

NOTE TO USERS

The original manuscript received by UMI contains pages with indistinct print. Pages were microfilmed as received.

This reproduction is the best copy available

UMI

La variation grammaticale en géolinguistique:
les pronoms sujet en roman central.

by

David John Heap

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
in Romance Languages and Linguistics
Graduate Department of French
University of Toronto

© David John Heap 1997



National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-28281-3

La variation grammaticale en géolinguistique: les pronoms sujet en roman central.

David John Heap

Ph.D. 1997, Romance Languages and Linguistics

Graduate Department of French

University of Toronto

This thesis deals with subject pronoun variation in the geolinguistic continuum which extends throughout Northern Italy and much of Southern France.

The subjects of finite verbs can be identified by pronouns (as in French *je parle* 'I speak') or by verbal morphology (as in Italian *parli* 'I speak'). It has been postulated by Perlmutter 1971 and Chomsky 1981 that this distinction reflects the two possible settings of an innate parameter: human grammars must be either [-null subject], like French, or [+null subject], like Italian. The nonstandard linguistic varieties studied here, however, include grammars which do not fit neatly into either of these two categories.

After general considerations concerning the importance of variation in linguistic theory, a methodology is developed to deal with data from traditional linguistic atlas surveys. Although almost completely ignored by 'theoretical' linguistics, the *Atlas Linguistique de la France* (Gilliéron & Edmont 1902-1908) and the *Sprach- und Sachatlas Italiens und der Südschweiz* (Jaberg & Jud 1928-1940) provide invaluable data for the investigation of morphosyntactic variation.

Data are coded from a total of 438 points appearing on 101 maps from each of

these atlases. These data are then synthesized in 15 different thematic maps, according to grammatical person and a variety of other linguistic factors. The quantitative analysis of this database also forms the foundation for some generalisations about the nature of subject pronoun systems.

First, the evidence does not point to an abrupt transition, as the parametric hypothesis would predict, but rather towards a gradual transition zone. Secondly, the typology of subject pronoun systems is not completely unconstrained, since certain clear tendencies emerge among the intermediate varieties. Subject pronouns are least likely to occur in the first person singular or plural, and the second person plural, and most likely to occur in the third person singular or plural and (especially) the second person singular. Such morphological asymmetries cannot be readily represented in syntactic or featural terms, much less by parametric theory.

Furthermore, many linguistic factors, including verb type and tense, clause type and position, other pronouns and negation, all contribute to the conditioning of subject pronoun usage. These facts strongly suggest that subject pronoun paradigms reflect complex morphosyntactic 'interface' phenomena (Chomsky 1995) rather than a straightforwardly syntactic parameter.

Table des Matières

	Page
Résumé	ii
Table des Matières	iv
Liste des Cartes	viii
Liste des Tableaux	viii
Abréviations	viii
Remerciements	ix
Chapitre 1: INTRODUCTION	1
1.0. ORIENTATION GÉNÉRALE	1
1.1. Préliminaires (méta-)théoriques	2
1.1.1. Approche théorique: éclectique	3
1.1.2. Théories sous-déterminées par les données	5
1.1.3. Note sur la terminologie	8
1.1.4. Variation et théorie linguistiques	12
1.2. LES CLITIQUES	12
1.2.1. Définitions et statut	12
1.2.2. Tests formels de cliticité	14
1.2.3. Statut ambigu des clitiques	17
1.3. APPROCHES ANTÉRIEURES	21
1.3.1. Classifications	22
1.3.2. Contraintes prosodiques	24
1.3.3. Analyses syntaxiques génératives	26
1.3.4. Modèles morphologiques	31
1.4. PRONOMS SUJET	36
1.4.1. De la philologie traditionnelle à la linguistique structuraliste	36
1.4.2. Le «paramètre du sujet nul»	39
1.4.3. Contre le paramètre du sujet nul	42
1.5. LA VARIATION GRAMMATICALE	46

CHAPITRE 2: LA VARIATION ET LES ATLAS LINGUISTIQUES	52
2.0. LA VARIATION LINGUISTIQUE	52
2.1. VARIATION <i>INTRAGRAMMATICALE</i> ET <i>INTERGRAMMATICALE</i>	52
2.1.1. Diachronie, synchronie et continuum	57
2.1.2. Continuums, zones de transition et typologie linguistique	63
2.1.3. La question du franco-provençal	66
2.1.4. La comparaison des grammaires	71
2.2. UN CORPUS D'ATLAS LINGUISTIQUES	75
2.2.1. À propos des corpus en linguistique	75
2.2.2. Pourquoi étudier les atlas linguistiques romans?	76
2.2.3. Notre corpus	81
CHAPITRE 3: LES PRONOMS SUJET ET LA VARIATION GRAMMATICALE	84
3.0. PRÉSENTATION	84
3.1. MÉTHODOLOGIE: UNE GRILLE HEURISTIQUE	84
3.2. GROUPES DE FACTEURS ET CODES CORRESPONDANTS	86
3.3. BASE DE DONNÉES	89
3.3.1. Numéro de carte	90
3.3.2. Numéro de point	91
3.3.3. Hypothèses de travail	95
3.3.4. Emploi catégorique ou variable des sujets pronominaux	96
3.3.5. Typologie des paradigmes	97
3.4. FACTEURS CONDITIONNANTS	100
3.4.1. Position du verbe	102
3.4.2. Temps et mode	103
3.4.3. Type de verbe et auxiliaire	105
3.4.4. Genre	107
3.4.5. Présence de négation	108
3.4.6. Présence d'interrogation	110
3.4.7. Présence de clitiques obliques	111
3.4.8. Type et position du sujet	113
3.4.9. Interactions entre facteurs	118

CHAPITRE 4: RÉSULTATS	121
4.0. PRÉSENTATION	121
4.1. PROBLÈMES PRATIQUES DANS LE DÉPOUILLEMENT DES DONNÉES	121
4.1.1. Problèmes de nature phonologique	121
4.1.2. «Réponses extorquées»	124
4.1.3. Rareté relative des quatrième et cinquième personnes	128
4.1.4. Échantillon final	129
4.1.5. Synchrétisme et oppositions morphologiques	130
4.1.6. Résultats «catégoriques»	131
4.2. RÉSULTATS PAR PERSONNES	133
4.2.1. Première personne	134
4.2.2. Deuxième personne	136
4.2.3. Troisième personne	138
4.2.3.1. Verbes météorologiques	141
4.2.3.2. Verbes impersonnels	146
4.2.4. Quatrième personne	148
4.2.5. Cinquième personne	151
4.2.6. Sixième personne	154
4.3. TYPOLOGIE DES PERSONNES	157
4.3.1. Variétés à six ou zéro pronoms sujet	157
4.3.2. Variétés à entre deux et cinq pronoms sujet	159
4.3.3. Le statut privilégié de la deuxième personne	167
4.3.4. La distribution typologique	169
4.3.5. Variation sociale	171
4.4. RÉSULTATS VARIABLES	175
4.4.1. Position du verbe	177
4.4.2. Temps et mode	178
4.4.3. Type de verbe et auxiliaires	179
4.4.4. Genre	180
4.4.5. Négation	180
4.4.6. Interrogation	181
4.4.7. Clitiques obliques	182
4.4.8. Type et position du sujet	183

CHAPITRE 5: CONCLUSIONS	185
5.0. PRÉSENTATION	185
5.1. CONCLUSIONS GÉOLINGUISTIQUES	185
5.1.2. Types de paradigmes et leur distribution géographique	191
5.2. CONCLUSIONS TYPOLOGIQUES	192
5.2.1. Généralisations 1 et 2	192
5.2.2. Autres implications typologiques	198
5.3. CONCLUSIONS VARIABLES	200
5.3.1. Variabilité inhérente	201
5.4. VARIATION CONDITIONNÉE	203
5.4.1. Position du verbe	203
5.4.2. Temps et mode	205
5.4.3. Type de verbe et auxiliaires	207
5.4.4. Genre	208
5.4.5. Négation	208
5.4.6. Interrogation	209
5.4.7. Clitiques obliques	209
5.4.8. Type et position du sujet	212
5.5. IMPLICATIONS THÉORIQUES	212
5.5.1. Paramètre ou paradigme?	213
5.5.2. Variation morphologique aux interfaces	215
Appendice I: Les Cartes de l' <i>ALF</i>	219
Appendice II: Les Cartes de l' <i>AIS</i>	225
Appendice III: Les points d'enquête	231
Bibliographie	249

Liste des Cartes

	Page
Carte 1: Les points d'enquête	93
Carte 2: Pronoms sujet à la première personne (contexte type)	135
Carte 3: Pronoms sujet à la deuxième personne (contexte type)	137
Carte 4: Pronoms sujet à la troisième personne (contexte type)	139
Carte 5: Pronoms sujet à la troisième personne (sujets argumentaux)	140
Carte 6: Pronoms sujet à la troisième personne (verbes météorologiques)	143
Carte 7: Pronoms sujet génériques du type [ka]	144
Carte 8: Pronoms sujet à la troisième personne (verbes impersonnels)	147
Carte 9: Pronoms sujet à la quatrième personne (contexte type)	149
Carte 10: Pronoms sujet à la quatrième personne (tous temps confondus)	150
Carte 11: Pronoms sujet à la cinquième personne (contexte type)	152
Carte 12: Pronoms sujet à la cinquième personne (tous temps confondus)	153
Carte 13: Pronoms sujet à la sixième personne (contexte type)	155
Carte 14: Pronoms sujet à la sixième personne (sujet argumentaux)	156
Carte 15: Typologie des systèmes pronominaux	158
Carte 16: Les systèmes pronominaux d'après Renzi & Vanelli	170
Carte 17: La variation sociale dans les doubles enquêtes de l' <i>AIS</i>	173

Liste des Tableaux

Tableau 1: Typologie selon les traits [\pm sujet clitique] >> [\pm sujet \emptyset]	43
Tableau 2: Typologie selon les traits [\pm sujet \emptyset] >> [\pm sujet clitique]	45
Tableau 3: Points à pronom sujet (quasi-)catégorique, par personne	159
Tableau 4: Hiérarchie des personnes dans les systèmes de sujets pronominaux	160
Tableau 5: Degré d'accord avec la Généralisation 1, selon le nombre de personnes dans les systèmes pronominaux	166
Tableau 6: Présence et absence de <i>tu</i> dans tous les paradigmes incomplets	167
Tableau 7: Présence et absence de <i>tu</i> dans les points en désaccord avec Généralisation 1	168

Abréviations

AIS= Jaberg, Karl et Jakob Jud. 1928-1940. *Sprach- und Sachatlas Italiens und der Südschweiz*. Zofingen: Ringier.

ALF= Gilliéron, Jules et Edmond Edmont. 1902-1910. *Atlas linguistique de la France*. Paris: Champion.

Les chiffres qui suivent immédiatement ces abréviations (par exemple, *AIS*10 ou *ALF* 87) correspondent aux numéros de cartes dans l'atlas respectif.

Remerciements

Je tiens à exprimer ici ma profonde reconnaissance envers diverses personnes qui ont contribué à rendre ce travail possible, et je m'excuse d'avance auprès de celles que j'aurais omises

Tout d'abord à mon directeur, Henry Schogt, dont les capacités intellectuelles sont seulement dépassées par ses qualités humaines, et qui m'a toujours rappelé qu'il y a des choses beaucoup plus importantes que la linguistique; ainsi qu'à Corrie Schogt, qui nous a gentiment poussés tous les deux quand il l'a fallu. À Ed Bursynsky, le premier linguiste que j'ai rencontré à l'Université de Toronto, et qui a su me guider comme nul autre à travers mes années d'études en linguistique romane. Et à Yves Roberge, d'avoir créé le Groupe de Recherches en Dialectologie Comparative, ce lieu d'échange et de débat dont tant de personnes ont bénéficié, et d'avoir toujours su combiner ses conseils avec l'humour, le pragmatisme avec la rigueur.

J'ai la chance d'avoir des collègues qui ont pris le temps dans leurs horaires déjà trop chargés pour lire et commenter des versions antérieures de ce travail: un grand merci donc à Julie Auger, Paul Bessler, Sarah Cummins, Terry Nadasdi et Jeff Tennant. Leurs regards critiques et leurs discussions perspicaces sont les signes d'amitiés qui me sont précieuses; je regrette seulement ne pas avoir pu suivre plus de leurs suggestions.

D'autres linguistes ont également contribué au processus de formation et de recherche qui a abouti à cette thèse, que ce soit de façon directe ou indirecte, par leurs écrits, par leurs cours ou par nos discussions. Je remercie donc Parth Bhatt, Eulàlia Bonet, Jack Chambers, Gianrenzo Clivio, Elizabeth Cowper, Joseph Gulsoy, Greg Guy, Ruth King, Monique Léon, Pierre Léon, Philippe Martin, Diane Massam, Keren Rice et Dieter Wanner.

Merci aussi à la foule de personnes qui ont rendu mes années dans ce programme agréables et intéressantes, aux niveaux intellectuel, administratif et personnel: Peter Avery, Ron Davis, Carrie Dyck, Vittorio Frigerio, Mirco Ghini, Jila Ghomeshi, Elaine Gold, Daniel Halliday, Michol Hoffman, Päivi Koskinen, Monique Lecerf, Silvana Mastro Monaco, Anna Moro, Paul Perron, Alexandre Sévigny, Júlia Todolí, Tom Wilson, et Hitay Yükseker.

Au niveau technique, l'aide de plusieurs individus m'a été indispensable. Je remercie Brian Merrilees, qui m'a initié dans les bases de données relationnelles, ainsi que Charles Doubt et Sean Hoppen, qui ont su me faire découvrir une partie du potentiel de cet outil. Le dépouillement et l'entrée des données a seulement été possible grâce à l'aide minutieuse de Nicla Gargano et d'Ivan Chow. Mark Daniels et Danny Heap ont toujours été disponibles pour répondre patiemment à mes questions d'informatique (même les plus bêtes). Au Département de Géographie de l'Université de Toronto, Byron Moldofsky et Siegfried Schulte m'ont facilité l'accès aux logiciels cartographiques et m'ont conseillé généreusement en cartographie électronique; Harold Reynolds m'a beaucoup aidé dans la mise au point des cartes. La localisation et la digitalisation des points d'enquêtes ont été consciencieusement effectuées par David Chong et par David Crosby, respectivement.

Au plan pratique, je reconnais l'appui financier d'une Bourse des Études Supérieures

d'Ontario, d'une bourse doctorale du Conseil des Recherches en Sciences Humaines (752-91-2167), et d'une bourse Queen Elizabeth II. Également importantes pour leurs contributions à rendre ma vie d'étudiant plus vivable et plus fructueuse sont les organisations suivantes: Graduate French Association, Graduate Students' Union, Linguistics Graduate Students' Union, Nancy's Part-Time Childcare Centre, et le Syndicat Canadien de la Fonction Publique section 3902 (anciennement CUEW/SCTTÉ section 2).

Finalement, je n'aurais jamais pu finir cette tâche sans l'appui moral, l'énorme patience et l'encouragement généreux de Claudia Vicencio. Cette thèse est dédiée à mes sources constantes d'inspiration: Nicolás Martín et Felipe Estevan.

CHAPITRE 1: INTRODUCTION

1.0. ORIENTATION GÉNÉRALE

Cette thèse se propose d'étudier la variation qui existe parmi les sujets pronominaux (notamment -- mais non pas exclusivement -- clitiques) des verbes finis, dans des variétés gallo-romanes et nord-italiennes. Pour ce faire, nous élaborons une méthodologie qui tient compte à la fois de la variabilité des faits et des hypothèses théoriques, et nous confrontons les résultats ainsi obtenus aux différentes interprétations et analyses possibles. Dans ce premier chapitre, nous considérons d'abord quelques questions d'ordre théorique et métathéorique, pour mieux définir le cadre dans lequel ce travail s'insère, la terminologie dont nous nous servons, ainsi que l'importance que nous accordons à la variation en linguistique. Nous passons ensuite au concept même de «clitique» et à une discussion de ce que nous entendons par ce terme dont le statut est souvent ambigu, avant de passer en revue différentes approches adoptées par le passé pour traiter les clitiques, soit comme éléments syntaxiques, soit comme éléments morphologiques. Nous nous concentrons par la suite sur la question des pronoms sujet et sur l'hypothèse du «paramètre du sujet nul», et nous terminons le chapitre en justifiant la présente étude sur la variation grammaticale.

1.1. Préliminaires (méta-)théoriques

Qu'il nous soit permis de commencer par quelques observations de nature métathéorique (qui ne sont pas, à vrai dire, strictement linguistiques), pour souligner l'inséparabilité heuristique des théories et des données linguistiques. C'est en effet cette dialectique entre les deux aspects de la recherche qui guide la méthode à suivre. Ce sont les différentes approches théoriques qui nous indiquent quelles sont les questions intéressantes à poser face à une masse de données, et l'analyse des données qui nous dit à son tour quelle approche théorique est la plus adéquate pour un phénomène déterminé:

Theory and data, collection and analysis or interpretation are not independent phases of successful research; the two components codetermine each other to a certain extent. (Wanner 1987:9)

Bien qu'il s'agisse ici d'un travail fondamentalement descriptif réalisé à partir d'un corpus empirique, les hypothèses que nous mettons à l'épreuve sont dérivées de propositions théoriques, et nous espérons aussi que nos résultats pourront à leur tour contribuer à la théorie linguistique.

1.1.1. Approche théorique: éclectique

Du point de vue de l'histoire des sciences du langage, le moment actuel ne nous semble pas promettre une grande percée nouvelle dans la théorie linguistique. Ce n'est pas que les théories linguistiques fassent défaut en ce moment; au contraire, on a plutôt affaire aujourd'hui à une surabondance de théories qui se font concurrence (il suffit de penser par exemple aux principales variantes de la grammaire générative: le cadre des Principes et Paramètres qui devient le Programme Minimaliste (Chomsky 1995), la grammaire relationnelle-fonctionnelle, la Grammaire Syntagmatique Généralisée (anglais «GPSG» et sa version plus récente, la «HPSG») pour ne mentionner qu'un petit échantillon des nombreux cadres théoriques disponibles dans un des sous-domaines de la linguistique). Une façon possible de faire progresser une science quelconque sous de telles conditions est de pratiquer ce que Kuhn (1963) appelle la science normale («normal science»), c'est-à-dire l'application méthodique des normes établies d'un cadre théorique afin de les confirmer sur ses propres termes, plutôt que de chercher à les dépasser. A la longue, les résultats de la science normale révèlent non seulement les possibilités d'un cadre donné, mais aussi ses limites et ses contradictions inhérentes. Une autre façon de s'approcher de ce même but (surtout face à une surabondance de théories en concurrence) serait de se pencher plutôt sur la comparaison et l'évaluation des différentes théories. Mais pour pouvoir évaluer les différentes approches théoriques, il est utile d'identifier préalablement des phénomènes qui présentent des difficultés dans tous les cadres théoriques, et ensuite de bien décrire les faits linguistiques pertinents. C'est justement ce genre de travail

descriptif préliminaire que nous nous proposons d'effectuer ici: nous cernons une zone empiriquement difficile à décrire -- les pronoms sujet variables -- et nous cherchons à décrire les faits de façon à pouvoir ensuite les confronter aux traitements théoriques possibles.

L'hypothèse de travail adoptée est donc que chaque théorie linguistique peut en effet illuminer un aspect de la réalité complexe qu'est le langage humain. En effet, la nature polyfacétique des langues naturelles est telle qu'il serait difficile de croire qu'un seul cadre théorique puisse réellement expliquer toutes ses caractéristiques (formelles et autres). Il semble évident qu'une approche «éclectique» -- qui se maintient ouverte (dans les limites du possible) aux divers courants de la pensée linguistique -- sera plus adéquate pour faire face à une telle situation qu'un cadre étroit qui exclut d'emblée d'autres approches:

the specific mix of linguistic views and traits employed here is clearly less committed to the frequently categorical choices in formalized grammar. . . . To take sides from the very outset -- e.g. for one or the other narrowly defined neo-transformational subtheory in current competition, . . . or for a given functionalist perspective to the exclusion of other viewpoints . . . -- could only mean to restrict the field of factual vision inordinately, given the essential interdependence of the theory and perceptible facts. Rather, most of these partially incompatible approaches provide alternative, and therefore complementary, insights for specific

questions if used conjunctively. That such an eclectic approach will lack at crucial points of the investigation the explanatory potential of a formalized theoretical fragment may appear as a serious drawback only in a puristic sense. At the present state of the investigation of clitics, and of our understanding of linguistic phenomena, the advantages of presenting a relatively colourless theoretical backdrop for the language specific clitic study far outweigh the lack of what appear to be premature formalizations. (Wanner 1987:10-11)

Il s'agit pour autant de juger les théories en fonction des données, plutôt qu'inversement, bien que les données ne prennent leur sens réel qu'à l'intérieur d'un cadre théorique donné. Cette attitude éclectique est partagée par d'autres linguistes contemporains qui refusent de sacrifier les données aux postulats théoriques (voir par exemple Mahmoudian 1993).

1.1.2. Théories sous-déterminées par les données

Idéalement, dans chaque cas qui nous intéresse on trouverait qu'il n'existe qu'une seule explication théorique adéquate; hélas, il est souvent le cas qu'un seul et même phénomène linguistique est susceptible d'être formalisé par deux ou plusieurs analyses théoriques, sans qu'on puisse vraiment faire un choix entre celles-ci à l'aide de critères indépendants. On peut dire alors que les théories sont sous-déterminées («underdetermined») par les données. Là où plusieurs analyses sont possibles, notre tâche

serait soit de choisir la plus apte, soit d'avouer que les deux analyses (ou plus) sont également utiles (au moins pour ce qui est des faits qui nous intéressent). Il faut bien souligner qu'on n'envisage pas ici de procéder à une évaluation hermétiquement théorique qui ne prend en compte que des critères internes aux théories mêmes, mais plutôt de tenter une évaluation empirique, qui vise à confronter les différentes théories à des masses de données susceptibles de les mettre à l'épreuve. C'est pour cela que nous trouvons préférable de rester agnostique sur la question délicate de savoir quel est le nombre et quelle est la nature des modules qui constituent un modèle de grammaire humaine:

... in principle, many of the same constructions could be described either by lexical rule or by a (non-lexical) syntactic rule. For example, passive sentences might be described as derived by movement (together with associated changes) from underlying active sentences. They might also be described, though, by relating active verbs lexically to passive participles. . . . It is in this sense that we can speak of syntactic and lexical accounts as covering the 'same' phenomena. Given a particular construction, described in a particular way by syntactic rules, we can ask whether a lexical account is available to cover the same range of properties, although, of course, the actual mechanics of the two descriptions will differ in important ways. (Anderson 1992:37-38)

Par exemple, même quand on est relativement sûr qu'un pronom, ou un paradigme entier de pronoms, doit être traité comme un clitique, d'après des critères formels (voir 1.2.2.), il est souvent loin d'être évident où, exactement, dans un modèle grammatical (c'est-à-dire, dans quel module) il faut représenter ce fait. C'est le cas, entre autres, des pronoms sujet en espagnol antillais, qui se comportent comme des clitics dans la mesure où ils ne sont pas affectés par les contraintes sur la position du sujet qui existent en espagnol standard. L'agrammaticalité de la forme standard en (1) montre qu'un pronom sujet ne peut pas apparaître entre un élément *WH* initial et un verbe dans une phrase interrogative (Torrego 1984):

- (1) *¿*Qué tú quieres?*
'que tu veux?'

La même forme en (1) est, en revanche, parfaitement grammaticale en espagnol antillais, parce que le pronom sujet *tú* dans cette variété n'a pas le même statut qu'en espagnol standard: c'est un clitique syntaxique, dans le sens qu'il n'est pas «visible» pour les contraintes qui régissent les questions à élément interrogatif initial en espagnol (Lipski 1977; Contreras 1989; Heap 1991; Benedicto 1993). Notons que ce contraste vaut (variablement, bien entendu) pour tous les pronoms sujet, au moins pour certaines variétés de l'espagnol antillais (Heap 1991).

Si l'existence des pronoms sujet clitiques en espagnol antillais semble assez certaine, la localisation et la formalisation de ce phénomène dans un modèle grammatical restent des questions ouvertes. S'agit-il d'une «règle» syntaxique de cliticisation? Si oui, à quel niveau de représentation est-ce qu'elle s'applique? Si l'on a affaire plutôt à un processus morphologique, il faut néanmoins décider s'il s'effectue avant ou après la syntaxe proprement dite. Et on ne peut pas exclure non plus la possibilité d'une simple alternance lexicale entre formes fortes et formes faibles. Les données disponibles ne sont pas suffisantes pour nous pousser à préférer une de ces options plutôt que les autres: il s'agit bien d'un cas où l'analyse est sous-déterminée par les faits.

1.1.3. Note sur la terminologie

Du point de vue scientifique, les termes «langue» et «dialecte» reflètent des différences d'ordre politique ou social mais non pas proprement linguistique. Comme le font remarquer Chambers et Trudgill, «The term 'language', then, is from a linguistic point of view a relatively nontechnical term.» (1980:5). De même, pour Roberge et Vinet: «Dans son acception courante, la langue se présente comme un concept sociopolitique obscur, fréquemment associé à quelque chose comme des frontières officielles ou des cartes géographiques. La réalité est le plus souvent tout autre.» (1989:16). Quelle est donc la différence réelle? Depuis au moins Sapir, la plupart des spécialistes de la linguistique ne voient pas de distinction utile entre ces deux termes: «To the linguist, there is no real difference between a "dialect" and a "language" which can be shown to be related,

however remotely, to another language.» (1931:123) Dans ce sens alors, toutes les «langues» seraient des «dialectes» dont certains, pour des raisons historiques, jouissent du statut d'une «langue standard» d'un pays (ou d'une autre unité politique): d'après la célèbre phrase de Max Weinreich «A language is a dialect with an army and a navy.» D'après ce critère, la seule différence entre langue et dialecte serait leur statut social: «D'une certaine façon, un dialecte est la contrepartie d'une langue; suivant ce point de vue, il serait tout simplement 'une langue sans drapeau et sans armée', ou encore, une langue dominée.» (Roberge et Vinet 1989) Mais si toutes les «langues» sont en fait des «dialectes» (et inversement), quel sens peut-on donc attribuer à l'idée qu'une «langue» puisse comporter plusieurs «dialectes»? Ces questions, qui mènent souvent à des polémiques chargées d'émotion et de préjugés, ne sont guère pertinentes pour la terminologie linguistique, qui ne cherche qu'un terme neutre pour désigner l'objet de son étude. Nous évitons donc, dans la mesure du possible, les termes «langue» et «dialecte»: quand nous nous référons à la documentation linguistique, nous employons les termes «dialecte», «dialectale» et «dialectologie» entre guillemets, pour les mettre en relief comme termes non scientifiques. Nous utilisons de préférence le terme plus précis de géolinguistique (adjectif et substantif) pour nous référer à l'étude de la variation linguistique à travers l'espace, et le terme géolecte pour désigner des variétés (ou leurs grammaires correspondantes) quand il est question de leur spécificité géographique.

Dans une perspective générativiste, l'objet des recherches linguistiques est la grammaire, c'est-à-dire l'ensemble des propriétés formelles qui peuvent rendre compte de

l'apprentissage du langage et de la productivité de la faculté langagière humaine. Sans adopter nécessairement les présupposés théoriques implicites dans cette idée, nous acceptons ici grammaire tout court comme terme neutre et scientifique pour désigner un parler quelconque (Roberge et Vinet 1989:16). On peut ainsi comparer la grammaire du français standard à celle du français ontarien, ou à celle du piémontais ou à celle du tagalog, sans aborder la question d'éventuelles relations (génétiques, typologiques ou autres) entre ces grammaires. Notons que cette acception du terme grammaire exclut évidemment l'idée d'une grammaire traditionnelle, soit une grammaire normative ou prescriptive, et que l'on peut donc se référer aux grammaires indépendamment du statut écrit ou non écrit, codifié ou non codifié, d'un parler déterminé. De la même façon, nous préférons le terme variation grammaticale à celui de variation dialectale, qui, lui, est trop chargé de possibilités de jugements de valeur.

La question est cependant compliquée par le fait qu'aucune grammaire n'est directement accessible à la recherche: nous ne pouvons connaître la faculté langagière que par le biais des comportements externes des sujets parlants, c'est-à-dire, par les énoncés réels. Dans un sens plus concret, donc, notre objet de recherche se limite aux variétés linguistiques que parlent les gens qui partagent une grammaire, c'est-à-dire une communauté linguistique. Nous comprenons donc par variété l'ensemble d'énoncés (réels ou virtuels) que peuvent produire et comprendre les personnes qui partagent une grammaire donnée. Celle-ci est, bien entendu, seulement empiriquement vérifiable d'après celle-là: seront considérées comme partageant une grammaire (et formant donc une

communauté linguistique) les personnes qui peuvent produire et comprendre le même ensemble d'énoncés. Nous nous référons donc aux différentes variétés romanes ou françaises comme les réalisations concrètes de leurs grammaires respectives. Suivant Chambers et Trudgill (1980:5), nous employons le terme variété de façon assez souple: «We shall use 'variety' as a neutral term to apply to any particular kind of language which we wish, for some purpose, to consider as a single entity. The term will be used in an ad hoc manner in order to be as specific as we wish for a particular purpose.» À un extrême, alors, une variété pourrait correspondre à un idiolecte ou la façon de parler idiosyncratique d'une seule personne, tandis que dans d'autres contextes une variété pourrait désigner un parler régional, qui comporterait éventuellement plusieurs sous-variétés et même des sociolectes.

Mais même les variétés sont des abstractions qui ne sauraient être que partiellement accessibles: aucune recherche réelle ne pourrait s'occuper de tous les énoncés susceptibles d'être produits et compris par les personnes partageant une grammaire donnée, car ceux-ci constituent, par définition, un ensemble infini. Nous devons donc nous contenter d'un corpus, soit un sous-ensemble (qui se veut représentatif) des énoncés d'une variété (voir 3.3.)

1.1.4. Variation et théorie linguistiques

Toute théorie linguistique devrait avoir comme un de ses buts de pouvoir rendre compte de la variation linguistique dans toutes ses dimensions -- sociolinguistique, géolinguistique, diachronique -- ou au moins de la décrire de façon adéquate. Il est impossible de nier que la variation existe effectivement dans le langage réel des humains, et la passer sous silence serait donc trahir la réalité linguistique. Nous privilégions donc les théories linguistiques qui sont susceptibles de rendre compte de la variation, ou au moins, qui ne l'excluent pas en la traitant comme quelque chose de marginal. Un cadre théorique adéquat doit au moins permettre la coexistence des formes standard et non standard; idéalement, il devrait aussi pouvoir exprimer les relations entre les différentes formes en termes de l'opposition marqué -- non marqué («markedness»). Sur les différents types de variation linguistique, voir 2.1.

1.2. LES CLITIQUES

1.2.1. Définitions et statut

Le terme *clitique* (du grec *klinein* «s'incliner, s'appuyer») s'applique à des éléments linguistiques qui sont phonologiquement (ou, plus précisément, prosodiquement) *dépendants*, c'est-à-dire, qui n'apportent pas leur propre accent de mot. Ils doivent donc s'attacher prosodiquement (et souvent aussi syntaxiquement) à un autre mot, leur «hôte»

(anglais *host*). Les *proclitiques* s'attachent au début de leur hôte, comme l'italien *lo* dans l'exemple en (2), tandis que les *enclitiques* s'attachent à la fin de leur hôte, comme en (3):¹

(2) *Lo voglio fare.*

le veux-1SING faire

'Je veux le faire.'

(3) *Voglio farlo.*

veux-1SING faire-le

'Je veux le faire.'

Dans les différentes langues du monde les clitiques peuvent remplir beaucoup de fonctions grammaticales, surtout celles des mots fonctionnels (Spencer 1991:350; Wanner 1987:43ff). Ceci dit, la plupart des clitiques romans sont pronominaux, comme notre exemple italien *lo*. Mais plutôt que de les définir à partir de leur fonction grammaticale, nous allons limiter ici la classe de clitiques à ceux qui se groupent autour des verbes dans les langues romanes modernes (Wanner 1987). Ce «complexe clitique» comporte surtout des clitiques pronominaux, mais non pas exclusivement: il inclut aussi les «adverbes pronominaux» (ou «pronoms adverbiaux») *y* et *en* du français, (ainsi que leurs homologues

1. Notons en passant que ces deux phrases sont équivalentes du point de vue du sens, et que l'attachement graphique de l'enclitique *lo*, tout comme le non attachement graphique du proclitique *lo*, relève d'une convention purement orthographique qui ne correspond à aucune distinction linguistique.

descendus du latin *IBI* et *INDE* dans d'autres langues romanes quand ils sont également clitiques) et la particule de négation *ne*, ainsi que ses congénères proclitiques. La liste des éléments susceptibles d'apparaître dans le complexe clitique est donc à établir empiriquement pour chaque variété étudiée.

En tant qu'éléments linguistiques, les clitiques sont d'un grand intérêt inhérent pour la linguistique, notamment pour ce qu'ils peuvent nous apprendre sur la notion de mot («wordhood»). On observe depuis longtemps (Meyer-Lübke 1895; Nyrop 1925; Schogt 1968b) que certains clitiques (par exemple, les pronoms sujet en français) se comportent comme des flexions verbales plutôt que des mots indépendants. Comme l'observe Gougenheim (1969:153), «Les pronoms personnels conjoints entrent dans la composition du groupe verbal, alors que les pronoms personnels disjoints constituent des groupes nominaux.» Cette difficulté dans la classification des pronoms sujet persiste dans toutes les écoles linguistiques.

1.2.2. Tests formels de cliticité

Dans le cadre générativiste, Kayne énumère plusieurs tests formels pour démontrer la différence entre ces deux types de pronoms; il s'agit de caractéristiques qui sont partagées par les clitiques sujet et objet (par opposition aux syntagmes nominaux et aux pronoms forts) en français (1975, ch.2). D'abord, les pronoms clitiques ne peuvent pas être séparés de leur verbe, d'où l'agrammaticalité des phrases en (4):

- (4) * Il, paraît-il, est fou.
* Il souvent mange du fromage. (Kayne 1975:84)
* Il le souvent fait.²

Au niveau phonologique, la liaison est obligatoire entre les clitiques et leurs verbes (Kayne 1975:80,85), ce qui explique les contrastes en (5):

- (5) * Il va le_s acheter.
Il va le_s_acheter.
* Vous_s allez à Marseille?
Vous__s allez à Marseille?

Les pronoms clitiques ne peuvent pas être mis en relief par un accent d'insistance (Kayne 1975:82); cette fonction pragmatique doit être remplie en français par une circonlocution et/ou un pronom fort, comme on voit dans les contrastes en (6), où l'accent d'insistance est représenté par des lettres majuscules:

2. La formule juridique figée *Je soussigné[e], déclare . . .* semblerait violer cette contrainte, mais elle relève justement d'un registre archaïque.

(6) * Jeanne LA voit.

C'est elle que Jeanne voit. / C'est ELLE que Jeanne voit.

* JE dois partir. / MOI, je dois partir.

C'est MOI qui dois partir. Je dois aller en classe, MOI.

* ILS vont à Aix.

EUX ils vont à Aix. / EUX vont à Aix. / Ils vont à Aix, EUX.

Enfin, il est impossible de coordonner des pronoms clitiques, à la différence des pronoms forts (Kayne 1975:83):

(7) * Dominique et je partirons demain.

* Je et Dominique partirons demain.

Dominique et moi partirons demain.

Dominique et Claude partiront demain.

Ces tests ou diagnostics sont donc des indices de «cliticité», pour ainsi dire, surtout dans le cas de certains pronoms français, mais certains (comme par exemple le test de la liaison) ne sont pas applicables aux clitiques dans d'autres langues. Ce sont donc des propriétés citées en tant qu'illustrations (plutôt que des définitions) des phénomènes qui caractérisent les clitiques. Nous comptons trouver dans d'autres langues romanes certains traits communs, tels l'inséparabilité des clitiques de leurs verbes.

1.2.3. Statut ambigu des clitiques

En général, les clitiques accusent une forte dépendance phonologique vis-à-vis de leur hôte; souvent, il existe des processus phonologiques particuliers (voir la liaison en (5) ci-dessus) qui sélectionnent les clitiques. Ainsi on trouve qu'en portugais les pronoms clitiques objet ont la forme /o, a, os, as/ mais deviennent /lo, la, los, las/ après une forme verbale qui se termine en /r/ ou /s/, et /no, na, nos, nas/ après une forme qui se termine en nasale:

(8)	<i>levo</i>	'je lève'
	<i>levas</i>	'tu lèves'
	<i>levam</i>	'elles lèvent'
	<i>levo-o</i>	'je le lève'
	<i>leva-lo</i>	'tu le lèves'
	<i>levam-no</i>	'elles le lèvent'

(Spencer 1991:363)

Ce genre d'allomorphie n'a rien d'extraordinaire dans le contexte des affixes morphologiques, où on doit s'attendre à trouver des formes spéciales qui changent selon les contextes phonologiques. Les formes contractées qui résultent de la combinaison des clitiques accusatif et datif en portugais renforcent cette impression d'«opacité» caractéristique de la morphologie (par opposition à la syntaxe):

(9)	<i>mostr-as</i>	‘montre-les’
	<i>mostra-me</i>	‘montre-moi
	<i>mostra-mas</i>	‘montre-les-moi’

(Spencer 1991:366)

Il existe également bien des processus phonologiques, comme le relâchement vocalique («vowel laxing») ou les métaphonies asturienne et sud-italienne, qui ont pour domaine le mot phonologique; ce domaine peut très bien équivaloir à des suites verbe+clitique ou clitique+verbe (Durand 1990:140), ce qui renforce de nouveau l'idée que les clitiques sont de nature plutôt affixale.

Si la forte relation phonologique qui existe entre un clitique et son hôte tend à souligner les ressemblances entre la cliticisation et l'affixation, la mobilité (même limitée) des clitiques fait ressortir leur statut distinct d'éléments doués d'une certaine liberté syntaxique. Les deux formes équivalentes qu'on voit en (2) et (3) correspondent au phénomène connu comme la montée des clitiques (*clitic climbing*), qui caractérise la plupart des langues romanes à part le français moderne (Vincent 1988:69). Notons que des phrases comme (10)b. étaient grammaticales en ancien et moyen français, et ne sont disparues qu'au XVIIIe siècle.

- (10) a. Je peux le voir.
b. * Je le peux voir.³

Nous évoquons ces différents phénomènes ici pour souligner le fait que les clitiques participent à des processus syntaxiques aussi bien que phonologiques. C'est justement cette double appartenance qui les rend difficiles à «caser» dans un modèle linguistique. C'est cet aspect ambigu des clitiques, relevant à la fois de la phonologie et de la syntaxe, de la morphologie et de la sémantique, qui permet donc une vue privilégiée sur l'interaction des différents «modules» (ou composantes) d'une grammaire. Mais la nature hétérogène de cette classe d'éléments linguistiques peut également pousser certains à rejeter l'idée de «clitique» comme un concept utile pour la théorie linguistique:

These difficulties of characterization have led some linguists to abandon the notion of clitic altogether as a theoretical primitive and regard the notion of clitic as simply a descriptive cover term. Other linguists, however, regard clitics as a separately identifiable morphosyntactic category. (Spencer 1991:350)

Et encore:

3. Cette construction est néanmoins employée à l'époque moderne à des fins stylistiques ou expressives, par des écrivains comme Gide.

The relevance of the notion of clitics has been taken for granted in the foregoing discussion. . . . Not all comprehensive category notions used in linguistics are thereby necessarily fully defined. . . . It remains to be seen where the notion of clitics belongs realistically, whether to the definable categories, to the meaningful prescientific notions, or to the pile of empty convenience labels.

(Wanner 1987:52-3)

S'il est vrai que les clitiques se placent au «carrefour», pour ainsi dire, des différentes sous-disciplines de la linguistique, la question des relations entre les différents modules d'une grammaire devient cruciale. Une erreur fondamentale qui est très répandue dans les études antérieures est la supposition que tous les phénomènes qui sont décrits informellement comme des clitiques sont nécessairement à traiter comme appartenant à un seul module; l'hypothèse de travail ici sera que les clitiques représentent une classe hétérogène, et que la distribution des données (plutôt que des *a priori* théoriques) déterminera le choix du module et du formalisme. La détermination de la place (ou des places) qu'occupent les clitiques dans la théorie linguistique représente donc un but de cette recherche plutôt qu'un point de départ:

The absence at this point of a well defined clitic notion is, however, not as serious as might appear at first glance. The general preoccupation with an explanatory level of scientific linguistic endeavors has overaccentuated the degree of failure of

a prescientific approach to a broad empirical phenomenon. The rigid definitional approach becomes fruitful only at the point where the comprehensive account of an existing amount of research becomes unavoidable. On the other hand, in a case such as clitic studies, where the uncertainties about clitic behavior, function and meaning still are overwhelming, the risks contained in the vague clitic notion propagated here may be far outweighed by the gains achieved through an open research orientation which is not constrained artificially into recognizing only situations conforming to the theory. (Wanner 1987:55-56).

Pour ce qui est des pronoms sujet qui nous intéressent tout particulièrement ici, le débat continue quant à leur statut. Rizzi (1986) affirme que les sujets clitiques ne constituent pas une classe homogène d'éléments linguistiques et leurs traitements formels doivent donc se faire de façon hétérogène, tandis que Roberge et Vinet (1989) plaident le cas d'une analyse unifiée de tous les clitiques (voir 1.3.4.).

1.3. APPROCHES ANTÉRIEURES

Bien qu'il soit impossible de traiter ici toute la documentation accumulée sur ce sujet, il est sans doute utile dès maintenant de jeter un coup d'oeil sur certaines des études les plus significatives.

1.3.1. Classifications

La tâche primordiale de toute science étant essentiellement taxinomique, Zwicky (1977) présente une typologie tripartite qui divise les clitiques en «simple clitics», «special clitics» et «bound words». Cette dernière catégorie, les mots liés, ne nous concernera pas dans la présente étude. Les clitiques simples sont des formes phonologiquement réduites qui apparaissent exactement dans les mêmes positions que les formes fortes (=accentuées) correspondantes. Les clitiques simples ne caractérisent pas, en général, les langues romanes: les pronoms clitiques dans les créoles à base lexicale romane constituent une exception à cette généralisation:

Perhaps the most salient difference in sentence structure between European Romance and creoles lies in the treatment of pronominal objects. In creole, these follow the verb and have the same privilege of occurrence as full object noun phrases. . . . Like nominal objects, pronouns can also be directly co-ordinated with other noun phrases. . . . Although pronouns are phonetically insubstantial and liable to morphophonemic reduction making them enclitic on the verb, . . . they are in principle free forms, and therefore no creole has an obligatory rule of clitic doubling. (Green 1988:458)

Les exemples fournis par Green démontrent que les pronoms objet créoles ont exactement la même distribution que les syntagmes nominaux normaux puisqu'ils sont en fait des

pronoms pleins qui sont parfois susceptibles d'être phonétiquement réduits. Il suffit de comparer les tests de cliticité de Kayne (1975, voir ci-dessus 1.2.2.) avec des phrases semblables dans différents créoles à base lexicale romane:

(11) *mi ta mira bo, . . . bo tata*

'je IMPERFECTIF voir toi, . . . ton père' (papiamentu)

u kaj ale we jo, . . . we kuzin jo

'tu FUTUR aller voir eux, . . . cousins-leurs' (saint lucien)

mwê te kônê jo epi kamwad jo

'je PERFECTIF connaître eux et camarades-leurs' (haïtien)

(Green 1988:458)

Dans le reste du domaine roman, la plupart des clitiques dans les variétés modernes sont des clitiques «spéciaux» d'après la taxinomie de Zwicky (1977), puisqu'ils se trouvent en distribution complémentaire avec les N^o (syntagmes nominaux) normaux: un clitique objet ne peut pas apparaître dans les contextes où un syntagme nominal le pourrait, et inversement. (voir 1.2.2.) Les pronoms sujet représentent un cas un peu exceptionnel à cet égard, puisque leur position normale est souvent celle qu'occupent aussi les syntagmes nominaux sujet. Le fait que l'on trouve des sujets clitiques qui redoublent un sujet nominal

suggère néanmoins qu'il s'agit bien de deux classes distinctes d'éléments qui se placent dans deux positions syntaxiques différentes. (voir 1.4.)

1.3.2. Contraintes prosodiques

Dans les vernaculaires romans médiévaux, les pronoms clitiques avaient plus de liberté de position, n'étant pas contraints à s'ancrer à un hôte verbal. Ils se situaient plutôt dans le domaine prosodique de la phrase (ou mieux, de la proposition) conformément aux «lois» dites de Wackernagel et de Tobler-Mussafia (1886). La loi de Wackernagel (Wanner 1991:314) spécifie que les pronoms atones (=clitiques) doivent apparaître en «deuxième position», c'est-à-dire, après le premier élément accentué de la phrase. Cette loi (ou, plus prudemment, cette tendance) s'appliquait indépendamment du temps ou du mode du verbe, et ne prenait pas en considération les conjonctions qui unissaient deux propositions coordonnées:

En quines condicions els pronoms àtons s'avantposaven o bé es postposaven al verb en l'Edat Mitjana és un tema a penes tractat Calia començar subratllant que aleshores la tria entre l'enclisi i la proclisi era independent del temps o mode del verb; el pronom es posava normalment, qualsevol que fos la forma verbal, immediatament darrere la primera paraula de l'oració (deixant de banda algunes

excepcions, i en particular, de vegades, les conjuncions unitives, que en rigor no pertanyen a cap de les dues oracions). . . (Coromines 1974:363-5)⁴

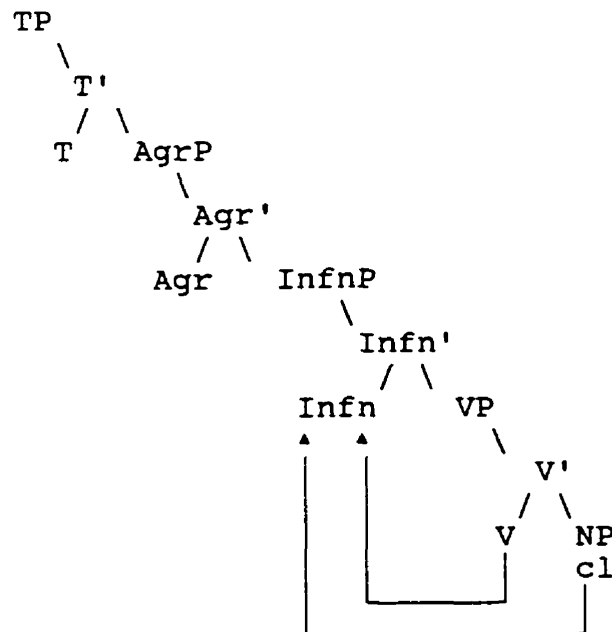
La loi de Tobler-Mussafia précise simplement qu'un pronom atone ne peut jamais apparaître en initiale absolue de phrase (Wanner 1991:314). Cette tendance s'appliquait dans les vernaculaires romans anciens, mais elle est disparue dans les langues romanes modernes. Elle persiste dans une forme atténuée en portugais européen (Parkinson 1988:158; Spencer 1990:363ff), où les pronoms objet apparaissent comme enclitiques aux verbes dans des phrases simples, positives, affirmatives. Les mêmes pronoms peuvent être proclitiques après un élément interrogatif, une négation, dans une subordonnée, quand le sujet du verbe est un syntagme nominal quantifié ou quand le verbe en question est un infinitif gouverné par une préposition: c'est-à-dire, l'enclise est obligatoire uniquement quand le verbe est en position initiale de proposition (voir 5.4.7).

4. «Dans quelles conditions les pronoms atones se plaçaient avant ou après le verbe au Moyen Age est un sujet à peine traité Il fallait commencer en soulignant qu'à cette époque-là le choix entre l'enclise et la proclise était indépendant du temps ou du mode du verbe; le pronom se plaçait normalement, quelle que soit la forme verbale, immédiatement après le premier mot de la phrase (faisant abstraction de quelques exceptions, et en particulier parfois des conjonctions de coordination, qui à la limite n'appartiennent à aucune des deux phrases). . . »

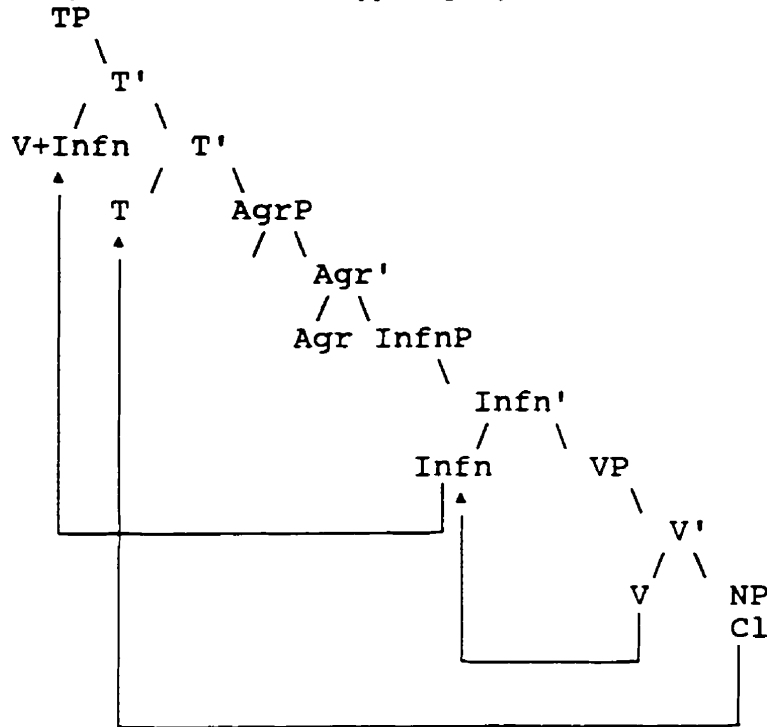
1.3.3. Analyses syntaxiques génératives

On voit depuis une quinzaine d'années un certain nombre de tentatives en syntaxe générative d'«expliquer» (ou plutôt, représenter formellement) l'ordre des éléments flexionnels et les différents clitiques qui s'attachent, par exemple, aux verbes de type roman, par le biais de mouvements syntaxiques («verb-raising» voir Kayne 1991, entre autres). Cette approche présuppose que les clitiques sont «générés» comme des NP et qu'ils se déplacent successivement à travers les noeuds d'accord (AGR), T[emps] et INFN (=infinitif) pour apparaître à la «bonne» position de surface. La différence entre le français *je veux le faire* et l'italien *voglio farlo* découlerait des différents déplacements représentés par les arbres syntaxiques en (12) et en (13) respectivement:

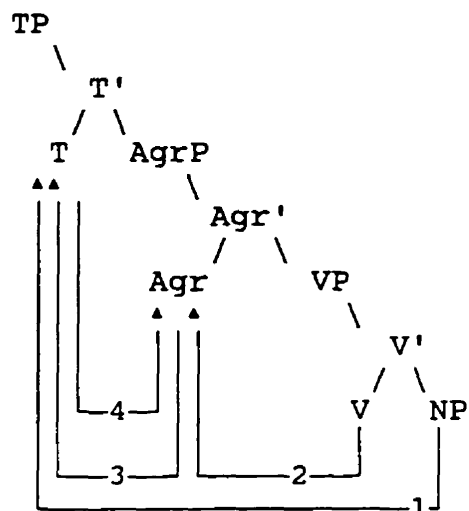
(12) français (infinitif enchâssé du type *je veux le faire*)



(13) italien (infinitif enchâssé du type *voglio farlo*)



Une telle approche risque toujours d'être téléologique, puisque ces déplacements ne sont postulés que pour rendre compte de l'ordre des éléments, n'ayant aucune motivation indépendante. On peut également démontrer formellement que la variation synchronique documentée parmi les clitiques romans ne peut pas être analysée de cette façon sans réduire sérieusement le pouvoir explicatif du modèle syntaxique (Bessler et. al. 1992). Pour ne donner qu'un exemple problématique: dans certaines variétés de l'espagnol, la désinence /-n/ du pluriel peut se «déplacer» de sa position canonique (voir (14)a.) pour apparaître soit après un pronom enclitique, comme en (14)b., soit aux deux endroits en même temps, comme en (14)c):

(15) *Siéntenseñ* (variante redoublée)

On peut facilement trouver d'autres données similaires, tirées d'autres variétés romanes, qui renforcent la conclusion que le mouvement syntaxique n'est pas le mécanisme le plus apte pour rendre compte de ce genre de faits. Dans chaque cas, l'ordre superficiel des morphèmes est tel qu'il serait difficile de l'obtenir par des déplacements syntaxiques. Par exemple, en milanais on trouve un clitique réfléchi «à l'intérieur», pour ainsi dire, de la désinence de la première personne du pluriel, comme (16), et des clitiques qui redoublent le pronom sujet fort à la fois avant et après le verbe, comme en (17):

(16) *Ripusèseme*

reposer + REFLEXIVE + 1 PLURIEL

'Reposons-nous.'

