Canadian languages in their social context, Edmonton: Linguistic Research Inc.

Sankoff, D & Thibault, P 1979 'Weak complementarity: tense and aspect in Montreal French', mimeo (also in Johns, B & Strang, D (eds) 1981 Syntactic change, Ann Arbor: University of Michigan.)

Sankoff, G 1973 'Above and beyond phonology in variable rules', in Bailey, C-J N & Shuy, R (eds) New ways of analyzing variation in English, Washington: Georgetown University Press (also in

Sankoff, G, 1980).

Sankoff, G 1974 'A quantitative paradigm for the study of communicative competence', in Bauman, R & Sherzer, J (eds) Explorations in the ethnography of speaking, Cambridge: Cambridge University Press (also in Sankoff, G, 1980).

Sankoff, G 1980 The social life of language, Philadelphia: University of Pennsylvania Press.

Sankoff, G & Cedergren, H 1971 'Some results of a sociolinguistic study of Montreal French', in Darnell, R (ed) Linguistic diversity in Canadian society, Edmonton: Linguistic Research Inc.

Sankoff, G & Thibault, P (1977) 'L'alternance entre les auxiliaires avoir et être en français parlé à Montréal', Langue Française, 34:81-108. (English version in Sankoff, G, 1980).

Sankoff, G & Vincent, D 1980 'The productive use of ne in spoken Montreal French', in Sankoff, G 1980 (first published in 1977 as 'L'emploi productif du ne dans le français parlé à Montréal', Le Français Moderne, 45:243-56).

Schulze, A 1888 Des altfranzösische direkte Fragesatz, Leipzig.

Schwegler, A 1983 'Predicate negation and word-order change: a problem of multiple causation', Lingua, 61:297-334.

Searle, J 1969 Speech acts, Cambridge: Cambridge University Press.

Shopen, T (ed) Language typology and syntactic description, vol. 1, Clause structure, Cambridge: Cambridge University Press.

Shuy, R & Schnukal, A (eds) 1980 Language use and the uses of language, Washington: Georgetown University Press.

Slobin, D 1977 'Language change in childhood and history', in Macnamara, J (ed) Language learning and thought, New York: Academic Press.

Söll, L 1982 'L'interrogation directe dans un corpus de langage enfantin', in Hausmann, F-J (ed) Etudes de grammaire française descriptive, Heidelberg: Groos. (originally published as 'Der neufranzösische direkte Fragesatz in einem Corpus der Kindersprache', in Sprache und Geschichte. Festschrift für Harri Meier zum 65. Geburtstag, 1971, München: Fink.)

Strang, B 1970 A history of English, London: Methuen.

Terry, R M 1970 Contemporary French interrogative structures, Montreal/Sherbrooke: Editions Cosmos.

Thibault, P (ed) 1979 Le français parlé: études sociolinguistiques, Edmonton, Alberta: Linguistic Research Inc.

Thibault, P 1983 Equivalence et grammaticalisation, thèse de doctorat, Université de Montréal.

Trudgill, P 1974 The social differentiation of English in Norwich, Cambridge: Cambridge University Press.

Trudgill, P (ed) 1978 Sociolinguistic patterns in British English, London: Arnold.

Tuailion, G 1975 'Analyse syntaxique d'une carte linguistique: ALF 25: "Où vas-tu?" ', Revue de Linguistique Romane, 39:79-96.

Valdman, A 1982 'Français standard et français populaire: sociolectes ou fictions?', French Review, 56:218-227.

van Kleeck, A, Maxwell, M & Gunter, C 1985 'A methodological study of illocutionary coding in adult-child interaction', Journal of Pragmatics, 9:659-681.

Vaugelas, C F de 1647 Remarques sur la langue françoise, (Slatkine Reprints, 1970), Geneva: Slatkine

Venneman, T 1974 'Topics, subjects and word order: from SXV to SVX via TVX', in Anderson, J & Jones, C (eds) Historical linguistics, vol.1, Amsterdam: North Holland.

Verdié, M 1985 'Colonies de vacances: 1 300 000 jeunes', in Potel, J-Y (ed) L'état de la France et de ses habitants, Paris: Decouverte.

Vigner, G 1978 Savoir-vivre en France, Paris: Hachette.

Wagner, R L & Pinchon. J 1962 Grammaire du français classique et moderne, Paris: Hachette.

Walker, A L et al 1986 Lyon à la une: communicative competence in French (Manuel de l'animateur), Edinburgh: Scottish Universities French Language Research Association.

Warren, J 1988 'A sociolinguistic study of complementiser QUE in Montreal French', paper presented at Sociolinguistics Symposium 7, University of York.

Weiner, E J & Labov, W 1983 'Constraints on the agentless passive', Journal of Linguistics, 19:29-58.

Winford, D 1984 'The linguistic variable and syntactic variation in creole continua', Lingua, 62:267-288.

Wolfram, W & Christian, D 1976 Appalachian speech, Washington, DC: Center for Applied Linguistics.