(Rohlf's 1966: 175)

(17) *ti te parlet*

toi tu parles-tu

(Nicoli 1983:138)

'Tu parles.'

De même en portugais, où les clitiques apparaissent entre la racine verbale et la marque de temps et aspect, on voit mal comment motiver les déplacements syntaxiques qui seraient nécessaires pour obtenir ces faits:

(18) *levá-lo-ei*

lever-le-FUTUR + je

'Je le leverai.'

levá-lo-ias

lever-le-CONDITIONNEL + tu

'Tu le leverais.'

(Spencer 1991:365-6)

En plus de constituer des problèmes sérieux pour une analyse par déplacement syntaxique, les données présentées ici ont en commun le fait qu'elles

. . . represent language- or dialect-specific properties of clitics. It must be noted, however, that these properties refer only to the realization of the clitics, and not to the argument positions which they represent. It is for this reason that a syntactic analysis, in which the syntactic and morphological properties of a clitic are inseparable, fails to account for the data straightforwardly. (Heap et al. 1993)

Les linguistes qui s'intéressent à ce genre de faits s'interrogent donc sur les autres mécanismes formels qui pourraient mieux représenter le comportement et la distribution des clitiques, et notamment, sur les formalismes proprement morphologiques.

1.3.4. Modèles morphologiques

Si, au lieu de voir le placement des pronoms comme une question de mouvements en syntaxe, on admet l'hypothèse que le type de variation dont il s'agit ici est plus caractéristique de la morphologie que de la syntaxe (c'est-à-dire, il ressemble plus à l'allomorphie flexionnelle qu'au mouvement d'éléments indépendants d'une phrase), on peut explorer la possibilité de représenter ces phénomènes par les formalismes disponibles actuellement en (morpho-) phonologie nonlinéaire ou associative. Un exemple prometteur de cette approche seraient les schèmes («templates») morphologiques, un mécanisme connu en phonologie nonlinéaire qu'on

peut exploiter en morphosyntaxe pour exprimer formellement l'ordre des éléments de façon économique, sans avoir recours aux mouvements en syntaxe. Le principal problème qu'implique cette approche est l'absence de contraintes sur ce formalisme, ce qui reflète le manque de travail théorique sur ce modèle schématique («templatic»). C'est une situation qui fait ressortir le besoin de nouvelles recherches dans le domaine, comme l'affirme Spencer (1991:212): «The problem of template morphology has not been widely discussed in the generative literature. One of the few contributions to debate the problem is Simpson and Withgott (1986).»⁵ Et encore, «. . . introducing the template concept into generative grammar leaves us with the question of how to formalize the notion, and how to accommodate it into a principled universal theory of word structure.» (1990:214)

Le travail de Bonet (1991) représente une tentative intéressante de décrire formellement la distribution et l'ordre des clitiques par le biais de schèmes morphologiques. Son projet est limité dès le départ par l'exclusion des pronoms sujet, bien qu'elle suggère qu'ils pourraient être traités par le même mécanisme. Il nous semble en fait qu'un tel formalisme devrait se poser comme but de traiter tous les

5. Spencer continue en note (1991:469, note 12) «Simpson and Withgott are noncommittal about how they would formalize the notion of template. The only formal models I know of which take the notion of template seriously are those of the Soviet linguist Revzin (e.g. Revzin and Juldaseva, 1969) and the model proposed by Grimes (1967, 1983).» De ces ouvrages, les travaux de Grimes fournissent des schèmes comme des «discovery procedures» sans s'adresser à la question de leur place dans un modèle grammatical, tandis que les travaux de Revzin nous sont inaccessibles parce qu'en russe.

clitiques, y compris les clitiques sujet, mais sans perdre de vue la nature hétérogène de la classe des clitiques (voir 1.2.3). Une autre idée qui est centrale au travail de Bonet est l'observation que les changements que peuvent subir des groupes de clitiques donnent toujours lieu à une forme qui correspond à un clitique déjà existant dans le système: «the output of the rule always coincides with an independently existing clitic.» (1991:11) Il est vrai qu'il existe une tendance marquée vers ce genre de solution, mais la généralisation telle qu'énoncée par Bonet est trop forte. En ancien espagnol, par exemple, une évolution phonétique normale fait qu'une suite d'un clitique objet indirect suivi d'un clitique objet direct, comme en (19), donne lieu à une forme de surface comme celle de (20), sans qu'il existe indépendamment une forme *ge* comme clitique.

(19) *le + lo*

'lui + le'

(20) *gelo* **ge*

'lui-le' 'lui'

Il semble donc plus prudent d'exprimer la généralisation de Bonet en termes de marque («markedness») la non-existence d'un clitique *ge* a sans doute exercé une pression sur la forme *gelo*, qui a fini par se convertir en *selo*, qui est une forme moins marquée dans

ce sens qu'elle est composée de deux clitiques (*se* et *lo*) déjà existants ailleurs dans le système.

Cette même tendance favorisant des suites de clitiques composées de formes existantes peut également exercer une pression en faveur de la création d'une nouvelle forme. En ancien galicien-portugais, par exemple, des suites de pronoms comme celles en (21) se réalisaient phonétiquement par les contractions en (22), ce qui donnait lieu (Parkinson 1988:148) aux nouveaux clitiques simples en (23):

(21)	<i>te + o</i> > [tjo]	<i>se + o</i> > [sjo]
	2SING INDIRECT + 3SING MASC	3SING INDIRECT + 3SING MASC
	'te le.'	'se le.'
(22)	<i>cho</i> [tʃo]	<i>xo</i> [ʃo]
(23)	<i>che</i> [tʃe]	<i>xe</i> [ʃe]
	2SING INDIRECT	3SING INDIRECT
	'te.'	'se.'

Encore une fois, on voit ici la généralisation de Bonet sous une forme atténuée: les suites en (21) sont permises, bien qu'elles ne soient pas composées de clitiques déjà existants dans la langue, mais elles motivent une réanalyse (*backformation*) qui donne lieu à de nouveaux clitiques qui, eux, sont conformes à la généralisation de Bonet. On

peut sans doute trouver d'autres exceptions à la généralisation de Bonet parmi les variétés du domaine roman: l'intérêt de cette généralisation est justement qu'elle explique la rareté relative de ces formes, ainsi que le fait qu'elles tendent à être réanalysées de façon à se conformer à la tendance générale (voir Heap 1996 pour une interprétation de cette tendance dans le cadre de la Théorie de l'Optimalité).

Il faut noter aussi que Bonet (1991:12ff) suppose l'existence d'une composante morphologique («morphological component»), d'après la conception de Halle (1973, 1989), qui applique des «règles d'ajustement» entre la sortie («output») du module syntaxique et l'entrée au module phonologique. Elle avoue cependant, que la composante d'interface («interface component») de Zwicky et Pullum (1986) pourrait accomplir les mêmes tâches entre le lexique et la structure-D qu'elle envisage dans sa composante morphologique.⁶

Bonet n'envisage pas cependant que cette composante morphologique puisse traiter les pronoms sujet, même quand ceux-ci sont des clitiques (1991:55). En fait, la plupart des travaux sur la morphologie pronominale font abstraction des pronoms sujet (quand ils ne les passent pas sous silence complètement): Auger (1995), qui applique les «processus morphologiques» d'Anderson (1991) aux sujets pronominaux du français

6. Cette même approche se retrouve plus récemment dans le programme minimaliste de Chomsky (1995:65), où les faits morphologiques, et notamment tout ce qui est susceptible de varier dans une grammaire, sont relégués aux «interfaces» entre les modules, et en particulier, à l'interface qui mène de la Composante computationnelle (=la syntaxe) à la forme phonologique. De cette façon, on obtient aussi un maximum d'universalité dans la syntaxe.

québécois familier, constitue une exception notable à cette tendance. Pour la plupart, on reconnaît implicitement (et ce depuis longtemps, voir 1.4.1.) que les pronoms sujet des variétés romanes centrales ont quelque chose d'assez particulier, quel que soit le cadre théorique que l'on adopte. Ce que nous retenons de cette discussion des pronoms objet dans le cadre génératif, c'est simplement le fait qu'il existe deux approches bien différenciées: soit on les traite comme des objets semblables aux mots, qui se déplacent donc en syntaxe, soit on les considère plutôt comme des affixes ou des désinences, qui sont à manipuler dans la composante morphologique de la grammaire.

1.4. PRONOMS SUJET

1.4.1. De la philologie traditionnelle à la linguistique structuraliste

Il s'agit évidemment d'une différence typologique connue depuis fort longtemps, et qui apparaît aussi (sous des formulations différentes) dans la plupart des manuels traditionnels de linguistique romane, par exemple Rohlfs:

. . . l'italiano (contrariamente al francese) dispone di desinenze distinte per le singole forme verbali, il pronome personale risulta superfluo all'espressione delle persone.⁷ (1966:145)

7. «L'italien (contrairement au français) dispose de désinences distinctes pour les différentes formes verbales, le pronom personnel s'avère superflu pour ce qui est de l'expression de la personne.»

Il est également bien connu (Grassi 1970:114) que les variétés de l'Italie septentrionale, bien que souvent pourvues de désinences verbales distinctes pour les différentes personnes grammaticales, se regroupent *grosso modo* avec le français plutôt qu'avec l'italien standard pour ce qui est de la présence des pronoms sujet:

Ben diversamente stanno le cose nel Settentrione. . . . Nei dialetti moderni invece il pronome è divenuto quasi ovunque complemento obbligatorio della forma verbale. Ora compare soltanto nelle forme toniche (ven. *mi son, lu ga*), ora solo nella forma ridotta (mil. *a sont* 'io sono'), ora il verbo è preceduto da entrambe le forme, tonica e atona (ven. *ti ti gavarà* 'tu avrai'). (Rohlf 1966:146)⁸

Dans la documentation philologique traditionnelle, on s'occupe surtout de la distribution géolinguistique de ce phénomène, qu'on met souvent en relation avec le superstrat germanique:

8. «Les choses sont bien différentes au Nord. . . Dans les dialectes modernes par contre le pronom est devenu presque partout un complément obligatoire de la forme verbale. Parfois il apparaît seulement sous la forme tonique (vénitien *mi son, lu ga*), parfois seulement sous la forme réduite (milanais *a sont* 'io sono'), parfois le verbe est précédé par les deux formes, toniques et atones (vénitien *ti ti gavarà* 'tu avrai').»

. . . d'où provient le fait que, dans une grande partie de la Romania, le sujet est exprimé à la fois par la désinence verbale et par le pronom personnel. Il constate que cette particularité est limitée au français, au franco-provençal, au rhéto-roman, et aux dialectes de l'Italie septentrionale auxquels se joint encore le coin N.-O. de la Toscane. . . . à l'origine de cette particularité doit se trouver la compénétration intime des peuples romans et germaniques dans ces régions. (Kuen, d'après von Wartburg, 1967:103-104).

Déjà chez les philologues et les dialectologues, ainsi que chez beaucoup de linguistes modernes, la propriété qu'ont les langues comme l'espagnol et l'italien standard de permettre à un verbe fléchi d'apparaître sans que la position structurale du sujet soit remplie, est souvent mise en relation avec la présence d'une morphologie «riche», c'est-à-dire, de terminaisons verbales suffisamment distinctes pour différencier les personnes grammaticales (Meyer-Lübke 1895:107; Nyrop 1925; Schogt 1968; Gougenheim 1969). Mais nous verrons que cette relation intuitive (souvent affirmée sans argumentation précise) ne peut pas rendre compte de tous les cas effectivement attestés dans les variétés que nous étudions.

1.4.2. Le «paramètre du sujet nul»

Comme il est souvent le cas, cette problématique traditionnelle se voit reprise et répétée dans le cadre générativiste, d'abord par Perlmutter (1971). Ensuite Chomsky (1981:240) propose que cette propriété de permettre des «sujets nuls» est en fait reliée à d'autres phénomènes syntaxiques.⁹ Il propose donc de réduire tous ces phénomènes à un seul aspect de la Grammaire Universelle (désormais GU), qu'on appelle le «paramètre du sujet nul».¹⁰ L'idée est assez séduisante: la faculté langagière innée fournirait un choix binaire abstrait dénommé un «paramètre», qui n'attendrait que des données concrètes d'une grammaire particulière pour être réglé soit positivement soit négativement chez l'enfant qui acquiert une langue. Ce réglage serait à son tour responsable de l'acquisition de toute une série de conséquences prédéterminées, soit des caractéristiques grammaticales qu'il serait autrement difficile de mettre en relation. L'italien standard (ainsi que l'espagnol standard, le latin, et le roumain) serait d'après cette typologie une grammaire «à sujet nul» ou [+sujet \emptyset] (Chomsky 1981:55ff); le français standard, en revanche, serait une grammaire «à sujet obligatoire» ou [-sujet \emptyset], tout comme l'anglais.

9. Les propriétés qui seraient reliées au paramètre du sujet nul sont: l'inversion libre, l'extraction d'éléments interrogatifs (QU-) à longue distance, et le filtre *que*. On n'examine pas ces propriétés en détail ici puisque leur association avec le paramètre du sujet nul se voit infirmée par des recherches plus récentes (voir 1.4.3.).

10. Les termes anglais «null subject parameter» et «pro-drop parameter» sont à peu près équivalents; pour éviter des confusions éventuelles, nous utilisons ici le premier.

Rizzi (1986) adopte cette typologie paramétrique des sujets pour considérer le statut des sujets clitiques dans les variétés romanes (en fait, dans les variétés standard de l'italien et du français, et dans quelques variétés non standard de l'Italie septentrionale et du Midi français). Il conclut que la notion de «sujet clitique» ne représente pas une catégorie réelle d'un point de vue grammatical, puisque les éléments qui reçoivent ce nom peuvent, d'après lui, apparaître dans des positions structurales diverses dans des variétés différentes. Concrètement, il propose une analyse d'après laquelle les pronoms sujet clitiques en français standard occupent la position du sujet, soit [N",I"], c'est-à-dire le syntagme nominal qui est la soeur du syntagme flexionnel, tandis que dans les variétés nord-italiennes (trentino, fiorentino) ils apparaissent sous le noeud INFL (= flexion, anglais *inflection*), la projection qui contient le verbe. Rizzi maintient donc comme primordiale la division typologique entre grammaires à sujet nul (tant l'italien comme les variétés de l'Italie septentrionale) et grammaires à sujet obligatoire (comme le français standard). Ce point de vue est repris dans Renzi et Vanelli, qui commencent leur étude des pronoms sujet dans les variétés nord-italiennes par une présentation assez tranchante de cette bifurcation typologique fondamentale:

Nelle frasi con soggetto pronominale la persona può essere espressa o da un pronome o dalla semplice flessione del verbo (o da entrambi). Concentrando la nostra attenzione sui pronomi, quindi, possiamo dire che ci sono *lingue che*

*fanno uso costante del pronome soggetto e lingue che non ne fanno un uso costante.*¹¹ (1983:122, notre emphase).

Pourtant, cette formulation strictement bipolaire est aussitôt nuancée par les auteurs qui observent qu'à l'intérieur de chaque type il existe des emplois «marqués» ou exceptionnels: des cas où les grammaires qui font un usage «constant» des pronoms peuvent néanmoins les omettre, et d'autres cas où les grammaires qui n'en font pas un usage constant les admettent (facultativement ou obligatoirement selon le contexte)¹². Mais l'essentiel pour leur propos, c'est la distinction robuste entre ces deux possibilités, qui peut même apparaître à l'intérieur d'une seule grammaire: «possono esistere lingue in cui c'è un uso alternato dei due modelli a seconda della persona» (1983:122).¹³ Ce sont en fait ces paradigmes pronominaux «mixtes» qui constituent l'un des objets

11. «Dans les phrases avec sujet pronominal la personne peut être exprimée soit par un pronom soit par la simple flexion verbale (soit par les deux). Concentrant notre attention sur les pronoms, donc, nous pouvons dire qu'il y a *des langues qui font un usage catégorique du pronom sujet et des langues qui n'en font pas un usage catégorique.*» (notre emphase)

12. Voir aussi Rohlfs, qui précise dans une note (1966:145) que l'italien standard exige exceptionnellement un pronom sujet dans certains contextes, par exemple pour lever l'ambiguïté des cas où des désinences verbales ne sont pas distinctes: *voglio che tu venga, voglio che lui venga, vuole che io venga*. Pour Renzi et Vanelli (1983:134, n.17), cette désambiguïssation serait obligatoire uniquement quand elle «force» une interprétation du sujet comme deuxième personne. Freedman (1983) examine également des cas où les pronoms sujet sont usuels en italien moderne.

13. «Il peut exister des langues dans lesquelles il y a un emploi alterné des deux modèles selon les personnes.»

principaux de leur article pionnier, et que nous examinons en détail dans la présente étude.

Notons que cette vision de la variation grammaticale s'éloigne déjà du modèle chomskien des paramètres «classiques», où le choix binaire [\pm sujet \emptyset] se ferait nécessairement une seule fois pour une grammaire tout entière. Il ne s'agirait plus de «grammaires à sujet nul» qui s'opposent uniquement aux «grammaires à sujet obligatoire», mais aussi de «grammaires à sujet nul pour certaines personnes grammaticales seulement, et à sujet obligatoire pour les autres», ce qui brouille considérablement l'élégance de la formulation originale du concept du dit «paramètre du sujet nul». Notons que d'après Borer (1984), ce paramètre serait fixé au niveau de chaque élément lexical, donc pour chaque pronom correspondant à une personne grammaticale différente; on voit mal, cependant, comment corrélérer de tels «réglages» individuels avec les autres propriétés qui feraient partie du même paramètre.

1.4.3. Contre le paramètre du sujet nul

Nous ne pouvons pas passer en revue ici toute la documentation relative aux changements et aux critiques qu'a subis la notion du paramètre du sujet nul au cours des quinze dernières années (voir Roberge 1990, pour une discussion plus détaillée). Il suffit pour notre propos de noter que la relation entre les différents phénomènes réunis sous la rubrique de ce «paramètre» s'avère en fait moins étroite que l'on ne le pensait au

début. Safir (1985) démontre notamment que l'inversion libre et les sujets nuls apparaissent de façon indépendante dans les variétés romanes. Pour Wanner, le paramètre du sujet nul ne correspond qu'à «la stipulation de la légitimation de la catégorie vide *pro* par l'élément lexical FLEX, sans autre conséquence immédiate». (1995:453)

Pour ce qui est des conséquences logiques de l'approche paramétrique à la variation grammaticale, Roberge et Vinet (1989) montrent que la combinaison des deux paramètres [\pm sujet \emptyset] et [\pm sujet clitique] produit une classification typologique comme celle qui suit:

Tableau 1: Typologie selon les traits [\pm sujet clitique] >> [\pm sujet \emptyset]:

+ sujet clitique		- sujet clitique	
-sujet \emptyset	+ sujet \emptyset	-sujet \emptyset	+ sujet \emptyset
français standard pied-noir québécois frioulan ...	?	anglais allemand romanche ...	italien espagnol occitan gévaudanais ...

(d'après Roberge et Vinet, 1989:57,[68])

Cette typologie met en relief la «non-existence» de grammaires qui soient simultanément [\pm sujet clitique] et [\pm sujet \emptyset], qui correspondraient à la case ? dans le Tableau 1. De l'existence de ce «trou typologique», Roberge et Vinet tirent la généralisation que, dans

une grammaire qui possède des sujets clitiques, on ne peut jamais trouver un verbe fléchi dans une proposition simple affirmative qui n'a ni un sujet lexical ni un sujet clitique, c'est-à-dire en termes formels que la configuration suivante n'apparaît jamais (1989:57,[69]):

- (24) * [SYNTAGME NOMINAL \emptyset] [_{cl} \emptyset] V
* 'vient.' (affirmatif)

Cette généralisation, qui présente les faits de façon négative en «interdisant» ou en «filtrant» une configuration non attestée, est reprise par Roberge (1990), qui la reformule positivement:

- (25) If the grammar of L includes subject clitic pronouns then, when [NP,IP] is phonetically null and filled by pro, a subject clitic must be present.
(1990:38,[6])

D'après cette analyse, les pronoms clitiques sujet en français standard seraient «générés à la base» dans la même position structurale que les clitiques sujet dans les variétés de l'Italie septentrionale, c'est-à-dire, dans le noeud INFL (1990:65). Roberge soutient donc, contrairement à l'analyse de Rizzi et d'autres, que les clitiques sujet représentent

en fait une classe grammaticale réelle et uniforme, et que la variation dans leur comportement distributionnel serait due aux «differences in the intrinsic properties of the subject clitics in each language.» (1990:56) Ces propriétés et leur interaction avec «other independently motivated principles of UG» (1990:65), seraient suffisantes pour rendre compte des différences comme l'inversion, le redoublement, la coordination etc., que Rizzi explique par le biais de différences paramétriques.

Notons qu'il ne serait pas nécessaire de stipuler la généralisation de Roberge donnée en (25), puisqu'elle découlerait de principes plus généraux de la GU, à savoir, les conditions sur l'apparition de *pro* et le principe de la récupérabilité. Cela veut dire que dans le modèle grammatical de Roberge, il n'existe pas de «paramètre du sujet nul» (1990:68): la propriété du sujet nul serait une conséquence des principes de la GU. Le problème du «trou typologique» dans le Tableau 1 disparaît alors, puisque les grammaires à clitiques sujet constitueraient un sous-type des grammaires à sujet nul:

Tableau 2: Typologie selon les traits [\pm sujet \emptyset] >> [\pm sujet clitique]:

+ sujet \emptyset		- sujet \emptyset
+ sujet clitique	- sujet clitique	anglais allemand romanche ...
français standard pied-noir québécois frioulan ...	italien standard espagnol occitan gévaudanais ...	

(d'après Roberge et Vinet 1989:58)

Il est important de noter cependant que dans ce modèle les dits «paramètres» [\pm sujet \emptyset] et [\pm sujet clitique] ne jouissent plus du même statut que dans celui de Rizzi et de Chomsky: ce ne sont que des paramètres «superficiels» qui servent à décrire les faits de façon systématique, et ne font aucunement partie des paramètres «profonds» ou «explicatifs» qui eux seraient inscrits dans la GU et réglés lors de l'acquisition (voir Roberge 1989).

1.5. LA VARIATION GRAMMATICALE

Pour ce qui nous concerne, la question de l'existence ou de la non-existence du paramètre du sujet nul conditionne notre façon d'envisager l'objet de cette étude. Si, d'une part, on accepte l'existence d'un paramètre du sujet nul, il faut expliquer comment un trait binaire peut rendre compte du type de variation grammaticale qu'on trouve dans les variétés étudiées. Or, une étude préliminaire (Heap 1993) suggère que la variation qui existe dans les paradigmes des pronoms sujet dans ces variétés dépasse ce que l'on peut capter à l'aide de cette simple opposition binaire [\pm sujet \emptyset], même si ce paramètre s'applique à chaque personne grammaticale d'un paradigme. Il s'agit d'un autre type de variation pronominale: les pronoms sujet qui apparaissent variablement dans une seule et même personne grammaticale d'une grammaire donnée, soit «les grammaires parfois à sujet nul pour certaines personnes grammaticales». Il serait bien sûr possible d'étendre l'application du trait binaire [\pm sujet \emptyset] au niveau de chaque

personne grammaticale plutôt qu'à une grammaire tout entière (c'est vraisemblablement ainsi qu'un tel trait pourrait être acquis par les enfants, voir Wexler et Manzini 1987), mais on voit mal comment un simple trait binaire pourrait rendre compte d'une grammaire où la présence d'un pronom sujet varie à l'intérieur d'une personne grammaticale. Dans cette optique donc, l'étude du comportement variable des pronoms sujet constitue en quelque sorte un test empirique de la théorie des paramètres binaires dans la variation grammaticale, et plus particulièrement, de l'adéquation descriptive du (désormais plutôt hypothétique) «paramètre du sujet nul».

Si en revanche on adopte un modèle qui rejette l'existence d'un «paramètre du sujet nul», la question empirique suivante se pose: quelles seraient ces «propriétés intrinsèques» des différentes variétés à clitiques sujet qui, en combinaison avec des principes de la GU, donnent lieu à la variation effectivement observée? Pour répondre à cette question, il faut préalablement établir quelle sorte de variation se produit effectivement dans des grammaires qui emploient des pronoms sujet (éventuellement entre autres mécanismes) pour récupérer les traits de *pro*. Les dialectologues reconnaissent aussi que les pronoms sujet dans ces variétés sont variables, mais sans préciser l'envergure de cette variation:

In Oberitalien bis zur bekannten Südgrenze der ligurisch-emilianisch-romagnolischen Mundarten ist die Setzung des Subjektpronomens meistens

obligatorisch, also Morphem. Indessen has Scheu. mit Recht sich darauf beschränkt, die Subjektspronomina bei einigen Verben beizufügen, aber sie sonst wegzulassen. Aus Raumrücksichten mussten wir uns begnügen, die Subjektspronomina, deren Trennung im Piemont und in der Emilia besonders grossen Schwierigkeiten begegnet, nur beim Präsens des Verbums der 1. Konjugation (v. unten K. 1683) einzusetzen. Natürlich variiert auch die lautliche Form der in tonloser Stellung gebrauchten Subjektspronomina je nach der Lautgestalt des folgenden Verbums: wer sich für diese pronominalen Varianten interessiert, dem sei die Einsicht in die zahlreichen Karten des AIS, die in Sätze eingekleidete Verbalformen bieten, empfohlen. (*AIS* 1682)¹⁴

La présente thèse poursuit en partie cette recommandation, non pas pour connaître les variantes phoniques des pronoms sujet, mais plutôt pour savoir où ils sont employés.

14. «En Italie septentrionale jusqu'à la limite méridionale bien connue des dialectes ligurien-émilien-romagnols, le placement du pronom sujet est le plus souvent obligatoire, donc morphème. Cependant Scheu[ermeier] a raison de se limiter à ajouter les pronoms sujet à quelques verbes mais les omettre autrement. Pour des considérations d'espace nous devons aussi nous contenter de mettre les pronoms sujet, dont la séparation pose des problèmes extrêmement sérieux en Piémont et en Emilie, seulement au présent du verbe de la première conjugaison (voir la Carte 1683). Évidemment la forme phonique des pronoms employés en position atone varie aussi en rapport avec la forme phonique du verbe qui suit: nous recommandons à celui qui s'intéresse à ces variantes pronominales de consulter les nombreuses cartes de l'*AIS* qui offrent des formes verbales présentées en des phrases.»

Pour être retenue comme valable, l'hypothèse de la variation paramétrique doit être applicable non seulement au niveau des différences typologiques grossières, mais aussi au niveau des différences plutôt subtiles entre variétés très similaires. Lorsqu'on ne considère que des grammaires comme celles de l'anglais, du français et de l'italien, on peut attribuer à la variation paramétrique certaines différences qui paraissent importantes (comme le fait d'avoir des sujets nuls ou des sujets clitiques). Or, lorsqu'on compare ces grammaires avec celles des variétés intermédiaires, comme celles de l'Italie septentrionale, ces différences deviennent moins nettes, ce qui peut mener soit à postuler des paramètres plus fins (Rizzi 1989), soit à appliquer les paramètres aux éléments lexicaux plutôt qu'aux grammaires tout entières (Wexler et Manzini 1987). Quand on analyse de près ces variétés intermédiaires, même ces différences qui sont encore attribuables aux paramètres ont tendance à se dissoudre vers un type de variation qui rappelle ce que certains dialectologues désignent comme «zones de transition» (voir 2.1.2.), c'est-à-dire des aires géographiques où les différences sont d'ordre quantitatif plutôt que binaire et donc abrupt. Cette étude se veut donc en partie une vérification empirique de l'utilité du concept de la variation «paramétrique».

Notre but dans cette étude est par conséquent double, ou même triple: nous espérons documenter la variation qui se produit entre les variétés et à l'intérieur de chaque variété (voir 2.1.). C'est-à-dire que d'une part, nous nous intéressons à la présence ou l'absence d'un pronom sujet, à chaque point des enquêtes géolinguistiques.

D'autre part, nous nous penchons également sur la variation qui se produit dans chaque variété, que ce soit entre les différentes personnes grammaticales ou à l'intérieur de chaque personne grammaticale. La variation entre les personnes grammaticales nous mènera à dresser une typologie des différents paradigmes pronominaux qui peuvent exister dans les variétés étudiées, tandis que la variation à l'intérieur des personnes grammaticales nous permettra de nuancer cette typologie en distinguant les aspects constants et variables des paradigmes possibles.

Pour résumer donc: ce travail se définit par son approche éclectique qui s'avère la plus utile face à la variation empiriquement attestée parmi les pronoms sujet. Au niveau théorique, les hypothèses que nous essayons de garder présentes à l'esprit sont les suivantes: d'une part, l'idée intuitive que les pronoms sujet sont conditionnés fonctionnellement par la «richesse» des désinences verbales; d'autre part, le débat dans le cadre générativiste entre la représentation syntaxique ou morphologique des pronoms clitiques. À l'intérieur de ce dernier débat, nous retenons un cas particulier s'appliquant aux pronoms sujet: le dit «paramètre du sujet nul», dont nous vérifions l'utilité par rapport aux faits d'un continuum morphosyntaxique.

Le reste de cette thèse s'organise de la façon suivante: le deuxième chapitre élabore notre approche à la variation linguistique, et en particulier aux atlas linguistiques romans. Dans le troisième chapitre nous présentons la méthodologie que nous suivons dans cette étude en la justifiant par rapport aux résultats et hypothèses

présentés dans des travaux antérieurs. Le quatrième chapitre contient les résultats de notre étude, et le cinquième est consacré aux analyses et interprétations possibles de ces résultats, ainsi qu'à nos conclusions.

CHAPITRE 2: LA VARIATION ET LES ATLAS LINGUISTIQUES

2.0. LA VARIATION LINGUISTIQUE

Dans ce chapitre nous partons d'une distinction fondamentale entre deux aspects de la variation linguistique pour examiner les approches théoriques qui correspondent à chacun d'eux. Ensuite nous situons les atlas linguistiques en général, et notre étude en particulier, par rapport à ces deux grands courants.

2.1. VARIATION *INTRAGRAMMATICALE* ET *INTERGRAMMATICALE*

Quand nous parlons de «variation grammaticale», il faut bien comprendre qu'il existe deux façons assez différentes de concevoir cet objet d'étude, qui, sans être complètement incompatibles, conditionnent fortement les approches méthodologiques. Pour le structuralisme (et notamment pour le structuralisme générativiste), les recherches doivent se concentrer fondamentalement sur la variation entre les grammaires des différentes communautés linguistiques, soit la variation *intergrammaticale*, tandis que la sociolinguistique variationniste s'intéresse surtout à la variation à l'intérieur d'une communauté linguistique, ou *intragrammaticale*. Examinons donc ces deux cadres heuristiques de plus près, pour mieux situer notre étude par rapport aux deux types de variation.

La dialectologie traditionnelle (pour un survol compréhensif, voir Pop 1950) se caractérise le plus souvent par une approche strictement descriptive et phonéticiste: qu'il

s'agisse d'une monographie sur un parler précis ou des atlas linguistiques, on pense plutôt en termes atomistiques, comme dans les sciences naturelles où chaque aspect d'un objet doit être décrit exhaustivement sans référence à tous les autres. C'est sans doute pour essayer de dépasser ce descriptivisme acharné de la dialectologie traditionnelle que le structuralisme essaie de systématiser les faits de chaque variété, entreprise qui se heurte à des problèmes quand il faut faire face à plusieurs systèmes linguistiques. Toute la problématique de la variation intergrammaticale dans le structuralisme prégénéraliste peut se résumer par le titre d'un article célèbre de Weinreich (1954): «Is a structural dialectology possible?» Si l'analyse linguistique mène à caractériser chaque variété comme un «système où tout se tient», les unités de chaque système seraient alors à proprement parler incommensurables et donc impossibles à comparer:

If structuralism were carried to its logical extreme, . . . it could only study relations within systems; and since in a perfect system all parts are interrelated ('tout se tient'), it is hard to see how systems could even be conceived of as partially similar or different. (Weinreich 1954:2)

Les systèmes phonologiques -- l'objet préféré (et souvent l'objet unique) des études structuralistes -- peuvent (même dans deux variétés voisines) différer dans le nombre et dans la distribution des phonèmes, ce qui met la comparaison des «dialectes» sur un pied d'égalité avec la comparaison des «langues» (un résultat peu satisfaisant pour les linguistes

qui veulent capter ce qu'ont en commun les différents «dialectes» -- que nous appelons ici les variétés -- d'une «langue»). Étant donné les difficultés que présente la description de la variation entre les grammaires, il n'est pas étonnant que le structuralisme passe sous silence le problème de la variation *intragrammaticale*.

En général, dans le cadre générativiste on fait également abstraction de toute variation interne aux communautés linguistiques, suivant presque toujours la phrase célèbre de Chomsky concernant l'objet principal de la linguistique: «the ideal speaker-hearer in a completely homogenous speech community . . . » (1965:3) mais le problème de la caractérisation des relations entre les «dialectes» et les «langues» reste intact. On a proposé (Halle 1962; Saporta 1965) que les divers «dialectes» d'une même «langue» partagent un seul inventaire phonémique et peuvent varier dans le nombre, la nature et l'ordre d'application des règles phonologiques:

. . . the grammatical description of a given dialect may be converted into an adequate description of a related dialect by the addition, deletion, or reordering of a relatively small number of rules. Indeed, it is tempting to propose that the degree of difference between dialects is nothing more than a function of the number and type of such changes. (Saporta 1965:291)

Mais cette approche comparatiste entraîne exactement la même contradiction entre variation et structure à laquelle se heurte la dialectologie structuraliste. Campbell (1972)

reprend la problématique de Weinreich (1954) dans un article paru presque vingt ans plus tard, intitulé «Is a generative dialectology possible?», où il fait ressortir les similitudes fondamentales des deux approches face à ces mêmes problèmes:

So far as "Saussure's dictum" (that structural units are defined within an idiolect by the web of relationships locking them into the system) is concerned, generative grammarians are structuralists too, and hence subject to the same criticism; we, too, as generative grammarians, must face the fact that it is impossible to compare two dialects. (1972:292)

Le problème semblerait également impossible à résoudre dans les deux cas. D'une part, si l'on décrit chaque variété indépendamment, les grammaires seront incomparables; si en revanche on postule une sorte de «méta-grammaire» qui puisse rendre compte des différences et des similarités entre deux ou plusieurs variétés, on aboutit à une abstraction dépourvue de réalité psychologique pour les variétés effectivement attestées: une compétence que personne ne possède. Mais Campbell (1972) conclut sur un ton un peu moins décourageant, envisageant une réconciliation possible entre structure et variation:

However, if psychological reality is attained first in independently written descriptions, and then the comparative method is applied to account for the variation, there is some hope. Competence can be defined off the independent

grammars, and variation can be read off the comparative method. The approach I am advocating presupposes that generative linguistics will be able to solve the problem of semantic equivalence, and semantic problems in general, before such an approach can become practical or even empirical. (1972:296).

La question de l'équivalence sémantique, dans le cadre générativiste comme dans n'importe quel autre, paraît encore loin d'être résolue, mais ceci risque de nous entraîner dans un débat qui nous éloigne trop de notre propos. Campbell note une autre difficulté bien plus significative pour nous, quand il continue en note en bas de la page:

In a strict sense, the approach I am advocating is ultimately doomed to failure. As Weinreich *et al.* [1968] have so aptly demonstrated, orderly heterogeneity, or variation, is an integral part of individual competence. . . . Thus I have no answer for socially conditioned variation within a single individual.

(Campbell 1972:296-7, n.1)

C'est justement ce type de variation *intra*grammaticale qui fait l'objet des recherches linguistiques variationnistes. En effet, depuis l'article programmatique de Weinreich, Labov et Herzog (1968), ce cadre théorique met l'accent sur l'*hétérogénéité ordonnée* comme clef pour comprendre la variation linguistique synchronique ainsi que l'évolution diachronique. Les recherches de Labov (1966, 1972 etc.) et d'autres linguistes

du courant variationniste montrent de façon indiscutable que les variables linguistiques (qu'elles soient des différences de nature phonétique, lexicale ou morphosyntaxique) entrent en corrélation statistique avec des variables sociales indépendantes, telles l'âge, le genre et la classe sociale.¹⁵ Ces corrélations permettent aussi de démêler les effets des différents facteurs linguistiques, qui se caractérisent souvent par des interactions complexes, comme nous le verrons plus loin.

Il convient de souligner que les deux approches que nous venons d'esquisser sont, au plan théorique, tout aussi valables l'une que l'autre. Elles diffèrent surtout en ce qu'elles visent des objets d'étude distincts, ce qui implique nécessairement des différences dans les méthodes qui seront appropriées pour chaque chacune d'elles. Les questions que se posent les linguistes dirigent leur façon de chercher des réponses, et par conséquent elles déterminent en grande mesure les réponses obtenues.

2.1.1. Diachronie, synchronie et continuum

Depuis ses débuts (Weinreich et al. 1968), le cadre conceptuel variationniste vise non seulement la dynamique de relations sociolinguistiques à une époque déterminée, en synchronie, mais voit aussi dans ces relations la force motrice du changement linguistique, ou la diachronie. On peut ainsi prétendre dépasser en quelque sorte la dichotomie

15. Nadasdi (1995) et Tennant (1995) démontrent clairement que la restriction linguistique peut également être un facteur décisif, surtout pour une communauté linguistique minoritaire.

saussurienne entre diachronie et synchronie pour arriver à une synchronie dynamique où le mouvement et la direction du changement sont implicites dans un portrait «instantané» des faits (Martinet 1945, 1990; Schogt 1968a; Jakobson 1976). Or, il est évident que notre étude des pronoms sujet à travers leur variation *intergrammaticale* constitue une description essentiellement synchronique, puisqu'elle vise à caractériser des états de langue momentanés plutôt que des évolutions linguistiques internes. C'est-à-dire que pour chaque variété étudiée, nous ne considérons que les données cueillies à un moment précis de son évolution (lors des enquêtes pour les atlas linguistiques). Ceci dit, il se peut toutefois que la comparaison de variétés étudiées synchroniquement soit pertinente pour la diachronie. Il est possible, par exemple, que les systèmes pronominaux que nous trouvons dans certaines des variétés étudiées correspondent à des étapes antérieures d'autres variétés. Il est cependant très peu probable que ce soit le cas pour toute l'histoire d'une grammaire donnée: la réalité linguistique est bien trop complexe pour qu'une variété synchronique reproduise exactement un état de langue antérieur. Mais si l'on ne considère qu'un seul aspect de chaque grammaire (en l'occurrence, le système des sujets pronominaux) il devient plus plausible que la variation synchronique puisse récapituler partiellement la diachronie, ou au moins fournir une idée générale des étapes intermédiaires entre deux stades historiques.

Un exemple prometteur de cette approche serait justement la grammaticalisation des pronoms sujet en français: optionnels en ancien français, ces pronoms deviennent obligatoires dans les phrases sans sujet lexical, lors de la transition au français moderne. Il

se peut que nos données tirées de la variation grammaticale synchronique puissent aider à évaluer deux hypothèses à propos de ce changement diachronique.

Les explications traditionnelles (voir par exemple Meyer-Lübke 1895) supposent qu'en ancien français l'érosion phonétique des désinences verbales entraîne la présence obligatoire des sujets pronominaux pour «(re)compenser» en quelque sorte cette perte d'information. Or Adams (1988) remet en cause cette explication traditionnelle, sur la base d'une chronologie comparative qui démontrerait que, contrairement à ce qu'on supposerait intuitivement, c'est la présence de plus en plus fréquente des pronoms sujet qui précède (et donc permet) l'érosion phonétique des désinences (voir aussi Dufresne 1993).

Les détails de ce débat (voir Price 1984:146; Roberge et Vinet 1989:43; Dufresne 1993; Nadasdi, 1995:43-45) ne sont pas directement pertinents pour la présente étude. Mais si, en revanche, cette étude nous aide à mieux comprendre comment on peut passer géographiquement d'un système sans pronoms sujet à un système avec sujets pronominaux dans toutes les personnes grammaticales, il se peut bien que ceci soit utile pour saisir comment le même changement aurait eu lieu historiquement. Si l'on accepte l'hypothèse que la variation grammaticale synchronique peut refléter le changement diachronique, il devient potentiellement intéressant de voir la gamme des stades intermédiaires entre ces deux pôles. Ces variétés peuvent, éventuellement, présenter un fonctionnement analogue aux étapes historiques qui ont mené d'une grammaire où les pronoms sujet sont facultatifs à une grammaire où les sujets pronominaux sont

obligatoires. Au pire, la typologie des systèmes pronominaux qui en résulte pourrait suggérer des contraintes qui limiteraient la portée des hypothèses historiques possibles à propos de l'évolution des systèmes pronominaux.

Notons néanmoins qu'une des géolinguistes les plus éminentes en matière de variation syntaxique préconise la prudence pour ce qui est d'étendre les analyses de la variation synchronique aux changements diachroniques:

. . . dalla distribuzione areale di fenomeni sintattici, per il momento almeno, non possiamo ricavare elementi per ricostruire l'applicazione di regole grammaticali in diacronia. La variazione sintattica, a questo livello di fenomeni, è di ragione grammaticale interna alle singole lingue. (Benincà, à paraître, 133)¹⁶

Elle est toutefois plus positive quant à l'importance des analyses synchroniques pour avancer des hypothèses diachroniques bien fondées:

16. «. . . nous ne pouvons pas, au moins pour le moment, obtenir des éléments pour reconstruire l'application des règles grammaticales en diachronie à partir de la distribution géographique des phénomènes syntaxiques. La variation syntaxique, à ce niveau-ci des phénomènes, est d'ordre grammatical interne à chaque langue.»

Tuttavia, vorremmo dire che i dati sono utilizzabili anche per la ricostruzione del mutamento sintattico, per migliorare e definire meglio questa area così mal conosciuta del mutamento linguistico. La ragione principale di questo stato di cose è, naturalmente, che è necessario capire bene in sincronia come funziona un fenomeno prima di poter ricostruire, con qualche verosimiglianza, il suo mutamento.¹⁷ (Benincà, à paraître, 133)

Il ne faut pas alors s'attendre à ce que la variation synchronique récapitule littéralement les changements diachroniques, mais il y a de bonnes raisons de croire qu'une typologie détaillée des systèmes pronominaux pourrait aider considérablement à comprendre l'évolution de ces mêmes systèmes, qui d'ailleurs posent encore problème pour la linguistique historique.

En fait, il s'agit ici d'une extension du concept du «continuum linguistique», dont on peut se servir pour décrire des changements non discrets entre variétés «adjacentes» (que ce soit géographiquement ou socialement). Ainsi on peut citer le cas du continuum géolinguistique de la Romania occidentale (Chambers et Trudgill 1980:6) qui allait du

17. «Toutefois, nous voudrions dire que les données sont utilisables aussi pour la reconstruction du changement syntaxique, pour améliorer et mieux définir ce champ si mal connu du changement linguistique. La raison principale de cet état de choses est, naturellement, qu'il est nécessaire de bien comprendre comment un phénomène fonctionne en synchronie avant de pouvoir reconstruire, avec quelque vraisemblance, son changement.» On rappellera que pour Saussure, la diachronie consiste en une succession d'étapes synchroniques: comme le disait déjà si bien Baudouin de Courtenay, «one cannot be a good paleontologist without previously studying biology.» (*apud* Schogt 1966).

portugais à l'italien et au wallon, en passant par des variétés du galicien, castillan, aragonais, catalan, français et occitan. Dans chaque cas, les variétés dites standard ne seraient pas mutuellement intelligibles, mais les variétés rurales formeraient un continuum où chaque géolecte serait compréhensible pour les géolectes limitrophes. D'autres exemples de transitions géolinguistiques graduelles existeraient en Europe (Chambers et Trudgill 1980:7), et de façon générale, partout où des langues similaires ou génétiquement apparentées coexistent adjacentes dans l'espace pendant longtemps.¹⁸

L'hypothèse que l'on peut donc retenir serait que les étapes successives d'un changement diachronique (tel la grammaticalisation des pronoms sujet) peuvent présenter des analogies intéressantes avec un continuum géolinguistique, en l'occurrence le continuum italo-gallo-roman. Pour faire ce genre de comparaison, il faut préalablement dresser un portrait assez détaillé de ce continuum géolinguistique pour ce qui est des éléments en question: les sujets pronominaux. Notre travail descriptif pourrait donc ouvrir de nouvelles voies de recherches pour les diachronistes qui s'intéressent à l'évolution des pronoms sujet. Mais la problématique du continuum géolinguistique soulève d'autres questions sur lesquelles il convient de nous pencher davantage: comment les géolectes

18. On observe également qu'un continuum peut être d'ordre sociolinguistique, dans ce sens que les différences linguistiques entre deux couches sociales seront en général beaucoup moins marquées, voire imperceptibles, si les deux sont très proches en termes socio-économiques, et plus marquées plus les deux sont éloignées dans la structure socio-économique.

limitrophes peuvent-ils se distinguer les uns des autres, et quelle importance ce type de variation peut-il avoir pour la théorie linguistique?

2.1.2. Continuums, zones de transition et typologie linguistique

Dans la dialectologie traditionnelle, les différences entre les variétés voisines sont représentées graphiquement par des isoglosses, soit des lignes tracées sur une carte linguistique pour séparer une région où un certain trait est présent d'une région voisine où le même trait est absent. Plus il y a d'isoglosses qui séparent des régions, plus il doit y avoir de différences linguistiques. Ainsi trouve-t-on entre la France septentrionale (l'aire de la langue d'oïl) et le Midi (l'aire de la langue d'oc) un important faisceau d'isoglosses qui correspondent, entre autres, à des différences telles la fermeture de la voyelle /a/ en syllabe tonique ouverte, l'élimination des occlusives intervocaliques, et la palatalisation du /K/ latin (Jochnowitz 1973: chapitre IV; von Wartburg 1967:65ss). Depuis de Tourtoulon et Bringuier (1876), on inclut la présence des pronoms sujet parmi les caractéristiques qui distinguent la langue d'oïl de la langue d'oc.

Il se trouve cependant que les isoglosses, bien qu'elles soient utiles pour résumer les différences entre des variétés, cachent en même temps une partie importante de la réalité linguistique. Les changements linguistiques sont en fait rarement abrupts, que ce soit en diachronie (voir Chen et Wang 1975) ou en géographie. C'est-à-dire qu'entre une variété où un trait A est présent catégoriquement et une autre où le trait B le remplace catégoriquement, il y a normalement au moins une variété où ces deux traits sont présents

variablement. L'étude de ces variétés intermédiaires mène à établir des «zones de transition» de plusieurs types: des variétés mixtes («mixed lects») où on trouve tantôt le trait A et tantôt le trait B, des variétés escamotées («fudged lects») où on trouve une forme intermédiaire entre le trait A et le trait B, et des variétés embrouillées («scrambled lects»), où le trait A alterne et avec le trait B et avec la forme escamotée intermédiaire (Chambers et Trudgill 1980:125-142). Évidemment, ces distinctions pourraient transformer les cartes linguistiques traditionnelles, munies d'isoglosses, en une représentation moins nette (mais peut-être plus fidèle) de la réalité linguistique, avec des symboles et même des pourcentages pour différencier chacun des cas intermédiaires. Il se trouve en plus que c'est dans les zones de transition que la variation intragrammaticale va de pair avec la variation intergrammaticale:

Mais c'est dans la zone intermédiaire, où les deux langues se rencontrent, que cette irrégularité atteint sa plus grande intensité. L'extension d'un trait linguistique n'y dépend pas seulement du mot qu'on choisit pour type, mais aussi de l'individu qui le prononce. (Rosenqvist 1919:110)

En dépit de l'intérêt évident de cette nouvelle conception de la cartographie dialectologique, les zones de transition n'ont pas reçu beaucoup d'attention de la part des spécialistes en géolinguistique (Chambers et Trudgill:142). À la différence de ces auteurs, nous croyons que les matériaux de base pour l'étude des zones de transition ne font pas

défaut: les atlas linguistiques romans contiennent beaucoup de données brutes qui attendent des analyses de ce type. La présente étude prétend être un exemple de ce qu'on peut effectivement faire avec les données d'atlas en terme d'analyses de variation linguistique.

Mais l'importance de l'étude des zones de transition ne réside pas uniquement dans la possibilité de représenter cette variation grammaticale plus fine sur des cartes géolinguistiques encore plus détaillées et chargées d'information. Ces cartes géolinguistiques ne sont que les représentations graphiques des données typologiques plus nuancées. La typologie linguistique (Sapir 1921:VI) s'occupe traditionnellement des «macrodifférences» qui servent à distinguer les systèmes grammaticaux nettement séparés sur le plan structural. Or, nous ne voyons aucune raison de croire que les différences moins grandes ne peuvent informer la typologie linguistique (voir 2.1.6.), et même les différences qui sont plutôt quantitatives. S'il est vrai que «there is no substantive difference between a significant "typology" of linguistic systems in a given domain and an explicit, articulated "theory" of that domain» (Anderson, 1992:5), alors une typologie détaillée des pronoms clitiques dans les variétés romanes constituerait une théorie de ces pronoms. Cette théorie ne serait pas, bien entendu, universellement applicable (puisque la typologie que l'on envisage ici n'est que locale), mais elle aurait néanmoins des implications significatives pour la façon dont ces systèmes linguistiques peuvent varier. Elle ne saurait non plus prétendre à l'exhaustivité, puisque notre corpus ne représente qu'un échantillon des emplois pronominaux possibles pour chaque variété étudiée. En revanche, il est certain

que toute théorie qui vise les systèmes des sujets pronominaux dans leur totalité doit tenir compte au moins de la variation qui sera documentée dans notre étude, puisque la variation réelle ne peut qu'être plus extensive encore.

2.1.3. La question du franco-provençal

Ce débat sur l'importance relative des divisions géolectales et des zones de transitions géolinguistiques est loin d'être nouveau: il date de plus d'un siècle, et constitue en fait une des forces motrices de la recherche scientifique en géolinguistique française. Car la zone de transition la plus célèbre (et la plus controversée) de la tradition romaniste est sans doute celle qui correspond à l'aire linguistique dénommée le franco-provençal.

Depuis l'affirmation de l'existence de cette variété par Ascoli (1874), sur la base de critères essentiellement phonologiques, les positions sur cette question se rangent en deux camps opposés. Ceux qui rejettent la division tripartite du gallo-roman (en langue d'oïl, langue d'oc et franco-provençal), nient l'existence même de telles divisions, en soulignant la nature continue et graduelle des transitions:

... dans une masse linguistique de même origine, comme la nôtre, il n'y a réellement pas de dialectes; il n'y a que des traits linguistiques qui entrent respectivement dans des combinaisons diverses, de telle sorte que le parler d'un endroit contiendra un certain nombre de traits qui lui sont communs, par exemple, avec le parler des quatre endroits les plus voisins, et un certain nombre de traits qui diffèrent du parler de ces quatre endroits. ... [A]ucune limite réelle ne sépare les Français du Nord de ceux du Midi, ... d'un bout à l'autre du sol national nos parlers populaires étendent une vaste tapisserie dont les couleurs variées se fondent sur tous les points en nuances insensiblement dégradées. (Paris, 1888:134-135).

Paris suit donc Meyer (1888) en s'opposant à la notion du franco-provençal proposée par Ascoli et, de façon plus générale, au concept même de «dialecte». Le parrain de l'*ALF* (car c'est à son collègue Paris que Gilliéron va concéder l'honneur d'épigraphier son atlas) affirme donc l'unité linguistique (et implicitement, politique) de l'Hexagone en rejetant comme arbitraire tout critère qu'on puisse avancer pour établir des frontières dialectales. Cette position est reprise par Hall (1949), qui utilise des données de l'*ALF* pour affirmer que le Midi constituerait tout simplement une énorme zone de transition entre le français et les autres langues romanes, sans qu'on puisse y définir des entités distinctes comme le franco-provençal. À cette position s'opposent, entre autres, Martinet (1939, 1956) et Haudricourt et Juilland (1949), qui présupposent tout simplement l'existence du franco-provençal comme unité linguistique, ainsi que les chercheurs et chercheuses qui essaient

de tracer les limites du franco-provençal (par exemple Gardette 1941; Lobeck 1945; Escoffier 1958). Pour ceux-ci, l'arbitraire des critères de division dialectale ne représente pas un inconvénient si ces critères permettent de cerner un objet d'étude intéressant.

La conclusion à laquelle aboutit Jochowitz, après un examen minutieux des correspondances possibles entre les différentes isoglosses et les divisions historiques et socio-culturelles (architecture, traditions légales, systèmes agricoles etc), nous semble assez juste: «all the disagreeing scholars, whose arguments on whether or not there are dialect boundaries in France were discussed in Chapter 2, are clearly right.» (1973:180) Autrement dit, si l'on cherche à trouver des variétés distinctes, on en trouve, et si l'on cherche à trouver des zones de transition, on en trouve aussi. Comme dans toute entreprise, les objectifs de départ déterminent en grande partie les résultats. Ceci dit, le choix de terrain peut également influencer la façon de concevoir cette problématique: dans l'Ouest de la France, on trouve un grand faisceau d'isoglosses qui partent ensemble de l'estuaire de la Gironde, où elles séparent assez nettement les variétés d'oïl et d'oc. Mais plus on va vers l'Est, plus ce faisceau se défait, les différentes isoglosses divergent à tel point qu'il est difficile de voir une quelconque division nette. Doit-on dire simplement qu'il existe des frontières nettes dans certaines régions et des zones de transitions dans d'autres?

Nous ne comptons pas répondre à cette question difficile que tant d'autres ont traitée sans la résoudre. Mais elle est pertinente pour notre étude dans ce sens que ce

faisceau recoupe la limite septentrionale de l'aire qui nous intéresse ici, soit l'aire des grammaires à sujets nuls:

... in Northern France subject pronouns are normally used in simple declarative sentences, and in Southern France they are normally omitted. Bundling is quite apparent here, despite the fact that different subject pronouns, different tenses, and a reflexive verb are being considered. Let us call Isogloss 7 the isogloss south of which the subject pronoun is omitted in *elle s'est couchée*, *ALF* map 1519.

(Jochowitz 1973: 129)

Nous reviendrons à cette isoglosse 7 plus loin (voir 3.2.4.) quand il est question de choisir les points de l'*ALF* à retenir pour notre étude. Notons cependant que Jochowitz trouve, de façon générale, que les isoglosses grammaticales sont plus solidaires que les isoglosses phonologiques:

The maps concerning grammatical phenomena show much clearer bundling than most of the maps concerning phonological phenomena. Perhaps this is because we are not dealing with bundling here at all, but with the area covered by the spread of a single morpheme. (1973:96)

Ce sont justement ces isoglosses grammaticales (ou plus précisément, morphosyntaxiques) que nous traitons dans cette étude. Il reste à voir (5.2.1.) dans quelle mesure cette zone de transition peut se réduire à la «propagation d'un seul morphème», comme l'affirme Jochnowitz. Pour Benincà (1992), il serait une erreur de chercher des différences typologiques significatives à partir d'isoglosses individuelles, car celles-ci représentent des réalisations particulières de règles plus générales, et sont donc sujettes à des influences idiosyncratiques. Elle rappelle que pour son auteur, Wenker, le *Deutscher Wortatlas* serait une expérience échouée, puisque ses données géolinguistiques ne récapitulent pas directement les *Lautverschiebungen*, les mutations phonétiques (1992:30):

Wenker contava di ritrovare una perfetta coerenza fra aree dialettali e isoglosse fonologiche relative a regole di fonologia diacronica. . . : egli si aspettava che le isoglosse . . . risultassero sovrapporsi esattamente, separando in modo coerente i dialetti alto-tedeschi dai dialetti basso-tedeschi. Risultò invece che le isoglosse non si sovrapponevano esattamente ma formavano un fascio di linee distinte.¹⁹

(Benincà 1992:30)

19. «Wenker comptait retrouver une parfaite cohérence entre les zones dialectales et les isoglosses phonologiques qui concernent les règles de phonologie diachronique. . . : il s'attendait à ce que les isoglosses . . . se superposent exactement, séparant de façon cohérente les dialectes haut-allemands des dialectes bas-allemands. Il s'avéra cependant que les isoglosses ne se superposaient pas exactement mais formaient un faisceau de lignes différentes.»

Or les *Lautverschiebungen* correspondraient en fait à des généralisations phonologiques, et seraient donc de niveau trop abstrait pour être reflétées directement dans la complexité des données géolinguistiques. Pour Benincà, Wenker se serait trompé non pas d'hypothèses mais de niveau d'analyse: les mutations phonétiques correspondraient aux faisceaux et non pas aux isoglosses individuelles. De même, pour ce qui est de la recherche des «données expérimentales en relation avec des hypothèses théoriques sur les propriétés syntaxiques des langues romanes», elle préconise non pas la recherche de «zones dialectales» compactes et homogènes, mais plutôt de phénomènes syntaxiques individuels (1992:31). C'est seulement à partir de ces phénomènes qu'on peut chercher des corrélations de solidarité entre eux, soit des abstractions de type paramétrique.

2.1.4. La comparaison des grammaires

Nous avons vu que la contradiction apparente entre structure et variation linguistiques réside dans le fait que la comparaison de grammaires différentes manque de base théorique et empirique dans les différentes écoles de structuralisme linguistique, étant donnée leur conception des règles linguistiques: «... there is no basis within the theory of generative grammar for comparing independently written grammars of separate dialects.» (Campbell 1972:292) Ce manque d'un mécanisme formel pour comparer des grammaires disparaît avec l'introduction du cadre théorique dit des Principes et Paramètres (Chomsky 1981), qui requiert entre autres choses un travail essentiellement comparatif pour identifier

les caractéristiques invariables du langage humain et les distinguer des aspects susceptibles de varier d'une grammaire à une autre:

Nous appelons *grammaire universelle* la théorie abstraite des propriétés linguistiques prédéterminées par la structure intrinsèque de l'esprit humain. Nous pouvons conceptualiser la grammaire universelle comme une fonction complexe, avec un certain nombre de variables indépendantes, ou *paramètres*. Les principes de la grammaire universelle, invariants, définissent la nature de la fonction. Les paramètres caractérisent l'espace limité de la variation possible. . . . Les paramètres de la grammaire universelle peuvent être conceptualisés comme les points de bifurcation fondamentaux du système grammatical général, les différences irréductibles entre les systèmes grammaticaux particuliers. Afin d'identifier empiriquement ces points de fracture primitifs, il est essentiel de focaliser le travail comparatif sur des systèmes grammaticaux assez proches. En effet, des langues dont la structure globale est très éloignée permettraient plus difficilement d'isoler des différences primitives plausibles, à cause de l'interaction complexe, parfois inextricable, d'une multiplicité de différences observables. (Rizzi 1989: 9)

Cette nouvelle perspective comparativiste fait que, depuis une quinzaine d'années maintenant, des recherches générativistes se penchent de plus en plus sur des variétés qui ne jouissent pas du statut de «langues officielles» (les objets presque exclusifs de

recherches en syntaxe générativiste avant l'introduction du programme comparativiste).

Les variétés non standard présentent des avantages importants pour ce type de recherches:

Nous avons affaire, dans l'étude comparative des dialectes, à des systèmes grammaticaux extrêmement proches, qui ne diffèrent que pour un nombre restreint de propriétés fondamentales; ces propriétés sont donc relativement faciles à isoler et à démêler de toute interférence cachée. (Rizzi 1989:9)

Là où la «dialectologie structurale» se heurtait inévitablement à l'incommensurabilité des grammaires, cette nouvelle «géolinguistique structurale» réussit en rejetant comme non scientifique la distinction entre «langues» et «dialectes». Elle se propose plutôt de traiter chaque grammaire, chaque variété géolinguistique, sur un pied d'égalité comme source de données sur la Grammaire Universelle (Roberge et Vinet 1989). Les relations entre «dialectes» sont donc de la même nature que les relations entre «langues» (c'est-à-dire qu'elles reposent sur un certain nombre de paramètres à régler) et ne posent pas de problèmes particuliers. Les typologies géolinguistiques se fondent sur les différences entre les grammaires, et non pas sur les relations possibles (mais difficiles à saisir en termes linguistiques) entre variétés standard et non standard. On abolit ainsi une partie de la distinction traditionnelle entre la «grammaire» qui étudie uniquement «la variété dite standard, sous l'angle de l'homogénéité», et la sociolinguistique, qui prend en compte différentes variétés «sous l'angle de l'hétérogénéité» (Baylon 1991:26): on traite

maintenant toutes les variétés linguistiques (même si l'on a encore du mal à faire face à l'hétérogénéité interne).

Pour résumer, notre étude peut se concevoir comme étant essentiellement intergrammaticale, puisque nous prétendons établir une typologie des pronoms sujet à partir de la comparaison des grammaires de variétés très similaires. Nous empruntons néanmoins au variationnisme labovien le fait de travailler sur un corpus pour chaque variété (même si ce corpus est limité et idiolectal) ainsi que l'approche quantitative qui risque de tempérer les généralisations typologiques en faisant ressortir le rôle des facteurs linguistiques internes (voir 3.2. et 3.4.) Évidemment, nous ne pourrions pas capter la variation sociolinguistique proprement dite, puisque notre corpus pour chaque variété se limite aux réponses d'une seule personne,²⁰ presque toujours choisie parmi la catégorie qu'on appelle des «NORMs», (pour «Nonmobile, Older, Rural Males», Chambers et Trudgill 1980:33). Les contraintes de cette étude nous obligent à prendre ces personnes comme représentatives des géolectes en question, avec la perte inévitable de la variation intergrammaticale.²¹ En récompense, nous obtenons un réseau géolinguistique assez dense,

20. C'est le cas à quelques exceptions près: voir 3.3.2. Même quand Edmont identifie deux ou plusieurs sujets, nous ne saurons établir aucune corrélation entre ce fait et les données correspondantes.

21. Les enquêtes doubles menées dans les cinq grandes villes de l'*AIS* (voir 3.3.2.) nous donnent cependant une indication du type de variation sociale qui pourrait exister dans ces communautés linguistiques (voir 4.3.5.).

puisque notre hypothèse de départ est que chaque point dans un atlas peut représenter, au moins en principe, une variété géolinguistique autonome.

La méthodologie choisie pour notre étude a donc l'avantage d'éviter l'hétérogénéité descriptive à laquelle doivent faire face d'autres survols ou études d'ensemble (voir par exemple Roberge et Vinet 1989 ou Cummins 1992: étant donné leurs buts, cette hétérogénéité ne représente pas un inconvénient important pour ces recherches). Le questionnaire unique utilisé pour chaque atlas garantit des centaines de réponses individuelles à des questions identiques. Il existe bien sûr des différences importantes entre les divers atlas nationaux, mais on peut trouver suffisamment de matériel en commun pour permettre une vision globale du roman central en ce qui concerne nos variables (voir 3.3.1.).

2.2. UN CORPUS D'ATLAS LINGUISTIQUES

2.2.1. À propos des corpus en linguistique

Il convient de souligner ici que l'importance, l'organisation et les méthodes de collecte des corpus varient énormément en fonction des types de recherches auxquels ils sont destinés (variation *inter-* ou *intra-*grammaticale). Pour la plupart des recherches en syntaxe générative, par exemple, il suffit souvent de quelques phrases d'une variété donnée, dont la grammaticalité ou l'agrammaticalité est vérifiée uniquement par l'introspection (soit directement fournie par les linguistes, soit indirectement par des informateurs et informatrices parlant la variété étudiée). Dans ce type de recherche, les

différences quantitatives ainsi que les influences éventuelles des méthodes d'enquête sont des problèmes qui ne se posent même pas. A l'autre extrême se trouvent les recherches en sociolinguistique variationniste qui visent surtout la validité statistique, ce qui implique nécessairement des corpus relativement importants qui se prêtent à l'interprétation quantitative, et dans lesquels les possibilités d'influences (directes ou inconscientes) de la part des chercheuses et chercheurs sont réduites à un minimum. Ces critères expliquent l'intérêt qu'attache cette méthodologie à l'élaboration de protocoles d'enquête qui réduisent au minimum l'interférence expérimentale, ainsi qu'à l'accumulation de corpus importants et bien structurés qui permettent l'emploi d'outils statistiques.

2.2.2. Pourquoi étudier les atlas linguistiques romans?

Les données des géolectes non standard (c'est-à-dire les données dites «dialectales») sont trop souvent exclues des analyses linguistiques (Roberge et Vinet 1989:21), d'une part à cause des préjugés extralinguistiques, et d'autre part à cause d'un manque de descriptions adéquates. Il existe de nombreuses monographies linguistiques décrivant des variétés précises, mais ces sources souffrent de deux désavantages sérieux. D'abord, à peu d'exceptions près, ces ouvrages consacrent la plus grande partie de leurs pages aux particularités d'ordre lexical et phonético-phonologique du parler en question, et parfois aux «déviations» de ce parler par rapport à la variété considérée comme standard. Et même là où l'on trouve des données morphosyntaxiques intéressantes et suffisantes pour nos analyses, la nature même de ces études précises et locales nous

empêche de comparer ces résultats sur l'axe géographique. C'est ainsi que nous disposons d'une excellente description morphologique et syntaxique du crémonais (Rossini 1975), qui est malheureusement difficile à mettre en contexte à cause du manque de descriptions analogues des variétés limitrophes.

L'incommensurabilité des données empiriques affecte même les études qui dépouillent systématiquement, dans les descriptions disponibles sur plusieurs variétés, toutes les données sur des questions grammaticales précises, par exemple Roberge et Vinet (1989:39-42) qui traitent une dizaine de variétés, ou Cummins (1992), qui en considère une trentaine. En fait, la seule façon d'obtenir des données comparables est d'appliquer la même enquête à des sujets parlant les variétés étudiées. C'est en effet ce que font les enquêteurs des atlas linguistiques, mais en général sans s'intéresser guère à la variation morphosyntaxique. La méthodologie des enquêtes postales (Renzi et Vanelli 1983, Benincà 1992) vise aussi le même but, mais elle n'est pas toujours pratique.

Tout choix de corpus limite nécessairement les questions qu'on peut poser dans une recherche donnée. Le fait que Renzi et Vanelli (1983) aient à leur disposition pour leur enquête postale des dialectologues qui sont aussi des locuteurs et locutrices d'une trentaine de variétés nord-italiennes (Benincà 1992:36), leur permet d'obtenir une série de traductions assez sophistiquées permettant d'illustrer les différents contextes syntaxiques où les pronoms sujet peuvent apparaître. De même, l'enquête postale décrite par Benincà (1992) en vue d'un «Atlas syntaxique de l'Italie septentrionale» propose aux informateurs et informatrices des phrases à traduire qui recouvrent toutes les constructions syntaxiques

pertinentes. Mais, à part le fait de ne pas nous être accessible, cette méthodologie a le désavantage sérieux de limiter le nombre de points géographiques enquêtés ainsi que la densité du réseau.

Les atlas linguistiques, en revanche, outre leur accessibilité générale comme sources de données relativement sûres et vérifiables, ont l'avantage d'offrir un réseau assez dense de points et une distribution géographique assez étendue. Le nombre et les types de questions posées sont évidemment fixés à l'avance, mais le dépouillement d'une quantité significative de constructions pertinentes permet, en principe, de tirer des conclusions statistiquement valables. Une autre différence concerne le type de sujet dans chaque cas: les enquêtes postales que nous venons d'évoquer s'adressent à des locuteurs et locutrices des variétés à l'étude qui sont des spécialistes en dialectologie, donc linguistiquement assez sophistiqués, tandis que les informateurs et informatrices des atlas linguistiques sont presque toujours des sujets «naïfs» du point de vue linguistique. Comme le note Labov:

The linguist who is alert to the widest range of dialect differences, who may construct the broadest pan-dialectal grammar, is often the worst informant on his own local dialect. There are of course exceptions -- some speakers show an extraordinarily tenacious hold on their original dialect. But we do not know who the exceptions are until we have studied the vernacular in the intact speech community itself. (1972:291)²²

Dans le cas des enquêtes postales envoyées aux spécialistes, la possibilité d'interférences involontaires est multipliée par le nombre d'individus qui réagissent aux questionnaires, ce qui ajoute une marge importante d'inconnu expérimental à la recherche. Il est bien sûr possible que les enquêteurs des atlas soient eux aussi des sources d'interférences linguistiques dans leurs données (voir 4.1.2.), mais au moins dans leur cas il s'agit d'interférences contrôlées: l'administration des questionnaires par un nombre restreint de personnes (une seule pour l'*ALF*, trois pour l'*AIS*, dont une seule pour la partie qui nous intéresse ici), suivant un protocole constant, doit réduire l'inconnu expérimental à un minimum.

22. Si à ce type d'interférences involontaires on ajoute la possibilité d'un parti pris professionnel (conscient ou inconscient) de la part des linguistes qui dépendent exclusivement des données obtenues par l'introspection, on risque d'aboutir à une situation où les différences «dialectales» correspondent aux postulats théoriques plutôt qu'aux données réelles: «If so, it will undoubtedly appear that no two linguists ever have the same dialect, unless they are colleagues jointly responsible for the same theory.» (Labov 1972:292)

Il faut aussi considérer l'effet de la présentation même des questionnaires: soit par voie orale, soit par voie écrite. La forme écrite des enquêtes postales risque elle-même de favoriser les réponses plutôt standardisantes. L'administration orale des enquêtes d'atlas, en revanche, favorise les réponses non standard: même si la situation d'interview manque de naturel et tend donc à produire un style plus formel (Chambers et Trudgill 1980:70), les données obtenues par cette méthode ne souffrent pas de l'interférence de l'écrit.

Nous ne prétendons pas bien entendu que notre méthode puisse capter toute la variation grammaticale présente dans les variétés en question: dépendre exclusivement des atlas linguistiques implique nécessairement accepter l'échantillon limité de données qu'ils offrent ainsi que les limitations propres aux méthodes d'enquête employées. La variation réelle présente dans ces variétés romanes au moment des enquêtes dépassait certainement la petite partie qui se trouve reflétée dans les atlas. Comme il s'agit dans plusieurs cas de variétés disparues depuis, nous devons nous contenter de ces données. Notre démarche permet néanmoins d'accumuler une masse de données plus importante et dont la méthode de collecte est plus homogène que celles des études antérieures. Toute analyse de ces variables doit alors tenir compte au moins de ce degré de variation, en reconnaissant que la variation réelle va sans aucun doute beaucoup plus loin.

On peut ainsi concevoir notre méthodologie comme une expérience scientifique «reconstituée» après coup, et donc non répétable. Nos données seraient les traductions par quelques centaines de sujets linguistiques (représentant des variétés romanes différentes) de la même série de phrases dans les questionnaires. Dans cette optique, le fait que les

enquêteurs originaux ne s'intéressaient guère à la morphologie, loin d'être un désavantage, tendrait à réduire les interférences expérimentales éventuelles.

2.2.3. Notre corpus

Pour réunir les matériaux nécessaires pour cette étude, nous examinons les divers atlas linguistiques des pays de langues romanes pour trouver les cartes comportant des formes pertinentes (c'est-à-dire, ayant potentiellement au moins un pronom clitique). Nous utilisons surtout l'*Atlas linguistique de la France* de Gillieron et Edmont (1902-1920), et le *Sprach- und Sachatlas Italiens und der Südschweiz* de Jaberg et Jud (1928-1940). Dans ces deux atlas, nous choisissons les zones où il semble y avoir de la variation parmi les sujets pronominaux. Les variétés qui nous intéressent dans cette étude sont en gros celles qui ne peuvent pas être classifiées simplement comme [+sujet \emptyset] ou [-sujet \emptyset]:

I dialetti italiani settentrionali si presentano come un tipo intermedio fra le lingue a soggetto nullo e le lingue a soggetto obbligatorio; nella maggior parte di questi dialetti infatti alcune persone del verbo hanno sempre un pronome soggetto che si accompagna al verbo flesso, anche se è presente un soggetto, alcune persone invece sono a 'soggetto nullo', altre persone del verbo infine sono come nelle lingue a soggetto obbligatorio, cioè con un pronome soggetto solo se manca il soggetto del verbo. I pronomi soggetto di cui si tratta sono pronomi clitici, chi costituiscono un paradigma parallelo a quello dei pronomi tonici. All'interno di

questo schema generale, si ha una minutissima variazione di tipi possibili. . .²³

(Benincà 1992:35).

Benincà illustre ensuite cette «minutissima variazione» en se référant schématiquement à deux cartes de l'*AIS* (IL CACCIATORE VA A CACCIA, *AIS*518-519, et PIOVE?, *AIS*366). Le but de notre étude serait donc essentiellement d'étendre cet examen schématique de façon systématique à toutes les cartes pertinentes, et ce pour tous les points pertinents (tant de l'*AIS* comme de l'*ALF*) pour arriver à une typologie générale des sujets dans ces variétés.

Le grand défi de la recherche effectuée à partir des atlas linguistiques est de trouver suffisamment de données pertinentes, précisément parce que les variables morphosyntaxiques ne font pas partie des buts explicites de la plupart des atlas linguistiques romans et sont par conséquent à rechercher parmi les questions posées originalement avec d'autres buts en vue:

23. «Les dialectes italiens septentrionaux se présentent comme un type intermédiaire entre les langues à sujet nul et les langues à sujet obligatoire; en fait, dans la plupart de ces dialectes certaines personnes du verbe ont toujours un pronom sujet qui accompagne le verbe fléchi, même si un sujet est présent, certaines personnes sont à 'sujet nul', et d'autres personnes sont comme dans les langues à sujet obligatoire, c'est-à-dire avec un pronom sujet seulement si le sujet du verbe est absent. Les pronoms sujet dont il est question sont des pronoms clitiques, qui constituent un paradigme parallèle à celui des pronoms toniques. À l'intérieur de ce schéma général, il existe une variation tout à fait minutieuse de types possibles. . .»

A parte la loro scarsità oggettiva, va sottolineato che i dati sintattici compaiono negli atlanti [linguistici] in modo *casuale*; un aspetto non marginale del problema sta quindi nel fatto che anche i dati presenti sono in genere difficilmente reperibili, in quanto si trovano nelle ‘pieghe’ di dati ricercati per scopi diversi.²⁴

(Benincà 1992: 29).

Dans ce chapitre, nous avons distingué nettement les deux facettes de la variation linguistique (intergrammaticale et intragrammaticale), ainsi que les deux approches méthodologiques correspondantes et leurs implications théoriques quant à l’emploi des corpus dans la recherche linguistique. Dans le chapitre qui suit il est question de la «rareté objective» des données syntaxiques dont parle Benincà: nous tâchons d’élaborer une méthodologie pour dépouiller et analyser les formes des atlas qui nous intéressent dans cette étude.

24. «À part leur rareté objective, il faut souligner que les données syntaxiques apparaissent dans les atlas [linguistiques] de façon *fortuite*; un aspect non marginal du problème réside donc dans le fait que même les données existantes sont en général difficilement repérables, dans la mesure où elles se trouvent dans les “plis” de données recherchées pour d’autres fins.»

CHAPITRE 3: LES PRONOMS SUJET ET LA VARIATION GRAMMATICALE

3.0. PRÉSENTATION

Dans ce chapitre nous élaborons une méthodologie qui nous permet de cerner les différents facteurs linguistiques susceptibles d'influencer la présence des sujets pronominaux. Évidemment, cette méthodologie détermine, et est déterminée par, les hypothèses que nous formulons au sujet de ce phénomène; ces hypothèses sont présentées et justifiées par rapport aux études antérieures.

3.1. MÉTHODOLOGIE: UNE GRILLE HEURISTIQUE

Comme point de départ pour les informations précises recherchées dans cette enquête, l'étude de Renzi et Vanelli (1983) possède l'avantage d'énumérer un certain nombre de propriétés relatives aux pronoms sujet dans les variétés étudiées. Ces propriétés constituent en quelque sorte autant de dimensions à mesurer, soit pour chaque carte soit pour chaque point sur chaque carte. Ces sept «aspects de la pronominalisation» qu'ils étudient pour chaque géolecte, sont (1983:123):

- a) se la pronominalizzazione sia tonica (del tipo dell'ingl. *he, she*, o del ted. *er, sie*), o atona, cioè clítico (del tipo del franc. *il*);
- b) se la pronominalizzazione riguardi tutte le persone o solo alcune, e in quest'ultimo caso, quali;

- c) se i pronomi abbiano forma diversa in ogni persona oppure se ci siano delle forme uguali, e in quest'ultimo caso, per quali persone questo avvenga, e se ci sia rapporto con la varietà delle desinenze del verbo che segue;
- d) se il pronome appaia anche con soggetto nominale;
- e) se appaia con verbi impersonali e meteorologici e con l'esistenziale ('c'è');
- f) se appaia con i quantificatori 'chi' e 'nessuno';
- g) se nell'interrogazione il posto del pronome sia dopo il verbo (fenomeno della 'inversione') e, se questo avviene, quali caratteristiche abbiano questi pronomi.²⁵

Pour notre étude, nous adaptons les propriétés visées par Renzi et Vanelli (1983) aux données disponibles dans les atlas linguistiques. Certaines de ces dimensions sont relativement constantes pour toute une carte donnée, tandis que d'autres varient pour chaque point d'atlas, c'est-à-dire, pour chaque géolecte. Certaines questions sont

25. «a) si la pronominalisation est tonique (du type de l'angl. *he, she*, ou de l'all. *er, sie*), ou atone, c'est-à-dire clitique (du type du fr. *il*);
 b) si la pronominalisation touche toutes les personnes ou seulement certaines, et dans ce dernier cas, lesquelles;
 c) si les pronoms ont des formes différentes dans chaque personne ou bien si il y a des personnes identiques, et dans ce dernier cas, pour quelles personnes cela se produit, et s'il y a un rapport avec la variété des désinences du verbe qui suit;
 d) si le pronom apparaît aussi avec un sujet nominal;
 e) s'il apparaît avec des verbes impersonnels et météorologiques et avec l'existentiel (*c'è* = 'il y a');
 f) s'il apparaît avec les quantifieurs *chi* ('qui') et *nessuno* ('personne');
 g) si dans l'interrogation la position du pronom est après le verbe (phénomène de l'«inversion») et, si ceci arrive, quelles caractéristiques possèdent ces pronoms.»

nécessairement élaborées ou ajoutées, mais il s'agit surtout d'adapter le noyau des questions posées par Renzi et Vanelli (1983) à une sorte de grille heuristique pour l'étude géolinguistique que nous proposons à partir des données des atlas linguistiques. Pour chaque carte d'un atlas il existe donc un certain nombre de facteurs constants en commun qui permettront de trier les formes susceptibles de présenter des sujets pronominaux. Nous croyons que cette façon de structurer les tâches rend la formulation des questions plus explicite, et le dépouillement des données plus pratique. La liste en 3.2. représente donc les renseignements que nous recherchons dans chaque carte, organisés en groupes de facteurs (qui correspondent aux «champs» dans notre base de données). Les facteurs, ainsi que les valeurs possibles, sont expliqués par la suite (voir 3.3. et 3.4.).

3.2. GROUPES DE FACTEURS ET CODES CORRESPONDANTS

Groupe de facteurs	Facteurs ²⁶
1. N° de Carte	Cartes de l' <i>AIS</i> et de l' <i>ALF</i> (voir Appendice I et II)
2. N° de Point	266 points de l' <i>AIS</i> et 172 points de l' <i>ALF</i> (voir Appendice III)

26. Dans tous les groupes, le code «/» sert à indiquer que ce facteur ne s'applique pas. Ainsi les phrases affirmatives positives ont toutes la valeur «/» pour 2.1.11. (Négation) et 2.1.12. (Interrogation). De même, quand nous n'avons pas pu déterminer avec certitude une valeur (par exemple, le temps précis de certaines formes verbales), nous utilisons ce code, qui a l'effet d'exclure la forme en question de l'analyse. Quand «/» apparaît dans le groupe de facteurs 2.1.3. «Pronom sujet», l'occurrence en question, c'est-à-dire la forme qui apparaît à un point sur une carte, est exclue toute entière de notre corpus.

3. Pronom sujet	Q aucun sujet pronominal I sujet pronominal k sujet générique ²⁷
4. Position du verbe	initiale de proposition principale non initiale de principale initiale de subordonnée non initiale de subordonnée relative
5. Temps / mode	présent de l'indicatif futur simple subjonctif présent futur périphrastique passé composé passé simple (=prétérit) imparfait de l'indicatif conditionnel passé antérieur imparfait du subjonctif subjonctif parfait

27. Il s'agit de formes équivalentes du français *ça*, voir 4.1.1. et 4.2.3.1. Nous ne retenons pas la distinction faite par Renzi et Vanelli entre pronoms toniques et atones. (1983:123).

6. Type de verbe	<u>t</u> ransitif
	in <u>a</u> ccusatif
	in <u>e</u> rgatif
	<u>m</u> étéorologique
	impersonnel
	mo <u>d</u> al
	ré <u>fl</u> échi
	<u>a</u> voir (verbe plein)
	copu <u>l</u> e (être verbe plein)
	7. Auxiliaire
	<u>e</u> ssere (être auxiliaire)
8. Type de sujet	<u>a</u> ucun
	in <u>d</u> éfini
	<u>d</u> éfini
	pronominal (fort)
9. Personne grammaticale	1,2,3,4,5,6 ²⁸
10. Genre	<u>a</u> mbigu
	<u>f</u> éminin
	<u>m</u> asculin

28. Nous employons la convention, assez répandue chez les romanistes (voir par exemple Jaberg et Jud, 1928-1940), selon laquelle les trois personnes au pluriel sont numérotées quatrième («nous»), cinquième («vous») et sixième («elles» ou «ils»).

11. Négation	n égation simple (préverbale seulement) n égation forte (pré et post-verbale) n égation postverbale (uniquement)
12. Interrogation	t otale p artielle i ndirecte
13. Clitiques obliques	0,1,2
14. Position du sujet	p réverbale i nvertie d isloquée
15. Memo	(toute information supplémentaire non codée)

3.3. BASE DE DONNÉES

Pour faciliter et systématiser le dépouillement des données dans les atlas linguistiques, les codes correspondant aux groupes de facteurs énumérés en 3.2. sont entrés dans une base de données spécialement conçue à l'aide du logiciel Paradox for Windows®. Cette base de données permet notamment la comparaison systématique et la corrélation des faits dans les différentes variétés, en vue d'une analyse statistique globale des phénomènes en question. C'est ainsi que nous pouvons, à l'aide de cette base de données plus ample et détaillée, vérifier ou infirmer entre autres les «généralisations

établies» qu'avancent Renzi et Vanelli à partir de leur enquête pionnière portant sur une trentaine de variétés (1983:143-144).

3.3.1. Numéro de carte

L'information dans ce groupe de facteurs (3.2.1.) sert à identifier tout simplement la carte en question, soit de l'*ALF*, soit de l'*AIS*. Bien sûr, nous considérons uniquement les cartes présentant un environnement possible pour les pronoms sujet: celles où figurent des verbes fléchis dans une proposition non impérative.

Les cartes retenues (101 de chaque atlas) sont répertoriées dans les Appendices I et II. Leur numérotation correspond à celle des atlas originaux, à l'exception des variantes qui sont ajoutées à certaines cartes (surtout dans l'*AIS*). Ainsi par exemple *AIS*887 «Ho messo. . . un chiavistello» (= *ALF*847 «J'ai mis. . . un verrou»), inclut une quinzaine de points où la réponse se trouve être l'équivalent de «J'ai verrouillé la porte» (voir 4.1.2.). Ces variantes correspondent à un numéro décimal dans l'Appendice II (par exemple 887,1), ce qui indique qu'elles ne correspondent pas à des cartes mais à des variantes de cartes.

Nous ne considérons en principe que les cartes qui apparaissent «en paires», c'est-à-dire les cas où il y a un équivalent (plus ou moins exact) d'une carte de l'*ALF* dans l'*AIS*. Puisque Jaberg et Jud reprennent et élaborent le questionnaire utilisé par Gilliéron et Edmont, l'*AIS* contient des équivalents pour la plupart des cartes pertinentes de l'*ALF*, ce qui nous donne automatiquement un corpus plus ou moins «équilibré».

3.3.2. Numéro de point

Ce facteur (3.2.2.) sert à identifier les points d'enquête dans les deux atlas qui fournissent les formes relevées pour notre étude. Pour ce qui est de l'*AIS*, la zone qui nous intéresse dans cette étude correspond en gros à la moitié septentrionale: les points 1 à 548 (juste au sud de Florence), soit 261 points d'enquête en tout. Cette zone comprend le territoire qui s'étend des frontières septentrionales de l'Italie jusqu'au centre de l'aire toscane, un peu au sud de la ligne «La Spezia-Rimini». Cette division dialectologique traditionnelle, qui se base sur des critères essentiellement phonologiques (voir von Wartburg 1967:IV), serait parallèle à la frontière septentrionale des variétés à sujet nul:

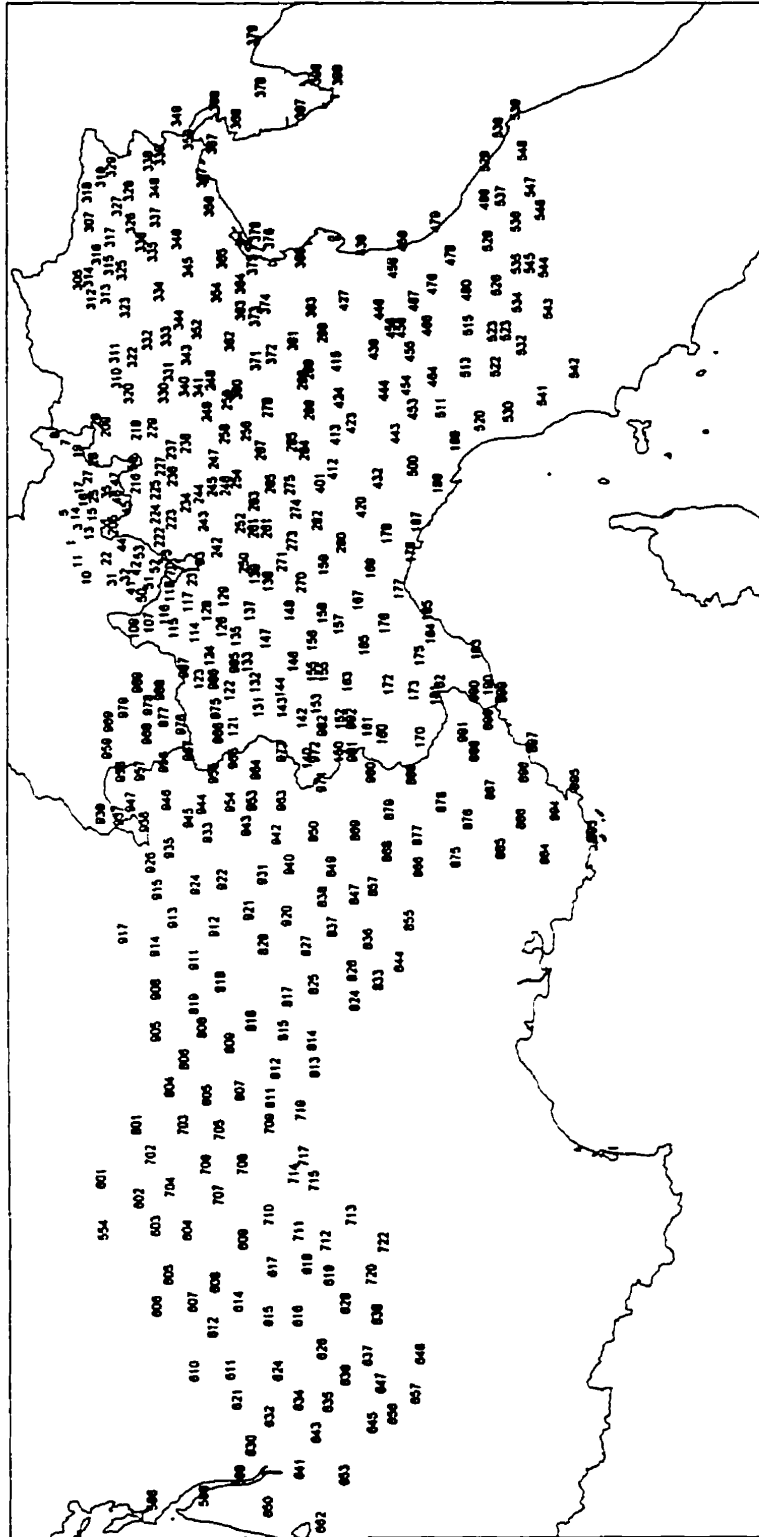
Da quanto si conosce, anche per la sintassi l'Italia dialettale sembra divisa in due, con una linea che, pur non corrispondendo alla linea La Spezia - Rimini che si basa su isoglosse fonologiche, non ne dista molto, dato che se ne discosta solo l'area di Firenze, che va con l'Italia settentrionale per la sintassi dei pronomi clitici soggetto.²⁹ (Benincà 1990:12)

29. «Autant que l'on sache, l'Italie dialectale semble se diviser en deux pour ce qui est de la syntaxe aussi, avec une ligne qui, sans vraiment correspondre à la ligne La Spezia - Rimini qui se base sur des isoglosses phonologiques, ne s'en éloigne pas beaucoup, étant donné que seule s'en détache la zone de Florence, qui va avec l'Italie septentrionale pour ce qui est de la syntaxe des pronoms clitiques sujet.»

La limite méridionale de notre étude se trouve donc bien ancrée dans une zone à sujet nul. Notons que la variété de l'italien dit «standard» n'a pas vraiment de localité géographique fixe, mais pour ce qui est de l'emploi des sujets pronominaux, l'usage standard correspondrait au toscan méridional plutôt qu'au florentin (qui emploierait plutôt des pronoms sujet «préfixés», Wanner 1995:453).

Il s'avère considérablement plus complexe de tracer la limite septentrionale de notre étude dans l'*ALF*. Évidemment, la partie sud-est continue (et chevauche même) les zones contiguës de l'*AIS*: nous devons sans doute inclure les départements des Alpes-Maritimes, Basses-Alpes, Hautes-Alpes, Savoie et Haute-Savoie, ainsi que les points en territoire italien. Nous déplaçant vers l'ouest, nous trouvons que la transition entre sujets nuls et sujets non nuls recoupe à peu près la tracée de l'important faisceau d'isoglosses divisant le tiers méridional de l'Hexagone des deux-tiers septentrionaux (Jochowitz 1973). Commençant à la frontière italienne près de Grenoble, ce faisceau monte légèrement autour du Massif Central vers Clermont-Ferrand, pour redescendre ensuite et rejoindre l'Atlantique à l'estuaire de la Gironde.

On retrouve ici la division traditionnelle du domaine gallo-roman: la frontière septentrionale de la langue d'oc (voir 2.1.4. et 2.1.5.). Comme cette frontière se trouve entourée d'une frange de variabilité, il semble prudent de laisser une bande des deux côtés de cette isoglosse (Isoglosse 7 de Jochowitz, 1973: 129) et d'inclure donc les points dans les départements suivants: Gironde, Dordogne, Corrèze, Loire, Rhône et Isère, ainsi que quelques points dans les départements voisins.



Carte 1: Les points d'enquête de l'ALF et de L'AIS (Nota: les points 504, 536, 548 et 549 de l'ALF sont renumérotés 554, 586, 598 et 599 respectivement. Voir 3.3.2. et Appendice III).

Dans les grandes villes les enquêteurs de l'*AIS* ont relevé leur questionnaire à deux reprises avec des sujets différents, correspondant à deux couches sociales différentes, une innovation par rapport aux atlas antérieurs (Jaberg 1936:21; Alvar 1960:56). Ces points sont donc «divisés» dans leur *Atlas*, avec des réponses séparées pour les deux couches. Dans notre corpus, les données de ces cinq villes correspondent aux points numérotés 155,1 et 155,2 (Turin); 261,1 et 261,2 (Milan); 376,1 et 376,2 (Venise); 456,1 et 456,2 (Bologne); et 523,1 et 523,2 (Florence). Ces doubles enquêtes portent le nombre total de points de l'*AIS* étudiés à 266.³⁰

Puisque les réalités linguistiques dépassent les barrières politiques, il est normal que ni l'un ni l'autre des deux atlas ne respecte les frontières nationales (ni actuelles ni de leur époque). Ainsi on trouve dans notre échantillon de points de l'*ALF* une douzaine de points dans la Suisse romande (canton de Vaud), ainsi que 8 points en territoire italien, essentiellement dans le Val d'Aoste et les Alpes piémontaises (et chevauchant donc légèrement les points 121 à 190 de l'*AIS*). Pour sa part, l'*AIS* commence dans la Suisse méridionale (comme l'indique d'ailleurs son titre complet), où il inclut 36 points dans les cantons de Tessin et des Grisons. Il embrasse également 5 points dans la péninsule

30. Le même problème ne se pose pas pour ce qui est de l'*ALF*: Edmont évitait soigneusement toute ville importante (1902:4) et il n'a pas relevé de questionnaire multiple. La liste des sujets (1902:29-55) indique qu'à certains points il lui a fallu deux ou même plusieurs sujets pour remplir un questionnaire, mais il s'agissait dans tous les cas d'un seul questionnaire, avec l'exception unique du point 284, St.-Pol-sur-Ternoise, où Edmont lui-même a servi de sujet pour St.Pol (ville) et St.Pol (faubourgs). Ce point tombe bien en dehors de la zone qui nous intéresse; dans le Midi, les sujets uniques prédominent.

istrienne, qui fait maintenant partie de la Croatie. Le répertoire complet des 438 points étudiés (172 de l'*ALF* et 236 de l'*AIS*) se trouve dans l'Appendice III, et la Carte 1 sert à localiser ces points dans le territoire pertinent des deux atlas.³¹

3.3.3. Hypothèses de travail

À l'intérieur de l'objectif général qui consiste à vérifier empiriquement la validité dudit «paramètre du sujet nul» face à la variation géolectale attestée dans les systèmes pronominaux du roman central, nous formulons un certain nombre d'hypothèses de travail plus précises, en tenant compte des travaux précédents (théoriques ou empiriques) sur les sujets pronominaux. Les informations codées dans les groupes de facteurs de la section 3.2. correspondent donc aux différentes hypothèses que nous présentons dans les sections qui suivent. Nos hypothèses concernent la question de l'emploi catégorique ou variable des sujets pronominaux dans les variétés à pronoms sujet, la typologie des paradigmes pronominaux dans les mêmes, et les divers facteurs linguistiques susceptibles de conditionner la présence ou l'absence des sujets pronominaux.

31. Quatre localités françaises sont renumérotées pour ne pas se confondre avec la série de l'*AIS*: les points 504, 536, 548 et 549 de l'*ALF* apparaissent dans cette étude comme 554, 586, 598 et 599 respectivement.

3.3.4. Emploi catégorique ou variable des sujets pronominaux

La méthodologie de Renzi et Vanelli (1983) est fondée sur un postulat crucial: soit les sujets pronominaux dans chaque variété étudiée sont catégoriquement présents, soit ils ne le sont pas («ci sono lingue che fanno uso costante del pronome soggetto e lingue che non ne fanno un uso costante.»³² 1983:122). Dans la présente étude, nous confrontons ce type de détermination tranchante avec des données qui risquent d'être beaucoup moins nettes: aucun des points dans le petit échantillon de l'*AIS* étudié dans Heap (1993) ne présente un emploi catégorique des sujets pronominaux. Nous nous attendons donc à ce que la notion de «variété à pronoms sujet» doive être formulée de façon moins absolue.

En fait, même Renzi et Vanelli (1983) nuancent leur position quant à la nature catégorique de cette opposition binaire [\pm pronom sujet], en citant le cas des variétés où «determinate persone presentano il pronome o assenza di pronome in *variazione libera*» (1983:129).³³ À partir de leurs données, ils formulent la généralisation que les pronoms sujet de la deuxième, troisième et sixième personne, s'ils sont présents dans une variété, seraient obligatoires, tandis que la première, quatrième et cinquième personne peuvent avoir des pronoms sujet facultatifs (1983:129). En revanche, Benincà et Poletto (1991:81) présentent des cas où «il clítico soggetto sembra essere completamente opzionale se la

32. «Il y a des langues qui font un usage catégorique du pronom sujet, et des langues qui n'en font pas un usage catégorique.»

33. «Certaines personnes présentent la présence ou l'absence du pronom en *variation libre*.»

frase ha un sogetto lessicale».³⁴ l'exemple qu'elles donnent est celui d'une variété vénitienne dans laquelle on peut dire soit *Nane le vien* soit *Nane vien* pour 'Jean vient'. Il est évident qu'une telle variabilité à la troisième personne va à l'encontre des généralisations de Renzi et Vanelli, pour qui la variation est exclue pour les pronoms sujet de la troisième personne (1983:129). Nous examinons donc ces contradictions apparentes à la lumière de nos données, en considérant les interactions éventuelles entre personne grammaticale et présence ou absence des sujets.

C'est donc à partir de notre base de données que nous distinguons les variétés à sujet pronominal catégorique (ou presque catégorique) de celles où leur emploi est variable. Les cas (quasi-)catégoriques sont traités typologiquement d'après les schémas de Renzi et Vanelli (voir 3.3.5. et 4.3.). Ensuite, pour les cas à sujets variables, nous essayons de déterminer quels facteurs favorisent ou défavorisent l'emploi du pronom sujet (voir 3.4.)

3.3.5. Typologie des paradigmes

Notre base de données nous permet d'élaborer une caractérisation plus détaillée des variétés *partiellement* à sujets pronominaux. Le groupe de facteurs 3.2.9. «Personne grammaticale» correspond à la question b) de Renzi et Vanelli (1983): il nous permet d'identifier les variétés qui seraient «à pronoms sujet» mais dans certaines personnes

34. «Le clitique sujet semble être complètement optionnel si la phrase a un sujet lexical.»

grammaticales seulement, et d'établir une typologie des paradigmes possibles. À titre d'exemple, Renzi et Vanelli concluent que dans une variété qui possède un seul sujet pronominal, ce dernier serait nécessairement de la deuxième personne (1983:128). Pour ce qui est des variétés où il existe des sujets pronominaux à deux personnes, ils avancent la généralisation (à une exception près) que ceux-ci seraient de la deuxième et troisième personne. Pour les variétés à trois pronoms sujet, ils seraient nécessairement à la deuxième, troisième et sixième personne; quatre pronoms seraient forcément aux première, deuxième, troisième et sixième personnes, tandis que dans une variété à cinq pronoms sujet, ils seraient de toutes les personnes sauf la première (1983:128).³⁵ Nos données permettent de confirmer ou d'infirmer la validité de ces généralisations implicationnelles, et de déterminer si cette typologie est effectivement exhaustive ou s'il existe en fait d'autres paradigmes attestés de pronoms sujet. Reste à voir si l'on peut ramener ce type d'observations à des principes plus profonds.

Un autre aspect pertinent des paradigmes pronominaux est la possibilité de syncrétisme, ou l'identité éventuelle entre les différents pronoms d'un paradigme. Renzi et Vanelli (1983) donnent une liste assez détaillée des différents cas de syncrétismes

35. À première vue, il peut surprendre que la première personne soit la seule absente dans le système à cinq pronoms, mais incluse dans le système à quatre pronoms. Retenons quand même que dans l'étude de Renzi & Vanelli, ces deux systèmes ne sont représentés que par une seule variété chacun (le provençal alpin de Rodoretto et le garfagnino di Vagli di Sopra, respectivement). Ces deux variétés étant assez éloignées géographiquement, il est possible qu'elles fassent partie d'un continuum typologique de paradigmes pronominaux dont les autres membres n'apparaissent pas dans l'étude de Renzi et Vanelli (voir 4.3.2.).

pronominaux, qui varient en fonction du nombre de sujets pronominaux qui existent dans chaque paradigme (1983:130-32). Ces cas de syncrétisme devraient être étudiés en conjonction avec le syncrétisme parmi les désinences verbales, qui peuvent également influencer la présence des pronoms sujet. Renzi et Vanelli rejettent l'hypothèse selon laquelle les pronoms sujet viennent remplacer les désinences verbales phonétiquement neutralisées. D'après eux, il serait difficilement explicable selon cette hypothèse que le pronom sujet le plus répandu soit celui de la deuxième personne (*tu*), étant donné que cette personne est également la seule qui est toujours bien marquée par une désinence verbale. Il paraît donc que les grammaires en question tolèrent un certain degré de redondance morphologique plutôt que de permettre l'identité formelle des personnes grammaticales:

In altre parole, nelle varietà esaminate la neutralizzazione di due o più persone non è ammessa. Dunque le persone della flessione verbale sono differenziate tra di loro con il morfema desinenziale o con il pronome o con entrambi.³⁶ (1983: 133-134).

Malheureusement, la vérification de cette généralisation, qui correspond à la question c) de Renzi et Vanelli (1983:123, voir 3.2.) dépasse la portée de notre étude, puisque nous

36. «En d'autres mots, dans les variétés examinées la neutralisation de deux ou plusieurs personnes n'est pas admise. Les personnes de la flexion verbale sont différenciées entre elles soit par le morphème désinentielle soit par le pronom soit par les deux.»

ne dépouillons pas les *formes* des pronoms sujet pour notre base de données, mais uniquement leur présence ou absence. Une étude du syncrétisme pronominal doit donc faire partie des recherches ultérieures.

3.4. FACTEURS CONDITIONNANTS

Comme on peut le constater dans la section 3.2., nos hypothèses de départ sur les facteurs conditionnants possibles sont assez ouvertes: nous admettons la possibilité qu'un grand nombre de facteurs linguistiques puissent éventuellement influencer la présence ou l'absence des pronoms sujet. Nous sommes convaincu de l'importance de prendre en considération la gamme la plus vaste possible de facteurs dans l'analyse du conditionnement des sujets pronominaux, puisqu'il est impossible de savoir à l'avance lesquels seront pertinents ou non pertinents. Évidemment, nous pourrions plus facilement éliminer plus tard des facteurs qui s'avèrent non pertinents qu'inclure après-coup de nouveaux facteurs. Benincà et Poletto soulignent l'importance de partir d'hypothèses ouvertes et d'inclure des facteurs dont le lien direct avec les pronoms sujet ne paraît pas évident:

Sembra quindi che ci siano molti fattori che intervengono nella distribuzione dei clitici soggetto. Oltre a fattori quali, ad esempio, la presenza di un soggetto lessicale, che ci possiamo attendere sia collegata con i clitici soggetto, dobbiamo prendere in considerazione anche altri elementi, che non sembrano avere a prima

vista un legame diretto con il soggetto, quali ad esempio il tipo di frase utilizzata o la presenza di un certo ausiliare.³⁷ (1991:84).

En fait, les seules limites sur les facteurs à considérer qui s'imposent de façon inéluctable sont celles de notre corpus: nous ne pouvons prendre en compte que les informations qui apparaissent dans les atlas linguistiques dont nous partons: nous n'avons presque pas de données sur les relatives et sur les sujets indéfinis, par exemple. À l'intérieur de ces limites, la liste de facteurs en 3.2. représente nos tentatives d'inclure tout facteur linguistique que l'on peut corréler de façon systématique avec les sujets pronominaux. Par exemple, dans les deux sections précédentes il s'agit de deux aspects différents de l'interaction du facteur de la personne grammaticale avec la présence des pronoms sujet: toutes nos données sont codées systématiquement pour la variable 3.2.9., «personne grammaticale». Nous donnons par la suite de brèves justifications et discussions des autres facteurs pour lesquels nous prévoyons des effets ou des corrélations éventuels.

37. «Il semble donc y avoir beaucoup de facteurs qui influencent la distribution des clitiques sujet. Outre les facteurs tels, par exemple, la présence d'un sujet lexical, pour lesquels une corrélation avec les sujets clitiques est prévisible, nous devons aussi prendre en considération d'autres éléments, qui ne semblent pas à première vue avoir un lien direct avec le sujet, tels, par exemple, le type de phrase utilisé ou la présence d'un certain auxiliaire.»

3.4.1. Position du verbe

Il s'agit ici du type de proposition dans laquelle se trouve le verbe fléchi dont il est question dans chaque forme relevée dans les atlas (3.2.4). Ce groupe de facteurs reprend la règle rythmique du roman primitif, connue sous le nom de «Loi de Tobler-Mussafia» (1.3.2), qui interdirait les éléments non accentués en position initiale de phrase (Mussafia 1886, 1898; Wanner 1987:25). On peut donc envisager que la position initiale de proposition aura un effet défavorable sur l'emploi des sujets pronominaux, en contraste avec les verbes en position non initiale, (que ceux-ci soient précédés de sujets nominaux, d'adverbes ou de conjonctions).³⁸ Ce groupe vise également la différence entre principales et subordonnées, puisque celles-ci seraient l'endroit préféré des pronoms sujet dans l'évolution du français (Moignet 1930; Franzén 1939; Price 1984).

Notons qu'au niveau pragmatique, une subordonnée entraîne souvent un changement de référence par rapport à la principale. Ce type de changement («switch reference») favorise l'emploi des pronoms sujet dans des langues où ceux-ci sont facultatifs: en espagnol, par exemple, un changement de référence entraîne souvent un pronom sujet pour marquer le changement (Cameron 1993). Malheureusement, les formes auxquelles nous avons affaire dans les atlas ne constituent pas du vrai discours mais consistent plutôt en de courtes phrases à traduire, où on perd ce type de dynamique

38. Ceci dit, dans une carte comme AIS1250 («Avreste dovuto vedere» *Vous auriez dû voir*), on trouve bien des points avec des réponses «extorquées» (voir 4.1.2.) du type «Si seulement vous aviez vu». Ces occurrences sont codées comme des subordonnées.

pragmatique. Nous ne pouvons donc pas coder les occurrences pour un facteur du type «changement de référence» par opposition à «référence continue», puisque ces informations ne sont pas disponibles dans notre corpus.

3.4.2. Temps et mode

À quelques exceptions près, les études sur les pronoms sujet se concentrent quasi-exclusivement sur les verbes au présent de l'indicatif. Or, étant donné les hypothèses évoquées ci-dessus quant aux effets possibles des auxiliaires *essere* et *avere*, la prise en considération des temps et des modes verbaux s'impose. D'une part, les temps composés présentent la possibilité de tester l'effet global des auxiliaires sur la présence des sujets pronominaux, indépendamment du choix lexical entre *essere* et *avere* (un choix qui est lui-même susceptible de varier géographiquement). D'autre part, les temps simples autres que le présent, ainsi que le subjonctif et le conditionnel, présentent plus de cas de synchrétisme dans les désinences verbales que l'indicatif présent. Comparons à titre d'exemple le paradigme des trois premières personnes d'un verbe régulier en italien, au présent de l'indicatif, à l'imparfait et au subjonctif présent:

(26) Verbe italien aux première, deuxième et troisième personnes:

présent de l'indicatif		imparfait de l'indicatif	subjonctif présent
(io)	<i>parlo</i>	<i>parlava</i>	<i>parli</i>
(tu)	<i>parli</i>	<i>parlavi</i>	<i>parli</i>
(lui, lei)	<i>parla</i>	<i>parlava</i>	<i>parli</i>

Comme on voit dans les formes en (26), des trois personnes à désinences distinctes au présent de l'indicatif, seulement deux sont distinctes à l'imparfait, et il y a synchrétisme complet entre les trois au subjonctif présent. Ce fait semble être relié à l'usage fréquent des sujets pronominaux forts en italien standard dans ces deux derniers cas,³⁹ en dépit de l'hypothèse qui voudrait que les sujets pronominaux soient «superflus» et donc facultatifs dans les langues «à sujet nul» comme l'italien et l'espagnol. À cet égard, Hochsberg (1986) démontre en fait que la fréquence des sujets pronominaux en espagnol varie en fonction du synchrétisme désinentiel (entre autres facteurs): dans les temps et les modes où les désinences verbales ne distinguent pas parfaitement les différentes personnes grammaticales, on emploie beaucoup plus de sujets pronominaux. Puisque dans les variétés qui nous concernent, la plupart des verbes distinguent généralement les différentes personnes grammaticales mieux au présent de l'indicatif, il est raisonnable de supposer que

39. Rohlfs (1966:145,n.) énumère certaines circonstances exceptionnelles où l'italien standard requiert un pronom sujet, par exemple pour désambigüer les formes verbales à désinences synchrétiques: *che io/tu/lui venga*. Voir aussi Freedman (1983).

plus un temps ou un mode présente des désinences identiques, plus on trouvera de pronoms sujet pour différencier les personnes, conformément à la généralisation de Renzi et Vanelli (1983:133-134) qui interdirait la neutralisation totale des distinctions entre les formes verbales.

Cette hypothèse fonctionnaliste est cependant loin d'être universellement acceptée: maintes études variationnistes trouvent justement le contraire, c'est-à-dire que les morphèmes de sujet apparaissent plus fréquemment dans des contextes où ils seraient en fait «redondants» du point de vue strictement fonctionnel.⁴⁰ Le groupe de facteurs en 3.2.5. nous permet d'évaluer ces deux hypothèses opposées quant à l'effet éventuel du temps et du mode sur la présence ou l'absence des sujets pronominaux.

3.4.3. Type de verbe et auxiliaire

Le type de verbe en question peut aussi avoir un impact sur la présence des pronoms sujet; le problème réside dans la difficulté de déterminer quelles catégories de verbes seront pertinentes pour notre analyse. Renzi et Vanelli (1983:123), dans leur question e), ne considèrent que les effets des verbes impersonnels, météorologiques et existentiels, ainsi que des auxiliaires. Nous avons préféré commencer par une classification plus élaborée des types de verbes, tout en sachant que certaines catégories de verbes

40. Voir Labov (1994:chapitre XIX) pour un survol des études sur les tendances contre-fonctionnelles et la préservation de structures parallèles, et Cameron (1993) pour une discussion nuancée des pronoms sujet espagnols sous cette optique.

devront sans doute être combinées, si notre analyse quantitative indique qu'il n'y a aucune différence significative entre leurs effets respectifs. Ceci dit, la liste donnée en 3.2.6. et 3.2.7. ne prétend nullement à l'exhaustivité, elle représente uniquement des hypothèses sur certains types de verbes qui s'avèrent pertinents pour cette problématique.

Dans certaines variétés la présence des sujets pronominaux est en corrélation avec les verbes aux temps parfaits accompagnés des auxiliaires *essere* et *avere* (Benincà et Poletto 1991:83). D'après Renzi et Vanelli (1983:129-130) ce facteur semble avoir des effets contradictoires: ces auxiliaires peuvent exiger des sujets pronominaux à toutes les personnes dans des variétés où ils ne sont normalement obligatoires qu'à certaines personnes, ou au contraire, ils peuvent faire disparaître les pronoms sujet autrement obligatoires à certaines personnes dans certaines variétés. Nous tâcherons de distinguer les différents effets possibles de ce facteur, en tenant compte de différents emplois de ces verbes (auxiliaires, existentiels, copulatifs), d'où la nécessité d'un autre groupe de facteurs, 3.2.7., consacré à la spécification des auxiliaires.

Renzi et Vanelli (1983:135-37) énoncent une hypothèse assez forte à propos des verbes impersonnels (du type *bisognare* «falloir»), météorologiques (du type *piovere* «pleuvoir»), et existentiels (du type *esserci* «y avoir»): d'après eux, seules les variétés à sujets pronominaux obligatoires à toutes les personnes auront aussi des pronoms sujet pour ces trois types de verbes, et **inversement**: il s'agirait d'un conditionnement **nécessaire, suffisant et bidirectionnel**. Nous essayons donc de vérifier si cette généralisation peut tenir, face aux faits des atlas linguistiques.

Pour leur part, Benincà et Poletto indiquent que leur questionnaire tient compte de la différence entre les verbes transitifs et intransitifs, et parmi ces derniers, les intransitifs qui sélectionnent l'auxiliaire *avere* «avoir» et ceux qui sélectionnent *essere* «être», mais sans donner des raisons concrètes pour cette distinction (1991:85,95). Dans son étude du redoublement du sujet dans le français ontarien, Nadasdi (1995:99-102) distingue trois catégories de verbes selon la position initiale des éléments qui apparaissent en position de sujet en surface: les transitifs, les inaccusatifs et les passifs. D'après son hypothèse, «puisque les [sujets des] inaccusatifs et les passifs sont des objets de base, leur sujet sera redoublé moins souvent par un CLS [clitique sujet] que les verbes dont le sujet est aussi un sujet de base.» (1995:100) Son analyse quantitative révèle que ce groupe de facteurs est effectivement pertinent: les transitifs favorisent les sujets pronominaux plus que les inaccusatifs, qui à leur tour les favorisent plus que les passifs (1995:113). Puisque son hypothèse de départ conduit à regrouper les inaccusatifs et les passifs, ce dernier résultat est plutôt surprenant: on voit mal pourquoi les passifs se distingueraient des inaccusatifs de cette façon. Il semble donc que ce groupe de facteurs peut nous aider à déterminer si (et dans quelle mesure) les différents types de verbes favorisent l'emploi des sujets pronominaux.

3.4.4. Genre

Il s'agit bien sûr du genre grammatical du sujet des verbes, qui n'a pas toujours de marque morphologique explicite. En général, ce ne sont que les verbes des troisième et

sixième personnes qui marquent systématiquement le genre du sujet, s'il y a un pronom; des fois, l'accord d'un adjectif prédicatif donne aussi le genre à la première ou deuxième personne. Autrement, nous indiquons que le genre est ambigu, ce qui arrive dans beaucoup de cas. Ce n'est pas une hypothèse très précise que nous vérifions ici, mais plutôt l'intuition que le genre pourraient influencer le taux d'emploi des pronoms sujet, et qu'il vaut donc mieux vérifier empiriquement cette possibilité, ce que nous faisons avec le groupe de facteurs 3.2.10.

3.4.5. Présence de négation

Benincà et Poletto (1991) présentent des données du frioulan qui indiquent que la négation préverbale peut «bloquer», pour ainsi dire, la présence d'un pronom sujet qui serait obligatoire ailleurs. Ainsi dans (27), on voit une phrase qui est grammaticale avec un pronom sujet ou avec la négation préverbale, mais non pas avec les deux en même temps:

(27) *I vin capit.*

CL avons compris.

No vin capit.

NÉG. avons compris.

**I no vin capit.*

CL. NÉG. avons compris. (1991:84).

Nadasdi aussi trouve que la présence de la négation préverbale défavorise quantitativement les pronoms sujet (1995:96, 118) en français ontarien. Renzi et Vanelli (1983:130) observent qu'il existe des variétés où la négation préverbale empêche les sujets pronominaux d'apparaître à certaines personnes mais non pas à d'autres.

Remarquons que pour Benincà et Vanelli l'effet de la négation est lié à celui des autres clitiques: « . . . un clítico oggetto o la negazione preverbale possono “bloccare” l'occorrenza di un clítico soggetto.»⁴¹ (1991:84) De même, Nadasdi regroupe la négation préverbale par le clitique négatif *ne* avec les clitiques objet et réfléchi dans un seul groupe de facteurs dénommé «présence de clitique» (1995:96). Il se trouve que le clitique *ne* a un effet défavorisant beaucoup plus fort que les autres clitiques (1995:112), ce que Nadasdi attribue au fait que la négation par *ne* est assez rare dans son corpus (1995:118), et peut être reliée au français standard écrit (par opposition au français ontarien parlé de son corpus, qui ne connaît presque pas le *ne* préverbal). Nous préférons néanmoins garder la négation séparée des autres clitiques dans notre groupe de facteurs 3.2.11., en distinguant soigneusement entre la négation simple (du type *ne, non*), la négation forte (du type *ne . . . pas, non . . . niente*), et la négation exclusivement postverbale (du type . . . *pas, . . . niente*). Nous espérons ainsi déterminer si l'effet de «blocage» est dû au fait que la négation occupe l'espace préverbal où un pronom sujet pourrait apparaître, ou si c'est

41. « . . . un clitique objet ou la négation préverbale peuvent “bloquer” l'occurrence d'un clitique sujet.»

plutôt la nature même des propositions négatives qui «bloque» les sujets pronominaux (voir 5.4.5.).

3.4.6. Présence d'interrogation

Renzi et Vanelli ne trouvent aucun rapport entre l'absence de l'interrogation et les systèmes de sujets pronominaux, mais ils constatent qu'une corrélation étroite semble exister dans l'autre sens: «Le varietà con pronome obbligatorio hanno tutte anche l'inversione interrogativa.»⁴² (1983:138) Ici aussi, ils décrivent différents types de paradigmes pronominaux selon les personnes grammaticales qui ont un pronom postposé au verbe dans les inversions interrogatives, et ils en tirent une autre généralisation:

Se una varietà forma l'interrogazione mediante l'inversione del pronome, il numero delle persone con il pronome costante è uguale o superiore al numero delle persone con pronome nell'assertiva. In ogni caso sono presenti le stesse persone dell'assertiva.⁴³ (1983:139)

42. «Toutes les variétés à pronom obligatoire ont aussi l'inversion interrogative.»

43. «Si une variété forme l'interrogative par le biais de l'inversion du pronom, le nombre des personnes avec le pronom catégorique est égal ou supérieur au nombre de personnes avec des pronoms dans les déclaratives. En tout cas, les mêmes personnes sont présentes que dans la déclarative.»

Cette généralisation, qu'on peut assimiler à l'ensemble des «universaux relatifs» (ou implicationnels), constitue une hypothèse que les données fournies par notre groupe de facteurs 3.2.12. devraient permettre de vérifier quantitativement. Nous distinguons l'interrogation totale de l'interrogation partielle parce que celle-ci comporte un élément interrogatif (QU-), le plus souvent en position initiale, qui pourrait avoir un effet de «blocage» sur les pronoms sujet, tout comme la négation ou les pronoms clitiques obliques (voir 3.4.5. et 3.4.7.)⁴⁴

3.4.7 Présence de clitiques obliques

Renzi et Vanelli (1983:130) remarquent que dans certaines variétés à pronoms sujet, ces pronoms disparaissent si le verbe est précédé d'un pronom clitique autre que le sujet. Des données semblables sont présentées par Benincà et Poletto (1991), ce qui justifie l'inclusion de ce paramètre dans leur questionnaire d'exemplification (1993:95). Ainsi on trouve dans la variété de Pederobba que le pronom sujet ne peut pas apparaître devant un clitique objet:

44. Notons en passant que Benincà et Poletto incluent l'interrogation comme paramètre dans leur questionnaire (1991:92), non seulement parce qu'elle peut influencer la présence des sujets pronominaux, mais aussi parce qu'elles étudient aussi la formation des phrases interrogatives (structures «clivées», double *che*), un thème de recherche passionnant qui dépasse malheureusement les limites du présent travail.

(28) *Nane l'a vist Toni.*

Nane CL. a vu Toni.

Nane m'a vist.

Nane m'a vu.

**Nane le m'a vist.*

Nane CL. m'a vu. (1991:83)

Pour Nadasdi, ce phénomène serait relié à des contraintes sur la longueur des suites de clitiques:

Ceci est sans doute dû à une tendance à limiter le nombre de clitiques qui peuvent paraître devant un verbe. À cet égard, Morin (1981:101) note “It is true that long sequences of clitics are less frequent than short ones. Single clitics are more frequent than sequences of two clitics and sequences of three clitics are rather exceptional.” Donc, vu que les pronoms sujet faibles sont des clitiques, il est probable que la présence d'autres clitiques défavorise le redoublement par un CLS [clitique sujet]. (1995:96)

Cette hypothèse de Nadasdi est partiellement confirmée par son analyse quantitative: outre l'effet fortement défavorable du clitique négatif *ne* (voir la 3.4.6), les clitiques objet direct et indirect défavorisent la présence d'un sujet clitique, mais les clitiques pronominaux ou

réfléchis ne la défavorisent pas (1995:116-117). Ces derniers n'auraient pas d'impact significatif sur la présence des sujets pronominaux puisqu'ils feraient partie de l'entrée lexicale du verbe plutôt que d'être introduits dans la composante morphologique.

Dans le groupe de facteurs 3.2.13. nous codons donc la présence ou l'absence d'objets clitiques, avec des valeurs qui vont de zéro à trois (puisque l'italien tolère des suites clitiques plus longues que le français) pour chaque carte. Le type de verbe (groupe de facteurs 3.2.6.) nous indique s'il s'agit d'un réfléchi; autrement, c'est le nombre de clitiques préverbaux que nous prenons en compte ici, sans nous préoccuper ni de leur nature (objet direct, indirect, réfléchi, locatif ou partitif) ni de leur linéarisation interne.⁴⁵

3.4.8. Type et position du sujet

Dans le groupe de facteurs 3.2.8. nous cherchons à distinguer les sujets argumentaux, qui peuvent être soit des syntagmes nominaux définis ou indéfinis soit des pronoms forts, des sujets non argumentaux, qui sont les pronoms sujet faibles. C'est un fait bien documenté que la présence d'un sujet lexical a un effet important (mais variable) sur la présence des sujets pronominaux. Par exemple, Benincà et Poletto (1983:81) contrastent les cas suivants:

45. La question de l'organisation des séquences de clitiques, ainsi que la variabilité de ces séquences, est un sujet fascinant mais très complexe, qui échappe depuis des décennies à une explication adéquate (voir entre autres Perlmutter 1971; Wanner 1974, 1978, 1996; Bonet 1991, 1995a, 1995b; Heap 1996)

- | | | | | |
|------|----|---------------------|----|----------------------|
| (29) | a. | <i>Al ven.</i> | b. | <i>Le vien.</i> |
| | | 'Il vient.' | | 'Il vient.' |
| | | <i>Meni al ven.</i> | | <i>Nane le vien.</i> |
| | | 'Meni il vient.' | | 'Nane il vient.' |
| | | * <i>Meni ven.</i> | | <i>Nane vien.</i> |
| | | 'Meni vient.' | | 'Nane vient.' |

Dans la série (29)a., du frioulan, le pronom sujet est obligatoire même après un sujet lexical, tandis qu'en (29)b., d'une variété vénitienne, le pronom sujet est complètement optionnel dans un cas parallèle.

À cette opposition entre sujets obligatoires et optionnels s'ajoute la différence entre sujets lexicaux définis et indéfinis. En trentino, par exemple, le pronom sujet est obligatoire avec un sujet lexical défini, mais impossible avec un sujet indéfini, comme le montrent les données en (30) de Benincà et Poletto (1991:82):

- (30) *Le Mario le vien.*
 'Mario il vient.'
 * *Le Mario vien.*
 'Mario vient.'
Nisun vien.
 'Personne vient.'

* *Nisun le vien.*

‘Personne il vient.’

Cette contrainte sur l’emploi des pronoms sujet serait reliée à celle qu’avance Roberge (1990:119), d’après laquelle les sujets indéfinis et quantifiés ne peuvent pas avoir un sujet redoublé en français populaire. Les exemples en (31) seraient donc agrammaticaux en français populaire à cause de cette contrainte sur le redoublement du sujet:

(31) * Tous les hommes ils aiment manger.

* Quelques gars ils ont trop bu hier.

* Une fille elle va chanter.

Or il existe des raisons pour croire que cette contrainte n’a pas le même statut dans toutes les variétés du français: Nadasdi (1995:81) trouve dans un corpus de français ontarien des exemples comme ceux de (32) qui violeraient la contrainte sur le redoublement du sujet avec les indéfinis et les quantifiés.

(32) Tous mes oncles ils vivent au Québec.

La majorité des professeurs ils s'entendent très bien.

Un homme il a sorti.

Un groupe de personnes il essaie de tuer un inspecteur.

Des personnes ils boit.

Nadasdi conclut de ceci qu'

On voit donc que les contraintes qui régissent le RS [=redoublement du sujet] en français populaire semblent être plus rigides qu'elles ne le sont en FO [=français ontarien]. Le FO n'est pas la seule variété de français canadien où on observe la violation des contraintes énoncées par Roberge (1990). Par exemple, Auger (1991:9) signale que le français montréalais permet aussi le RS après les sujets quantifiés et indéfinis. (1995:81)

Comme le disent Benincà et Poletto (1991:82) «Qui emerge una chiara differenza tra soggetti definiti e quantificatori soggetto. . . . Nella nostra identificazione dei parametri di variazione dovremo quindi distinguere all'interno dei soggetti lessicali le proprietà

connesse con lo statuto di quantificatori.»⁴⁶ Pour ce qui est de notre étude, nous gardons l'hypothèse de Nadasdi d'après laquelle ces contraintes s'appliquent variablement aux sujets pronominaux dans les différentes variétés étudiées. Ainsi nous distinguons les sujets absents (qui correspondent à la valeur par défaut dans le groupe de facteurs 3.2.8.), les sujets lexicaux définis ou indéfinis (Renzi et Vanelli question d), et les sujets pronominaux forts⁴⁷. Cette dernière catégorie comprend les syntagmes nominaux indéfinis, les syntagmes nominaux quantifiés et les quantifieurs du type *nessuno* «personne» et *chi* «qui» (qui reprennent la question f) chez Renzi et Vanelli), et les éléments interrogatifs (soit des syntagmes QU-) en position de sujet.⁴⁸

Également significative pour les sujets pronominaux est la **position** du sujet (si celui-ci est exprimé) par rapport au verbe: préposé ou postposé (Renzi et Vanelli 1983:134-135). Benincà et Poletto donnent les exemples en (33) comme illustration de cette opposition dans des phrases interrogatives (1991:82):

46. «Ici surgit une différence nette entre les sujets définis et les sujets-quantifieurs. . . . Dans notre identification des paramètres de variation nous devons donc distinguer des propriétés reliées au statut de quantifieur, à l'intérieur des sujets lexicaux.»

47. Nous ne considérons ici que les cas où la glose inclut un sujet pronominal «fort» ou tonique, du type *Io dico* ou bien: *Celui-là est bien*. Notons que les pronoms sujet clitiques sont considérés, *ex hypothesi*, comme faisant partie de la morphologie verbale (voir Chapitre 1.): ainsi les formes qui ne comportent qu'un pronom sujet faible sont rangées ici sous la rubrique «aucun sujet».

48. Il s'agit évidemment d'une catégorie assez hétérogène, mais le fait que ces types de sujets soient très peu attestés dans notre corpus justifie leur regroupement ensemble.

- (33) *Cossa votu?* (trevigiano)
'Que veux-tu?'
- Coss 'ti vol?*
'Que tu-veux?' (veneziano)

Dans notre base de données, cette différence positionnelle sera encodée dans le groupe de facteurs 3.2.14. «Position du sujet», où nous codons la position des sujets argumentaux comme préverbale, invertie ou disloquée. Mais les différences éventuelles des *formes* des pronoms sujet en proclise et en enclise (Renzi et Vanelli 193:123,140) ne sont pas prises en compte.

3.4.9. Interactions entre facteurs

En général, Renzi et Vanelli (1983) ne considèrent qu'incidemment la possibilité de conditionnement complexe par deux ou plusieurs facteurs à la fois: ils commentent certains cas où, par exemple, les différentes personnes grammaticales agissent différemment face aux autres contraintes sur la présence des pronoms sujet sans vraiment se poser la question systématiquement pour chaque variable. Pour leur part, Benincà et Poletto présentent dans leur questionnaire (1991:94,97) des paires de phrases conçues pour tester chaque combinaison binaire des «paramètres» qu'elles étudient:

Una volta creata una coppia di frasi per ogni valore di ogni parametro . . . è possibile anche cercare di determinare se due diversi parametri interagiscono tra di loro. . . . La combinazione di più parametri tra di loro è interessante, perché permette di determinare quale dei due parametri presi in considerazione è il più forte.⁴⁹ (1991:87)

Si l'on examine de près leur questionnaire, on trouve que beaucoup des possibilités considérées dans les sections précédentes y sont également vérifiées: la combinaison des différents types de phrases et les différents types de verbes, par exemple. Puisqu'elles ont l'avantage de pouvoir inventer toutes les phrases à traduire qu'elles veulent pour leur questionnaire, on y trouve des contrastes entre des structures bien précises, telles que les relatives appositives, relatives restrictives etc. qui sont soit absentes, soit très rares dans notre corpus. Cependant, leur questionnaire semble passer sous silence la question de la variation entre les différentes personnes grammaticales, un facteur qui est systématiquement contrôlé et vérifié dans notre étude.

Au delà de ces différences reliées à la nature des corpus utilisés, Benincà et Poletto (1991) ne semblent pas envisager la possibilité d'interactions complexes entre plus de deux facteurs ou paramètres. Pour notre part, nous tentons, dans la mesure du possible, de

49. «Une fois créée une paire de phrases pour chaque valeur de chaque paramètre . . . il est aussi possible de chercher à déterminer si deux paramètres différents interagissent entre eux. . . . La combinaison de plusieurs paramètres entre eux est intéressante, parce qu'elle permet de déterminer lequel des deux paramètres en considération est le plus fort.»

contrôler chacune des variables linguistiques indépendamment. Il faudrait également les considérer en combinaison, pour déterminer si, par exemple, l'effet de la personne grammaticale varie selon le temps verbal, ce qui serait plausible, étant donné les syncrétismes qui paraissent à certains temps verbaux.⁵⁰ Nous devons donc garder à l'esprit que les effets qui seraient dus aux interactions des facteurs peuvent échapper à notre grille heuristique. Mais à part cette réserve, nous croyons avoir cerné tous les facteurs récupérables dans notre corpus qui sont susceptibles d'influencer l'emploi des sujets pronominaux. Les facteurs que nous codons sont donc bien justifiés dans la mesure où ils correspondent aux hypothèses que nous formulons à propos de la variation parmi les pronoms sujet, hypothèses qui à leur tour renvoient aux travaux antérieurs sur cette problématique. Dans le chapitre qui suit nous présentons les résultats que nous tirons de l'application de cette grille à notre base de données.

50. Des difficultés d'ordre technique nous empêchent de réaliser une analyse de régression multiple du type VARBRUL: voir 4.4.

CHAPITRE 4: RÉSULTATS

«Envoyez vingt dialectologues dans le même endroit et vous aurez vingt résultats différents.» (Gilliéron)

4.0. PRÉSENTATION

Dans ce chapitre, il est question exclusivement de la présentation des résultats que nous tirons de notre base de données: l'interprétation et l'analyse de ces résultats sont présentées dans les sections indiquées du chapitre qui suit. Ici, nous considérons d'abord certaines contraintes d'ordre pratique qui déterminent et limitent la portée du corpus qui forme la base de nos analyses. Ensuite, nous établissons les critères qui servent à cerner les résultats que nous considérons «catégoriques» (ou presque), qui sont présentés séparément pour chaque personne grammaticale. Ces résultats fournissent une base pour proposer quelques généralisations typologiques relatives aux pronoms sujet pris dans leur ensemble. Finalement, nous considérons les résultats variables, et les facteurs qui peuvent influencer la présence ou l'absence non catégoriques des sujets pronominaux.

4.1. PROBLÈMES PRATIQUES DANS LE DÉPOUILLEMENT DES DONNÉES

4.1.1. Problèmes de nature phonologique

En principe, nous ne nous intéressons dans cette étude qu'à la présence ou à l'absence d'un pronom sujet dans chaque contexte possible, et nullement à la forme phonologique que peut prendre ce pronom. Or il s'avère impossible d'éviter un certain

nombre de phénomènes phonologiques qui rendent difficile, sinon impossible, la tâche d'identifier les formes qui nous intéressent. Nous citons ici quelques cas à titre d'exemples.

Dans la carte *LO MANDERÒ A MILANO* (*AIS* 11), il se trouve que le pronom sujet de la première personne peut souvent prendre la forme de [a] ou [ɔ] dans certaines variétés, et le pronom objet peut prendre, entre autres, la forme [l], [ɔl] ou [al]. Dans ces conditions, il devient donc très difficile de décider si une forme donnée [al] ou [ɔl] correspond à un pronom sujet suivi d'un pronom objet, soit /a + l/, ou bien à un pronom objet /al/. Dans tous les cas présentant une forme comme [al] ou [ɔl], nous vérifions d'abord si des pronoms similaires pertinents existent effectivement dans la variété en question. Dans les cas douteux où il nous manque des données suffisantes pour éliminer une des deux analyses possibles, nous optons pour l'élimination des points ambigus de notre corpus (environ 70 dans le cas cité de *AIS* 11) plutôt que de risquer de fausser les résultats en tranchant arbitrairement d'un côté ou de l'autre. Dans certains cas, la prudence nous oblige à écarter des cartes toutes entières qui pourraient mener à des interprétations arbitraires. Ainsi, malgré l'intérêt inhérent d'étudier les tournures existentielles du type *c'è* et *il y a* (Renzi et Vanelli), l'examen des cartes *IL Y AVAIT* et *IL Y A EU* (*ALF* 95 et 103) démontre la difficulté que présente la distinction du clitique locatif «y» du pronom sujet «il» dans les transcriptions phonétiques. Dans une seule forme [i] ou [j], il s'avère impossible de savoir s'il s'agit de l'un ou de l'autre ou des deux en même temps (puisque /i/ + /i/ donnerait facilement [i] dans la parole normale). Nous préférons

donc supprimer ces cartes de notre corpus plutôt que de risquer de faire des découpages morphologiques peu sûrs. Comme les auteurs de l'*ALF*, qui ont voulu suivre le plus possible le modèle du français dans la séparation des mots (1902:13-14), force nous est de reconnaître que cette tâche nous dépasse parfois:

Où *gar chompêtre* cesse-t-il d'exister et fait-il place à *garchon pêtre*? Il est impossible de le préciser pour certains patois, surtout si on considère que *gar* peut équivaloir à *garde* et à *gars*. (Gilliéron et Edmont 1902:13)

Sans que ces problèmes précis de l'étymologie populaire entrent en ligne de compte, il faut reconnaître que l'homonymie potentielle a fait ses ravages dans notre corpus, surtout du côté des variétés italiennes. En effet, dans une carte comme *ALS* 1638 *VOLETE CHE CI VADA IO*, comment faire le tri entre les [i] (<IBI), équivalents de «ci», et les [i], pronoms sujet du type «io», rendu souvent par /i/? La note en légende de Jaberg et Jud se contente de faire remarquer le problème: «Das Ortsadverb 'ci' ist von den Suj. nicht immer wiedergegeben worden»⁵¹ sans pourtant offrir une solution. Et encore, dans *ALS* 1641 *CHE NON LA TROVASSIMO*, une forme comme [nu] devant le verbe peut équivaloir tantôt à une particule de négation tantôt à un pronom sujet, et le choix entre les deux

51. «L'adverbe de lieu 'ci' n'a pas toujours été rapporté par le sujet [d'enquête].»

dépasse parfois nos possibilités d'analyse. Résistant à la tentation d'appliquer des critères arbitraires, nous sommes obligé d'exclure beaucoup de ces formes du corpus.

Mais il reste que parfois l'homonymie, même d'apparence assez grave à première vue, peut se résoudre sans que nous soyons obligé d'éliminer des points ou d'écarter des formes de notre analyse. Il s'agit de formes comme [ke] qui apparaissent là où nous nous attendons à trouver aussi des sujets pronominaux, c'est-à-dire devant des verbes fléchis. Cette forme peut être soit un pronom sujet générique, l'équivalent du «ça» français, soit le complémenteur gascon /ke/, qui introduit (quasi obligatoirement) les propositions finies. Pour compliquer l'affaire, les deux éléments en question peuvent apparaître dans une position morphosyntaxique identique: devant des verbes à la troisième personne. La seule façon de trancher la question est de recourir à des verbes qui n'admettent pas un sujet générique, c'est-à-dire, qui sélectionnent un sujet spécifique. Les points de notre étude qui sont limitrophes de l'aire gasconne présentent donc des complémenteurs devant des verbes qui n'admettent pas un sujet générique, et nous pouvons donc trancher entre les deux [ke], bien que les zones en question soient contiguës (voir 4.2.3.1. et la Carte 8).

4.1.2. «Réponses extorquées»

Très conscients des graves inconvénients d'un questionnaire par traduction directe, Gilliéron et Edmont attachent une importance capitale à ce que nous pouvons appeler la méthode impressionniste: «Les réponses que nous reproduisons dans nos cartes représentent toujours l'inspiration, l'expression première de l'interrogé, une traduction de

premier jet.» (1902:7). Le choix même d'Edmont, épicier sans formation de philologue, comme enquêteur répond en fait à la grande préoccupation de Gilliéron d'obtenir des «instantanés» de chaque parler qui seraient «francs de toute retouche» (voir également Pop 1950), et obtenus en laissant la «plus grande liberté» aux sujets. Or, les traductions faites sur-le-champ sont nécessairement très variables et parfois inexactes (surtout si ces traductions doivent être répétées par plus de six cents non spécialistes dans tout l'Hexagone). Il s'avère donc difficile sinon impossible de suivre la méthode impressionniste de transcription et en même temps relever exactement les mêmes tournures de tous les sujets. Les traductions qui ne correspondent pas au questionnaire, que Gilliéron et Edmont appellent les «réponses extorquées», deviennent particulièrement fréquentes lorsqu'on passe des mots isolés aux phrases:

Sans doute, il peut paraître regrettable qu'à: *Vous trouverez bien quelqu'un qui vous ira*, on nous réponde: . . . *quelqu'un qui vous ira bien*, . . . *qui vous conviendra, à votre goût*, . . . *qui fera bien votre affaire*, . . . *qui vous plaira*, . . . *qui vous agréera*, et même . . . *qui vous voudra*. (Dans ce dernier cas, nous mettons la forme entre parenthèses, comme ne répondant point à *qui vous ira*).

Mais, que faire à cela? (1902:7,n1)

Que faire, en effet. Le problème se pose inéluctablement pour notre étude, puisque ce sont justement les propositions finies (celles qui nous intéressent) qui sont les plus susceptibles

de fournir des traductions «extorquées». Dans l'exemple donné, on voit le verbe en question changer d'un inaccusatif (*ira*) à un transitif (*fera*) à un modal (*voudra*), sans parler des tournures non verbales (*à votre goût*). Dans le cas des verbes, nous pouvons toujours faire varier les codes correspondant aux différents types de verbes: c'est un des avantages de la base de données. Mais dans le cas des tournures non verbales, ou encore là où un infinitif se substitue à un verbe fléchi, la seule solution est d'éliminer la forme en question du corpus, en substituant le code «/» (qui se lit «ne s'applique pas») pour les variables 1 ou 0. Nous croyons gagner en exactitude plus que nous ne perdons en nombre d'occurrences par ce pis-aller nécessaire.

Si la méthode impressionniste est une préoccupation capitale pour les collaborateurs de l'*ALF*, elle devient une véritable obsession chez l'équipe de l'*AIS*, qui la considère comme une garantie contre les interférences éventuelles dues à la formation ou aux théories préconçues des enquêteurs:

Le danger de l'autosuggestion ne peut être évité que si l'on pratique la *méthode impressionniste*, employée par les enquêteurs de l'*AIS*. Celle-ci consiste à reproduire le plus fidèlement possible toutes les nuances phonétiques saisies sur le coup. . . . On enregistre, en ce cas, toutes les irrégularités et même les inconséquences observées dans la prononciation des informateurs.

(Pop 1950:570).

La dialectologie traditionnelle se méfie surtout des interférences en ce qui concerne la transcription phonétique, mais cette méthode mène aussi à transcrire bien des «irrégularités» et «inconséquences» dans la morphosyntaxe, domaine qui ne figure guère parmi les préoccupations majeures de cette école. Du coup, on voit se multiplier le nombre de traductions «extorquées».

On pourra apprécier un peu l'amplitude des difficultés que ce problème pose pour nous en comparant l'exemple de l'*ALF* cité ci-dessus avec la carte 1590 de l'*AIS*: ...CHE ME DOVETE («...que vous me devez»). À cette proposition apparemment simple correspondent les traductions les plus variées: «que vous avez à me payer», «que vous devez me donner», «que tu as à me donner», «que j'ai à avoir», «que j'avance», «qui m'est dû», «ce qui me vient», «qui m'est à moi» etc. Et ceci sans compter les difficultés que présente la segmentation des transcriptions phonétiques: ici comme ailleurs, il faut déterminer si l'équivalent du complément «che» est fusionné ou non avec un pronom: «Nicht immer war es möglich, das unbetonte Subjektspronomen 'a' (= 'ihr' oder = 'es') von 'che' zu lösen.» (Jaberg et Jud, *AIS* 1590).⁵² Si les différents types de verbes peuvent être triés par la base de données (suite à un minutieux travail de codage), il en reste que le partage de ces occurrences entre les quatre personnes grammaticales figurant dans les traductions diminue le nombre de données qui seront comparables entre elles (et ceci pour

52. «Il n'a pas toujours été possible de séparer le pronom sujet atone 'a' (= 'vous' ou 'il' impersonnel) de 'che'» Il en est de même pour les autres formes du pronom sujet, souvent réduites à une seule voyelle.»

une carte précieuse qui représente une des rares propositions relatives de nos atlas). Inutile de dire que ce problème se répète et se multiplie partout dans les données de l'*AIS*. Il faut avouer que parfois nous avons de la chance, et les traductions différentes ne correspondent pas à des différences dans notre grille d'analyse. Ainsi, le fait que pour *AIS* 394 . . . È CASCATO . . . (« . . . est tombé . . . ») on obtient des formes comme «est venu bas» et «est allé vers le bas», ne pose aucun problème de codage, puisqu'il s'agit de trois verbes inaccusatifs, et les différents compléments qui peuvent éventuellement suivre un verbe ne nous concernent pas dans ce cas.

4.1.3. Rareté relative des quatrième et cinquième personnes

Pour des raisons d'ordre pragmatique, les personnes grammaticales les moins fréquentes dans ce corpus sont la quatrième (*nous, noi*) et la cinquième (*vous, voi*). Il suffit de consulter les Appendice I et II pour se rendre compte de la rareté relative et absolue de ces formes. À ce problème s'ajoute le fait que ce sont précisément ces deux personnes qui se voient souvent remplacées par d'autres. Ainsi le vouvoiement se voit traduit souvent par le tutoiement là où le pluriel n'est pas clairement requis (par exemple dans l'*AIS* 1113 *COSA NE FARESTE* («qu'en feriez-vous?»), on trouve 42 formes à la deuxième personne. La quatrième personne est menacée par deux substitutions: d'une part, elle se traduit par l'équivalent du «on» habituel du français familier (par exemple dans *ALF* 91 *NOUS AVONS PRIS UNE PURGE* où on trouve 11 formes à la troisième personne); d'autre part certaines variétés toscanes, dont notamment le florentin, remplacent la

quatrième personne par une forme réfléchie à la troisième personne (*AIS* 1613

LAVOREREMMO DI PIÙ «nous travaillerions davantage» donne 22 réponses du type *noi si lavorebbe*, en gros l'équivalent du français populaire «nous on travaillerait»). Dans tous ces cas, le nombre d'occurrences de ces personnes grammaticales, déjà rares dans le corpus, se voit réduit au profit d'autres personnes qui sont amplement représentées ailleurs dans le corpus.

4.1.4. Échantillon final

Tous les problèmes pratiques que nous évoquons dans les sections précédentes, ainsi que bien d'autres facteurs ponctuels surgis en cours de route, déterminent la portée réelle de la base empirique de cette étude. Les 172 points examinés dans les 101 cartes de l'*ALF* donnent potentiellement 17 711 formes, dont 531 sont écartées avec le code “/” parce qu'elles sont inexistantes ou inapplicables pour diverses raisons.⁵³ Prévisiblement, la proportion de formes exclues est plus forte pour l'*AIS*: les 266 points dans les 101 cartes fournissent 26 465 cellules dans la base de données, dont 3,739 reçoivent le code “/”.⁵⁴ Le

53. Ce chiffre n'équivaut pas à la simple multiplication du nombre de points par le nombre de cartes, étant donné qu'un certain nombre de celles-ci ne correspondent qu'au coin sud-est de l'Hexagone (Gilliéron et Edmont, 1902).

54. Encore une fois, ce chiffre cache le fait qu'à un certain nombre de sujets (par exemple dans les grandes villes) on fait relever le questionnaire «réduit» de seulement 800 questions (Pop 1950:566), alors qu'à d'autres points les sujets fournissent des réponses multiples (voir 5.3.1.).

total s'élève donc à 44 176, moins 4 474 écartés (soit environ dix pourcent), ce qui nous laisse une masse de 39 703 occurrences utiles pour servir de base à nos analyses.

4.1.5. Synchrétisme et oppositions morphologiques

De façon générale, nous ne considérons dans chaque occurrence que la présence du pronom sujet par opposition à son absence. C'est-à-dire que nous n'avons aucune information quant à la forme phonologique de chaque pronom, et par conséquent, nous ne pouvons avancer aucune hypothèse quant aux synchrétismes éventuels entre les pronoms des différentes personnes grammaticales. Des raisons d'ordre technique nous empêchent de le prendre en considération en dépit du rôle important que joue le synchrétisme dans des analyses antérieures.

Poletto divise les systèmes des pronoms des première, deuxième, quatrième et cinquième personnes par le critère du synchrétisme (1993:130). Elle distingue donc les systèmes où le même morphème sert de pronom pour les quatre personnes en question de ceux qui ont des pronoms distincts aux quatre personnes. Cette subdivision binaire ne recoupe pas la typologie plus complexe proposée par Renzi et Vanelli, d'après laquelle le synchrétisme peut frapper diverses combinaisons de personnes: première, quatrième et cinquième; première, quatrième et sixième; première et quatrième; première, troisième et sixième; troisième et sixième (1983:131). Il est à noter que d'après Renzi et Vanelli, le seul cas de synchrétisme dont parle Poletto, celui des première, deuxième, quatrième et cinquième personnes, ne se produirait jamais, pour la simple raison que la deuxième

personne serait toujours distincte des autres (1983:130). L'analyse de ces faits complexes aurait demandé une base de données permettant d'emmagasiner et de comparer non seulement la présence ou l'absence d'un pronom (comme le fait la nôtre), mais aussi sa forme phonologique. Étant donné les phonétismes hautement variables et minutieusement transcrits des atlas en question, une telle entreprise ne semble pas faisable dans le cadre de cette étude. Nous préférons plutôt laisser cette question importante pour des recherches ultérieures.

4.1.6. Résultats «catégoriques»

Les cas les plus clairs parmi nos résultats sont ceux où l'emploi des sujets pronominaux est catégorique, ou presque catégorique. À la différence des «généralisations établies» de Renzi et Vanelli (1983), et les différents systèmes pronominaux discernés par Poletto (1993), nos conclusions à cet égard ne nous permettent pas de parler d'emploi «constant» ou «toujours obligatoire». En fait, la nature même de notre corpus nous empêche de retrouver une catégoricité absolue dans l'emploi pronominal, là où les dialectophones spécialistes consultés par Renzi et Vanelli (1983) et par Poletto (1993) peuvent fournir des jugements plus tranchants. Bien que les transcriptions phonétiques des atlas soient sûrement très fidèles à ce que Edmont et Scheuermeier entendent, nous ne pouvons jamais savoir si les enquêteurs entendent effectivement tous les sujets pronominaux prononcés, surtout quand ceux-ci sont atones et exposés donc à tous les phénomènes de réduction phonétique que cette position amène. Qui plus est, un cas isolé

de non prononciation du pronom sujet peut toujours représenter un fait de parole ou une erreur momentanée plutôt que de refléter le système grammatical des sujets: même une variété à sujet obligatoire comme le français standard admet l'omission occasionnelle du pronom sujet, (Renzi et Vanelli 1983:129n1; Sandfeld 1970:18-19). Nous retenons donc comme «quasi-catégoriques» les cas où le taux d'emploi des sujets pronominaux dépasse le seuil de 80%. De même, les points dont le taux d'emploi ne dépasse pas 20% sont considérés comme faisant un usage sporadique des sujets pronominaux: un tel emploi sporadique du pronom sujet ne saurait changer le statut de ces variétés qui sont essentiellement à sujet nul.⁵⁵

Afin de faciliter la comparaison entre les différents contextes grammaticaux, nous nous servons d'un contexte «type» qui sert à niveler les effets des diverses variables. Pour chaque personne grammaticale nous donnons donc le taux brut d'occurrences des sujets pronominaux dans tous les contextes, mais aussi dans un «contexte type» bien précis: les phrases affirmatives au présent de l'indicatif, sans objets proclitiques ni particule de négation en position préverbale. Nous varions les contextes là où cela semble justifié, et d'autres facteurs sont indiqués quand ils semblent avoir un effet sur le taux d'emploi des sujets pronominaux.

55. Ces deux seuils -- 20% et 80% -- ne représentent pas des choix complètement arbitraires, puisqu'ils correspondent statistiquement à des divisions naturelles («natural breaks») dans les données (voir aussi 4.3.2.) pour la plupart des facteurs.

4.2. RÉSULTATS PAR PERSONNES

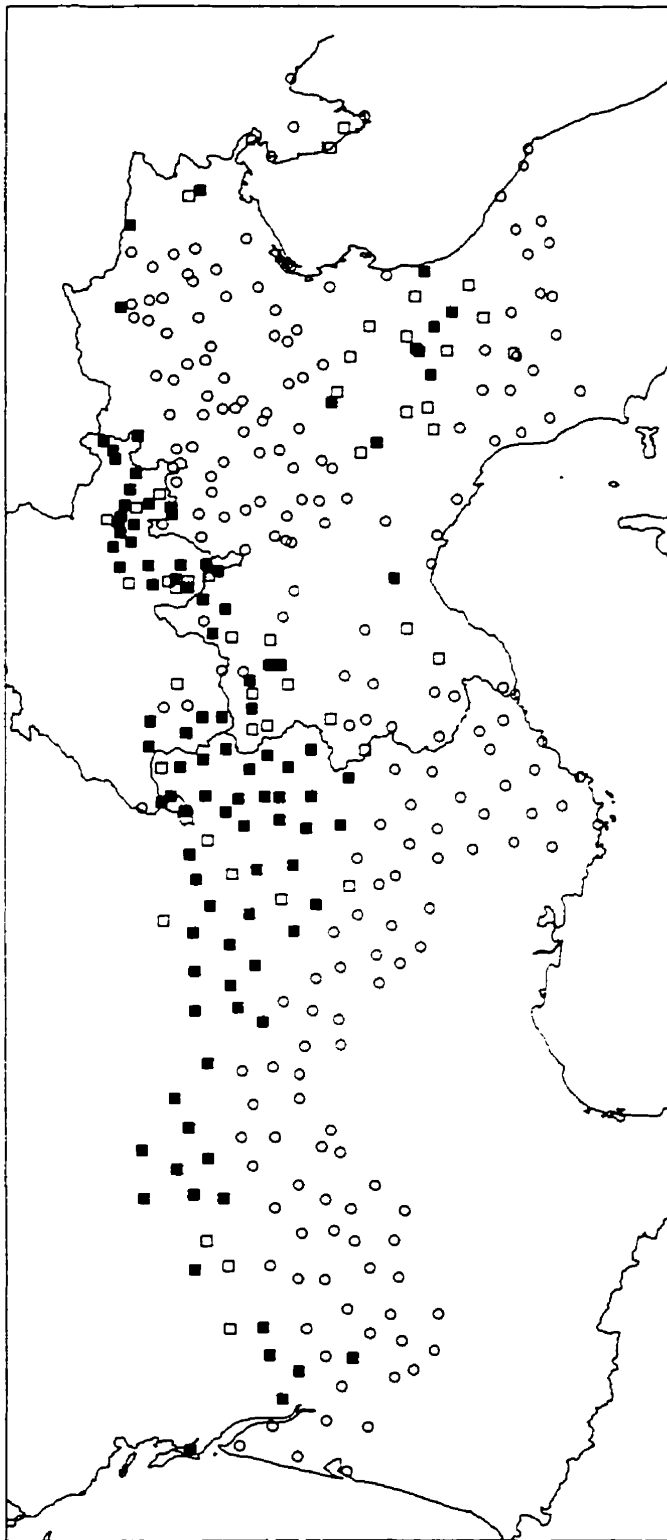
Dans les sections qui suivent nous présentons l'étendue de l'emploi des sujets pronominaux pour chacune des six personnes grammaticales. Nous reconnaissons évidemment qu'il existe de bonnes raisons théoriques pour regrouper certaines personnes sous une même rubrique: Benveniste (1967:225-236,251-257) souligne par exemple la différence fondamentale entre les «vraies personnes», celles qui participent directement à un acte de communication, et les «non personnes» qui en sont exclues (soit la troisième et la sixième). Poletto reprend une distinction similaire, en traitant les troisième et sixième personnes comme un groupe, et les autres personnes comme un groupe à part.⁵⁶ Nous préférons cependant l'approche neutre qui considère chaque personne séparément d'abord pour ensuite voir quels regroupements typologiques peuvent se faire entre les différentes personnes.

56. Pour Poletto, les personnes 1, 2, 4, et 5 seraient celles qui sont susceptibles de présenter un emploi variable du pronom sujet, et des paradigmes pronominaux où les quatre personnes ne sont pas distinguées. Comme nous remarquons plus haut, cette étude ne tient pas compte des formes des pronoms; pour ce qui est de l'emploi variable du pronom sujet, nous le trouvons à toutes les personnes.

4.2.1. Première personne

Si nous considérons tous les contextes présentant des verbes à la première personne, soit une vingtaine de cartes dans les deux atlas,⁵⁷ nous trouvons 39 points qui ont un emploi catégorique, 87 points à emploi quasi-catégorique du pronom sujet, et 186 points à usage sporadique. Si en revanche nous nous limitons au contexte type, ces chiffres s'élèvent à 101 points catégoriques, 44 points quasi-catégoriques et 203 points à emploi sporadique du pronom sujet. Ces résultats sont présentés dans la Carte 2. Les points intermédiaires ont un emploi variable (entre 20% et 80%) des pronoms sujet à la première personne dans le contexte type.

57. Pour chaque personne et chaque facteur conditionnant, il se peut qu'une carte donnée n'ait pas de formes applicables à tel ou tel point: c'est bien pourquoi nous donnons une approximation du nombre de cartes impliquées dans chaque résultat, plutôt qu'un chiffre précis qui ne serait exact que pour une partie des points.



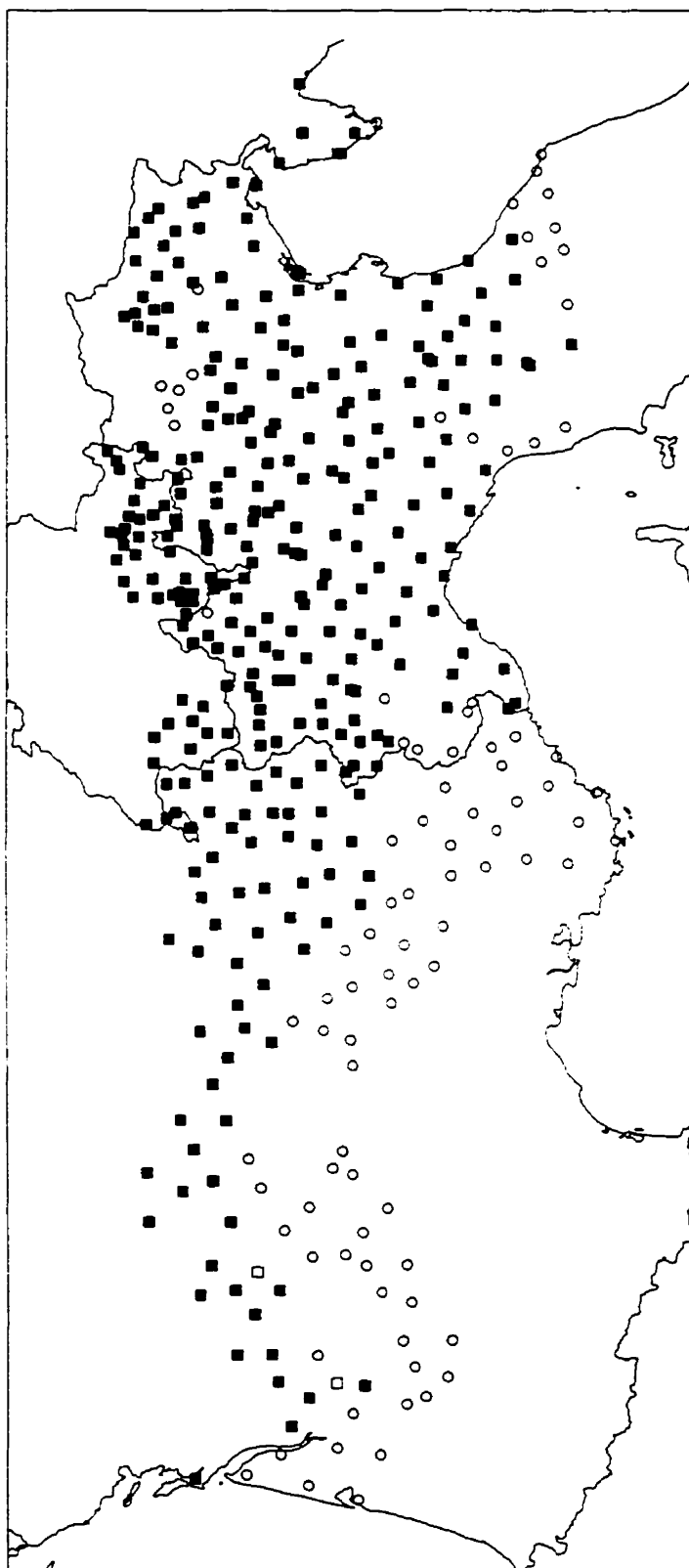
Carte 2. Pronoms sujet à la première personne.
Dans le contexte type (voir 4.2.1.) Taux d'emploi et n° de points.

■ 100%	pronoms sujet.	(101)
□ 80 à 100%	pronoms sujet.	(44)
○ 0 à 20%	pronoms sujet.	(203)

4.2.2. Deuxième personne

Dans la dizaine de cartes présentant des verbes à la deuxième personne, il y a 171 points où les pronoms sujet sont catégoriques, 81 où ils sont quasi-catégoriques, et 92 où leur taux d'usage est nul ou sporadique. Mais si nous ne considérons que le contexte type (voir la Carte 3), le total monte à 305 points à sujets pronominaux catégoriques et seulement 2 points quasi-catégoriques. Sur cette carte, 92 points ont un emploi de pronoms de la deuxième personne de 20% ou moins dans ce contexte.

L'écart frappant entre l'ensemble des occurrences et le contexte type peut s'attribuer à l'influence des phrases interrogatives. L'emploi plus répandu du pronom sujet dans les phrases affirmatives du contexte type serait dû au fait que dans les interrogatives le pronom sujet est souvent obligatoirement inversé, et donc morphologisé comme élément de la flexion verbale. Il serait intéressant d'examiner l'étendue de ce phénomène des sujets enclitiques (Jaberg 1932) et leur interaction avec les pronoms, puisque les deux marques de sujets peuvent se présenter en cooccurrence: dans *A/S359* ...VIENT?, on voit par exemple au point 263 la forme [ta véñat], avec le pronom sujet redoublé (voir également 4.2.5.). Mais ceci impliquerait l'examen détaillé des transcriptions à chaque point en question (tant les formes du pronom comme celles de la flexion verbale), ce qui dépasse le cadre de la présente étude. De toute façon, on voit bien ici qu'il s'agit de la personne la plus répandue dans les systèmes de pronoms sujet, ce qui semble confirmer la généralisation de Renzi et Vanelli: s'il existe un seul pronom sujet dans une variété, ce sera



Carte 3: Pronoms sujet à la deuxième personne.
 Dans le contexte type (4.2.2.). Taux d'usage et n° de points.

■ 100%	pronoms sujet.	(305)
□ 80 %	pronoms sujet.	(2)
○ 0	pronoms sujet.	(92)

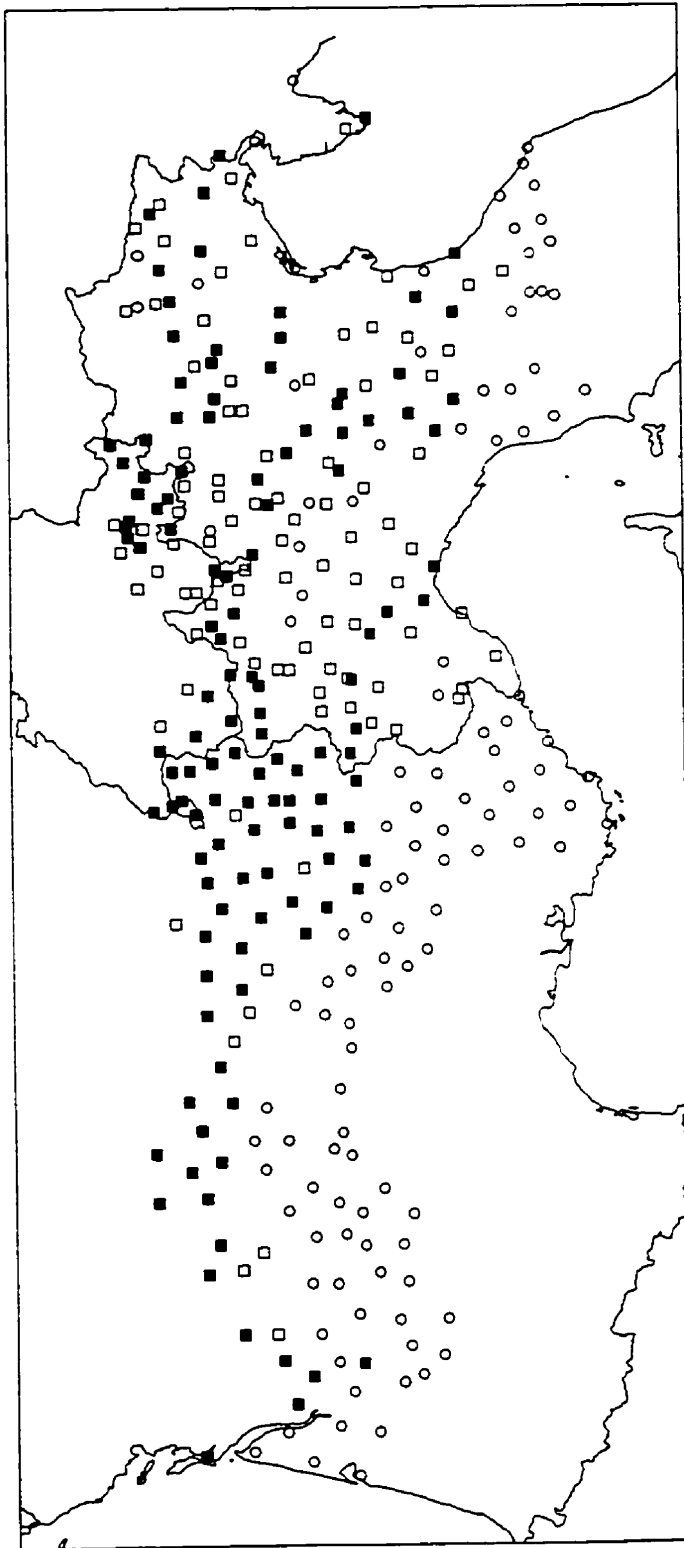
celui de la deuxième personne. (Sur le statut privilégié de la deuxième personne, voir 4.3.2. et 5.2.1.).

4.2.3. Troisième personne

La troisième personne est celle qui présente le plus de contextes grammaticaux différents, et le plus grand nombre aussi d'occurrences dans notre corpus. Nous traitons d'abord de la troisième personne «personnelle», et ensuite des cas particuliers des verbes météorologiques et impersonnels.

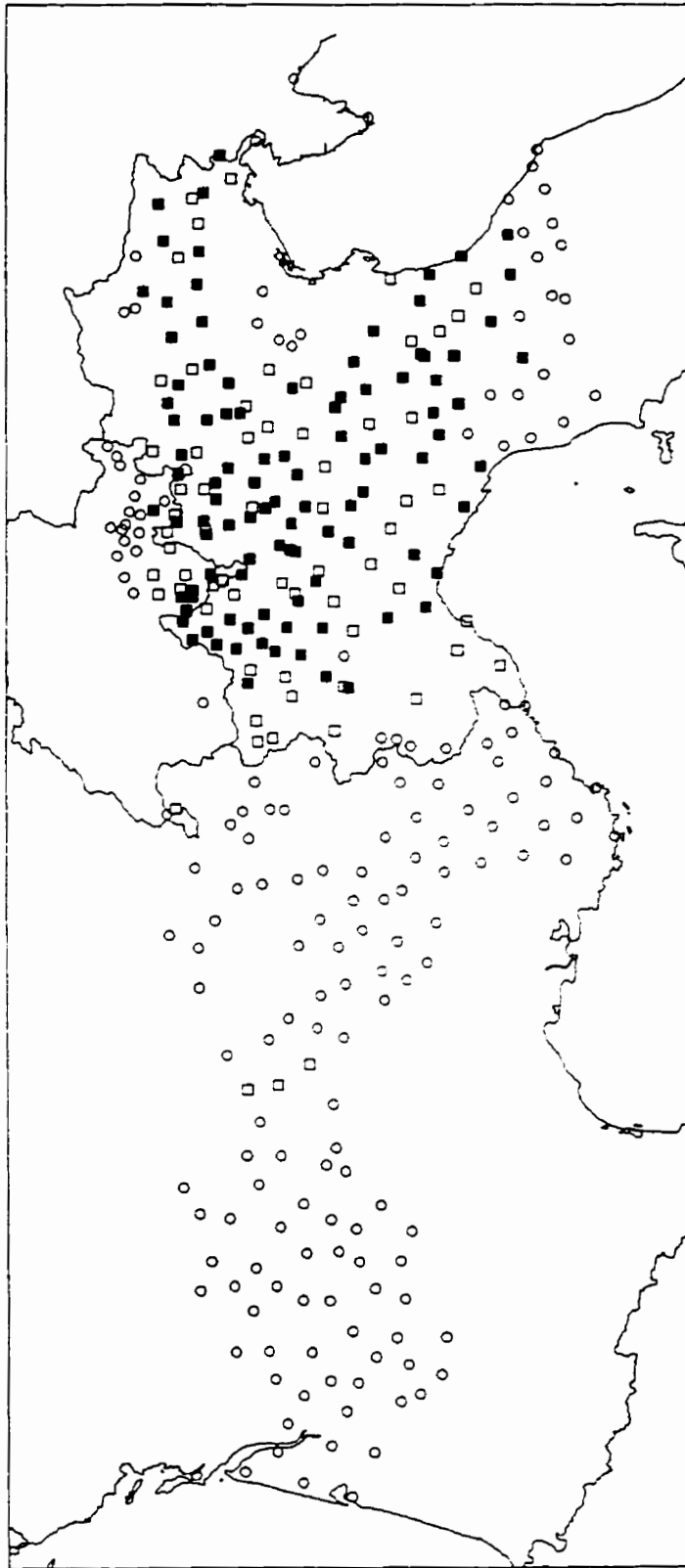
Si on prend ensemble tous les contextes présentant un verbe personnel à la troisième personne (soit une vingtaine de cartes), on trouve qu'il n'y a que 6 points où le pronom sujet est employé catégoriquement, 131 points où l'emploi est quasi-catégorique et 102 où il ne dépasse pas 20%. Pour rendre les résultats plus clairs, nous appliquons le contexte type en séparant les cas sans sujet explicite des cas à sujet argumental (syntagme nominal ou pronom fort).

Dans la Carte 4 on voit les résultats pour le contexte type sans sujet explicite: il y a 125 points à emploi catégorique du sujet pronominal, 91 quasi-catégoriques, et 115 à emploi sporadique. Si en revanche nous considérons les cas de sujets argumentaux en



Carte 4 : Pronoms sujet à la troisième personne.
 Dans le contexte type (4.2.3.). Taux d'usage et n° de points.

■ 100 %	pronoms sujet. (125)
□ 80 à 99%	pronoms sujet. (91)
○ 0 à 20%	pronoms sujet. (115)



Carte 5: Pronoms sujet à la troisième personne.
Sujets argumentaux (4.2.3). Taux d'usage et n° de points.

■ 100 %	pronom sujet.	(100)
□ 80 à 100%	pronom sujet.	(64)
○ 0 à 20%	pronom sujet.	(172)

contexte type (la Carte 5), on trouve 100 points à emploi catégorique des sujets pronominaux, 64 à emploi quasi-catégorique et 172 à emploi sporadique. On voit ici que l'emploi du pronom sujet cède du terrain en combinaison avec les sujets argumentaux. C'est-à-dire que les variétés qui emploient un pronom sujet avec un sujet lexical ou un pronom fort doivent aussi l'employer quand il n'y a pas de sujet explicite, mais l'inverse n'est pas vrai: il y a des variétés qui n'emploient un pronom sujet que dans l'absence d'un sujet argumental. C'est la distribution de ce phénomène qui est frappante: l'emploi du pronom avec les sujets argumentaux est assez répandu en Italie et presque absent en France (voir 4.2.6., 4.4.8. et 5.1.).

Notre corpus ne contient pas suffisamment de cas de sujets argumentaux qui sont des pronoms forts, ni de ceux qui se trouvent en position postverbale, pour tirer des conclusions à propos de ces contextes grammaticaux en isolation, comme le fait Poletto (1993).

4.2.3.1. Verbes météorologiques

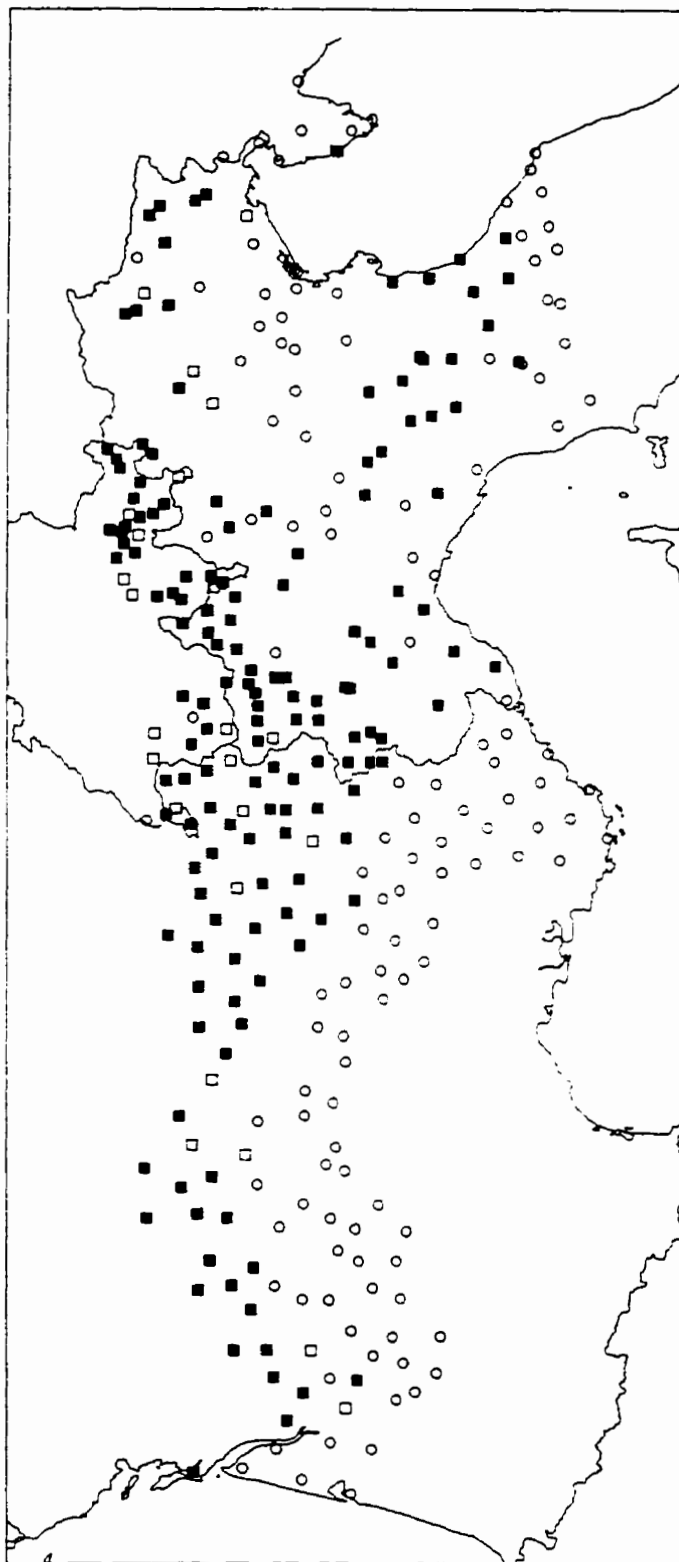
Les verbes météorologiques (soit une dizaine de cartes dans chaque atlas) représentent un cas particulier des verbes à la troisième personne, celui des sujets explétifs «atmosphériques». À l'intérieur de ce groupe il convient de faire quelques précisions: étant donné le caractère quasi-spontané des réponses aux questionnaires (voir 4.1.2). on trouve bien des réponses avec des verbes qui ne sont pas à strictement parler «météorologiques»: ainsi pour *AIS 377 NEVICA* (il neige) on obtient des équivalents locaux du type «(il) vient

de la neige» (point 424), «tombe neige» (point 140), où le sujet des verbes inaccusatifs se confond avec l'objet superficiel.⁵⁸ Ensuite il y a des cas comme «(il) fait la neige» (point 499), ou encore «il donne de la grêle» (*AIS* 372 GRADINA, point 1), où les verbes, tout en ayant vraisemblablement le même sujet explétif que les «vrais» météorologiques, sont clairement transitifs et donc à séparer des verbes météorologiques classiques.

Si nous excluons ces cas de verbes «pseudométéorologiques», nous trouvons 16 points qui ne manquent jamais de pronom sujet, 82 qui en font un emploi quasi-catégorique, et 156 dont l'emploi est de 20% ou moins. Ces chiffres changent considérablement si l'on ne considère que les verbes météorologiques au contexte type: on trouve 154 points à pronom catégorique, 24 quasi-catégoriques, et 120 à emploi sporadique. La Carte 6 présente la distribution de ces points. Les points intermédiaires ont un emploi variable des pronoms avec les verbes météorologiques.

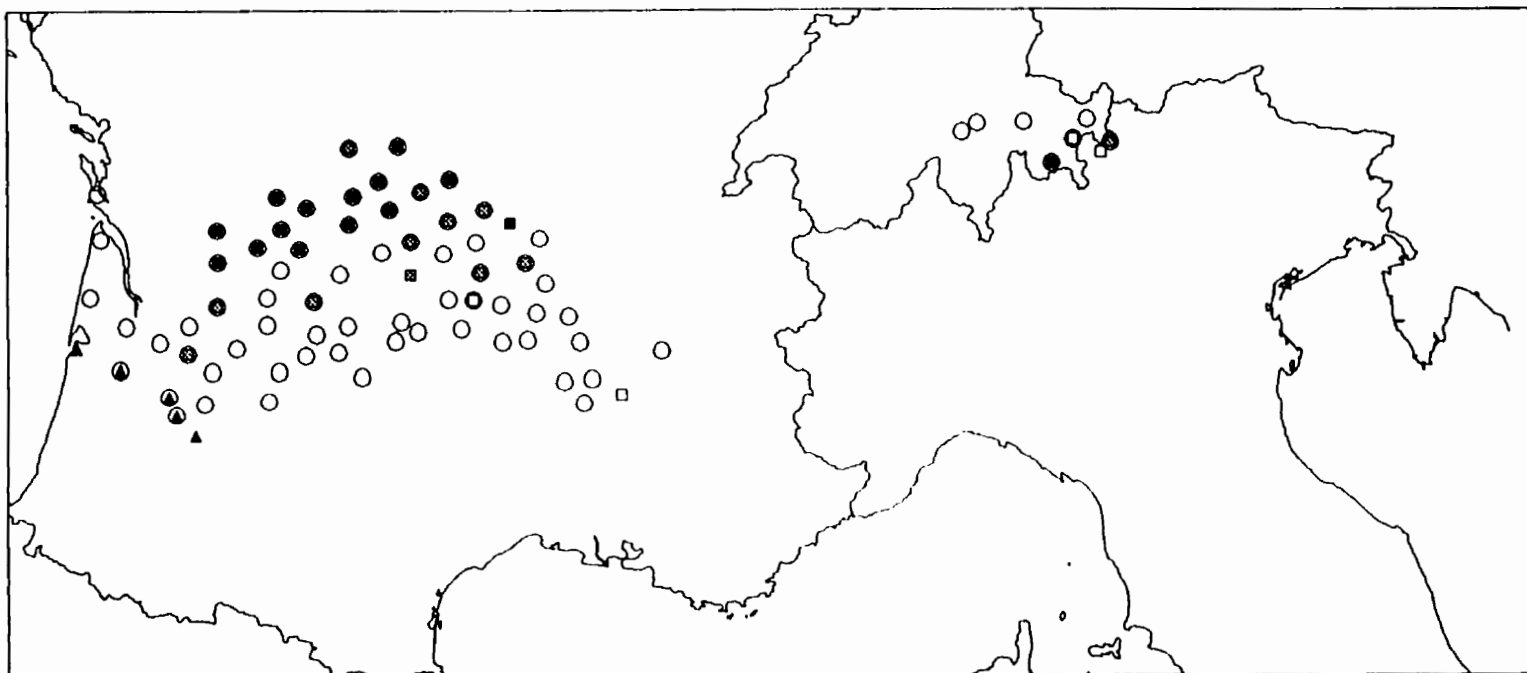
On notera que les 31 points représentés sur la Carte 7 ont comme sujet des verbes météorologiques des pronoms du type [ka] ou [ke] ou [ku], formes occitanes que nous pouvons rapprocher au sujet générique «ça» du français (sans la palatalisation des vélaires qui caractérise les parlers d'oïl): «ça pleut» plutôt que «il pleut». À part le problème pratique de distinguer ces formes du complémenteur [ke] qui introduit toute proposition principale finie en gascon, (voir Campos, 1986:II, et les références qui s'y trouvent), ces

58. Nous citons ces formes par leur glose française plutôt qu'une transcription particulière, puisqu'elles représentent souvent des réponses types correspondant à plusieurs points différents.



Carte 6: Pronoms sujet à la troisième personne.
 Verbes météorologiques (4.2.3.1.). Taux d'usage et n° de points.

■	100 %	pronoms sujet.	(154)
□	80 à 100%	pronoms sujet.	(24)
○	0 à 20%	pronoms sujet.	(130)



Carte 7: Pronoms génériques du type [ka] (4.2.3.1.).

- 80 à 100% [ka] avec verbes météorologiques (27 points).
- 0 à 20% [ka] avec verbes météorologiques (4 points).
- [ka] avec autres verbes.
- ▲ [ke] énonciatif gascon (5 points.).

formes ne présentent aucune difficulté particulière. Comme le «ça» français (voir Auger 1995), ce sont des sujets clitiques avec certaines propriétés sémantiques spécifiques mais qui se comportent morphosyntaxiquement comme les autres sujets clitiques. Nous comptons donc ces formes (qui se présentent à 33 points, surtout dans l'*ALF*) comme l'équivalent des différentes formes du pronom masculin de la troisième personne «il» qui apparaît ailleurs comme sujet des verbes météorologiques. Il est à noter que le sujet générique /ka/ apparaît variablement: des 33 points où il est attesté, il n'y a que 10 où il est catégorique et 5 où il est quasi-catégorique. Aux autres points, il alterne tantôt avec le sujet nul, tantôt avec un autre pronom sujet du type «il».

La Carte 7 montre aussi les cinq points où le [ke] énonciatif du gascon est attesté dans notre corpus,⁵⁹ ainsi que les 74 points où le pronom sujet générique apparaît avec des verbes non météorologiques dans des contextes non définis, notamment comme sujet de «être», par exemple *AIS407* (L'ORO) È PESANTE et *ALF511* SI C'ÉTAIT BIEN CUIT, mais aussi comme sujet de «falloir» (*ALF534-537*) ou «faire» (*ALF729*).⁶⁰ Nous ne le distinguons pas

59. L'emploi du complémenteur n'est catégorique à aucun de ces cinq points. Il semble varier en fonction de facteurs similaires à ceux qui conditionnent l'emploi du pronom sujet, pour devenir plus catégorique plus on s'approche du «noyau» gascon dans les Pyrénées. La variation géolinguistique de ces formes mérite une étude approfondie qui dépasse les limites de la présente.

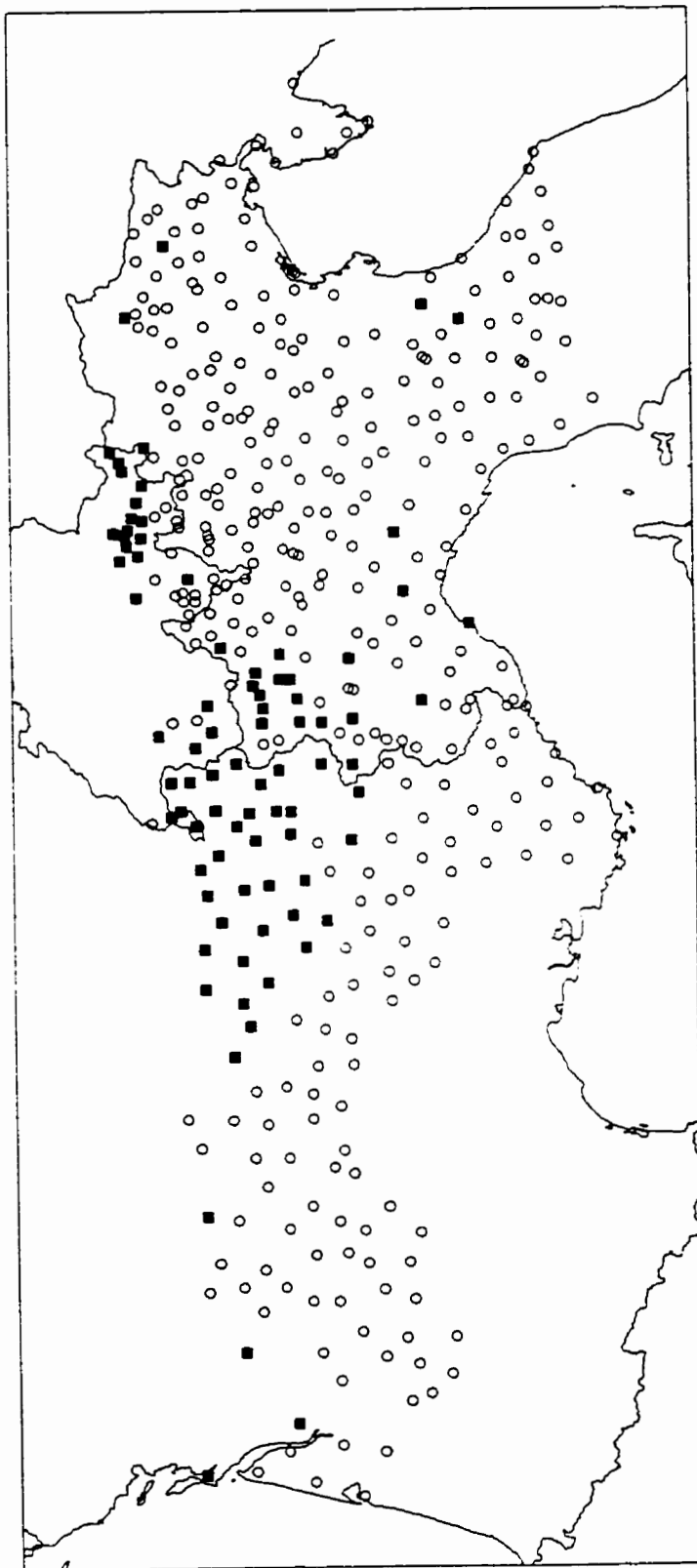
60. Ces deux cas se chevauchent: trois des points où le [ke] gascon est documenté présentent également des sujets génériques pour la carte *ALF511* SI C'ÉTAIT BIEN CUIT. La forme du pronom ici est du type [ako], bien distincte du complémenteur [ke].

des autres pronoms sujet de la troisième personne dans les cas de verbes non météorologiques. Sur la distribution géographique du pronom générique, voir 5.2.2.

4.2.3.2. Verbes impersonnels

L'autre cas particulier de la troisième personne est le cas des sujets explétifs des verbes impersonnels «falloir» et «bisognare».⁶¹ La Carte 8 montre les 85 points qui ont des sujets pronominaux catégoriques avec les verbes impersonnels. Il n'y a pas de point quasi-catégorique ici, et le nombre de points à emploi nul s'élève à 322 (comme on peut s'y attendre, il n'y a pas d'emploi sporadique dans ce cas). Il faut cependant traiter ces chiffres avec une certaine prudence: ce petit échantillon ne comprend que 2 cartes de l'*AIS* (deux fois «bisogna» au présent), et 4 cartes de l'*ALF* («(il) faut» au présent). Le total assez restreint tend évidemment à polariser les résultats; étant donné les contextes presque identiques, ce qui est frappant est le fait que nous trouvions quand même des points intermédiaires variables.

61. Le verbe «bisognare» pris dans son sens impersonnel, évidemment: on exclut les occurrences du type *bisogniamo fare* «nous avons besoin de faire», ainsi que les réponses extorquées du type «on doit faire» pour «il faut faire», ce qui réduit davantage l'échantillon déjà trop petit de ce type de verbe.

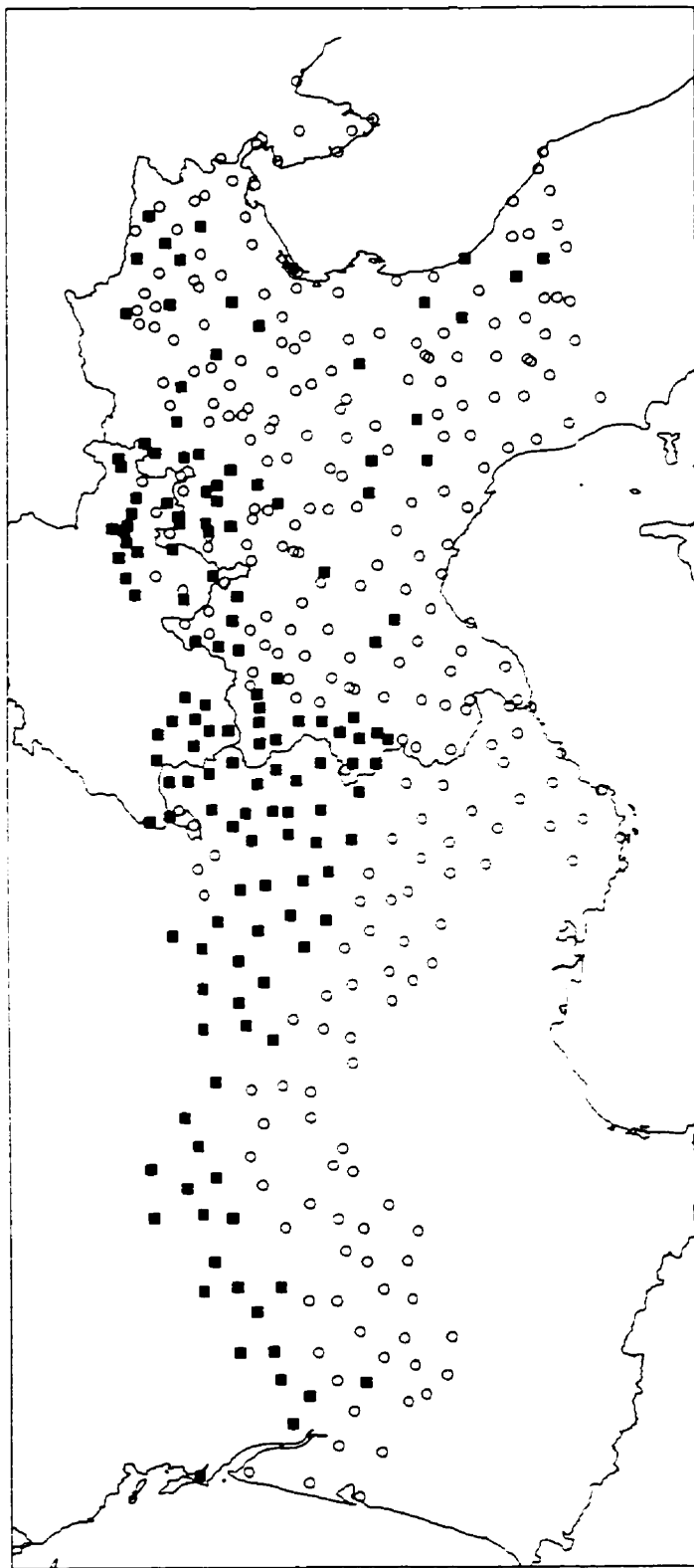


Carte 8: Pronoms sujet à la troisième personne.
Verbes impersonnels. (4.2.3.2.). Taux d'usage et n° de points.

■ 100%	pronoms sujet.	(85)
□ 80 à 100 %	pronoms sujet.	(0)
○ 0 à 20%	pronoms sujet.	(322)

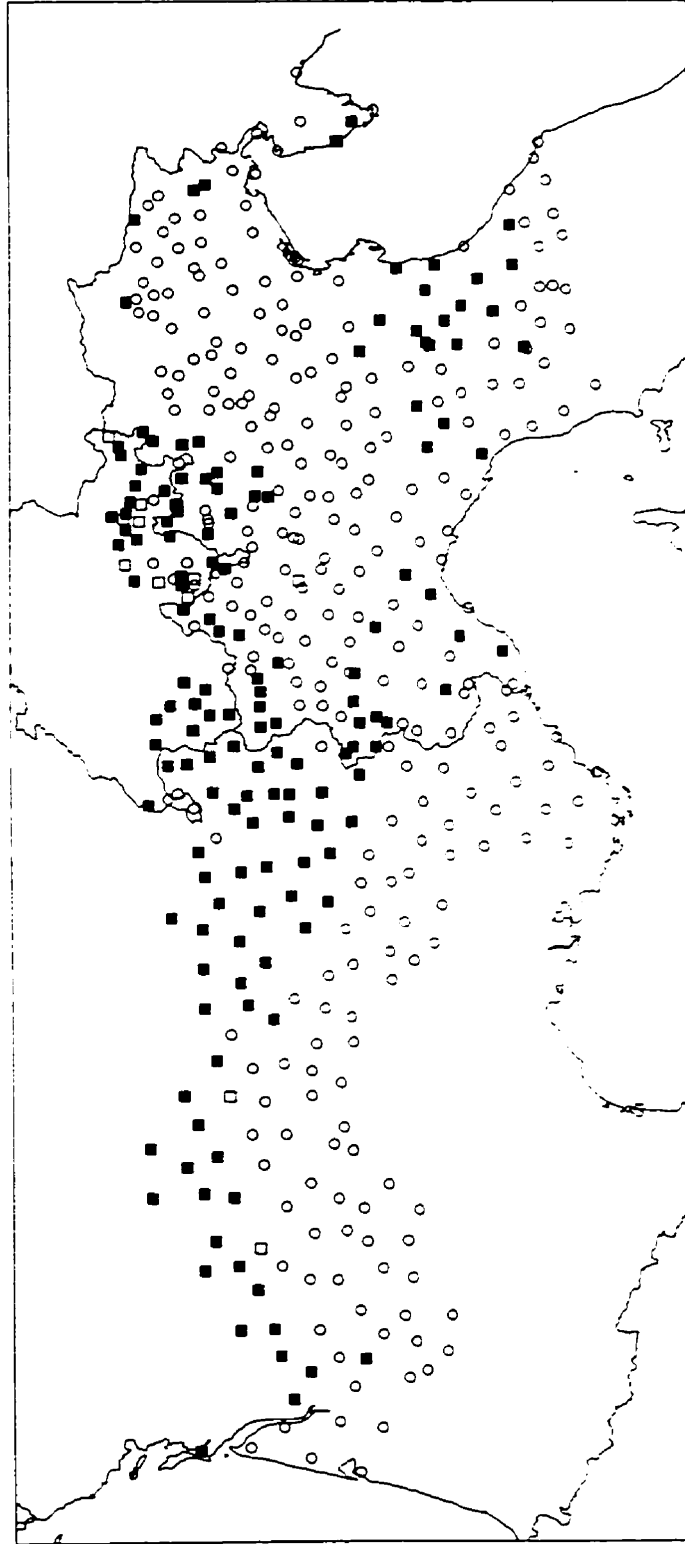
4.2.4. Quatrième personne

Pour des raisons présentées plus haut (voir 4.1.3.), la quatrième personne est la moins représentée dans tout ce corpus: elle ne compte en fait que 6 cartes dans l'*AIS* et 8 dans l'*ALF*. En prenant tous les contextes confondus, nous trouvons 113 points à sujets pronominaux catégoriques, 55 à emploi quasi-catégoriques, et 177 à emploi sporadique (dont 143 sans aucun pronom). Nous limitant au contexte type, nous obtenons la Carte 9, où il y a 145 points à pronom sujet catégorique et 263 points sans pronoms: il n'y a pas de point quasi-catégorique, sans doute à cause du total assez réduit de cartes (seulement 1 ou 2) qui satisfont aux critères du contexte type. Si nous éliminons le critère du présent de l'indicatif de notre contexte type, pour admettre tous les temps verbaux, le résultat est un échantillon un peu moins réduit (5 ou 6 cartes) qui nous donne la Carte 10 avec 151 points à emploi catégorique des sujets pronominaux, 11 points quasi-catégoriques et 173 points à pronoms sporadiques. Il est évident donc que les éléments du contexte type autres que le temps présent -- essentiellement, l'exclusion des éléments préverbaux -- rend la distribution plus claire dans ce cas: on retrouve des blocs compacts dans le bolognais, le piémontais et le frioulan, et on perd quelques points dispersés intermédiaires.



Carte 9: Pronoms sujet à la quatrième personne.
Contexte type (4.2.4). Taux d'usage et n° de points.

■ 100%	pronoms sujet.	(145)	
□ 80	à 100%	pronoms sujet.	(0)
○ 0	à 20%	pronoms sujet.	(263)

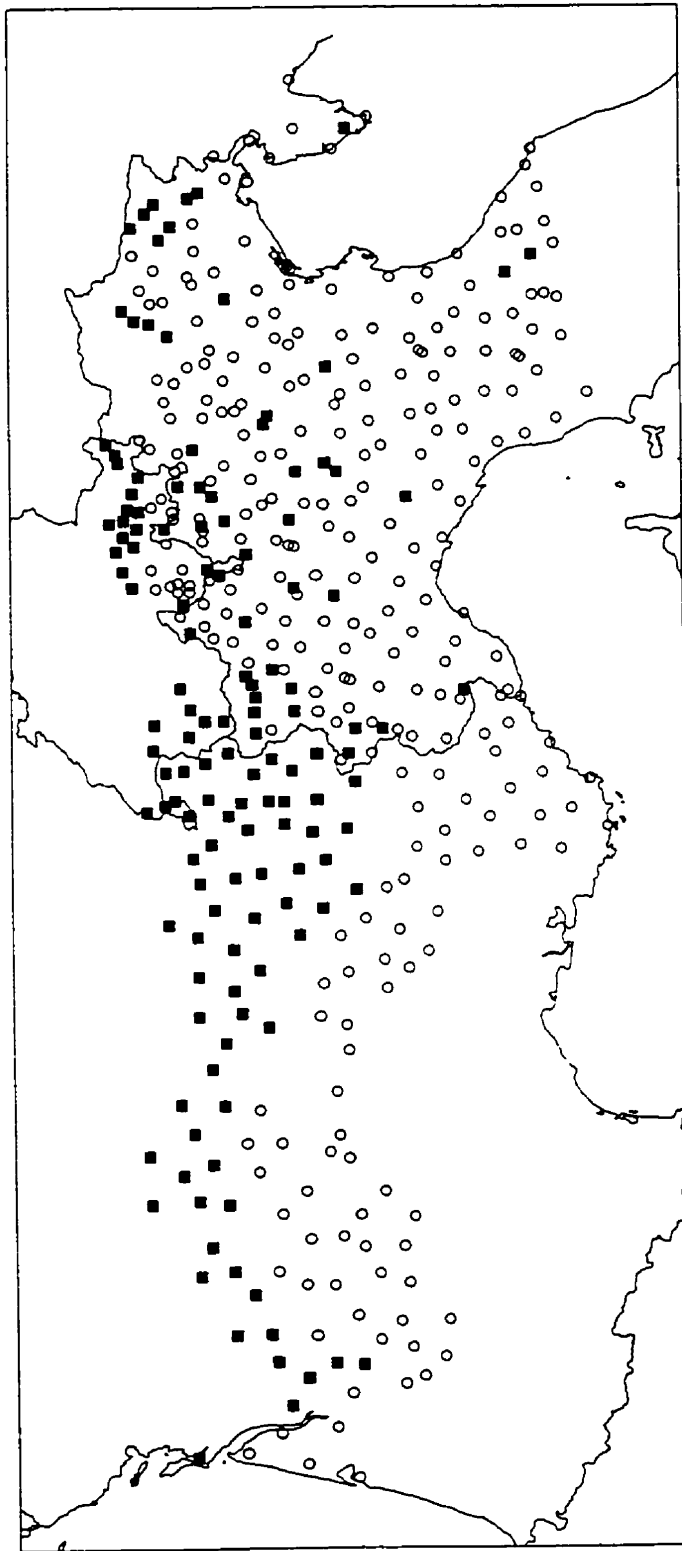


Carte 10: Pronoms sujet à la quatrième personne.
Tous temps confondus (4.2.4). Taux d'usage et n° de points.

■ 100%	pronom sujet.	(151)
□ 80 à 100%	pronom sujet.	(11)
○ 20 à 80%	pronom sujet.	(103)

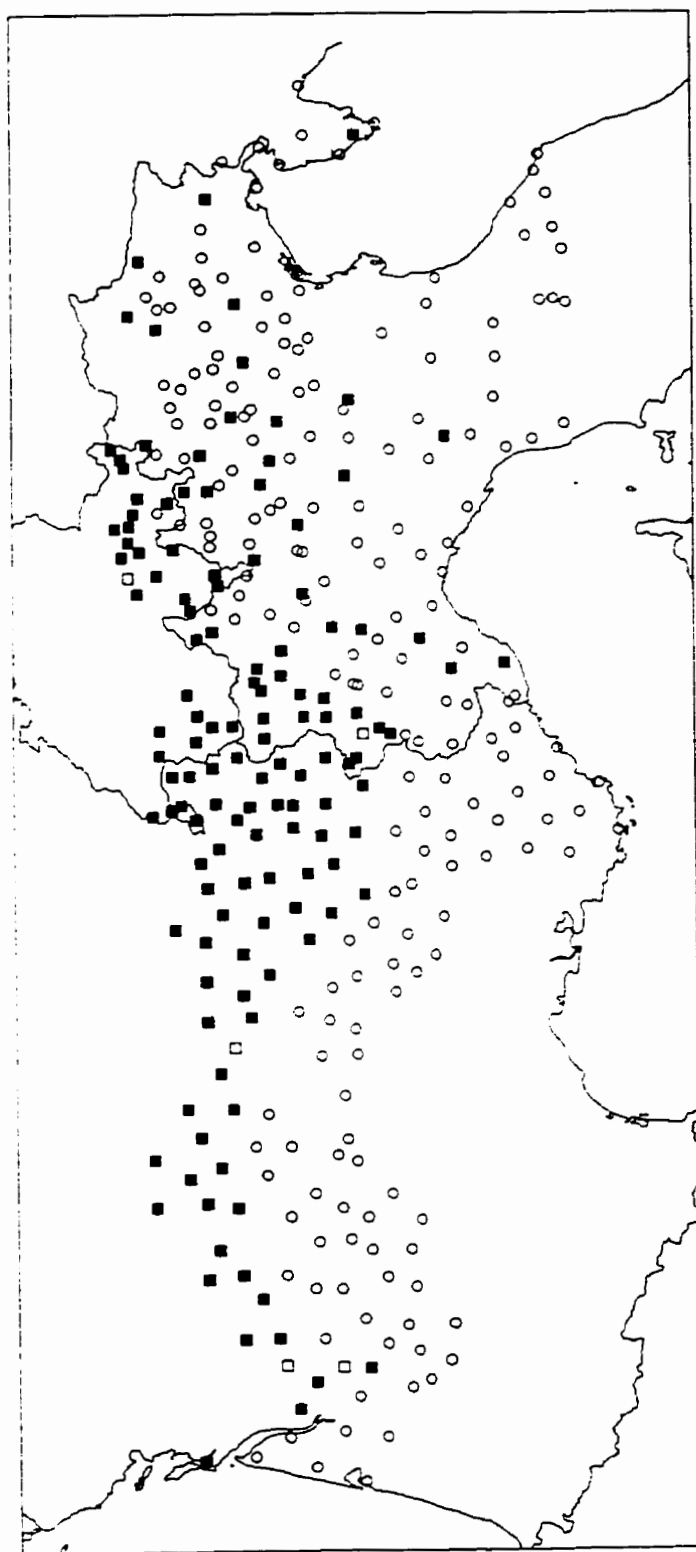
4.2.5. Cinquième personne

À la cinquième personne correspondent en principe 8 à 10 cartes dans notre corpus. Pris dans l'ensemble, cet échantillon tout entier donne 79 points à pronoms catégoriques, 31 quasi-catégoriques et 135 points à emploi sporadique des sujets pronominaux. Et comme dans le cas de la deuxième personne, l'interrogation amène la complication importante du pronom enclitique morphologisé sous la forme de /-f/ ou /-v/ (Jaberg, *Aspects*; voir également 4.2.2.). Avec le contexte type qui élimine entre autres les phrases interrogatives, nous obtenons un sous-échantillon de seulement 2 ou 3 cartes, avec les résultats que l'on peut voir sur la Carte 11: 140 points à pronom sujet catégorique et 276 points avec moins de 20% de sujets pronominaux. En modifiant de nouveau les caractéristiques du contexte type en enlevant le critère du présent de l'indicatif, l'échantillon atteint 4 à 6 cartes. Ainsi sur la Carte 12 nous avons 133 points dont l'usage du pronom sujet est catégorique, seulement 5 où il est quasi-catégorique et 199 où l'emploi ne dépasse pas le seuil de 20%.



Carte 11 : Pronoms sujet à la cinquième personne.
Contexte type (4.2.5). Taux d'usage et n° de points.

■ 100%	pronoms sujet.	(140)	
□ 80	à 100%	pronoms sujet.	(0)
○ 0	à 20%	pronoms sujet.	(276)

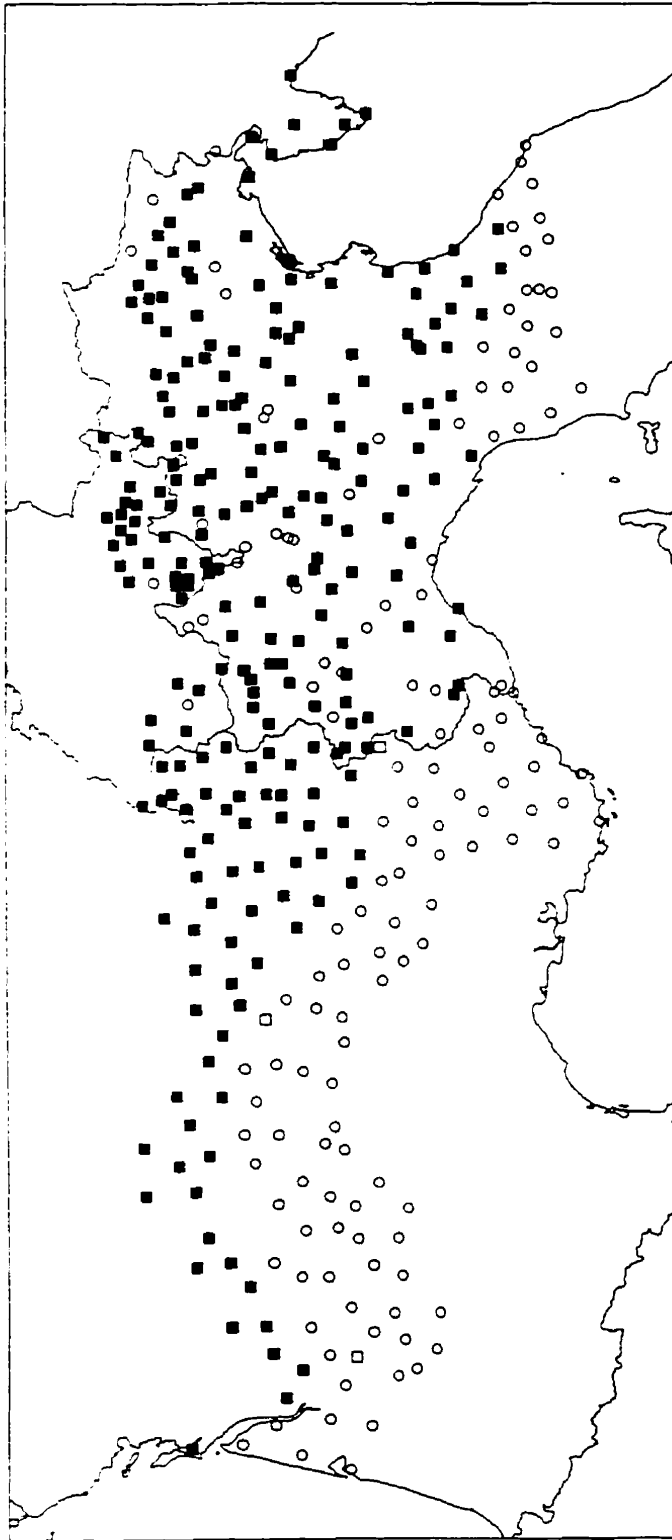


Carte 12. Pronoms sujet à la cinquième personne.
Tous temps confondus (4.2.3.). Taux d'usage et n° de points.

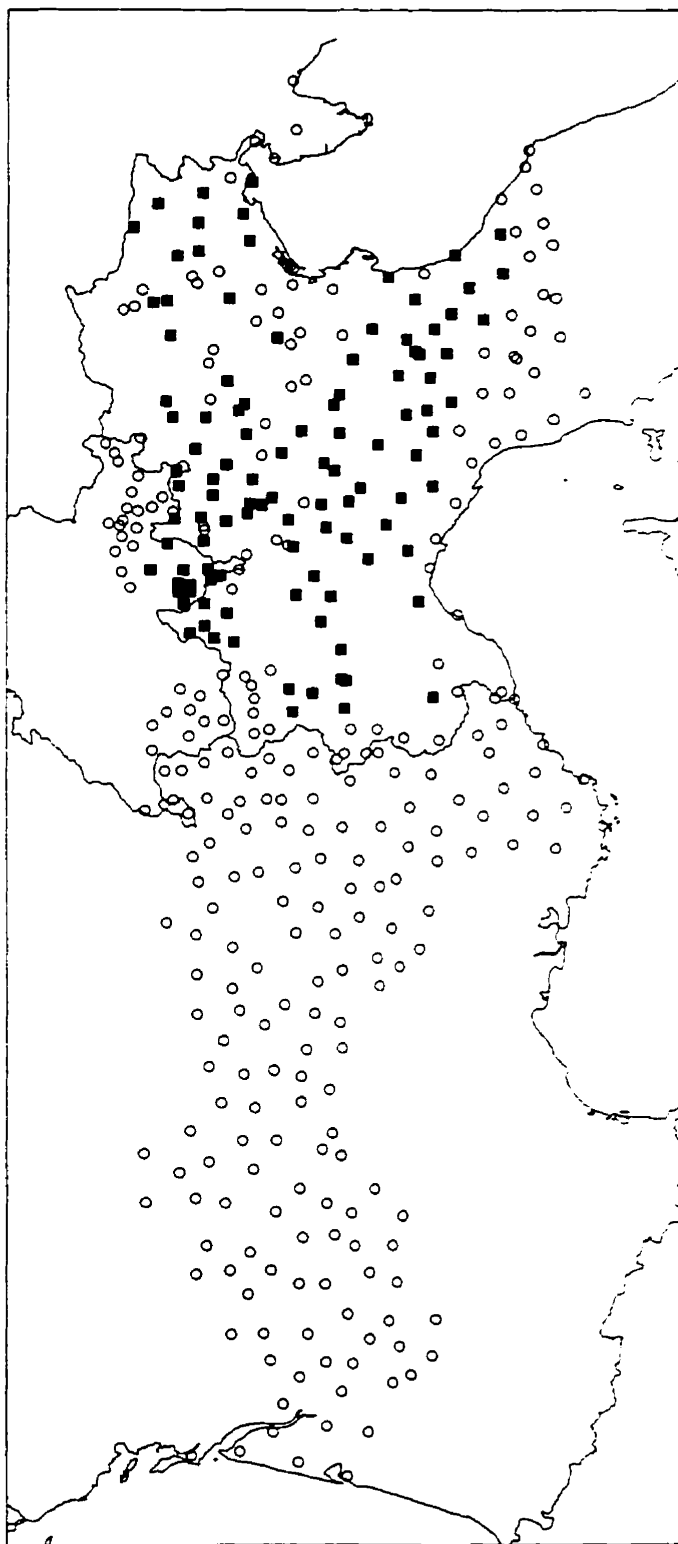
■	100%	pronoms sujet.	(133)
□	80 à 100%	pronoms sujet.	(5)
○	0 à 20%	pronoms sujet.	(199)

4.2.6. Sixième personne

Le corpus contient une dizaine de cartes à la sixième personne, dans lesquelles on trouve 39 points à emploi catégorique des sujets pronominaux, 54 à emploi quasi-catégorique, et 118 avec un emploi de moins de 20%, dont 95 sans aucun pronom sujet. La sixième personne partage avec la troisième l'opposition entre sujets argumentaux (syntagmes nominaux ou pronoms forts) et sujets absents, qui conditionne de façon significative les sujets clitiques. Si nous nous penchons uniquement sur les cas sans sujet explicite ni particule de négation ni objet proclitique (soit le contexte type sans le critère du temps présent), la proportion d'emploi des sujets pronominaux change de façon dramatique: sur la carte 13 on trouve alors 236 points où l'emploi des sujets pronominaux est catégorique, seulement 3 points quasi-catégoriques et 10 points où le taux d'emploi se situe entre 0 et 20%, mais 130 points sans pronom sujet. Encore une fois, le filtre des facteurs conditionnants aurait un effet polarisant sur la distribution des faits, entre autres parce qu'il réduit le sous-échantillon à entre 3 et 8 cartes.



Carte 13 : Pronoms sujet à la sixième personne.
 Contexte type (4.2.6.). Taux d'usage et n° de points.
 ■ 100% pronom sujet. (236)
 □ 80 à 100% pronom sujet. (3)
 ○ 0 à 20% pronom sujet. (140)



Carte 14: Pronoms sujet à la sixième personne.
Sujets argumentaux (4.2.6.). Taux d'usage et n° de points.

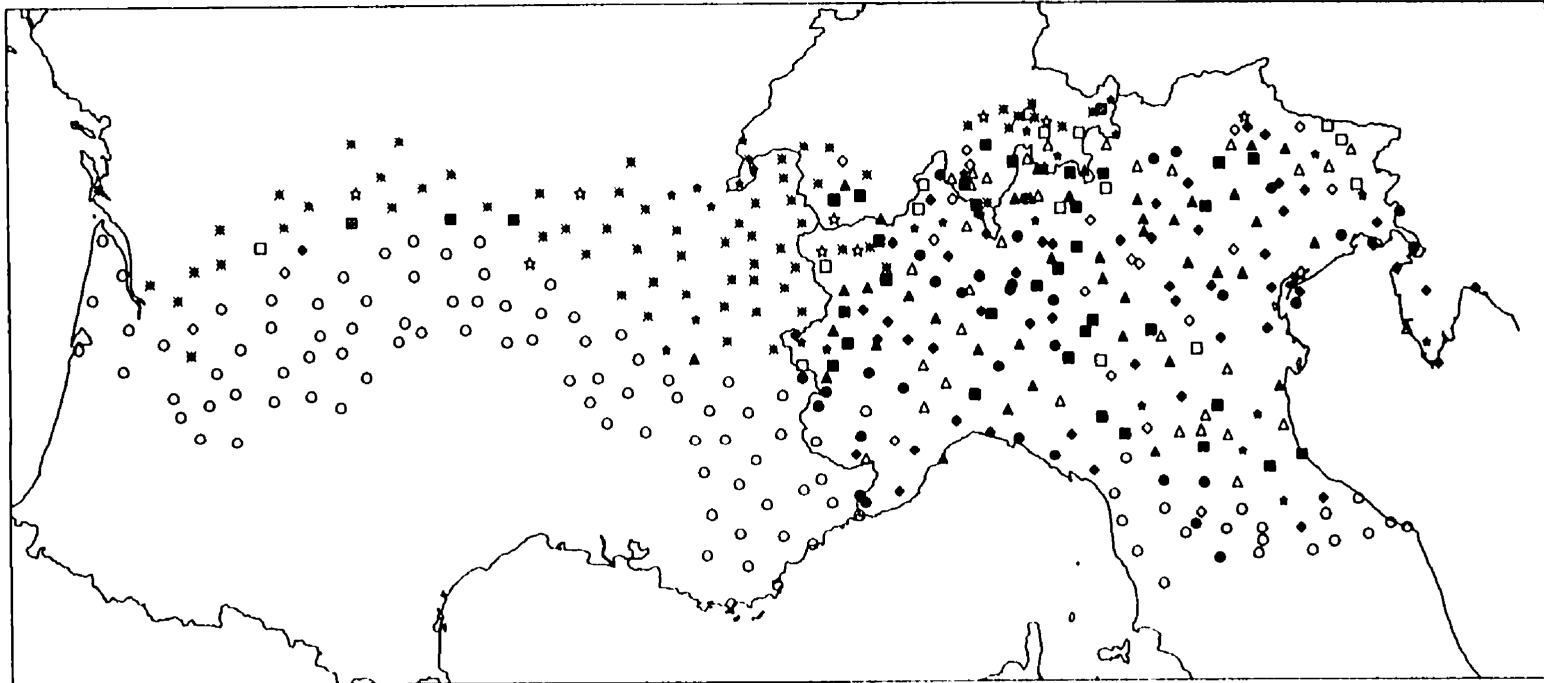
■	100%	pronom sujet.	(108)
□	80 à 100%	pronom sujet.	(0)
○	0	pronom sujet.	(267)

Les cartes avec des sujets lexicaux donnent en revanche un emploi catégorique des sujets pronominaux à 108 points, alors que les sujets pronominaux sont totalement absents à 267 points. Comme on peut voir sur la Carte 14, il n'y a aucun point à emploi quasi-catégorique. Encore une fois, la distribution est inégale: tous les points où le pronom sujet est compatible avec un sujet argumental se trouvent en Italie (voir 4.2.3., 4.4.8. et 5.1.).

4.3. TYPOLOGIE DES PERSONNES

4.3.1. Variétés à six ou zéro pronoms sujet

Considérons d'abord les systèmes à six pronoms, c'est-à-dire des variétés qui seraient considérées comme [-sujet nul] dans une description paramétrique. La Carte 15 montre 68 points où nos données indiquent ce système avec des formes au contexte type. Aux troisième et sixième personnes, cet échantillon distingue les variétés où les sujets pronominaux se combinent avec les sujets argumentaux de celles où les deux types d'éléments s'excluent. De façon générale, ces variétés se rangent sur la frange septentrionale de la zone étudiée, comme on pourrait le prévoir d'après la distribution macrogéographique des variétés françaises et italiennes. Sur la même Carte 15 on voit également les 103 points où l'emploi du pronom sujet reste inférieur à 20% dans toutes les personnes grammaticales: ce seraient les variétés où il existe un système du type «à sujet nul», ce qui veut dire ici tout simplement que l'emploi du pronom sujet reste sporadique.



Carte 15: Typologie des systèmes pronominaux (voir 4.3.).

Types (N°)

• Type 6	(68)
☆ Type 5b	(9)
• Type 5a	(27)
□ Type 4c	(15)
▨ Type 4b	(3)
■ Type 4a	(33)
△ Type 3c	(33)
▲ Type 3b	(9)
▲ Type 3a	(32)
○ Type 2c	(20)
◆ Type 2b	(1)
◆ Type 2a	(51)
● Type 1	(34)
○ Type 0	(103)

Nota: Chaque type est composé d'un chiffre entre 0 et 6, indiquant le nombre de pronoms dans le paradigme, suivi d'une lettre, indiquant si les pronoms de ce paradigme: a) suivent la Généralisation 1, b) la contredisent, ou c) la suivent partiellement.

4.3.2. Variétés à entre deux et cinq pronoms sujet

Ayant écarté ainsi les 171 points qui constituent les deux extrêmes du territoire étudié (absence totale et présence catégorique des pronoms sujet), considérons maintenant les 267 points qui restent, soit ceux où y il y a entre un et cinq pronoms dont l'emploi dépasse 80%. Si nous reprenons des sections précédentes les chiffres correspondant à chaque personne dans le contexte type, nous obtenons le Tableau 3 (les chiffres pour les troisième et sixième personnes correspondent à l'échantillon sans sujet argumental, plus comparable aux échantillons des autres personnes):

Tableau 3: Points à pronom sujet (quasi-)catégorique, par personne (voir Cartes 2-14):

Personne	Points à 100%	Points à entre 80% et 100%	Total des points ≥ 80	% sur le total de 438 points	Rang
Première	101	44	145	33%	5
Deuxième	305	2	307	70%	1
Troisième	125	91	216	49%	3
Quatrième	145	0	145	33%	4
Cinquième	100	0	100	23%	6
Sixième	236	3	239	55%	2

La dernière colonne du Tableau 3 classe les différentes personnes, par ordre décroissant, selon le pourcentage des points à emploi catégorique ou quasi-catégorique des pronoms sujet. La première personne a le même pourcentage que la quatrième, mais nous la classons en cinquième rang puisque son total comprend moins de points catégoriques (deuxième colonne). Des chiffres de la cinquième colonne, on peut dégager facilement une

division naturelle (dans le sens technique du terme: deux groupes dotés d'un maximum de cohérence interne qui se différencient maximalelement l'un de l'autre). Cette division naturelle sépare les trois premiers rangs (dont les totaux dépassent 215 points ou 49%) des trois derniers (dont les totaux sont inférieurs à 146 points ou 33%). Une telle division de la hiérarchie recoupe en partie la classification proposée par Renzi et Vanelli (1983:143), où les personnes grammaticales se rangent par ordre inverse d'«obligatorité» dans les systèmes du pronom sujet. Le Tableau 4 compare leur classification à la notre (la double ligne indique la division naturelle commune):

Tableau 4: Hiérarchie des personnes dans les systèmes de sujets pronominaux:

Rang	Renzi et Vanelli 1983	Notre classification	Bloc
1	Deuxième personne	Deuxième personne	A
2	Troisième personne	Sixième personne	
3	Sixième personne	Troisième personne	
4	Cinquième personne	Quatrième personne	B
5	Quatrième personne	Première personne	
6	Première personne	Cinquième personne	

On remarque que dans les deux hiérarchies on retrouve le même regroupement en deux blocs, avec la deuxième personne en tête de toutes. Si l'ordre à l'intérieur de chaque groupe varie légèrement, c'est bien ce à quoi s'attendent ces auteurs à propos de leur hiérarchie:

Questa gerarchia è provvisoria e condizionata dalla campionatura ritagliata in uno spazio geografico arbitrario. Ma alcuni elementi (per es. la posizione privilegiata della 2. persona, e la divisione in due gruppi) dovrebbe resistere a un esame più approfondito. (1983:143,n.1)⁶²

Quant aux différences de classement à l'intérieur de chaque groupe, il faut remarquer que les troisième et sixième personnes ne sont pas séparées par une différence très importante dans les deux hiérarchies: 239 points comparés à 216 chez nous, par rapport à 90% et 86,7% chez eux. On pourrait donc à la limite admettre qu'elles sont à peu près à égalité en deuxième rang. Plus significatif est l'écart dans le classement des personnes qui incluent le locuteur (première et quatrième): pour Renzi et Vanelli ces deux personnes se trouvent au rang le plus bas, à l'autre extrême de la deuxième personne.

Mais la hiérarchie présentée dans le Tableau 4 est purement fréquentielle: elle ne nous dit que le nombre de points à chaque personne, et rien du tout sur l'identité des points ni les regroupements possibles des personnes à pronom sujet. On peut supposer par exemple que tous les points qui ont un pronom pour au moins une personne l'auront forcément à la deuxième personne, puisqu'elle est la plus répandue numériquement, mais rien dans les chiffres du Tableau 3 n'oblige les points à suivre une hiérarchie

62. «Cette hiérarchie est provisoire et conditionnée par l'échantillonnage restreint dans un espace géographique arbitraire. Mais certains éléments (par ex. la position privilégiée de la 2ème personne, et la division en deux groupes) devraient résister à un examen plus approfondi.»

implicationnelle: pourrait-on trouver également des variétés où le seul pronom sujet à atteindre la quasi-catégoricité est celui de la troisième personne (ou une autre encore) plutôt que la deuxième?

La liste de six systèmes de pronoms sujet fournie par Renzi et Vanelli (1983:127-128) suggère une typologie qui se voudrait peut-être exhaustive. Or, il semble d'emblée contradictoire qu'un système à cinq pronoms doit exclure la première personne alors qu'un système à quatre pronoms se passerait de la quatrième et de la cinquième (voir 3.3.5. et note 35). Notons que chacun de ces deux types est représenté dans leur corpus par une seule variété, le provençal et le garfagnanino respectivement, qui ne sont même pas proches sur le plan géographique. Si nous admettons l'hypothèse que l'absence d'autres systèmes à quatre ou cinq pronoms est fortuite plutôt que typologique, quelle généralisation est-ce que nous pouvons tirer des systèmes documentés par Renzi et Vanelli? Il est à noter que les pronoms qui manquent dans ces deux cas font partie du «bloc B» établi plus haut, c'est-à-dire des pronoms nettement moins communs dans les deux études. Nous avançons donc l'hypothèse qui suit:

Généralisation 1: Dans les systèmes où le paradigme des pronoms sujet est «incomplet» (comportant de 1 à 5 personnes), les pronoms présents seront préférablement du bloc A (deuxième, troisième et sixième personnes), tandis que les pronoms absents seront préférablement du Bloc B (la première, quatrième ou cinquième personnes).

Il s'agit donc de vérifier cette généralisation face aux faits des différents systèmes attestés dans nos données. Les résultats de cette vérification typologique sont présentés dans la Carte 15.

Parmi les 36 points avec un système à cinq pronoms sujet, il y en a 27 du type 5a, où le pronom absent provient du bloc B, résultat qu'on peut voir sur la Carte 15. Sur la même carte on voit aussi les neuf points du type 5b, où il y a un système à cinq pronoms et dont celui qui manque est du bloc A: une minorité de points qui contredisent donc la Généralisation 1.

Pour ce qui est des systèmes à quatre pronoms, les possibilités se multiplient, puisqu'il peut y avoir (théoriquement) jusqu'à quinze combinaisons possibles de pronoms qui peuvent manquer. Heureusement, on n'a pas affaire à toutes ces quinze possibilités: certaines paires de pronoms dont l'absence est typologiquement possible ne sont pas réalisées. Des cinq «trous» dans la typologie (systèmes possibles qui ne sont pas attestés), dans deux cas les deux pronoms absents sont du bloc A tandis que dans les trois autres cas un des deux pronoms absents est du bloc A, ce qui constitue un appui indirect à la Généralisation 1. D'autre part, des 51 variétés attestées où on trouve des systèmes à quatre pronoms, 33 sont du type 4a, prédit par la Généralisation 1, c'est-à-dire où les pronoms absents sont du bloc B. Parmi les combinaisons qui vont à l'encontre de la Généralisation 1, il y a seulement trois du type 4b où les pronoms absents sont tous les deux du bloc A (il s'agit de cas où la troisième et la sixième personnes sont absentes: il n'y a aucun cas où la deuxième personne manque). Dans les quinze autres variétés on voit le

type 4c, une sorte de «demi-violation» de la Généralisation 1, où sont absents un pronom du bloc A et un autre du bloc B.

C'est dans les systèmes à trois pronoms (donc à trois pronoms absents) qu'on voit une véritable floraison de possibilités: les six personnes fournissent vingt triplets théoriquement possibles. Le système 3a, en accord complet avec la Généralisation 1, se rencontre 32 fois sur les 74 points avec des systèmes de trois pronoms. D'autre part, les triplets virtuels non réalisés (systèmes non attestés) incluent la combinaison des trois pronoms du bloc B, ainsi que cinq cas où deux des trois pronoms sont du bloc B, ce qui constitue un appui indirect de la Généralisation 1. Entre ces deux extrêmes, on trouve que les autres systèmes 3b, «à violation majoritaire» de la Généralisation 1, soit avec deux pronoms du bloc B, sont attestés à seulement neuf points. Les systèmes du type 3c, en «accord majoritaire» avec la Généralisation 1 -- deux des trois pronoms sont du bloc A -- sont les plus nombreux après le cas 3c, étant attestés à un total de 33 points. Et il semble y avoir une influence aussi du rôle spécial de la deuxième personne (voir 4.3.3.): parmi les systèmes du type 3b (systèmes où deux des trois pronoms présents sont du bloc B) le plus commun (attesté à cinq points) est le système composé des deuxième, quatrième et cinquième personnes, tandis que le système le plus fréquent parmi ceux du type 3c est composé des première, deuxième et sixième personnes.

Parmi les 72 points présentant des systèmes à deux pronoms sujet, on en trouve 51 du type 2a où les deux pronoms sont du bloc A, soit en accord complet avec la Généralisation 1. Parmi ces 51 points, le système le plus commun est celui qui combine la

deuxième et la sixième personnes, attesté à 26 points. La combinaison des troisième et sixième personnes n'est attestée qu'à cinq points. Il y a un seul point avec un système du type 2b, soit deux pronoms du bloc B (quatrième et cinquième personnes), les autres combinaisons de ce type étant non attestés. Le reste des cas, soit vingt points, ont des systèmes du type 2c, soit un pronom de chaque bloc. Presque tous ici (18 points) sont des systèmes qui combinent la deuxième personne avec un pronom du bloc B.

Finalement les systèmes à un seul pronom confirment la Généralisation 1 sans exception. Les 34 points où ce type de système est attesté ont donc un pronom tiré du bloc A, et il n'y en a aucun avec un pronom du bloc B. Le système que Renzi et Vanelli documentent en franco-provençal (deuxième personne présente) est le plus fréquent, étant présent à 22 des 34 points, suivi par la sixième personne.

Nos résultats typologiques, loin d'être aussi catégoriques et tranchants que les travaux antérieurs, reposent cependant sur des tendances centrales discernées dans une masse de faits atomistiques, tendances qui sont beaucoup plus marquées que ce que le hasard seul pourrait produire. Le Tableau 5, qui reprend les résultats pour chaque type de système pronominal, montre bien que la Généralisation 1 se voit confirmée dans la grande majorité des cas.

Tableau 5: Degré d'accord avec la Généralisation 1, selon le nombre de personnes dans les systèmes pronominaux.⁶³

N° de Pronoms	Total points	% accord total (n°)	% accord partiel (n°)	% désaccord (n°)
5	36	75 (27)	n/a	25 (9)
4	51	65 (33)	29 (15)	6 (3)
3	74	43 (32)	45 (33)	12 (9)
2	72	71 (51)	28 (20)	1 (1)
1	34	100 (34)	n/a	0
Total	267	66 (177)	26 (68)	8 (22)

Cette généralisation, bien que formellement moins «forte» que les prévisions de Renzi et Vanelli, inclut néanmoins l'essentiel de leur observation en le combinant avec les cas attestés qui tombent en dehors de leur typologie. Il est intéressant de noter que le degré d'accord est plus élevé aux deux extrêmes de l'échelle, et plus bas au milieu: les systèmes à trois pronoms sont les plus répandus, et aussi ceux où on observe le plus souvent l'accord partiel avec la Généralisation 1.

63. Évidemment, la généralisation en question ne peut s'appliquer qu'aux paradigmes partiels, c'est-à-dire aux systèmes ayant moins de six pronoms sujet mais plus de zéro. De même, l'accord «partiel» avec cette généralisation ne peut se produire que dans les systèmes de deux, trois ou quatre pronoms.

4.3.3. Le statut privilégié de la deuxième personne

Il y a une autre généralisation de Renzi et Vanelli qui revient dans la discussion qui précède: la place privilégiée de la deuxième personne (1983:143). Encore une fois, il s'agit d'une tendance plutôt que d'une règle inviolable, puisqu'il existe des variétés à sujets pronominaux où manque celui de la deuxième personne. Mais si nous modifions légèrement la généralisation de Renzi et Vanelli, comme en (2):

Généralisation 2: dans un système qui emploie plus d'un pronom sujet de façon catégorique ou quasi-catégorique, la deuxième personne doit être présente.

on peut vérifier que cette généralisation est respectée elle aussi dans la majorité de cas.

Considérons donc les pourcentages dans le Tableau 6:

Tableau 6: Présence et absence de *tu* dans tous les paradigmes incomplets.

N° pronoms	Total points	% <i>tu</i> présent (n°)	% <i>tu</i> absent (n°)
5	36	94 (34)	6 (2)
4	51	100 (51)	0
3	74	92 (68)	8 (6)
2	72	89 (64)	11 (8)
1	34	65 (22)	35 (12)
Total:	267	90 (239)	10 (28)

En fait, les deux généralisations semblent se compléter: dans les systèmes à un seul pronom, où *tu* n'est présent qu'à 65% des points, la Généralisation 1 obtient un respect catégorique. Bien sûr, la Généralisation 1 inclut partiellement la Généralisation 2, puisque la deuxième personne fait partie du Bloc A. Et si nous prenons tous les points du Tableau 5 qui sont en désaccord complet ou partiel avec la Généralisation 1, nous trouvons que la plupart de ces variétés respectent quand même la Généralisation 2:

Tableau 7: Présence et absence de *tu* dans les points en désaccord avec Généralisation 1:

N° pronoms	Total points en désaccord	% <i>tu</i> présent (n°)	% <i>tu</i> absent (n°)
5	9	78 (7)	22 (2)
4	18	100 (18)	0
3	42	86 (36)	14 (6)
2	21	86 (18)	14 (3)
total	90	88 (79)	12 (11)

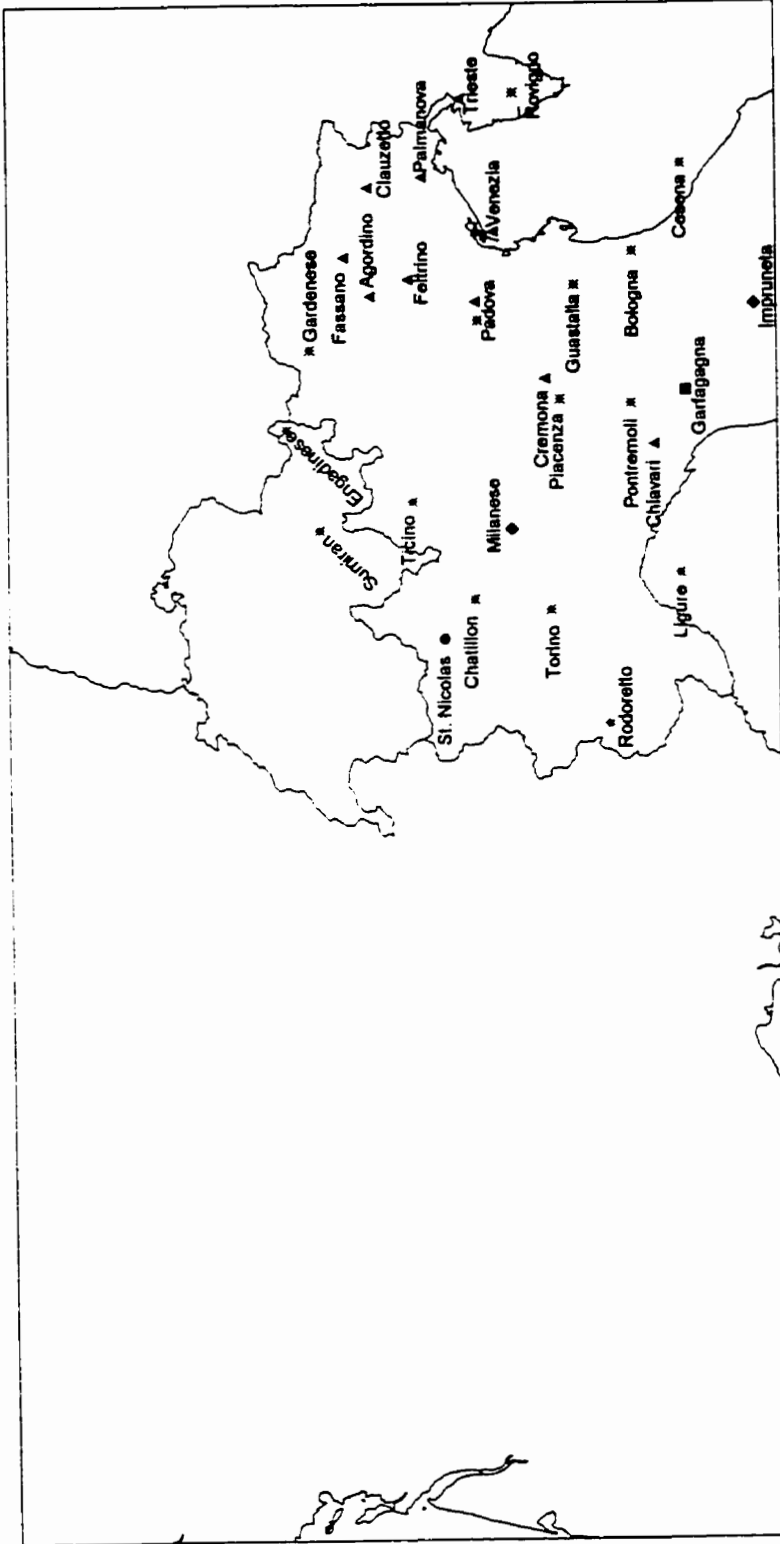
La combinaison des deux généralisations recouvre donc la presque totalité des points: seulement onze points (soit moins de 3% du total de 438 localités) ont des systèmes de pronoms sujet qui ne respectent ni la Généralisation 1 ni la Généralisation 2. Même si elles ne sont pas parfaites, ces deux généralisations représentent des tendances significatives qui ne peuvent pas être le résultat du hasard. Les implications typologiques de ces deux généralisations sont reprises dans nos conclusions (voir 5.2.1.).

4.3.4. La distribution typologique

Si on compare la distribution des systèmes à six pronoms dans notre Carte 15 aux variétés que Renzi et Vanelli (1983) citent comme ayant des systèmes à six pronoms avec sujets non argumentaux, dans la Carte 16, on voit ressortir immédiatement certains décalages intéressants. Plusieurs variétés à six pronoms dans leur étude ne correspondent pas à des systèmes comportant six pronoms mais plutôt à des systèmes qui en ont entre deux et cinq dans notre Carte 15.⁶⁴ Comme leur étude porte sur des variétés italiennes, il est évident que nos résultats pour la France tombent en dehors de la portée de leur travail, mais pour la partie comparable, notre Carte 15 présente une Italie où l'étendue des systèmes à six personnes se voit sévèrement réduite au profit des systèmes intermédiaires. En fait, les seuls points de coïncidence seraient dans les variétés suisses et valdôtaines, soit les marges septentrionales de la dialectologie italienne. Nos résultats divergent aussi de ce qu'on trouve dans les monographies descriptives sur les variétés nord-italiennes, qui décrivent en général plus de pronoms sujet que n'en constate notre Carte 15.

Passant aux variétés à moins de six pronoms, nous trouvons 36 points du type 5a (c'est-à-dire avec cinq pronoms obéissant à la Généralisation à 1) sur la Carte 15, face à une seule variété de ce type d'après Renzi et Vanelli (1983), sur la Carte 16. Le système qu'ils documentent (celui où le pronom absent est de la première personne) n'a pas de

64. Nous situons ici géographiquement l'information qui est donnée sous forme de listes de variétés par Renzi et Vanelli (1983). Il convient de noter que Poletto (1993) donne aussi des listes de variétés, mais ses classifications sont difficilement réductibles à une typologie par nombre de personnes.



Carte 16: Les systèmes pronominaux d'après Renzi et Vanelli (1983)

Les types se réfèrent au nombre de pronoms sujet (voir aussi 4.3.):

- Type 6 *
- Type 5 •
- Type 4 ■
- Type 3 ▲
- Type 2 ◆
- Type 1 ●

place privilégiée, ne correspondant qu'à sept de ces 27 points. Mais la situation géographique de cette variété documentée par Renzi et Vanelli (1983) -- le provençal alpin de Rodoretto -- correspond assez bien à certains des points que nous localisons.

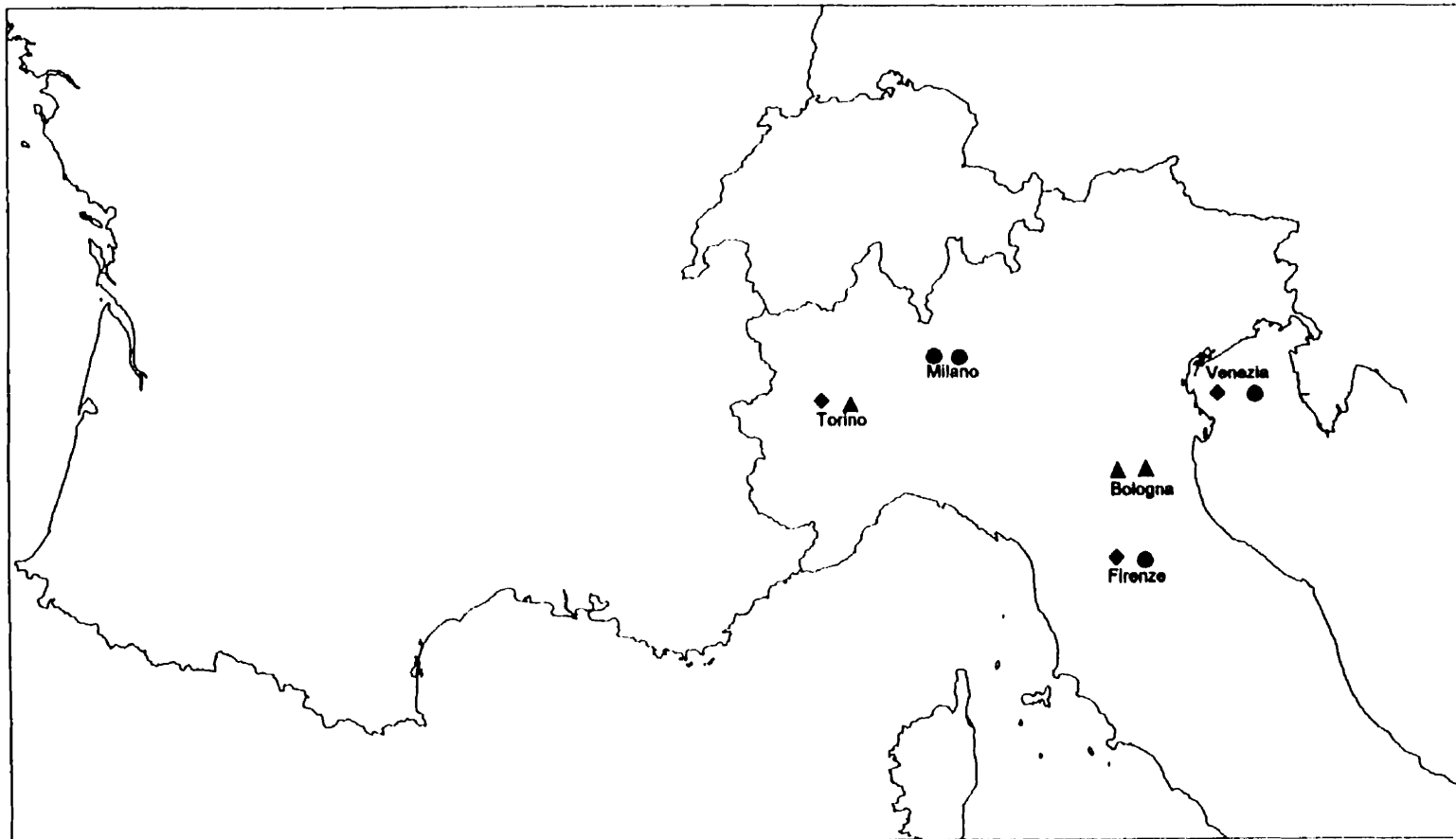
Des 33 points du type 4a (paradigmes de quatre pronoms obéissant à la Généralisation 1), seulement dix points correspondent au système documenté par Renzi et Vanelli, où les pronoms absents sont de la quatrième et cinquième personnes, et de ces dix, aucun ne s'approche géographiquement de Garfagnana, le seul point où ce système se localise d'après leur enquête (voir la Carte 16). Les 32 points du type 3a (trois pronoms suivant la Généralisation 1) sur la Carte 15 correspondent à dix variétés où Renzi et Vanelli documentent des paradigmes de ce type, sur la Carte 16. On rappellera qu'aucun paradigme n'est documenté avec trois pronoms contredisant la Généralisation 1.

Les 51 points du type 2a (en accord avec la Généralisation 1) sur la Carte 15 correspondent à une seule variété dans l'enquête de Renzi et Vanelli, sur la Carte 16. Des 34 points du type 1a (obéissant tous à la Généralisation 1) qu'on voit sur la Carte 15, un seul point se trouve chez Renzi et Vanelli, le franco-provençal. Nos systèmes 2b e 2c ne sont pas attestés dans leur étude (voir 4.3.2.).

4.3.5. Variation sociale

Il est évident que cette étude ne saurait vraiment viser la variation à l'intérieur d'une communauté linguistique, puisque nous n'avons en général qu'une seule personne comme sujet pour chaque localité: en d'autres mots, la seule variable sociale dont nous

disposons dans notre corpus est la localité même (voir 2.1.4.). Ceci dit, il y a quand même une suggestion de ce qui pourrait être des différences sociolinguistiques quand on considère les écarts entre les doubles enquêtes relevées dans les cinq grandes villes, qui sont présentées dans la Carte 17. Ainsi à Florence il se trouve que l'informateur 523,1 est maître cordelier tandis que l'informateur 523,2 est employé de bibliothèque (Jaberg et Jud 1928:100-101). Il n'est donc pas surprenant de trouver que l'artisan a un système à deux pronoms sujet, alors que l'individu plus instruit a un système à un seul pronom sujet, qui diverge donc moins par rapport à la langue standard. Le même écart social se trouve à Venise, où l'informateur 376,1 (un charpentier en fer) possède un système à deux pronoms sujet, alors que 376,2 (un employé de bureau, Jaberg et Jud 1928:90) a un système à un seul pronom. Notons que dans le cas de Florence, c'est le locuteur de la classe moyenne, plus proche du standard, qui est en accord avec la région circonvoisine, tandis qu'à Venise c'est celui de la classe ouvrière, divergente par rapport au standard, qui ressemble le plus à ce qu'on trouve aux points voisins. À Turin, la situation est moins nette: l'informateur 155,1 (musicien et commerçant) a un système à deux pronoms sujet, qui ressemble à ce qu'on trouve aux points du côté nord-est de la ville, tandis que 155,2 (employé de bibliothèque, Jaberg et Jud 1928:56) possède un système de trois pronoms sujet, ce qui est plutôt en accord avec les points au sud-est et à l'ouest de la ville. Ici donc ni l'écart social entre les deux individus ni la relation avec la campagne voisine ne présente une indication claire des tendances sociolinguistiques.



Carte 17: La variation sociale dans les doubles enquêtes de l'AIS (4.3.5.).

Types (voir 4.3.1. et Carte 15):

▲ Type 3a.

◆ Type 2a.

● Type 1.

Enfin, à Milan et à Bologne, les doubles enquêtes donnent des résultats identiques du point de vue typologique.

L'information sociale fournie par les enquêteurs de l'*AIS* peut donc nous suggérer des tendances qui vont dans le sens de la variation sociolinguistique, mais sans qu'il y ait suffisamment de données pour en inférer des corrélations robustes. Du côté de l'*ALF*, il n'y a pas de double enquête (voir 3.2.3.2., note 4.), mais nous possédons néanmoins quelques indications précieuses sur les informateurs. Ainsi au point 635 on trouve un système à six pronoms sujet, entouré de points sans pronoms ou à deux pronoms sujet seulement. Or ce même point ressort comme un îlot d'éléments septentrionaux, qu'il s'agisse de variantes lexicales ou phonétiques (voir les cartes de Jochowitz 1973 *passim*). On ne s'étonne donc nullement d'apprendre que le sujet de cette enquête est une servante d'hôtel âgée d'une trentaine d'années (Gilliéron et Edmont 1902: 45). Il s'agit donc d'une personne réunissant justement les caractéristiques qui favorisent l'innovation linguistique (Gauchat 1902; Chambers et Trudgill 1980): une femme relativement jeune et en contact avec des voyageurs, dont un certain nombre parleraient des variétés plus prestigieuses venues du nord. Il nous est impossible ici de corréler toutes les informations fournies par les enquêteurs d'atlas avec les résultats de cette étude, mais cette tâche devrait faire l'objet de recherches ultérieures.

4.4. RÉSULTATS VARIABLES

Dans cette section nous considérons l'effet global des facteurs conditionnants énumérés dans la section 3.4. sur la présence ou l'absence des pronoms sujet dans le corpus pris dans son ensemble. Il convient d'être plutôt prudent dans l'interprétation des chiffres cités dans les sections qui suivent, puisque le nombre total de cartes sur lesquelles apparaissent certains facteurs est parfois assez réduit, ce qui tend aussi à favoriser les résultats catégoriques. D'autre part, l'ensemble du corpus qui sert de groupe de comparaison aura toujours moins de catégoricité simplement en fonction du fait qu'il est plus grand et plus hétérogène.

Évidemment, les différents facteurs considérés dans les sections qui suivent n'agissent pas en isolation: l'effet de chacun se combine avec les effets de tous les autres pour déterminer la probabilité d'emploi d'un pronom sujet dans chaque cas. L'outil statistique adéquat pour mesurer ce type d'interactions serait une analyse par régression multiple, du type appelé «de règle variable», réalisé par les logiciels VARBRUL . L'inconvénient de ce genre d'analyse pour nous est qu'elle présuppose que toutes les occurrences examinées viennent de locuteurs qui partagent la même grammaire. Or notre corpus est justement composé de locuteurs de variétés distinctes, ce qui rend ce type d'analyse assez douteuse (Heap 1995). Il existe une version du logiciel en question, VARBRUL 3, qui a la capacité de regrouper les données par sous-groupes linguistiques: l'algorithme «pèse» dans chaque cas la probabilité que les locuteurs partagent la même grammaire avec les mêmes contraintes, ou bien qu'ils possèdent des grammaires distinctes.

Ce serait donc le type d'analyse qui pourrait potentiellement nous indiquer à la fois le poids relatif des facteurs et les sous-groupes de locuteurs par grammaire. Hélas, cette analyse reste «potentielle» pour l'instant, puisque les caractéristiques de notre corpus nous empêchent de soumettre nos données à ce type d'analyse. Les raisons techniques qui rendent cette option impossible concernent le nombre de locuteurs (nos 438 points dépassent de loin la limite prévue par le logiciel) et le nombre de sous-divisions parmi les locuteurs.⁶⁵ L'analyse et le regroupement de nos données par règles variables (ou par d'autres techniques statistiques plus sophistiquées) devra donc faire l'objet de recherches ultérieures.

Dans les sections qui suivent, nous n'estimons pas qu'il soit utile de cartographier les résultats pour chaque facteur. La présentation de chaque facteur se termine par un résumé des effets favorisant ou défavorisant l'emploi du pronom sujet. Ces résultats sont repris dans le chapitre qui suit, où nous examinons des interprétations possibles que ces faits suggèrent.

65. Les créateurs de cette version du logiciel, Rousseau et Sankoff, soulignent que «Another problem associated with the procedure involves the number of groups to be constructed. There is no feasible, strictly mathematical way now available for judging, for example, whether the increase in likelihood obtained by allowing four groups instead of three is significant enough to warrant considering the extra distinction meaningful.» (1978:102) Étant donné que nous avons vraisemblablement affaire ici à au moins six groupes (d'après notre typologie) il est probable que l'application de cette analyse à notre corpus est trop complexe pour être incluse ici.

4.4.1. Position du verbe

Il est difficile de voir un effet important pour ce facteur: la position la plus favorable aux sujets pronominaux est à l'initiale des subordonnées, ce qui donne 37 points catégoriques et 105 points quasi-catégoriques (dans une quinzaine de cartes). Les positions initiale et non initiale des principales ne donnent chacune que deux points catégoriques, avec 122 et 107 points quasi-catégoriques chacune. Il n'y a pas suffisamment de données pour la position non initiale des subordonnées pour en tirer des conclusions. Il n'y a pas de différences claires entre les subordonnées et les principales globalement (toutes positions confondues).

Il aurait été peut-être désirable de coder ces occurrences pour un facteur du type «information nouvelle / non nouvelle», mais ce type de précision est assez difficile à déterminer dans le contexte d'un questionnaire de ce genre (voir 3.4.1). On pourrait croire qu'il s'agit en fait de l'influence cachée d'un autre facteur: est-ce le fait que les principales sont plus souvent interrogatives que les subordonnées qui est responsable de cette différence? Mais si par exemple nous éliminons les interrogatives du sous-échantillon, la position relative de ces deux facteurs reste la même: 131 points catégoriques ou quasi-catégoriques pour les initiales des principales affirmatives par rapport à 155 points catégoriques ou quasi-catégoriques pour les initiales de subordonnées affirmatives.

La position initiale de propositions subordonnées est la plus favorable aux sujets pronominaux. (voir 5.4.1.)

4.4.2. Temps et mode

Plus de la moitié des cartes du corpus ont des verbes au présent de l'indicatif. En comparant tous les verbes au présent avec tous les verbes à un temps autre que le présent, nous obtenons des résultats peu différenciés: les verbes au présent (une cinquantaine) ne présentent aucun point à emploi catégorique et 101 points à emploi quasi-catégorique, tandis que les verbes non présents (une quarantaine) ne présentent aucun point catégorique et 111 points quasi-catégoriques. Il faut procéder à une séparation plus fine des temps et modes pour trouver des contrastes.

Certains temps verbaux, comme par exemple le passé simple et le passé antérieur, ne sont pas suffisamment fréquents dans le corpus pour en tirer des résultats auxquels l'on peut se fier. En combinant certaines autres classes de temps, on obtient des sous-échantillons intéressants. Ainsi on trouve les différents subjonctifs -- présent, imparfait, parfait -- dans un certain nombre de cartes (cinq à huit), avec 68 points catégoriques et 29 points quasi-catégoriques, un total légèrement inférieur à celui des verbes au présent. Le passé composé et la combinaison de l'imparfait et le conditionnel représentent des sous-échantillons d'une vingtaine de cartes chacun, et sont un peu plus favorable aux sujets pronominaux que le présent: ils ont respectivement un total de 122 et 124 points catégoriques et quasi-catégoriques. Le plus grand contraste se trouve dans les 5 à 7 cartes avec le futur simple ou le futur périphrastique, qui donnent 64 points catégoriques et 68 points quasi-catégoriques.

Les futurs sont les temps les plus favorables aux sujets pronominaux. L'imparfait et le passé composé sont plus favorables aux sujets pronominaux que le présent de l'indicatif, et les subjonctifs sont moins favorables. (voir 5.4.2.)

4.4.3. Type de verbe et auxiliaires

Les verbes transitifs (une trentaine de cartes) et inergatifs (environ 6 cartes) montrent à peu près la même distribution d'emplois catégoriques ou quasi-catégoriques des pronoms sujet: 107 ou 108 points au total pour chaque type. Les inaccusatifs en ont un peu moins, soit 84 points au total. Les types de verbes qui favorisent les pronoms sujet sont: les modaux (seulement 5 cartes, avec 113 points catégoriques et 31 quasi-catégoriques), les réfléchis (7 cartes, avec 89 catégoriques et 55 quasi-catégoriques), «être» ou «essere» comme copule (une quinzaine de cartes avec 24 points catégoriques et 117 quasi-catégoriques), et «avoir» ou «avere» comme verbe plein (une dizaine de cartes, avec 50 points catégoriques et 105 points quasi-catégoriques).

Pour ce qui est des auxiliaires, «être» ou «essere» en fonction d'auxiliaire n'est attesté que dans cinq ou six cartes, dans lesquelles on trouve 66 points catégoriques et 60 points quasi-catégoriques. Dans la quinzaine de cartes avec «avoir» ou «avere» en fonction d'auxiliaire on trouve 13 points à emploi catégorique du pronom sujet et 121 à emploi quasi-catégorique. Les deux auxiliaires ont donc presque le même total de points où l'emploi des pronoms sujet dépasse 80%, un taux qui est légèrement moins élevé que celui des verbes pleins correspondants.

Le verbe le plus favorable aux sujets pronominaux est «avoir» verbe plein, suivi par les verbes réfléchis, les modaux et «être» verbe plein, et ensuite par les deux auxiliaires. Les inaccusatifs ont un effet légèrement défavorisant par rapport aux transitifs et aux inergatifs. (voir 5.4.3.)

4.4.4. Genre

Pour ce qui est du genre, il y a une quarantaine de cartes où le sujet est clairement identifié comme masculin et une dizaine où le sujet est féminin (nous en excluons une autre quarantaine où le genre morphologique n'est pas marqué). De ces deux cas, le masculin n'a que 2 points qui font un usage catégorique du pronom sujet et 76 qui en font un usage quasi-catégorique. Le féminin présente 25 points catégoriques et 90 quasi-catégoriques.

Le genre masculin défavorise l'emploi des sujets pronominaux, par rapport au genre féminin. (voir 5.4.4.)

4.4.5. Négation

Les cartes avec des formes négatives ne sont pas très nombreuses dans ce corpus, et en plus, il n'y a aucune forme de négation qui soit commune à toutes les variétés étudiées. La comparaison de toutes les cartes avec négation et toutes les cartes sans négation ne nous renseigne pas beaucoup: celles-ci comptent 101 points à emploi quasi-

catégorique des pronoms sujet et aucun à emploi catégorique, mais celles-là représentent un mélange de trois types de négation. Si nous ne considérons que le sous-échantillon de variétés présentant la négation simple (préverbale seulement), soit 161 points en tout, nous ne trouvons que 24 points catégoriques et 5 quasi-catégoriques. Pour la négation forte (pré- et post-verbale) on a 126 points, dont 85 à emploi catégorique du sujet pronominal et aucun à emploi quasi-catégorique. Pour ce qui est de la négation postverbale, elle apparaît à 271 points, dont 83 ont un emploi catégorique du pronom sujet et 18 un emploi quasi-catégorique, pour un total de 101, identique au total pour les formes sans négation.

La négation a un effet défavorisant sur l'emploi des sujets pronominaux uniquement quand elle inclut un élément préverbal. (voir 5.4.5.)

4.4.6. Interrogation

Comme point de comparaison nous pouvons prendre ici l'ensemble des cartes sans aucun type d'interrogation, soit plus de 80 cartes, sur lesquelles on ne trouve aucun point à emploi catégorique du sujet pronominal, et 107 à emploi quasi-catégorique. Il y a cinq ou six cartes avec l'interrogation partielle, qui donnent 114 points à emploi catégorique des sujets pronominaux et 43 points à emploi quasi-catégorique. L'interrogation totale produit des résultats similaires, bien qu'un peu moins frappants: sur un échantillon de six cartes, on trouve 112 points où l'emploi des sujets pronominaux est catégorique et 30 où il est quasi-catégorique.

L'interrogation, qu'elle soit partielle ou totale, favorise le pronom sujet, par rapport aux phrases sans interrogation. (voir 5.4.6.)

4.4.7. Clitiques obliques

Si l'on considère toutes les cartes sans aucun objet clitique, soit environ 80 cartes, on y trouve 102 points à usage quasi-catégorique des pronoms sujet et aucun à emploi catégorique. Si l'on compare ce résultat à l'ensemble de toutes les cartes comportant au moins un pronom objet clitique (entre 10 et 20 cartes), on trouve que le total des points catégoriques et quasi-catégoriques est de 96, un chiffre très similaire au total des cartes sans clitique. Mais il serait une erreur d'en conclure que les pronoms clitiques objet n'ont aucun effet sur la présence d'un pronom sujet, car il est fort possible que les pronoms réfléchis se comportent d'une façon différente des autres pronoms, ce qui masquerait partiellement l'effet de ce groupe. L'exclusion des réfléchis du groupe des clitiques est appuyée par le fait que, considérés à part, ils montrent la tendance inverse: dans les 5 cartes avec un pronom réfléchi, on trouve 103 points à sujet pronominal catégorique et 43 quasi-catégoriques, soit plus que les contextes sans aucun clitique.

En effet, si nous choisissons l'ensemble des cartes (une dizaine) qui ont un pronom clitique objet non réfléchi, nous obtenons un échantillon légèrement réduit de 434 points, dont 11 font un emploi catégorique du pronom sujet et 79 un usage quasi-catégorique, pour un total de 90, soit un peu moins que le total sans aucun clitique. Pour ce qui est des suites plus longues, il n'y a qu'un ou deux cartes qui ont deux pronoms clitiques sans

réfléchis, et ceci à seulement 114 points, dont 50 emploient le pronom sujet de façon catégorique et aucun ne le fait de façon quasi-catégorique. Bien que le nombre total de points ici soit trop petit pour que ce résultat soit statistiquement fiable, il semble néanmoins que les suites de deux clitiques objet ont un effet encore plus défavorisant sur l'emploi du pronom sujet.

Plus il y a d'objets proclitiques (hormis les réfléchis), moins on tend à employer les sujets pronominaux. (voir 5.4.7.)

4.4.8. Type et position du sujet

Comme nous l'avons fait remarquer plus haut, il n'y a pas suffisamment de sujets indéfinis dans ce corpus pour les faire corrélérer avec quoi que ce soit. Quant aux autres types de sujets, les pronoms forts sont peu nombreux aussi et semblent se comporter comme les sujets nominaux définis. Ces deux types de sujets pris ensemble constituent les sujets argumentaux, dont la distribution diffère substantiellement de celle des sujets nuls, comme on le voit dans notre discussion des troisième et sixième personnes (voir 4.2.3. Carte 5, et 4.2.6. Carte 14).

Quant à la position du sujet, elle n'entre en jeu que par rapport aux sujets argumentaux. Nous avons suffisamment de données de ce type pour opposer les sujets en position préverbale (une quinzaine de cartes) à ceux qui sont en position invertie ou disloquée (une dizaine de cartes): ces deux dernières sont prises ensemble comme à peu

près équivalentes en termes syntaxiques. La position préverbale défavorise fortement l'emploi des sujets pronominaux: on relève 131 points à emploi catégorique ou quasi-catégorique avec cette position, par rapport à 259 catégoriques ou quasi-catégoriques avec les sujets disloqués et invertis.

La position préverbale défavorise fortement l'emploi des sujets pronominaux. Les sujets argumentaux excluent l'emploi du pronom sujet dans les variétés gallo-romanes. (voir 5.4.8.)

C'est ce système qui domine parmi les variétés françaises, où l'on ne trouve qu'un emploi sporadique du pronom sujet avec les sujets lexicaux, mis à part quelques points assez isolés dans le Massif Central qui en font un emploi qui atteint la quasi-catégoricité. Cette différence frappante dans la distribution du phénomène est examinée dans nos conclusions (voir 5.1.).

Ayant vu les différents types de systèmes pronominaux qui existent selon les personnes grammaticales, ainsi que les effets quantitatifs des différents facteurs conditionnants, nous passons dans le chapitre qui suit aux interprétations possibles de ces faits empiriques.

CHAPITRE 5: CONCLUSIONS

«Quel travail ai-je au fond accompli à Charmey? J'ai étudié, d'une façon sommaire, environ 50 langues individuelles et je n'y ai rien trouvé d'individuel.»

(Gauchat 1902:231)

5.0. PRÉSENTATION

Nos résultats peuvent se réduire fondamentalement à trois classes de faits: la distribution géographique des différents systèmes de pronoms sujet, la typologie des systèmes possibles, et la variation qui se produit à l'intérieur de ces systèmes de pronoms. Les deux premiers types de résultats peuvent à leur tour se comparer aux résultats de l'étude antérieure de Renzi et Vanelli (1983). La variation parmi ces systèmes, qui ne fait pas partie de leurs préoccupations, suggère néanmoins des analyses qui peuvent aider à comprendre le fonctionnement des sujets pronominaux. Voyons donc d'abord la distribution géographique des faits, et ensuite les implications de la typologie tracée dans le chapitre précédent, pour en venir enfin aux différents types de variation et leurs implications.

5.1. CONCLUSIONS GÉOLINGUISTIQUES

Dans une étude de l'envergure de celle-ci, il serait inutile de vouloir commenter les résultats de chaque point individuellement. Vue de loin, la Carte 15 reconferme la généralisation de laquelle nous partons: les variétés sans pronom sujet se situent au sud

(provençal, toscan et italien central) et les variétés avec sujets pronominaux au nord (français, franco-provençal, romontsch, piémontais, lombard, avec l'exception de l'émilien-romagnol qui constitue une déviation au sud). Mais à l'intérieur de cette généralisation macrogéographique, on peut se demander s'il est possible de trouver des bandes régulières de variation. Un continuum morphosyntaxique idéal donnerait des faisceaux d'isoglosses pronominales, avec les variétés à six pronoms à côté des variétés à cinq pronoms, qui seraient à leur tour à côté des variétés à quatre pronoms, et ainsi de suite. La réalité est bien plus complexe: tout dépend d'où on cherche le continuum. De façon générale, on peut dire que les variétés sans pronom sujet forment un bloc plutôt homogène dans le sud des régions étudiées, tandis que les variétés avec des pronoms forment un ensemble hétérogène et irrégulier, avec des frontières floues et des inconsistances internes.

Dans la partie de la Carte 15 qui correspond à l'*ALF*, la transition entre variétés à six pronoms et variétés sans pronom sujet est assez abrupte: tantôt on passe directement de six à zéro (par exemple à l'ouest, où l'estuaire de la Gironde forme la division géolectale la plus nette de la France, et aussi dans la vallée du Rhône), tantôt on trouve quelques variétés intermédiaires à cinq ou quatre pronoms. Les variétés à un, deux et trois pronoms sont assez rares parmi les données de l'*ALF*; quand on les trouve, elles ne se placent pas nécessairement dans une transition graduelle ou un continuum.

Le panorama est en revanche assez bigarré du côté italien de la même Carte 15. Les deux pôles de la typologie (systèmes à six pronoms par opposition aux systèmes à

pronom nul) sont toujours aux deux extrêmes de la zone étudiée (Toscane sauf Florence, par rapport aux Grisons et au Val d'Aoste). Mais même ce pôle septentrional n'est pas homogène: on trouve aussi des points à quatre et cinq pronoms aux Grisons. Plus significativement, aucune des régions intermédiaires ne présente un seul type homogène ni même majoritaire: l'aire piémontaise compte des variétés qui vont d'un pronom à cinq, l'aire lombarde d'un à quatre, l'aire frioulane d'un à cinq, l'aire vénitienne d'un à trois et l'aire émilien-romagnole de deux à cinq. Il est difficile de voir dans ceci des «bandes» typologiques, ou des faisceaux d'isoglosses. Cette situation rappelle les mots de Jaberg quand il compare la diversité linguistique italienne à l'unité relative des variétés du territoire français: «En Italie, on n'a jamais douté, comme en France, de l'existence des dialectes. . . : la réalité était trop évidente.» (1933:32). La relative homogénéité linguistique de la France que Gaston Paris (1888:134-135, voir aussi 2.1.3.) prend comme preuve de l'inexistence des dialectes est pour Jaberg plutôt le résultat de l'histoire particulière de la France: la «centralisation d'abord politique et administrative ensuite» qui a nivelé les anciennes aires dialectales (1933:32). Mais là où Jaberg voit des aires dialectales bien définies sans beaucoup de transitions graduelles, notre Carte 15 présente des aires assez floues, sans frontières nettes. Cette différence est due peut-être au type de variables considérées: lexicales et phonétiques dans son cas, morphosyntaxiques dans le nôtre. Il se peut donc que les transitions en morphosyntaxe soient plutôt graduelles alors que les variations lexicales et phonético-phonologique se concentrent mieux dans des zones bien

définies. Mais il faudrait davantage d'études comparatives pour pouvoir affirmer cette différence avec certitude.

Le manque de diversité du côté français peut donc être vu (au moins partiellement) comme le produit des conditions socio-historiques. L'existence de l'État français avec toutes les forces centralisatrices que cela implique, y compris la langue standard, favoriserait une situation de transition abrupte du système provençal à sujet nul au système français à six pronoms. Il est significatif que les rares cas de variétés transitionnelles se trouvent surtout dans des régions relativement isolées. Le fameux Croissant limousin-auvergnat est protégé du contact externe par le Massif Central, les variétés provençales et franco-provençales des Alpes et du Jura par leur terrain montagneux. Hormis ces cas d'isolation géographique, on peut affirmer qu'un fort État centralisateur crée des conditions sociolinguistiques où on peut trouver une transition géolinguistique syntaxiquement abrupte du type [\pm sujet nul], sans beaucoup de variation interne.

Pour les variétés italiennes, en revanche, l'absence jusqu'à date très récente d'un état central permet une plus grande diversité dans les systèmes de sujets pronominaux. Pour la majorité des sujets de l'*AIS*, l'état italien représenterait une entité politique nouvelle, surgie de leur vivant, en contraste net avec les sujets de l'*ALF*, pour qui l'existence de l'état français serait une réalité permanente et incontestable.

L'évolution des systèmes pronominaux en Italie, plus ou moins libres de pression normative, mènerait naturellement à la morphologisation du pronom sujet comme marqueur préverbal, en contraste avec les variétés attestées dans la partie française de

notre étude, où le pronom occuperait la position syntaxique du sujet. Wanner voit dans ce contraste l'effet du degré de standardisation dans les deux pays:

«The non-NS [null subject] phenomenon in its free development approaches a morphological solution, while a consciously standardized variety with primary written direction tends towards the syntactic solution (1995:17).⁶⁶

Notre Carte 15 suggère que l'influence standardisante du français favoriserait une évolution syntaxique qui s'approche à ce que prédirait le modèle paramétrique, avec une transition géolinguistique relativement abrupte, tandis que l'absence d'une norme unique en Italie favoriserait la morphologisation du pronom sujet, une situation qui échappe à la paramétrisation.

D'après nos résultats, la «forte divergence» que Wanner (1995:8) voit dans les pronoms sujet des variétés italiennes est encore plus radicale que l'on ne croirait d'après les travaux antérieurs (voir 4.3.4.) Ces différences sont attribuables en partie aux méthodes employées: là où celle de Renzi et Vanelli met en relief l'invariabilité et les

66. L'acadien de Terre-Neuve contredirait peut-être cette corrélation: à la différence du français canadien parlé hors Acadie, cette variété ne connaît pas le redoublement du sujet, en dépit d'une isolation presque totale de la norme métropolitaine. D'après Nadasdi et King (1996), la conservation du sujet en position syntaxique serait due à la morphologie verbale plus riche dans cette variété. Et comme le remarque Wanner (1995:8-9), la grammaticalisation ne s'applique pas nécessairement à l'évolution interne d'un cas particulier, qui peut toujours contredire les tendances globales.

distinctions binaires, la nôtre permet une marge de tolérance plus large qui permet de voir la pénombre quasi-catégorique qui entoure le noyau dur de la catégoricité. La leur se prête donc mieux à l'obtention des données catégoriques aux dépens de la variation, alors que la nôtre vise la variation et perd donc en généralisations nettes et abstraites.

À part les différences d'ordre épistémologique entre les programmes de recherche, ces différences de méthode répondent surtout aux types de données disponibles pour chaque projet. Les personnes qui fournissent les données de Renzi et Vanelli, tout comme celles qui écrivent les monographies descriptives, seraient soucieuses de donner un portrait cohérent de leur variété comme entité distincte de l'italien, et mettraient donc en relief tout élément qui la différencie, comme par exemple le système pronominal. Les sujets de l'*AIS*, en revanche, devaient «traduire en patois» sans trop se poser la question de savoir si leur emploi des pronoms sujet correspondrait au système géolectal «pur» ou non. Outre la question des sujets d'enquête, il y a des différences qui découlent du réseau de points employé pour chaque étude. Un filet à mailles plus fines attrape nécessairement plus de poissons, que ceux-ci soient désirables ou non. Comme notre réseau de points est considérablement plus dense (262 points de l'*AIS*, comparés à 30 variétés chez Renzi et Vanelli), il est normal que nos résultats incluent des faits qui leur échappent. D'autre part, Renzi et Vanelli se servent souvent d'une étiquette unitaire comme «Bologne» qui correspondrait à plusieurs points d'enquête dans notre corpus. Ces étiquettes cachent très probablement un certain degré de nivellement local, alors que nos résultats ne correspondent qu'aux points individuels.

5.1.2. Types de paradigmes et leur distribution géographique

Pour ce qui est des autres types de systèmes pronominaux à paradigmes incomplets documentés par Renzi et Vanelli (1983:127-128), nos résultats indiquent à la fois une typologie plus riche et une distribution plus large. C'est-à-dire qu'en décomposant les résultats présentés dans la Carte 15, nous trouvons un ou deux sous-types à chacun de leurs systèmes, et que ces systèmes ont une distribution géographique en général plus vaste que celle qui est indiquée par leur étude (voir 4.3.2. et 4.3.4.).

On notera que les systèmes non attestés dans l'étude de Renzi et Vanelli mais qui sont captés par notre réseau sont majoritairement des types 2a, 3a, 4a et 5a, soit ceux qui obéissent à la Généralisation 1. Que nous en trouvions plus qu'eux pour chaque type est en partie une conséquence naturelle de la densité de notre réseau (à part le fait évident que notre étude recouvre aussi les points français qui tombent en dehors de leur étude). On doit y voir aussi la contrepartie du fait que nous trouvons moins de systèmes à six pronoms, qui sont plus répandus dans leur étude (comparer les Cartes 15 et 16).

5.2. CONCLUSIONS TYPOLOGIQUES

5.2.1. Généralisations 1 et 2

Nous reproduisons ici les deux généralisations présentées dans le chapitre précédant (voir 4.3.2. et 4.3.3.) quant aux systèmes des pronoms sujet possibles:

Généralisation 1: Dans les systèmes où le paradigme des pronoms sujet est «incomplet» (comportant de 1 à 5 personnes), les pronoms présents seront préférablement du bloc A (deuxième, troisième et sixième personnes), tandis que les pronoms absents seront préférablement du Bloc B (première, quatrième ou cinquième personnes).

Généralisation 2: dans un système qui emploie plus d'un pronom sujet de façon catégorique ou quasi-catégorique, la deuxième personne doit être présente.

Si elles n'atteignent pas la précision d'une hypothèse fortement prédictive comme celle de Renzi et Vanelli, elles représentent néanmoins une régularité assez robuste qui ne peut aucunement être attribuée au hasard: les deux généralisations décrivent plus de 95% des paradigmes partiels (ceux qui comptent entre un et cinq personnes). Mais voilà le problème: elles ne font que les décrire, et elles resteront des stipulations descriptives tant qu'on ne pourra pas trouver une explication plus approfondie du fait que certaines personnes sont privilégiées dans les systèmes de pronoms sujet.

Si l'on commence par la Généralisation 2, le statut particulier de la deuxième personne semble assez évident d'un point de vue intuitif. Dans un certain sens, la deuxième personne représente beaucoup plus que la personne à qui on adresse un énoncé: pour Benveniste,

. . . toute *personne* qu'on se représente est de la forme «tu», tout particulièrement -
- mais non pas nécessairement -- la personne interpellée. . . . On pourra donc définir le «tu» comme la *personne non-subjective*, en face de la *personne subjective* que «je» représente; et ces deux «personnes» s'opposeront ensemble à la forme de «non-personne» (=«il»). (1966:232)

Et en effet, la deuxième personne est souvent «surdéterminée», comme le disent Renzi et Vanelli (1983:), avec un pronom et une désinence distincte.⁶⁷ Mais alors, comment expliquer l'écart important entre la deuxième et la première personne: s'il importe de distinguer les deux participants essentiels d'un acte de communication, pourquoi est-ce le *tu* qui est toujours bien marqué, même sur-marqué, alors que le *je* ne l'est que relativement rarement? On voit mal ce qui fait que la «personne subjective» *je* peut se

67. La deuxième personne générique aurait vraisemblablement un comportement différent à cet égard (voir Cameron 1993), mais nous n'avons repéré aucun cas de ce type dans notre coprus (voir Appendices I et II).

passer d'être aussi bien identifiée, puisqu'elle fait également partie de l'acte de communication.

Pour Poletto, la distribution exceptionnelle des pronoms sujet de la deuxième personne se réduit à une question de sa position syntaxique: le *tu* serait toujours réalisé sous le noeud d'accord AGR, comme les clitiques objets, alors que les pronoms des autres personnes peuvent se réaliser variablement sous AGR ou sous MOD. Mais cette façon de représenter la différence ne fait que déplacer la question sans y répondre, comme cette chercheuse l'avoue elle-même:

Resta comunque misterioso perché tutti i dialetti debbano realizzare un clitico di questo tipo solo per la seconda persona singolare, mentre nel caso delle altre persone la scelta resta variabile.⁶⁸ (1993:140)

Elle suggère par la suite que, si les personnes grammaticales sont définies par les traits binaires [\pm déictique] [\pm deuxième personne] et [\pm singulier], ce serait alors les trois valeurs positives du *tu* qui en font la personne qui est toujours réalisée dans les systèmes de pronoms sujet. Cette explication nous semble bien trop *ad hoc*: le choix des traits et des valeurs n'est motivé que par les faits dont il est supposé rendre compte. Pour ne prendre

68. «Il reste cependant mystérieux pourquoi tous les dialectes doivent réaliser un clitique de ce type seulement pour la deuxième personne du singulier, alors que dans le cas des autres personnes le choix reste variable.»

que le dernier trait, [\pm singulier], il est généralement admis qu'en termes de marque morphologique, le singulier serait la valeur «par défaut» et le pluriel serait donc la valeur à marquer positivement. Plutôt que de façonner exprès un système de traits pour les besoins de la cause, nous préférons garder une simple observation qui restera «mystérieuse», et pourra faire l'objet de recherches ultérieures.

Ce refus de l'«analyse» par traits nous semble d'autant plus justifié quand on pense à la Généralisation 1: les deux blocs sur lesquels elle se base ne peuvent pas se traduire facilement en «classes naturelles», dans un quelconque système de traits. Dans le système de Poletto, le bloc A (deuxième, troisième et sixième personnes) serait composé des valeurs [+déictique] [+deuxième personne] et [+singulier] d'une part, et [-déictique] [-deuxième personne] et [\pm singulier] d'autre part. Dans le même système, le bloc B comprendrait les valeurs [+déictique] [+deuxième personne] et [-singulier] (cinquième personne), et [+déictique] [-deuxième personne] et [\pm singulier] (première et quatrième personnes). Des problèmes similaires semblent se produire avec n'importe quel autre système de traits que l'on peut inventer pour classer les six personnes grammaticales: les «géométries» des traits pronominaux proposées récemment, entre autres par Bonet (1991, 1995a, 1995b) et Ritter (1996) auraient la même difficulté, puisqu'elles emploient également la catégorie de «personne grammaticale» sous une forme ou une autre (que ce soit un trait privatif ou un noeud dans la géométrie), ce qui pousse inévitablement à regrouper ensemble les deuxième et cinquième personnes.

La composition des deux blocs de la Généralisation 1 est pourtant assez robuste, puisque les mêmes blocs reviennent dans nos résultats et ceux de Renzi et Vanelli, mais nous voyons mal quel statut attribuer à ces deux blocs à part celui d'une stipulation descriptive. Le cœur du problème réside dans la séparation de *tu* et *vous*: dans notre Généralisation 1, comme dans la hiérarchie proposée par Renzi et Vanelli, la deuxième et la cinquième personnes se classent dans des positions assez éloignées, alors que les systèmes de traits (y compris les géométries pronominales) tendent toujours à les mettre ensemble. Pourquoi est-ce que *tu* seul se regrouperait avec les «non personnes» (troisième et sixième), alors que la combinaison de *tu* avec une ou des non personnes [*tu + il(s) ou elle(s) = vous*] se rangerait du côté de *je* et *nous*? Peut-être parce que la distinction entre singulier et pluriel est moins évidente que l'on croirait. Pour Benveniste (1966:236), seule la troisième personne (la non-personne) aurait un vrai pluriel: quand on «pluralise» les première et deuxième personnes, on obtient des groupes mixtes qui contiennent ces personnes au singulier accompagnées d'autres personnes, soit des personnes «amplifiées». ⁶⁹

69. Également intéressant pour nous est l'absence de symétrie entre la troisième et la sixième personne: pour Poletto, ces deux personnes grammaticales iraient toujours de pair, tandis que pour Renzi & Vanelli ce ne serait qu'en milanais (leur système 5) qu'on trouverait un pronom sujet à la troisième personne sans le trouver aussi à la sixième (1983:128,13). Or nous trouvons que ces deux «non personnes» ne sont pas associées aussi souvent que l'on croirait d'emblée: il existe bien des points avec la deuxième et la sixième personne, combinaison qui respecte la Généralisation 1 mais qui introduit une autre asymétrie de nombre.

Bien que nous n'ayons pas de données précises sur les désinences verbales à tous les points de cette enquête, nous pouvons affirmer que la composition des blocs A et B ne semble pas s'expliquer non plus en termes fonctionnels de «richesse morphologique». S'il est vrai d'une part que les quatrième et cinquième personnes sont celles qui gardent en général les terminaisons verbales les plus distinctes (et qui auraient donc moins «besoin» du pronom sujet), la première ne l'est que dans les cas où le /-o/ final est conservé. D'autre part, la terminaison distincte que conserve presque partout la deuxième personne ne l'empêche pas d'avoir aussi un pronom sujet dans la grande majorité des cas où il y a un pronom sujet. La seule observation fonctionnelle que l'on peut faire est celle de Renzi et Vanelli: les personnes sont toujours distinctes, soit par les terminaisons, soit par les pronoms, soit par les deux. Mais ce fait ne saurait nous aider à motiver la composition des deux blocs de la Généralisation 1. Quel que soit le cadre théorique que l'on adopte, il y a des faits qui résistent à nos efforts d'analyse, et l'existence de ces deux généralisations semble en être un (du moins pour l'instant). Il semble aussi douteux qu'on puisse réduire la variation attestée dans ces zones de transition à la «propagation d'un seul morphème», comme le suggère Jochnowitz (1973:69, voir aussi 2.1.3.).

Si nous acceptons tout simplement l'existence de ces deux généralisations, sans pouvoir les expliquer, il devient intéressant d'examiner leur façon d'interagir. Nous ne pouvons pas dire catégoriquement que l'une des deux soit plus forte: les Tableaux 5 et 6 (voir 4.3.2.) montrent que dans certains cas c'est la Généralisation 1 qui est respectée le plus souvent, et dans d'autres cas c'est la Généralisation 2. Cette flexibilité d'application

suggère que ces généralisations doivent être implantées dans un cadre théorique qui permette l'interaction libre des contraintes, plutôt qu'une théorie à dérivation déterministe et catégorique. À certains points, notamment avec les paradigmes des types 3 et 4, nos deux généralisations sont contredites plus souvent, ce qui suggère qu'elles interagissent alors avec d'autres contraintes (que nous ne pouvons pas encore identifier). Quand nous aurons cerné clairement l'ensemble des contraintes qui contribuent à la constitution de ces paradigmes partiels, il est fort probable qu'il faudra les implanter dans une théorie morphologique qui admette l'interaction des contraintes dans une hiérarchie (voir 5.2.2.).

5.2.2. Autres implications typologiques

Dans cette section nous traitons des relations d'implication qui peuvent se dégager de nos résultats outre celles concernent les personnes grammaticales. Ces conclusions sont comparées ensuite à certaines des conclusions de Renzi et Vanelli.

Pour ce qui est des verbes météorologiques et impersonnels, on remarquera une relation d'inclusion presque parfaite entre les Cartes 6 et 8: presque tous les points catégoriques ou quasi-catégoriques sur la Carte 8 le sont aussi sur la Carte 6, mais l'inverse n'est pas vrai. C'est-à-dire que si les verbes impersonnels ont un pronom sujet, les verbes météorologiques doivent l'avoir aussi, mais les météorologiques peuvent l'avoir sans qu'il apparaisse avec les impersonnels.

Renzi et Vanelli proposent une hypothèse forte et testable à cet égard, selon laquelle seules les variétés à sujets pronominaux à toutes les personnes emploient

catégoriquement des pronoms avec des sujets argumentaux pré- et post-posés, ainsi qu'avec des verbes météorologiques, impersonnels et existentiels (1983:136-137). Nous ne pouvons pas répondre à la dernière partie de cette question, puisque les tournures existentielles du type *il y a* et *c'è* sont précisément des cas où la prudence nous empêche de nous aventurer à une segmentation tranchante entre pronom sujet et clitique locatif (voir 4.1.1.). Pour ce qui est du reste du «bloc» postulé par Renzi et Vanelli, il est vrai que la présence du pronom sujet avec les impersonnels l'implique chez les météorologiques. Mais nos résultats suggèrent qu'il n'y a pas de relation nécessaire entre la présence des pronoms sujet avec ces deux types de verbes et le nombre de pronoms sujet qui existent dans une variété. En effet, nous trouvons des sujets pronominaux avec les météorologiques et les impersonnels dans des variétés qui n'ont pas un système à six pronoms, mais plutôt un système comportant entre deux et cinq pronoms, ainsi que des variétés à six pronoms qui n'ont pas de pronom sujet pour les verbes impersonnels (voir la Carte 15). Qui plus est, il existe des variétés qui emploient un pronom sujet avec les météorologiques sans s'en servir de façon régulière avec les autres verbes sans sujet argumental à la troisième personne (voir la Carte 4). Il ne peut pas y avoir donc de relation d'implication forte entre le type de paradigme pronominal et l'emploi du pronom sujet avec les verbes impersonnels et météorologiques.

La Carte 7 montre une autre relation d'implication qui ne figure pas, à notre connaissance, dans les travaux antérieurs: presque tous les points où on voit le pronom générique comme sujet des verbes météorologiques l'emploient aussi dans d'autres

contextes génériques, sans que l'inverse soit vrai. Une variété doit nécessairement avoir le pronom générique comme sujet des verbes «normaux» si elle va s'en servir comme sujet des verbes météorologiques: l'emploi de ce pronom irait donc du cas général au cas plus spécifique. Le fait que cette distribution inclusive se reproduise dans deux régions séparées de notre étude (centre-ouest de la partie française, est de la partie suisse) souligne que cette relation d'implication est réelle et non pas casuelle.

5.3. CONCLUSIONS VARIABLES

Il convient de rappeler ici que les résultats obtenus quant à la variation sont contraints en grande partie par les techniques statistiques que nous employons, qui sont assez élémentaires dans l'ensemble. Des analyses plus sophistiquées pourront sans doute faire ressortir des inférences beaucoup plus nuancées et des corrélations sous-jacentes que nos critères assez grossiers de «catégoricité» et «quasi-catégoricité» ne permettent pas de capter. Ceci dit, si les tendances centrales qui se dégagent de la grande variété typologique des systèmes pronominaux peuvent ressortir même dans notre analyse statistique rudimentaire, on peut être sûr qu'elles représentent des tendances réelles et non pas des effets purement statistiques. De même pour la variation: nous perdons sans doute les corrélations et les interactions plus subtiles, mais nous croyons cerner l'essentiel de la variation dans les systèmes des pronoms sujet. Il faut quand même bien distinguer les cas de variation inhérente de ceux qui sont conditionnés par un des facteurs que nous retenons

comme pertinents. Nous examinons d'abord la variation inhérente, pour l'exclure ensuite de notre discussion des cas de variation conditionnée.

5.3.1. Variabilité inhérente

Pour la dialectologie traditionnelle, la variabilité concerne surtout les divergences dans les transcriptions phonétiques. Scheuermeier, par exemple, partage l'avis de Gilliéron à propos des enquêteurs multiples: «Envoyez vingt dialectologues dans le même endroit et vous aurez vingt résultats différents.» (chez Pop 1950:597) Il dit qu'on constatera de telles divergences «immanquablement quand plusieurs dialectologues -- et même des plus expérimentés -- transcriront ensemble les mots qui sortiront de la bouche du même sujet» (1932:108). Il paraît que pour la morphosyntaxe, il suffit en fait de beaucoup moins de vingt dialectologues pour obtenir des réponses multiples: en l'occurrence, il suffit d'une seule personne pour chaque localité. Notre corpus, recueilli par un seul enquêteur pour chaque pays, dont un non spécialiste (Edmont) et un philologue entraîné en phonétique (Scheuermeier), fait preuve que la variabilité est la norme plutôt que l'exception. Et ceci dans un domaine -- la morphosyntaxe -- où il est peu probable que l'autosuggestion de la part des enquêteurs ait faussé les transcriptions. Il faut le reconnaître: ni le critère de l'enquêteur unique, ni le critère de l'enquêteur non spécialiste ne garantit une réponse homogène et invariable, car les faits sont presque toujours hétérogènes et variables. Comme Gauchat (1902:178-179) le remarque, l'erreur de transcription est seulement une

des sources des divergences entre les données recueillies de la même localité: la variation contextuelle et individuelle y est aussi pour beaucoup.

Quelques exemples précis suffiront pour indiquer en quoi cette variabilité échappe à toute rigueur analytique. Dans *AIS943* (LA NOSTRA CUCINA) È TROPPO STRETTA. . . , au point 467 on trouve [ɛ trɛpa štrét] sans pronom sujet à côté de [l è trɔpa štrét] avec pronom sujet. Dans *AIS900* SI È NASCOSTO. . . on trouve au point 363 la forme [sé ga škónto] sans pronom et aussi [ɛl sé ga škónto] avec pronom. Dans ce dernier cas, le point en question peut être considéré comme faisant partie d'une zone de transition, puisqu'il se trouve entre des points qui utilisent le pronom sujet dans cette même carte et d'autres qui ne l'utilisent pas. Mais dans notre premier exemple, tous les points adjacents au point variable emploient le pronom sujet, et la variation n'a pas d'autre motivation que la variabilité inhérente. Que ces deux exemples soient tirés de l'*AIS* n'est pas un hasard, puisque cet atlas conserve soigneusement toutes les hésitations et les répétitions produites lors de l'entrevue (voir 4.1.2.). Mais ceci ne veut pas dire que l'*ALF* soit complètement libre de ce genre de variation: dans une carte comme *ALF87* IL A (MAL À LA TÊTE), IL A (LES SOURCILS ÉPAIS), on observe que les deux formes, bien qu'identiques pour la plupart, diffèrent à deux points. Ainsi on trouve au point 611 [ó] sans pronom et [ol ò] avec pronom, et au point 978 on trouve [à] sans pronom et [y à] avec pronom. Bien que dans ces deux cas le pronom soit présent dans la deuxième phrase plutôt que dans la première, nous ne croyons pas que l'on puisse y voir une différence de sens qui motiverait cette alternance. De même, dans l'équivalent italien de cette carte, *AIS123* HA (LE SPALLE

LARGHE), HA (LE GAMBE TORTE), on trouve également trois points (109, 190 et 216) où le pronom sujet est tantôt présent, tantôt absent. Dans ces cas où les conditions morphosyntaxiques sont à toutes fins utiles identiques, il faut tout simplement reconnaître que les faits sont plus variables que ne voudrait la description formelle. Cette variabilité est la raison principale pour laquelle les résultats tirés de ce corpus sont moins catégoriques que les descriptions des mêmes variétés réalisées à partir de jugements intuitifs.

5.4. VARIATION CONDITIONNÉE

Si nous considérons maintenant les cas où les conditions morphosyntaxiques ne sont pas identiques, nous pouvons identifier des facteurs qui favorisent ou défavorisent l'emploi du pronom sujet. Nous reprenons par la suite chacun des résultats du chapitre précédent (voir 4.4.) pour en considérer des analyses possibles.

5.4.1. Position du verbe

La position initiale de propositions subordonnées est la plus favorable aux sujets pronominaux. (voir 4.4.1.)

Il est un peu surprenant de trouver plus de sujets pronominaux précisément à l'initiale des subordonnées: à première vue, le sujet d'une subordonnée semblerait avoir moins «besoin» d'être spécifié morphologiquement, puisqu'il est souvent déjà exprimé dans la principale ou ailleurs dans le discours. Comme notre système de codage n'inclut

pas un facteur du type «information nouvelle / non nouvelle» qui correspondrait aux changements de référence («switch reference», voir 3.4.1), nous ne pouvons pas savoir si ce genre de contrainte pragmatique est active ici.⁷⁰ Le problème de savoir comment expliquer ce décalage est atténué mais non pas éliminé par le fait que la différence n'est pas très marquée quantitativement. À part l'absence apparente de motivation fonctionnelle, ce résultat surprend aussi puisqu'il contredit ceux de Nadasdi (1995:121) et de Sankoff (1982) pour le redoublement du sujet: ces études trouvent plus de sujets redoublés dans les principales que dans les subordonnées.⁷¹ Bien qu'apparemment très similaire au redoublement du sujet, l'emploi des sujets pronominaux dans notre corpus agit de la façon contraire, étant plus fréquent dans les subordonnées. Ainsi par exemple dans *AIS1645 SENTO UN RUMORE*, au point 252, on trouve [sénti um frekás] sans pronom sujet, mais aussi [k a sénti um frékas] où le pronom sujet est exprimé dans la variante subordonnée, après le complémenteur [k].

70. Comme remarqué plus haut (voir 3.4.1.), il est difficile de trouver dans un corpus d'atlas de véritables changements de référence («switch reference»), puisque les traductions des phrases des questionnaires n'ont pas les connections naturelles du discours réel. Ceci dit, nous croyons que l'effet pragmatique des changements de référence se fait sentir surtout dans des cas comme l'espagnol (voir Cameron 1993), où l'emploi des pronoms sujet est suffisamment rare pour servir comme un recours stylistique marqué. Quand l'emploi des pronoms sujet s'approche à la quasi-catégoricité, comme dans les cas qui nous intéressent, il est difficile d'y voir un effet stylistique ou pragmatique: le pronom sujet devient alors trop commun pour attirer l'attention ou «focus».

71. Selon Nadasdi, ce serait une manifestation quantitative du «Penthouse principle» de Ross (1973), d'après lequel il se passe plus de choses plus souvent dans les principales que dans les subordonnées.

Ce résultat représente peut-être un parallèle avec l'évolution diachronique de l'ancien français au moyen français. Selon Foulet (1919), les sujets nuls en ancien français apparaîtraient surtout dans les principales, et les sujets pronominaux préverbaux dans les subordonnées. Pour Franzén (1939) et Price (1966), la différence serait quantitative: les sujets non nuls en ancien français sont plus fréquents dans les subordonnées.⁷² Les propositions subordonnées constitueraient donc un des points de départ à partir desquels les sujets préverbaux se répandent à travers la syntaxe du français pour devenir obligatoires. Le parallèle diachronique est assez évocateur: dans nos données, les variétés où les pronoms sujet sont plus communs (sans être catégoriques) dans les subordonnées que dans les principales pourraient fonctionner donc comme des étapes intermédiaires dans l'évolution de l'ancien français au moyen français. Mais en attendant des études plus détaillées des variétés en question et des comparaisons systématiques de ces états de langue, ce parallèle diachronique restera une spéculation intéressante.

5.4.2. Temps et mode

Les futurs sont les temps les plus favorables aux sujets pronominaux. L'imparfait et le passé composé sont plus favorables aux sujets pronominaux que le présent de l'indicatif, et les subjunctifs sont moins favorables. (voir 4.4.2.)

72. Ces analyses sont reformulées et remises en question dans le cadre du gouvernement et liage, notamment par Adams (1987, 1988), et Vance (1988). Voir Dufresne (1993) pour un survol de ces développements.

Parmi ces résultats, seule la différence entre l'imparfait et le présent peut s'expliquer en termes fonctionnels: l'imparfait et le conditionnel ont en général des terminaisons plus syncrétiques et auraient donc davantage besoin de l'«appui» pronominal pour exprimer la catégorie morphologique de la personne grammaticale. Mais à peu près le même degré de syncrétisme se trouve chez les subjunctifs (Hochsberg 1986:613), qui ont un taux d'emploi des pronoms sujet plus faible que tous les autres temps, et ce en dépit du fait que les subjunctifs se trouvent nécessairement en subordonnées, la position qui favorise en général l'emploi du pronom sujet. Le passé composé, dont le taux d'emploi des pronoms sujet ressemble approximativement à celui des imparfaits, a autant (sinon plus) de différenciation que le présent parmi ses terminaisons morphologiques (mais voir la section suivante sur les auxiliaires). Et les temps les plus favorables aux sujets pronominaux, les deux futurs, distinguent les personnes grammaticales au moins aussi bien que le présent. Il est donc clair que le temps des verbes a une influence sur l'emploi des sujets, mais cette influence ne peut pas s'expliquer comme une motivation fonctionnelle en termes de «richesse» morphologique, puisque celle-ci prédirait exactement le contraire de ce que l'on trouve dans ces résultats.

Il est toutefois possible que ces résultats soient aussi des artefacts de l'analyse statistique ou des effets cachés d'autres facteurs, possibilité que seule une analyse par règle variable permettra d'écarter.

5.4.3. Type de verbe et auxiliaires

Le verbe le plus favorable aux sujets pronominaux est «avoir» verbe plein, suivi par les verbes réfléchis, les modaux et «être» plein, et ensuite par les deux auxiliaires.

Les inaccusatifs ont un effet légèrement défavorisant par rapport aux transitifs et aux inergatifs. (voir 4.4.3.)

De ces résultats nous pouvons dégager la généralisation que les verbes à faible contenu sémantique -- les modaux, «être» et «avoir» -- semblent favoriser l'emploi des pronoms sujet. Ceci est peut-être dû au fait que ces verbes sont eux-mêmes partiellement grammaticalisés, et se combinent donc plus facilement avec d'autres éléments en voie de grammaticalisation, comme les pronoms sujet. L'emploi élevé du pronom sujet avec les réfléchis reste plutôt mystérieux, pour deux raisons: d'abord, le pronom réfléchi lui-même identifie partiellement le sujet avec lequel il s'accorde, et ensuite, on pourrait s'attendre à ce que la présence d'un pronom quelconque devant le verbe ait tendance à «bloquer» l'emploi du pronom sujet (voir 5.4.7.). Il n'est pas clair non plus pourquoi l'emploi du pronom sujet avec les auxiliaires serait moins élevé qu'avec les verbes pleins correspondants.

En termes syntaxiques, on peut rapprocher ce résultat aux travaux de Emonds (1978), Pollock (1989) et Cummins (1996), qui postulent que certains types de verbes montent à une tête fonctionnelle, INFL, qui serait peut-être plus proche de la position des pronoms sujets. Nous laissons aux recherches ultérieures la tâche de déterminer dans

quelle mesure cette hypothèse théorique est vraiment compatible avec les tendances que nous identifions ici, surtout parce que la classe de verbes qui semblent favoriser l'emploi du pronom sujet est assez hétérogène, et donc difficilement réductible à une seule catégorie qui «monterait» dans la représentation syntaxique.

5.4.4. Genre

Le genre masculin défavorise l'emploi des sujets pronominaux, par rapport au genre féminin. (voir 4.4.4.)

Il semble donc que le masculin, en tant que cas par défaut ou «non marqué», a moins besoin de l'information morphologique apporté par le pronom que le féminin. Ce contraste est particulièrement clair dans la carte *AIS76 SONO MORTI*, où on trouve au point 261,2 que le pronom sujet est absent au masculin pluriel [im mórt] mais une note en marge indique qu'au féminin singulier le pronom sujet est présent: [l e mórtá].

5.4.5. Négation

La négation a un effet défavorisant sur l'emploi des sujets pronominaux uniquement quand elle inclut un élément préverbal. (5.4.5.)

De ces résultats nous pouvons dégager que les deux types de négation qui s'expriment par des particules négatives préverbaux ont un effet défavorisant sur les sujets

pronominaux, tandis que la négation postverbale n'a aucun effet particulier. Ce résultat appuie donc l'hypothèse selon laquelle la présence d'un autre élément dans l'espace préverbal peut «bloquer» les sujets pronominaux (voir 5.4.7.).

5.4.6. Interrogation

L'interrogation, qu'elle soit partielle ou totale, favorise le pronom sujet, par rapport aux phrases sans interrogation. (voir 4.4.6.)

Ce résultat rejoint en partie l'effet des subordonnées: dans les interrogatives partielles, l'élément interrogatif (QU-) précède la proposition dans laquelle se trouve le verbe et remplit une position sous COMP, tout comme le complémenteur dans les subordonnées. Pour ce qui est des interrogatives totales, le fait qu'elles favorisent l'emploi du pronom sujet confirme la généralisation de Renzi et Vanelli selon laquelle le nombre de personnes avec un pronom dans les interrogatives inverties doit être égal ou supérieur au nombre de personnes avec pronom dans les phrases déclaratives (1983:139). Une combinaison de ces deux effets serait vraisemblablement responsable de l'effet favorisant des interrogatives sur l'emploi des pronoms sujet.

5.4.7. Clitiques obliques

Plus il y a d'objets clitiques non réfléchis, moins on tend à employer les sujets pronominaux. (voir 4.4.7.)

Les clitiques objets non réfléchis, comme la négation préverbale, défavorisent l'emploi du pronom sujet. Ce résultat suggère qu'il existe en fait un effet de «blocage» selon lequel un élément atone en proclise -- un pronom objet, une particule de négation -- «occupe» en quelque sorte la place destinée au pronom sujet, et cette «occupation» empêche la réalisation de ce dernier.⁷³ Cette façon de concevoir la réalisation des éléments proclitiques, comme en concurrence pour un nombre limité de positions, implique bien sûr que la structure syntaxique n'y est pour rien: ce qui compte, c'est l'espace linéaire et non pas les relations hiérarchiques. Plutôt que des têtes fonctionnelles auxquelles des éléments pourraient se déplacer ou non, ce seraient des contraintes linéaires sur la longueur maximale des suites proclitiques qui déterminent le nombre d'éléments préverbaux permis dans une variété déterminée. Ces contraintes pourraient évidemment varier selon la grammaire: dans certaines variétés elle laisserait passer des suites d'un élément proclitique seulement, dans d'autres deux, et ainsi de suite, jusqu'au cas où elle permettrait des suites de n'importe quelle longueur (soit que la contrainte ne serait plus active).

Des contraintes sur la longueur des suites de clitiques semblent en fait être assez communes. Perlmutter (1970:50, n. 2) note que certaines variétés de l'espagnol n'admettent qu'un ou deux clitiques, tandis que d'autres tolèrent aussi des suites de trois;

73. Les clitiques réfléchis ne semblent pas compter dans cet effet de blocage, puisqu'ils ne défavorisent pas les pronoms sujet comme les autres clitiques obliques. Les réfléchis seraient alors «invisibles» en ce qui concerne les contraintes sur la longueur des suites proclitiques, ce qui appuie l'idée selon laquelle les réfléchis font partie de l'entrée lexicale des verbes et comme tels ne comptent pas dans certains processus morphosyntaxiques (Nadasdi 1995:96).

Bastida (1974, 1976) cite des cas où des suites avec quatre clitiques objet peuvent être acceptables dans certaines variétés de l'espagnol (notamment chez des bilingues hispano-galicieus). Le catalan barcelonais dépasse les autres variétés romanes, avec des suites de jusqu'à six clitiques objet (Bonet 1991), mais ces dernières ne sont pas attestées dans d'autres variétés du catalan (par exemple en valencien, voir Todolí 1995). A propos de l'italien standard, Wanner (1987:59, note 5) dit qu' «In practice, more than three clitics in a row are very rare or cannot be processed.» Sur l'incompatibilité de ce genre de contrainte arbitraire sur la longueur avec une analyse syntaxique de la cliticisation, Heap et al. (1993) notent que les exemples de ce type de variation

... represent language- or dialect- [sic] specific properties of clitics. It must be noted, however, that these properties refer only to the realization of clitics, and not to the argument positions which they represent. It is for this reason that a syntactic analysis, in which the syntactic and morphological properties of a clitic are inseparable, fails to account for the data straightforwardly. (1993:118)

L'existence de ces contraintes sur la longueur des suites clitiques constitue donc une preuve qui appuie l'hypothèse d'une «interface» morphologique entre la syntaxe et la forme phonologique (voir 5.5.2.).

5.4.8. Type et position du sujet

La position préverbale défavorise fortement l'emploi des sujets pronominaux. Les sujets argumentaux excluent l'emploi du pronom sujet dans les variétés gallo-romanes. (voir 4.4.8.)

Ce résultat est peu surprenant: les sujets en position préverbale occupent la même position syntaxique que les sujets pronominaux, et on s'attend donc à ce qu'ils se trouvent en distribution complémentaire, alors que les sujets disloqués ou invertis laissent cette position libre pour un pronom qui remplirait le rôle de véritable sujet syntaxique. Ce qui est à expliquer alors c'est la cooccurrence des sujets argumentaux et pronoms sujet dans les variétés italiennes, alors que cette combinaison est presque inconnue parmi les variétés gallo-romanes (comparer les Cartes 4 et 5, ainsi que 13 et 14). Si l'on admet que les variétés italiennes évoluent plutôt vers un traitement morphologique des sujets pronominaux (voir 5.1.), il serait alors normal qu'elles tolèrent la cooccurrence des pronoms avec les sujets nominaux. Les variétés gallo-romanes en revanche opteraient d'après cette analyse pour un traitement syntaxique des pronom sujets, ce qui exclut cette cooccurrence.

5.5. IMPLICATIONS THÉORIQUES

Ayant vu les différentes analyses que suggèrent les résultats de cette étude, nous tentons de résumer ici l'essentiel de ces conclusions en considérant leurs implications pour

la variation grammaticale dans le contexte d'un continuum morphosyntaxique. Avouer que certains des faits échappent à l'analyse, c'est se rendre à l'évidence: aucun cadre théorique formel, ni générativiste ni fonctionnaliste, n'est à même de capter tous les faits de langage, surtout quand nous prenons la variation linguistique au sérieux. Le mieux que nous pouvons faire, c'est présenter la synthèse de ces faits de variation que nos outils théoriques nous permettent de capter, et espérer que ce bilan contribuera quelque chose à notre compréhension des phénomènes à l'étude, et suggérera peut-être des pistes d'analyses à suivre dans des recherches ultérieures.

5.5.1. Paramètre ou paradigme?

Tout d'abord, il doit être très clair que l'idée du «paramètre du sujet nul» n'est plus viable du tout. Il ne s'agit pas là d'une conclusion originale -- on annonce sa mort depuis une décennie déjà -- mais il vaut la peine de se convaincre de ce fait. D'autres études le démontrent mieux que la nôtre ne peut le faire: le paramètre au sens large, comme corrélation de propriétés apparemment indépendantes, ne peut pas tenir debout (Safir 1985; Roberge 1989; Wanner 1993). Il n'est plus question ici que du paramètre dans le sens le plus étroitement immédiat, comme outil descriptif permettant d'opposer deux classes de grammaires naturelles. Or même à ce niveau descriptif, le «paramètre du sujet nul» n'a qu'une utilité très limitée: en plus des réglages binaires [\pm sujet nul], il faudrait au moins une bonne douzaine de réglages intermédiaires correspondant aux paradigmes partiels dans notre typologie (voir 4.3. et 5.2.). Appeler «paramètre» une dimension

grammaticale qui admet toutes ces valeurs scalaires, c'est garder la terminologie d'un modèle sans y attribuer un contenu réel.⁷⁴

Nous ne voudrions pas cependant enterrer le paramètre du sujet nul sans proposer quelques idées sur la question de savoir comment la transition entre les deux pôles de cette dimension scalaire s'agence. En rejetant le paramètre du sujet nul, nous renonçons aussi à l'idée qu'un principe syntaxique quelconque puisse rendre compte de cette variation. Le principe syntaxique pertinent -- la récupération du sujet -- reste invariant à travers ce continuum, que cette récupération se fasse par le biais de la morphologie verbale ou par les pronoms ou par les deux. Ce qui varie, nous semble-t-il, ce sont les paradigmes d'éléments morphologiques disponibles pour cette récupération.⁷⁵ Or ce concept de «paradigme», bien qu'il soit tout à fait nécessaire et intuitivement utile, ne reçoit pas de définition formelle dans la linguistique actuelle.

La tâche préalable est donc celle de bien cerner la nature des paradigmes dans un modèle de la morphologie. Quand nous saurons quel statut théorique attribuer à cet objet linguistique qui est un paradigme pronominal, nous serons en mesure d'explorer la

74. Il convient cependant de s'interroger sur le fait que l'illusion du paramètre binaire soit beaucoup plus convaincante si l'on ne considère que la partie gallo-romane de nos résultats: les relations entre la variation grammaticale et les conditions sociolinguistiques restent une question ouverte.

75. Nous sommes en désaccord ici avec Wanner (1995:8), pour qui la variation grammaticale ne touche que les aspects syntagmatiques d'une grammaire (accord, lexique, paramètres) et non pas les aspects systématiques et paradigmatiques (au sens saussurien). Pour nous les sujets pronominaux appartiennent à l'axe paradigmatique, et en même temps sont impliqués pleinement dans la variation grammaticale.

question de savoir pourquoi les paradigmes varient de grammaire en grammaire, et quelles contraintes s'appliquent à cette variation. Il est clair que des contraintes sont en jeu: la distribution des personnes dans les paradigmes partiels est loin d'être aléatoire. Si notre typologie des paradigmes partiels peut se réduire à deux contraintes qui rendent compte de 97% des types attestés, il faut s'interroger sur la meilleure façon de représenter ce type de contrainte dans un modèle de grammaire humaine.

5.5.2. Variation morphologique aux interfaces

S'il est vrai que nous avons affaire ici à des contraintes sur les paradigmes morphologiques, il reste beaucoup à faire en matière de théorie morphologique pour que nous puissions exprimer formellement ces généralisations descriptives. Il est clair cependant qu'un module morphologique qui permette l'interaction des contraintes dans une hiérarchie paraît beaucoup plus prometteur qu'un modèle catégoriquement dérivationnel. Pour l'instant, la Théorie de l'Optimalité (McCarthy et Prince 1993; Prince et Smolensky 1993) constitue le cadre le plus développé dans ce sens, mais elle s'applique surtout en phonologie. Il n'est pas clair qu'elle puisse s'adapter à ce genre de problème morphologique: ce qui nous semble en revanche très clair, c'est qu'un cadre qui admet une certaine flexibilité et des réalisations multiples et variables répondra mieux au genre de faits que nous documentons ici.

L'analyse de nos résultats nous pousse à nous méfier donc des explications syntaxiques de ce type de variation. Qu'il s'agisse des contraintes sur la longueur des

suites proclitiques et l'effet du blocage, ou de l'interaction des différentes contraintes sur la composition des paradigmes des pronoms sujet, nous avons affaire à des objets qui n'ont aucun lien direct avec les relations hiérarchiques caractéristiques de la syntaxe. Les contraintes sur l'espace proclitique font appel uniquement à la concaténation linéaire, un processus typique de la morphologie; même s'ils sont mal définis en linguistique formelle, les paradigmes doivent nécessairement se constituer à partir de traits morphologiques.

En excluant la variation de la syntaxe, nous rejoignons un des aspects du modèle «minimaliste» de Chomsky: pour rendre la composante computationnelle (soit la syntaxe) la plus universelle possible, Chomsky propose que toute la variation intergrammaticale serait d'ordre lexical (1995:169). Mais nous ne pouvons pas admettre que les paradigmes soient de simples items lexicaux, des «mots» comme tous les autres. La nature des paradigmes c'est qu'ils n'ont pas de contenu propre mais servent à organiser d'autres éléments dotés de contenu. Autrement dit, le fait qu'il existe un pronom sujet «tu» de la deuxième personne doit être inscrit dans le lexique comme tout autre fait idiosyncratique; le fait que seulement «tu» (ou seulement «tu» et «elles», et ainsi de suite) doit s'employer (quasi-)catégoriquement, doit être inscrit dans une autre partie de la grammaire, nous semble-t-il.

La position naturelle pour faire intervenir les paradigmes dans un modèle générativiste, à notre avis, est justement dans l'interface entre la syntaxe et la forme phonologique. C'est ici qu'un vocabulaire d'items morphologiques serait mis en jeu pour épeler les noeuds terminaux de la syntaxe: les paradigmes constitueraient en quelque sorte

le système d'emmagasinage et de classement des ces items morphologiques. Toutes les grammaires auront donc un ou des paradigmes de pronoms sujets, mais les caractéristiques de ces paradigmes varieront: le nombre de pronoms, leur nature obligatoire ou facultative, et leur comportement quant aux diverses contraintes qui déterminent leur emploi.

Le choix de situer la variation au niveau de l'interface morphologique entraîne des implications intéressantes pour l'histoire des pronoms sujet, même si nous savons que les extrapolations de la synchronie à la diachronie sont périlleuses. Comment passe-t-on (que ce soit à travers l'espace ou le temps) d'une grammaire à pronoms sujet facultatifs à une grammaire où leur emploi est obligatoire? Le début et la fin du processus sont fortement conditionnés par la typologie: les paradigmes à un seul pronom et à cinq pronoms sont ceux qui se conforment le mieux à nos Généralisations 1 et 2. Vers le milieu du processus, les paradigmes possibles se multiplient, les généralisations typologiques sont violées plus fréquemment, et d'autres contraintes deviennent actives. De ces contraintes, nous identifions avec certitude le blocage par les éléments proclitiques, l'effet des subordonnées, ainsi que ceux des temps et modes et des types de verbes. À cette liste pourront s'ajouter sans aucun doute d'autres contraintes -- morphologiques, syntaxiques, phonologiques, même pragmatiques -- que nous ne pouvons pas encore identifier. Le changement en question n'est apparemment «syntaxique» que si l'on en considère le début et la fin: quand on regarde de près les étapes intermédiaires, tous les sous-systèmes de la grammaire sont impliqués. Comme le remarque Anderson:

. . . the solution to the problem of accounting for what is found in the morphology of natural languages does not lie entirely within the synchronic theory of morphology itself. . . . If the most nearly 'explanatory' theory of the content of synchronic morphologies is thus to be found in their inter-relation with diachronic processes, the study of morphological change acquires an importance well beyond its interest as an object in itself. (1988: 325-6)

L'interface morphologique est donc en même temps le nexus des autres modules d'une grammaire et locus de variation, qui est à son tour une condition nécessaire du changement diachronique. Et de manière dialectique, le changement diachronique finit par transformer la morphologie même. En ce qui concerne l'évolution vers les pronoms sujet, le phénomène commence et finit par des manifestations qui sont «visibles» et actives en syntaxe, mais l'implantation même du changement implique la morphologie et ses relations avec les autres composantes d'une grammaire.

Appendice I: Les cartes de l'ALF.⁷⁶

N° de Carte	Personne Grammaticale	Lemme française	=AIS	Contexte immédiat	Volume
10	6	ils s'agenouilleraient	(949)	... devant lui.	1
23	1	je vais	822	... acheter deux chevaux à la foire.	1
25	2	où vas-tu?	821		1
26	3	va	519	Le chasseur ... à la chasse	1
32	6	sont allés	(522)	Les garçons... dénicher des nids d'oiseaux.	1
34	1	j'allume	760	... ma pipe.	1
83.1	1	...que j'ai achetés...	(1203)	et vendre les deux...	1
83.2	1	j'en ai	913	... plein la tête.	1
83.3	1	je l'ai	1107	... déjà entendu.	1
84	2	tu as	1649	... oublié que vous deviez...	1
85	2	l'as-tu?	770	lu? (le journal)	1
86	2	as-tu?	1145	Quel âge...	1
87.1	3	il a	123	... mal à la tête.	1
87.2	3	il a	123	... les sourcils épais.	1
91	4	nous avons	1635	... pris une purge.	1
92.1	5	vous avez	(1599)	... là un beau chien.	2

76. Les numéros de cartes entre parenthèses dans la quatrième colonne correspondent à des équivalents approximatifs plutôt qu'à des traductions exactes (voir 3.3.1.).

N° de Carte	Personne Grammaticale	Lemme française	=AIS	Contexte immédiat	Volume
92.2	5	vous avez	1599	... gagné quelquechose	2
94	2	tu avais	1017	... raison; j'ai...	2
99	5	vous auriez	1250	... dû voir comme les arbres en étaient chargés.	2
102	1	j'ai eu	(83)	... de la peine...	2
136	3	le blé est	(922)	... mûr, mais l'avoine...	2
311	6	commencent	1261	Les pommiers ... à fleurir.	3
343	3	est couvert	(1195)	Le ciel ... de nuages.	3
358	2	crois-tu	1650,1	... qu'elle tienne?	3
359	1	que je croyais	1672	Il me serrait si fort ... qu'il m'étranglerait.	3
383	6	ont déjà	1264	Les cerisiers ... commencé à fleurir.	3
400	5	que vous deviez	1590	Tu as oublié ... nous faire signe...	3
406	2	si tu me dis	1628	Tant pis pour toi ... des mensonges.	3
408	5	vous dites	(1595)	tantôt ... oui, tantôt...	3
439	3	il fait des éclairs	391		3
465	1	j'entends	465	Attention! ... du bruit.	3
469	1	je l'enverrai	11	Quand mon fils sera grand, ... à Paris.	4
470	1	ou que j'envoie	1639	Voulez-vous que j'aïlle ... quelqu'un?	4
498	3	qu'il m'étranglerait	1672	Il me serrait si fort que je croyais ...	4

N° de Carte	Personne Grammaticale	Lemme française	=AIS	Contexte immédiat	Volume
500	1	je me suis	694	... assis sous un arbre.	4
501	2	que tu es	53	Tu ne vois donc pas ... aussi vieux que moi.	4
502	3	il est bon	(407)	Celui-ci ... mais il ne vaut pas le mien.	4
507	5	vous êtes	1598	... venu sans rien; ... perdu, mon vieil ami.	4
509	6	ils sont	509	... morts.	4
511	3	si c'était bien cuit	1018	J'en mangerais...	4
513	6	en étaient	1252	Comme les arbres ... chargés (de fruits).	4
514	3	sera grand	10	Quand mon fils ..., je l'enverrai à Paris	4
515	4	et nous serions	(1613)	Tu aurais dû te taire, ... loin.	4
516	5	quand vous seriez	1113	Tu as oublié que vous deviez nous faire signe ... en haut.	4
522	4	nous avons été	1248	Quand nous étions jeunes, ... riches et heureux.	4
531	3	il fait	(363)	... si chaud!	4
532	6	ils feront	1661	... ce qu'ils voudront.	4
534	3	il faut	351	... savoir; ... que nous ayons; ... qu'il vienne; ... que vous finissiez.	4
536	3	il fallait	(351)	... le laisser où il était.	4
537	3	il faudrait	(351)	... être aveugle ou fou pour...	4
563	6	ils sont fiancés	(146)		4

N° de Carte	Personne Grammaticale	Lemme française	=AIS	Contexte immédiat	Volume
578	6	ont fini	1261	Les poules ... de pondre.	4
632	3	il gèle	382		4
654	3	quand elle est gonflée	(943)	Il faut savoir bien nager pour passer outre ... (la rivière)	4
667	3	il gèle	372		5
729	3	il y a huit jours	1265	Les cénistes ont commencé à fleurir ...	5
764	5	si vous vous levez	1590	Je me lève...	5
774	3	il lit	768	... le journal.	5
785	4	et que nous lui	(1641)	Il faut qu'elle vienne ... rendions son argent.	5
786	3	il lui a	65	... donné un baiser (un sous).	5
803	1	je suis malade	700	... plus que toi.	5
806	4	si nous ne mangeons pas	1278	... nos prunes, elles...	5
807	1	j'en mangerais	1019	Si c'était bien cuit, ... bien.	5
817	5	pourquoi vous ne mariez-vous pas?	69		5
847	1	j'ai mis	887	... un verre à la porte.	6
869	6	elles se moisiront	1280		6
904	3	il neige	377		6
1035	3	il pleut	(366)		7
1036	3	il a plu	367	, surtout quand ...	7

N° de Carte	Personne Grammaticale	Lemme française	=AIS	Contexte immédiat	Volume
1064	6	portent	592	Ils ... le deuil.	7
1064.2	6	portent	592	Les chênes... des glands.	7
1084	1	je ... pouvais	1669	... ne ... ni avancer ni reculer.	7
1090	1	j'ai pris	512	... un moineau au piège.	7
1095	3	est profonde	(209)	A cet endroit, la rivière ..., surtout...	7
1103	2	puisque tu as	1015	Mange, ... faim.	7
1144	1	je le remplis	(1601)		7
1145	2	tu remplis	(1680)		7
1154	4	nous ne le revîmes plus	(1641)	Il partit au bout d'une semaine, ...	7
1203	1	j'ai su ça	(390)	Tu avais raison, ... plus tard.	7
1223	3	il me serrait	1671	... si fort que ...	8
1264	1	Je sue	(177)	Quelle chaleur! ... à grosses gouttes.	8
1291	3	quel temps fait-il?	363		8
1312	3	m'est tombée	1617	Une branche pourrie ... sur la figure.	8
1315	3	il tonne	396		8
1341	5	vous trouverez bien	(1633)	... quelqu'un qui vous ira.	8
1362	2	si tu viens	359	Je pars si ... avec moi.	8
1409	2	Tu ne vois donc pas	52	que tu es aussi vieux que moi.	9
1415	1	je veux	1086	l'attacher au poteau, qu'il le reçoive.	9

N° de Carte	Personne Grammaticale	Lemme française	=AIS	Contexte immédiat	Volume
1417	5	Voulez-vous	1638	que j'aïlle ou que j'envoie quelqu'un?	9
1419	1	Je voudrais	1519	... bien avoir de celle-ci.	9
1433	6	qu'ils aillent	1144	Il faut ...	9
1448	1	(que) j'aie	1603	Il faut ...	9
1470	3	ça bout	953		9
1519	3	elle s'est couchée	(922)		9
1545	3	est-elle déjà	41	... baptisée	9
1559	4	nous sommes	(1646)	... tous ici.	9
1662	3	elle s'est pendue	(900)		9
1678	6	elles ont perdu	1254		9
1696	4	nous nous reverrons	660		9
1712	1	j'ai sommeil	643		9
1872	3	il faut	1391	... payer	9

Appendice II: Les cartes de l' AIS (voir 3.3.1.).⁷⁷

N° de Carte	Personne grammaticale	Lemme italienne	=ALF	Glose française	Contexte immédiat	Volume
7	3	come somiglia	310	comme (il) ressemble	... sua madre!	1
10	3	sarà grande	514	... sera grand	Quando mio figlio	1
11	1	lo manderò	469	je l'enverrai	... a Torino, Milano ecc.	1
38	3	è ancora	(1095)	il est encore	... piccolo.	1
41	3	è già battezzata?	1545	est-elle déjà baptisée?	La vostra piccola bambina...	1
52	2	non vedi	1409	ne vois-tu	... che sei..	1
53	2	che sei	501	que tu es	... vecchio quanto me?	1
65	3	lui l'ama molto	786	lui l'aime beaucoup		1
69	5	... non vi sposate?	817	ne vous mariez-vous pas?	Perchè...	1
74	3	è gravida	(1095)	est enceinte	La sua moglie...	1
76	6	sono morti	509	ils sont (morts)	(check subject)	1
112	2	Perchè hai	(84)	pourquoi as(-tu?)	... la guancia gonfia?	1
123.1	3	Ha	87	(Il) a...	... le spalle larghe.	1
123.2	3	Ha	87	(Il) a...	... le gambe torte.	1
146	6	sono rotte.	1702	sont cassées.	Le due braccia...	1

77. Les numéros de cartes entre parenthèses dans la quatrième colonne correspondent à des équivalents approximatifs plutôt qu'à des traductions exactes (voir 3.3.1.).

N° de Carte	Personne grammaticale	Lemme italienne	=ALF	Glose française	Contexte immédiat	Volume
209	3	è acuta	(1095)	(Le burin) est pointu.	La lesina...	2
220	3	casco	(1312)	(Le charpentier) tombe ...	Il falegname ... dal letto	2
351	3	Bisogna	534	Il faut...	... restare dentro,	2
359	2	vieni?	1362	(L'ou) viens(-tu)?	Donde...	2
363	3	che tempo fa?	1291	que temps fait(-il)?		2
366	3	piovere; piove?	(1035)	pleuvoir; pleut(-il)?		2
367	3	è piovuto	1036	il a plu		2
372	3	grandinare; grandina 3	667	grêler; il grêle 3		2
377	3	nevicare; nevicare	904	il neige; neiger		2
382	3	gelare; gela 3	632	geler; il gèle 3		2
388	1	Ho le mani	83	J'ai les mains	... intirzite.	2
390	1	Ho visto	(1203)	J'ai vu (éclairer).	... balenare.	2
391	3	balenare; balena 3	439	il fait des éclairs		2
391.1	3	ha balenato	(439)	il a fait des éclairs		2
394	3	è cascalo	(1312)	(La foudre) est tombée	Il fulmine ... sulla nostra casa.	2
396	3	tuonare; tuona 3	1315	tonner; il tonne		2
397	3	ha tonato	(1315)	il a tonné		2
407	3	è pesante	(1095)	est lourd.	L'oro...	2

N° de Carte	Personne grammaticale	Lemme italienne	=ALF	Glose française	Contexte immédiat	Volume
512	1	Ho preso	1090	J'ai pris	un uccello al laccio.	3
519	3	va a caccia	26	(Le chasseur) va à la chasse.	Il cacciatore...	3
522	3	è andato	(32)	est allé (à pecher)	... a pescare.	3
592	6	portano	1064	(les chênes) portent delle ghiande	3
643	1	ho sonno	1712	j'ai sommeil		4
660	4	Ci leviamo,	1696	nous nous levons,...		4
661	5	se voi vi levate	764	... si vous vous levez		4
694	1	Mi sono raffreddato;	500	j'ai pris froid;	... 695	4
697	1	ho la febbre	83	j'ai la fièvre		4
698	1	ho la voce rauca	83	j'ai la voix rauque		4
704	3	andava spesso	30	(elle) allait souvent	... dal medico; i medici.	4
760	1	accendo la pipa	34	j'allume la pipe		4
821	2	dove vai?	25	où vas(-tu)?		4
822	1	Vo a comprare	23	Je vais acheter	... due cavalli alla fiera,	4
887	1	Ho messo	847	J'ai mis (un verrou).	.. un chivistello.	5
887.1	1	Ho incatenacciato	1679	J'ai verrouillé (la port)	... la porta	5
900	3	Si è nascosto	(1519)	Il s'est caché	dietro l'armadio.	5
913	1	l'ho accesa	83	Je l'ai allumée	(la lanterna)	5
922	3	... è spento	(136)	est éteint	Il fuoco...	5

N° de Carte	Personne grammaticale	Lemme italienne	=ALF	Glose française	Contexte immédiat	Volume
943	3	è troppa stretta	(654)	(Notre cuisine) est trop étroite	La nostra cucina...	5
949	6	poi si asciugnano	(10)	puis on les essue ...	con un cencio.	5
953	3	bollire; bolle 3	1470	bouillir; bouillir 3		5
975	3	è fesso	(1095)	(L'assiette) est fendue.	Il tondo...	5
976	3	è intero; intera	(1095)	est entier; entière		5
1005	3	la pappa è buona	(1095)	la bouffe est bonne		5
1013	1	sono digiuno...	(1095)	je suis à jeûne...		5
1015	2	piochè hai fame!	1103	puisque tu as faim!		5
1017	2	...se avessi fame.	94	si tu avais faim.		5
1018	3	Se fosse ben colto,	511	si c' était bien culti...	... ne mangerei	5
1019	1	ne mangerei	807	j'en mangerais (volentiers).	... volentieri.	5
1086	2	voglio attaccarla	1415	je veux l'attacher		6
1107	1	L'ho ricevuto	83	Je l'ai reçu (de mon ami).	... dal mio amico.	6
1113	5	cosa ne fareste ?	516	qu'est-ce vous en feriez ?	1112 ...	6
1144	6	non vadano nel giardino.	1433	n'ailent pas dans le jardin!	1143 ...	6
1145	2	hai venduto le uova ?	86	as-tu vendu les oeufs ?		6
1195	3	la vacca è munta	(343)	la vache est traite		6
1250	5	Avreste dovuto vedere	99	Vous auriez dû voir	... come gli alberti	7
1252	6	ne erano carichi	513	en étaient chargés	come gli alberti ...	7

N° de Carte	Personne grammaticale	Lemme italienne	=ALF	Glose française	Contexte immédiat	Volume
1254	6	ne hanno abbastanza.	578	en ont beaucoup.	Loro pure ...	7
1261	6	cominciano a	311	commentent	I pomi ... fiorire.	7
1264	6	hanno già cominciato	383	ont déjà commencé	I ciliegi ...	7
1265	3	otto giorni fa	729	il y a huit jours	1264 ...	7
1278	4	Se non mangiamo	806	Si non ne mangions pas le nostre susine	7
1280	6	marriranno ben presto.	869	... moriront bientôt.	le nostre susine ...	7
1340	3	Il vino è buono.	(1095)			
1391	3	bisogna	1872	Il faut	... tagliare (il fieno).	7
1519	1	Vorrei di questa qui,	1419	Je voudrais de celle-ci, non di quella lì.	8
1537	2	hai cucito bene	(85)	tu as bien cousu		8
1588	1	(è impossibile) che io abbia detto ciò	1448	(il est impossible) que j'aie dit ça		8
1590	5	che mi doveti	400	que vous me devez!	1589 ...	8
1595	5	credere; Credete	(408)	croire; Croyez-vous	... 1596	8
1598	5	Siete venuta senza niente,	507	Vous êtes venue sans rien,	... 1599	8
1599	5	avete guadagnato qualchecosa.	92	Vous avez (gagné qq' chose)		8
1601	1	lo dico di nuovo	(1144)	Je le dis de nouveau		8
1613	4	Lavoreremo di più,	(515)	Nous travaillerons davantage,	... 1614	8
1628	2	se lo trovassi.	406	si tu le trouvais.	1627 ...	8

N° de Carte	Personne grammaticale	Lemme italienne	=ALF	Glose française	Contexte immédiat	Volume
1633	5	Voi lo trovereste (in qualche luogo)	(1341)	Vous le trouverez (quelque part)	... 1634	8
1635	4	abbiamo cercato insieme (dappertutto)	91	nous avons cherché ensemble (partout)		8
1638.1	5	Volete (che ci vada io)...	1417	Voulez-vous (que j'y aille)	... 1639	8
1639	1	(o) che ci mandi qualcheduno?	470	(ou) que j'y envoie quelqu'un?	1638 ...	8
1641	4	(mi rincresceva) che non la trovassimo	(1154)	(je regrettais) que nous ne la trouvions pas		8
1645	1	sento un rumore.	465	j'entends un bruit.		8
1646.1	4	siamo arrivati (poco fa)	(1559)	nous sommes arrivés.		8
1650	2	credi (che lo troviamo)?	358	crois-tu (que nous le trouvons)?		8
1661.1	6	faranno (ciò che voranno)	532	ils feront (ce qu'ils voudront)		8
1669	1	io non potevo andare	1084	Moi je ne pouvais marcher (ni en avant ni en arrière).	né avanti né indietro.	8
1672.1	1	credevo (che mi strozzasse).	359	je croyais qu'il m'étranglait		8
1672.2	3	(credevo) che mi strozzasse.	498			8

Appendice III: Les points d'enquête dans l'*AIS* et dans l'*ALF* (voir 3.3.2. et Carte 1).⁷⁸

Numéro de Carte	Nom de la localité	Population	2me nom de localité	Département / canton / région
1	Brigels (Breil)	1055		Grisons
3	Pitasch	101		Grisons
5	Ems (Domat)	1779		Grisons
7	Ardez	636		Grisons
9	Remûs (Ramosch)	579		Grisons
10	Tavetsch (Tuietsch)	867	Camischollas	Grisons
11	Somvix (Sumvitg)	1397	Surrhein	Grisons
13	Vrin	369		Grisons
14	Prâz	219	Dalin	Grisons
15	Mathon	65		Grisons
16	Scharans	410		Grisons
17	Lenz (Lansch)	302		Grisons
19	Zernez	760		Grisons
22	Lavorceno	(748)	Olivone	Tessin
25	Reams (Riom)	224		Grisons
27	Bergün (Bravuogn)	710	Latsch	Grisons
28	Zuoz	749		Grisons
29	Santa Maria	438		Grisons
31	Osco	287		Tessin
32	Chironico	79		Tessin
35	Bivio (Beiva)	121		Tessin
41	Cavergno	359		Tessin

78. Les chiffres entre parenthèses dans la troisième colonne correspondent à la population de la commune qui apparaît dans la quatrième colonne (le deuxième nom que fourni l'*AIS* pour certaines localités). Voir *AIS1*.

Numéro de Carte	Nom de la localité	Population	2me nom de localité	Département / canton / région
42	Sonogno	293		Tessin
44	Mesocco	1163		Grisons
45	Soglio	331		Grisons
46	Stampa	479	Coltura	Grisons
47	Sils (Segl)	295	Fex-Platta	Grisons
50	Cimalmotto	(256)	Campo	Tessin
51	Vergeletto	304		Tessin
52	Aurigeno	212		Tessin
53	Lodrino	616	Prosito	Tessin
58	Sommaino	3695	Poschiavo	Grisons
70	Indemini	306		Tessin
71	Breno	342		Tessin
73	Corticiasca	252		Tessin
93	Ligornetto	847		Tessin
107	Trasquera	1035		Piémont
109	Albogno	(1044)	Premia	Piémont
114	Borgone	84 (524)	Ceppomorelli	Piémont
115	Antronapiana	973		Piémont
116	Domodossola	7598		Piémont
117	Ornavasso	1724		Piémont
118	Malesco	700		Piémont
121	Rhêmes-Saint-Georges	449		Piémont
122	Saint-Marcel	1225		Piémont
123	Brusson	1477		Piémont
124	Riva Valdobbia	549	Selveglio	Piémont
126	Borgosesia	5788	Pianezza	Piémont

Numéro de Carte	Nom de la localité	Population	2me nom de localité	Département / canton / région
128	Nonio	801		Piémont
129	Borgomanero	11432		Piémont
131	Noasca	842		Piémont
132	Ronco Canavese	1679		Piémont
133	Vico Canavese	711		Piémont
135	Piana	276 (1735)	Pettinengo	Piémont
137	Carpignano	2470		Piémont
138	Novara	56260		Piémont
139	Galliate	9501		Piémont
140	Rochemolles	503		Piémont
142	Bruzolo	1326		Piémont
143	Ala di Stura	596		Piémont
144	Corio	5073		Piémont
146	Montanaro	3717		Piémont
147	Cavaglia	2724		Piémont
149	Desana	1898		Piémont
150	Sauze di Cesana	295		Piémont
152	Ruata	98 (1071)	Pramollo	Piémont
153	Giaveno	9850		Piémont
155.1	Torino I	502274		Piémont
155.2	Torino II	502274		Piémont
156	Castellnuovo d'Asti	3141		Piémont
157	Asti	40597		Piémont
158	Ottiglio	2435		Piémont
159	Isola Sant'Antonio	1358		Piémont
160	Maddalena	97 (847)	Pontechianale	Piémont

Numéro de Carte	Nom de la localité	Population	2me nom de localité	Département / canton / région
161	Ostana	886		Piémont
163	Pancalieri	2209		Piémont
165	Corneliano d'Alba	2284		Piémont
167	Mombaruzzo	3163		Piémont
169	Gavi	5930		Piémont
170	Pietraporzio	365		Piémont
172	Villafalletto	3797		Piémont
173	Cuneo	31221		Piémont
175	Fiamenga	461 (2911)	Vicoforte	Piémont
176	Cortemilia	3430		Piémont
177	Sassello	4050		Ligurie
178	Genova	316217		Ligurie
179	Isola	142 (1639)	Rovegno	Ligurie
181	Valdieri	2040		Piémont
182	Limone	2053		Piémont
184	Calizzano	2817		Ligurie
185	Noli	2098		Ligurie
187	Zoagli	3008		Ligurie
189	Cassana	479 (1598)	Borghetto di Vara	Ligurie
190	Airole	1123		Ligurie
193	Borgomaro	704		Ligurie
199	Castelnuovo	4477		Ligurie
205	Campodolcino	1381 (237)	Prestone	Lombardie
209	Val di Dentro	1869 (427)	Isolaccia	Lombardie
216	Vetto	212 (1037)	Lanzada	Lombardie
218	Grosio	4144		Lombardie

Numéro de Carte	Nom de la localité	Population	2me nom de localité	Département / canton / région
222	Germasino	338		Lombardie
223	Colico	3900		Lombardie
224	Colico	3900 (399)	Curcio	Lombardie
225	Mello	1496		Lombardie
227	Moia	1014 (2743)	Albosaggia	Lombardie
229	Sonico	1561		Lombardie
231	Arcumeggia	232		Lombardie
234	Introbio	788		Lombardie
236	Branzi	1032		Lombardie
237	Gromo	998		Lombardie
238	Borno	3598		Lombardie
242	Como	48066		Lombardie
243	Canzo	1943		Lombardie
244	Selino	1131		Lombardie
245	Stabello	353		Lombardie
246	Bergamo	62303		Lombardie
247	Monasteroio	706		Lombardie
248	Limone	482		Lombardie
249	Bagolino	4556		Lombardie
250	Magnago	5371 (1337)	Bienate	Lombardie
252	Monza	57060		Lombardie
254	Martinengo	5814		Lombardie
256	Brescia	100168		Lombardie
258	Mosniga	220 (2093)	Lumezzane-Sant'Apollonio	Lombardie
259	Toscolano	3047		Lombardie
261.1	Milano I	836046		Lombardie

Numéro de Carte	Nom de la localité	Population	2me nom de localité	Département / canton / région
261.2	Milano II	836046		Lombardie
263	Rivolta d'Adda	6109		Lombardie
265	Crema	11874		Lombardie
267	Dello	1628		Lombardie
270	Cozzo	1619		Lombardie
271	Vigevano	30583		Lombardie
273	Beregardo	2303		Lombardie
274	Sant'Angelo	9474		Lombardie
275	Castiglione d'Adda	4285		Lombardie
278	Solferino	1910		Lombardie
282	Montù Beccana	3913		Lombardie
284	Cremona	59056		Lombardie
285	Pescarolo	2435		Lombardie
286	Bozzolo	4222		Lombardie
288	Mantova	36633		Lombardie
289	Bagnolo San Vito	6521		Lombardie
290	Godiasco	2312		Lombardie
299	Sérmide	8989		Lombardie
305	Marebbe (Enneberg)	1353 (428)	San Vigilio	Vénétie
307	Comelico Superiore	3935 (856)	Padola	Vénétie
310	Rabbi	2541 (533)	Piazzola	Vénétie
311	Castelfondo	889		Vénétie
312	Selva in Gardena (Wolkenstein)	912		Vénétie
313	Canazei	791 (177)	Penia	Vénétie
314	Colfosco in Badia (Colfuschg)	182		Vénétie
315	Livinallongo	1960	Arabba	Vénétie

Numéro de Carte	Nom de la localité	Population	2me nom de localité	Département / canton / région
316	Cortina d'Ampezzo	3343 (198)	Zuel	Vénétie
317	Pieve di Cadore	3661 (722)	Pozzale	Vénétie
318	Forni Avoltri	1473		Vénétie
319	Arta	3112 (421)	Cedarchis	Vénétie
320	Pejo	444		Vénétie
322	Tuenno	1755		Vénétie
323	Predazzo	3692		Vénétie
325	Cencenighe	2048		Vénétie
326	Claut	2271		Vénétie
327	Forni di Sotto	1699		Vénétie
328	Tramonti di Sotto	2798		Vénétie
329	Travasans	(4082)	Moggio	Vénétie
330	Mortaso	529		Vénétie
331	Stenico	917		Vénétie
332	Faver	814		Vénétie
333	Citadella	(897)	Viarago	Vénétie
334	Canal San Bovo	4052		Vénétie
335	Belluno	26753		Vénétie
336	Ponte nelle Alpi	6259		Vénétie
337	Aviano	9772		Vénétie
338	Adorgnano	1190 (5853)	Tricesimo	Vénétie
339	Udine	56041		Vénétie
340	Fontanedo	320 (1719)	Roncone	Vénétie
341	Tiarno di Sotto	735		Vénétie
343	Volano	1774		Vénétie
344	Rocegnò	3625		Vénétie

Numéro de Carte	Nom de la localité	Population	2me nom de localité	Département / canton / région
345	Vas	1718		Vénétie
346	Tarzo	5015		Vénétie
348	Sant'Odorico	1874		Vénétie
349	Gorizia	28154		Vénétie
352	Tonezza	2305		Vénétie
354	Romano	4162		Vénétie
356	Santo Stino di Livenza	7050		Vénétie
357	Ronchis	2405		Vénétie
359	Ruda	1675		Vénétie
360	Torri del Benaco	1720 (293)	Albisano	Vénétie
362	Crespadoro	2883		Vénétie
363	Vicenza	60267		Vénétie
364	Campo San Martino	4248		Vénétie
365	Istrana	4759		Vénétie
367	Grado	5219		Vénétie
368	Pirano	14158		Vénétie
369	Trieste	238655		Vénétie
371	Verona	92536		Vénétie
372	San Giovanni Lupatoto	8279 (1699)	Raldon	Vénétie
373	Montebello	5387		Vénétie
374	Teolo	6296		Vénétie
375	Mira	18133	Gambarare	Vénétie
376.1	Venezia I	192306		Vénétie
376.2	Venezia II	192306		Vénétie
378	Montona	6321		Vénétie
379	Fiume	45000		Vénétie

Numéro de Carte	Nom de la localité	Population	2me nom de localité	Département / canton / région
381	Frescá	(8933)	Cerea	Vénétie
385	Cavarzere	21248		Vénétie
393	Fratta Polesine	3895		Vénétie
397	Rovigno	10863		Vénétie
398	Dignano	10536		Vénétie
399	Cherso	7214		Vénétie
401	Piacenza	58190		Émilie-Romagne
412	Carpaneto	7892		Émilie-Romagne
413	San Secondo	6135		Émilie-Romagne
415	Concordia	10851		Émilie-Romagne
420	Coli	3653		Émilie-Romagne
423	Parma	62603		Émilie-Romagne
424	Poviglio	6168		Émilie-Romagne
427	Ferrara	107618 (3725)	Baura	Émilie-Romagne
432	Bardi	6674		Émilie-Romagne
436	Nonantola	8820		Émilie-Romagne
439	Comacchio	12181		Émilie-Romagne
443	Tizzano	5037		Émilie-Romagne
444	PiscerottoMontericco	1808 (5619)	Albinea	Émilie-Romagne
446	Tintoria	(7414)	Minerbio	Émilie-Romagne
453	Villa Minozzo	7915 (734)	Sologno	Émilie-Romagne
454	Prignano	5372		Émilie-Romagne
455	Savigno	4371 (444)	Merlano	Émilie-Romagne
456.1	Bologna I	211157		Émilie-Romagne
456.2	Bologna II	211157		Émilie-Romagne
458	Fusignano	6126		Émilie-Romagne

Numéro de Carte	Nom de la localité	Population	2me nom de localité	Département / canton / région
459	Ravenna	71877		Émilie-Romagne
464	Sestola	3663		Émilie-Romagne
466	Loiano	5782		Émilie-Romagne
467	Dozza	3018		Émilie-Romagne
476	Brisighella	14599		Émilie-Romagne
478	Meldola	7268		Émilie-Romagne
479	Cesenatico	9813		Émilie-Romagne
490	Molino	(2974)	Portico e San Benedetto	Émilie-Romagne
499	Saludecio	5113		Émilie-Romagne
500	Pontremoli	15022 (249)	Arzengio	Toscane
511	Castiglione di Garfagnana	3389 (146)	Campori	Toscane
513	Piteglio	3832 (214)	Prunetta	Toscane
515	Barberino di Mugello	12845		Toscane
520	Carnaiore	20102		Toscane
522	Vinci	9753		Toscane
523.1	Firenze I	253565		Toscane
523.2	Firenze II	253565		Toscane
526	Stia	3718		Toscane
528	Sant'Agata Feltria	5421		Les Marches
529	Fano	28538		Les Marches
530	Pisa	68933		Toscane
532	Montespertoli	12140		Toscane
534	Incisa	4592		Toscane
535	Valboncione	(2841)	Caprese Michelangelo	Toscane
536	Mercatello	2558		Les Marches

Numéro de Carte	Nom de la localité	Population	2me nom de localité	Département / canton / région
537	Urbino	19932		Les Marches
538	Montemarciano	5284		Les Marches
539	Ancona	66291		Les Marches
541	Fauglia	4691		Toscane
542	Montecatini	4602		Toscane
543	Radda	3260		Toscane
544	Arezzo	51862		Toscane
545	Subbiano	4834	Chiavaretto	Toscane
546	Pietralunga	4695		Ombrie
547	Foce	416 (2085)	Frontone	Les Marches
548	Montecarotto	3538		Les Marches
554 ⁷⁹	Dun-le-Palleteau	1643		Creuse
586	La Tremblade	3664		Charente-Inférieure
598	S.-Vivien	1511		Gironde
599	Cissac	1202		Gironde
601	Lavaufranche	629		Creuse
602	Cressat	1546		Creuse
603	S.-Dizier	2428		Creuse
604	Eymoutiers	4192		Haute-Vienne
605	Limoges	77703		Haute-Vienne
606	S.-Junien	9376		Haute-Vienne
607	Châlus	2589		Haute-Vienne
608	Coussac-Bonneval	3672		Haute-Vienne
609	Seilhac	2035		Corrèze

79. Les points 504, 536, 548 et 549 de l'ALF sont renumérotés 554, 586, 598 et 599 respectivement, pour éviter le chevauchement avec la série de l'ALS.

Numéro de Carte	Nom de la localité	Population	2me nom de localité	Département / canton / région
610	Chazelles	1115		Charente
611	La Tour-Blanche	574		Dordogne
612	S.-Pardoux-la-Rivière	1761		Dordogne
614	Excideuil	1935		Dordogne
615	S.-Pierre-de-Chignac	952		Dordogne
616	Le Bugue	2651		Dordogne
617	Larche	783		Corrèze
618	Souillac	3118		Lot
619	Gourdon	4834		Lot
621	Yviers	892		Charente
624	Bourgnac	407		Dordogne
626	Issigeac	851		Dordogne
628	Villefranche-de-Belvès	1543		Dordogne
630	S.-Savin	1769		Gironde
632	Abzac	1620		Gironde
634	Vélines	881		Dordogne
635	Andraut	1497		Gironde
636	Seyches	1084		Lot-et-Garonne
637	Sainte-Livrade	2644		Lot-et-Garonne
638	Tourmon-d'Agenais	1077		Lot-et-Garonne
641	Pessac	692		Gironde
643	Targon	1100		Gironde
645	S.-Côme	397		Gironde
647	Aiguillon	3094		Lot-et-Garonne
648	Layrac	2609		Lot-et-Garonne
650	Lacanau	1106		Gironde

Numéro de Carte	Nom de la localité	Population	2me nom de localité	Département / canton / région
653	Hostens	1133		Gironde
656	Houeillès	1252		Lot-et-Garonne
657	Mézin	2640		Lot-et-Garonne
662	La Teste-de-Buch	6680		Gironde
702	Auzances	1471		Creuse
703	Pontgibaud	1067		Puy-de-Dôme
704	S.-Quentin	1146		Creuse
705	Mont-Dore	1866		Puy-de-Dôme
706	Merlines	913		Corrèze
707	Meymac	3964		Corrèze
708	Bort	3858		Corrèze
709	Allanche	1885		Cantal
710	La Roche-Canillac	533		Corrèze
711	Beaulieu	2259		Corrèze
712	Gramat	2996		Lot
713	Figeac	6680		Lot
714	Pleaux	2426		Cantal
715	S.-Mamet	1905		Cantal
717	Vic-sur-Cère	1701		Cantal
719	Les Ternes	741		Cantal
720	Cahors	14502		Lot
722	Promilhanes	627		Lot
801	S.-Eloy-les-Mines	4508		Puy-de-Dôme
804	Ennezat	1220		Puy-de-Dôme
805	Monton	1807		Puy-de-Dôme
806	Thiers	17135		Puy-de-Dôme

Numéro de Carte	Nom de la localité	Population	2me nom de localité	Département / canton / région
807	S.-Germain-Lembron	2181		Puy-de-Dôme
808	Sail-sous-Couzan	1260		Loire
809	Ambert	7709		Puy-de-Dôme
811	Massiac	2069		Cantal
812	Paulhaguet	1503		Haute-Loire
813	Monistrol-d'Allier	1082		Haute-Loire
814	Solignac-sur-Loire	1355		Haute-Loire
815	Chamalières	1050		Haute-Loire
816	S.-Bonnet-le-Château	2311		Loire
817	Riotord	2992		Haute-Loire
818	S.-Symphorien-sur-Coise	2307		Rhone
819	Nérondes	1379		Loire
824	Burzet	2734		Ardèche
825	S.-Agrève	3217		Ardèche
826	Les Ollières	1872		Ardèche
827	Vion	666		Ardèche
829	Clonas	456		Isère
833	Vogüé	876		Ardèche
836	Marsanne	1390		Drôme
837	Chabeuil	3076		Drôme
838	S.-Nazaire-en-Royans	700		Drôme
844	Pierrelatte	3184		Drôme
847	Die	3729		Drôme
849	Monestier-de-Clermont	606		Isère
855	Nyons	3611		Drôme
857	Luc-en-Diois	1124		Drôme

Numéro de Carte	Nom de la localité	Population	2me nom de localité	Département / canton / région
866	Orpierre	667		Hautes-Alpes
868	Veynes	2185		Hautes-Alpes
869	S.-Firmin	1110		Hautes-Alpes
875	S.-Etienne-les-Orgues	838		Basses-Alpes
876	Mezel	695		Basses-Alpes
877	Châteaufort	168		Basses-Alpes
878	La Javie	466		Basses-Alpes
879	Chorges	1480		Hautes-Alpes
884	S.-Maximin	2419		Var
885	Gréoux	1163		Basses-Alpes
886	Aups	1892		Var
887	Castellane	1780		Basses-Alpes
888	Puget-Théniers	1224		Alpes-Maritimes
889	Barcelonnette	2223		Basses-Alpes
893	Hyères	17708		Var
894	Le Luc	2875		Var
895	S.-Tropez	3533		Var
896	Seillans	1522		Var
897	Le Cannet	2572		Alpes-Maritimes
898	Plan-du-Var	1647		Alpes-Maritimes
899	Menton	9050		Alpes-Maritimes
905	Ambierle	2563		Loire
908	Cours	5994		Rhone
911	Bully	1455		Rhone
912	S.-Priest	2587		Isère
913	Villars-en-Dombes	1426		Ain

Numéro de Carte	Nom de la localité	Population	2me nom de localité	Département / canton / région
914	S.-Lager	1000		Rhone
915	Lent	1184		Ain
917	Replonges	1642		Ain
920	Le Grand-Serre	1438		Drôme
921	S.-Jean-de-Bournay	3308		Isère
922	Morestel	1440		Isère
924	Torcieu	605		Ain
926	Brion	274		Ain
931	Charavines	1042		Isère
933	La Biolle	1229		Savoie
935	Surjoux	255		Ain
936	Bernex	1004		Vaud
937	Gingins	377		Vaud
939	Le Brassus	821		Vaud
940	Sassenage	1549		Isère
942	Theys	1916		Isère
943	Chignin	760		Savoie
944	Thônes	2935		Haute Savoie
945	Pringy	405		Haute Savoie
946	S.-Pierre-de-Rumilly	1126		Haute Savoie
947	Bons	1009		Haute Savoie
950	Le Bourg-d'Oisans	2375		Isère
953	Espierre	577		Savoie
954	Verrens-Arvey	637		Savoie
955	Hauteluce	1198		Savoie
956	Sixt	1192		Haute Savoie

Numéro de Carte	Nom de la localité	Population	2me nom de localité	Département / canton / région
957	Le Biot	622		Haute Savoie
958	Meillerie	888		Haute Savoie
959	Vevey	11781		Vaud
963	S.-Martin-de-la-Porte	741		Savoie
964	Bozel	1124		Savoie
965	Séaz	1327		Savoie
966	Courmayeur	1201		Piémont
967	Chamonix	2447		Haute Savoie
968	S.-Maurice	2162		Valais
969	L'Etivaz	341		Vaud
971	Monétier-les-Bains	2068		Hautes-Alpes
972	Oulx	1886		Piémont
973	Lanslebourg	971		Savoie
975	Aoste	7376		Piémont
976	Bourg-S.-Pierre	355		Valais
977	Châble	557		Valais
978	Nendaz	2289		Valais
979	Lens	2504		Valais
980	Guillestre	1403		Hautes-Alpes
981	Aiguilles	515		Hautes-Alpes
982	Maïsette	826		Piémont
985	Champorcher	1167		Piémont
986	Châtillon	2903		Piémont
987	Ayas	1753		Piémont
988	Evolène	1208		Valais
989	Vissoye	307		Valais

Numéro de Carte	Nom de la localité	Population	2me nom de localité	Département / canton / région
990	Fontan	1158		Alpes-Maritimes
991	S.-Sauveur	686		Alpes-Maritimes
992	Bobi	1613		Piémont

BIBLIOGRAPHIE

- Adams, Marianne. 1987. «From Old French to the Theory of Pro-drop.» *Natural Language and Linguistic Theory*. 5:1-32.
- Adams, Marianne. 1988. «Les effets V2 en ancien et en moyen français.» *Revue québécoise de linguistique théorique et appliquée*. 7,3:13-39.
- Alvar López, Manuel. 1960. *Los nuevos atlas lingüísticos de la Romania*. Granada: Universidad de Granada.
- Anderson, Stephen R. 1988. «Morphological Change.» *Linguistics: The Cambridge Survey*. F. J. Newmeyer (éd.). Cambridge: Cambridge University Press. 234-262.
- Anderson, Stephen R. 1992. *A-morphous Morphology*. Cambridge: Cambridge University Press.
- Ascoli, Graziadio I. 1874. «Schizzi franco-provenzali.» *Archivio glottologica italiano* III:61-120 (paru en 1878).
- Auger, Julie. 1994. *Pronominal Clitics in Québec Colloquial French: A Morphological Analysis*. Thèse de doctorat, University of Pennsylvania.
- Baylon, Christian. 1991. *Sociolinguistique, langue, société et discours*. Paris: Nathan.
- Benedicto, Helena. 1994. «AGR, ϕ -features and V-movement: Identifying *pro*.» *Functional Projection*. Amherst: University of Massachusetts Working Papers in Linguistics, Benedicto et Runner (réds.). 1-18.

- Benincà, Paola. 1989. «Introduction.» *Dialect Variation and the Theory of Grammar*.
Dordrecht: Foris.
- Benincà, Paola. 1992. «Geolinguistica e sintassi.» *Atlanti linguistici italiani e romanzi: esperienze a confronto*. G. Ruffino (éd.). Palermo: Centro di Studi Filologici e Linguistici Siciliani.
- Benincà, Paola. 1993. «Note introduttive a un atlante dialettale sintattico.» *Dialettologiae varia linguistica*. M. Cortelazzo (éd.). Padua: Unipress.
- Benincà, Paola. À paraître. «I dati dell'ASIS e la sintassi diacronica.» *L'Italia settentrionale, crocevia di lingue e culture*. E. Banfi et al. (réds.). Tübingen: Niemeyer.
- Benincà, Paola, et Cecilia Poletto. 1991. «Il modello generativo e la dialettologia: un'indagine sintattica.» *Rivista Italiana di Dialettologia*. XV:77-97.
- Benveniste, Émile. 1967. *Problèmes de linguistique générale*. vol. I. Paris: Galimard.
- Bessler, Paul. 1994. *Une analyse morphosyntaxique de l'accord grammatical en français*.
Thèse de doctorat, Université de Toronto.
- Bessler, Paul, Sarah Cummins, David Heap, Terry Nadasdi et Yves Roberge. 1992.
«Cliticisation morphologique ou catégories fonctionnelles?» *Actes du Congrès annuel de l'Association canadienne de linguistique*. C. Dyck, J. Ghomeshi et T. Wilson (réds.).29-42.
- Bonet, Eulàlia. 1991. *Morphology After Syntax: Pronominal Clitics in Romance*. Thèse de doctorat, Massachusetts Institute of Technology.

- Bonet, Eulàlia. 1994. «The Person-Case Constraint: A Morphological Approach.» *MIT Working Papers in Linguistics 22: The Morphology-Syntax Connexion*. 33-52.
- Bonet, Eulàlia. 1995a. «The Feature Structure of Romance Clitics.» *Natural Language and Linguistic Theory* 13:607-617.
- Bonet, Eulàlia. 1995b. «The Where and How of Clitic Ordering.» *Revue québécoise de linguistique* 24,1:61-81.
- Borer, Hagit. 1984. *Parametric Syntax: Case Studies in Semitic and Romance languages*. Dordrecht: Foris (Studies in Generative Grammar, 13).
- Bracco, Claudio, Luciana Brandi et Patrizia Cordin. 1985. «Sulla posizione soggetto in italiano e in alcuni dialetti dell'Italia centro-settentrionale.» *Sintassi e morfologia della lingua italiana d'uso: Teorie e applicazioni descrittive*. A.F. de Bellis et L.M. Savoia (réd.) Rome: Bulzoni. 185-209.
- Brandi, Luciana et Patrizia Cordin. 1989. «Two Italian Dialects and the Null Subject Parameter.» *The Null Subject Parameter*. O. Jaeggli et K.J. Safir (réds.). Dordrecht: Kluwer. 111-142.
- Cameron, Richard. 1993. «Ambiguous agreement, functional compensation and nonspecific *tú* in the Spanish of San Juan, Puerto Rico, and Madrid, Spain.» *Language Variation and Change*, 5:305-334.
- Campbell, Lyle. 1972. «Is a Generative Dialectology Possible?» *Orbis* 21:289-298.
- Campos, Héctor Ramiro. 1986. *Inflectional elements in Romance*. Thèse de doctorat, University of Southern California at Los Angeles.

Chambers, Jack, et Peter Trudgill. 1980. *Dialectology*. Cambridge: Cambridge University Press.

Chen, Matthew Y. et William Y.S. Wang. 1975. «Sound change: actuation and implementation». *Language* 51:255-281.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge: MIT Press.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam. 1995. *A Minimalist Program for Linguistic Theory*. Cambridge: MIT Press.

Contreras, Heles. «Closed Domains.» *Probus* 1.2:163-180.

Coromines, Joan. 1974. «De gramàtica històrica catalana: a propòsit de dos llibres.» *Studia philologica et litteraria in honorem L. Spitzer*. Bern. 123-148. Nous citons du recueil de 1974: *Lleures i converses d'un filoleg*. Barcelona: Club. 245-275.

Cummins, Sarah. 1992. «La position du pronom objet dans quelques dialectes français et italiens.» MS inédit, Université de Toronto.

Cummins, Sarah. 1996. *Meaning and Mapping*. Thèse de doctorat, Université de Toronto.

Cummins, Sarah et Yves Roberge 1994. «Romance Inflectional Morphology In and Out of Syntax.» *MIT Working Papers in Linguistics* 22: *The Morphology-Syntax Connexion*. 53-70.

Dufresne, Monique. 1993. *L'articulation syntaxique et phonologique de la cliticisation: le cas des pronoms sujets en moyen français*. Thèse de doctorat, Université du Québec à Montréal.

- Durand, Jacques. 1990. *Generative and Nonlinear Phonology*. London: Longman.
- Emonds, Joseph. 1978. «The Verbal Complex V-V' in French.» *Linguistic Inquiry* 9.2:151-175.
- Escoffier, S. 1958. «La rencontre de la langue d'oïl, de la langue d'oc et du francoprovençal entre Loire et Allier.» *Publications de l'Institut linguistique de Lyon* XI:1-269.
- Foulet, Lucien. 1930. *Petite syntaxe de l'ancien français*. Paris: Champion.
- Franzén, Torsten. 1939. *La syntaxe des pronoms personnels sujets en ancien français*. Uppsala: Almqvist & Wiksells.
- Freedman. 1983. «'Vuoi tu murare?' The italian subject pronoun.» *Studi di grammatica italiana*. Firenze: Presso de l'Accademia della Crusca.
- Gardette, Pierre. 1941. *Géographie phonétique du Forez*. Mâcon (=Publications de l'Institut linguistique de Lyon, IX:1-288).
- Gauchat, Louis. «L'unité phonétique dans le patois d'une commune». *Festschrift Heinrich Morf: Aus romanischen Sprachen und Literaturen*. Halle: M. Hiemeyer. 175-232.
- Gilliéron, Jules et Edmond Edmont. 1902. *Atlas linguistique de la France: Notice servant à l'intelligence des cartes*. Paris:Champion.
- Gilliéron, Jules et Edmond Edmont. 1902-1910. *Atlas linguistique de la France*. Paris:Champion.
- Gougenheim, Georges. 1969. *Système grammatical de la langue française*. Paris: D'Artrey.

- Grassi, Corrado. 1970. *Elementi di dialettologia italiana*. Torino: Giappichelli.
- Green, John. 1988. «Romance Creoles.» *The Romance Languages*. M. Harris et N. Vincent (réds.). New York: Oxford University Press. 420-473.
- Grimes, Joseph E. 1967. «Positional Analysis.» *Language*. 43,437-444.
- Hall, Robert. 1949. «The Linguistic Position of Franco-Provençal.» *Language XXV*:1-14.
- Halle, Morris. 1962. «Phonology in generative grammar.» *Word* 18:54-72. Recueilli dans *The Structure of Language*. J. Fodor et J. Katz (réds.) Englewood Cliffs N.J.: Prentice Hall. 334-352.
- Halle, Morris. 1973. «Prolegomena to a theory of word formation.» *Linguistic Inquiry* 4:3-16.
- Halle, Morris. 1989. «An Approach to Morphology.» *Proceedings of the Northeastern Linguistics Society*. 20. Amherst: Graduate Linguistic Students' Association, University of Massachusetts.
- Haudricourt, André-Georges, et Alphonse G. Juillard. 1949. *Essai pour une histoire structurale du phonétisme français*. Paris: Klincksieck.
- Heap, David. 1991. «Questions with Preverbal Subjects in Dominican Spanish: Morphological Variation and Syntactic Restructuring.» *Proceedings of the Annual Meeting of the Atlantic Provinces Linguistics Association*. J. Black (éd.) 59-67.
- Heap, David. 1993. «Subject Clitics in the Northern Italian Geolinguistic Continuum.» Réunion annuelle de la Société canadienne pour les études italiennes, Université Carleton. (à paraître dans *Quaderni d'Italianistica*).

- Heap, David. 1995. «Subject pronoun variation in Central Romance.» *(N)Waves and Means: a selection of papers from NWAVE 24*. M. Meyerhoff (éd.). Philadelphia: University of Pennsylvania Working Papers in Linguistics 3,1:43-50.
- Heap, David. 1996. «Optimalizing Romance Clitic Sequences.» Linguistic Symposium on Romance Languages XXVI, Universidad Autónoma Metropolitana, Mexico.
- Heap, David, Silvana Mastromonaco, Terry Nadasdi et Paul Bessler. 1993. «If a Template Fits...: Case Studies in Morphological Cliticization». *Proceedings of the Fourth Annual Canadian Workshop on Lexical-Syntactic Relations*. C. Dyck (éd.). Toronto: Toronto Working Papers in Linguistics, 12.1:109-124.
- Hochsberg, Judith. 1986. «Functional compensation for /s/ deletion in Puerto Rican Spanish.» *Language* 62,3:609-621.
- Jaberg, Karl et Jakob Jud. 1928. *Der Sprachatlas als Forschungsinstrument*. Halle: Niemeyer.
- Jaberg, Karl et Jakob Jud. 1928-1940. *Sprach- und Sachatlas Italiens und der Südschweiz*. Zofingen: Ringier.
- Jaberg, Karl. 1932. *Aspects géographiques du langage*. Paris: Droz.
- Jakobson, Roman. 1976. *Six leçons sur le son et le sens*. Paris: Minuit.
- Jochnowitz, George. 1973. *Dialect Boundaries and the Question of Franco-Provençal*. La Haye: Mouton.
- Kayne, Richard. 1975. *French Syntax: the Transformational Cycle*. Cambridge: MIT Press.

- Kayne, Richard. 1991. «Romance Clitics, Verb Movement, and PRO.» *Linguistic Inquiry* 22:647-686.
- King, Ruth, et Terry Nadasdi. 1996. «La puissance des pronoms faibles en français terre-neuvien.» *Proceeding of the Annual Meeting of the Atlantic Provinces Linguistics Association* 19. D. Lillian (éd.). 129-138.
- Kuhn, Thomas. 1962. *The Structure of Scientific Revolutions*. Chicago: Chicago University Press.
- Labov, William. 1966. *The Social Stratification of English in New York City*. Washington, D.C.: Center for Applied Linguistics.
- Labov, William. 1972. *Language in the Inner City: Studies in the Black English Vernacular*. Philadelphia: University of Pennsylvania Press.
- Labov, William. 1994. *Principles of Linguistic Change: Internal Factors*. Oxford: Blackwell.
- Lipski, John M. 1977. «Preposed Subjects in Questions: Some Considerations.» *Hispania* 60:61-67.
- Lobeck, Konrad. 1945. «Die französisch-frankoprovenzalische Dialektgrenze zwischen Jura und Saône.» *Romanica helvetica* XXIII:i-xii, 1-3 18.
- Mahmoudian, Mortéza. 1993. *Modern Theories of Language: The Empirical Challenge*. Durham and London: Duke University Press.
- Martinet, André. 1945. *La prononciation du français contemporain*. Paris: Droz.

- Martinet, André. 1956. *La description phonologique, avec application au parler franco-provençal d'Hauteville (Savoie)*. Genève-Paris: Droz (=version augmentée de l'article dans la *Revue de linguistique romane* XV, daté 1939 mais paru en 1945).
- Martinet, André. 1990. «La synchronie dynamique.» *La Linguistique*. 26,2:13-23.
- McCarthy, John et Alan Prince. 1993. «Generalized Alignment.» MS, University of Massachusetts at Amherst et Rutgers University.
- Meyer, Paul. 1888. Compte-rendu sans titre dans *Romania* IV:293-296.
- Meyer-Lübke, Wilhelm. 1895. *Grammaire des langues romanes* (tome II). Paris: Welter.
- Moignet, Gérard. 1965. *Le pronom personnel français: essai de psycho-systématique historique*. Paris: Klincksieck.
- Morin, Yves-Charles. 1981. «Some myths about pronominal clitics.» *Linguistic Analysis*, 8:95-109.
- Mussafia, Adolfo. 1886. «Una particolarità sintattica della lingua italiana dei primi secoli.» *Miscellanea di filologia e linguistica in memoria di N. Caix et U.A. Canello*. G.I. Ascoli et al. (réds.). Florence: LeMonnier. 255-261.
- Mussafia, Adolfo. 1898. «Enclisi o proclisi del pronome personale atono quale oggetto.» *Romania* 27.145-146.
- Nadasdi, Terry. 1995. *Variation morphosyntaxique et langue minoritaire: le cas du français ontarien*. Thèse de doctorat: Université de Toronto.
- Nicoli, Franco. 1983. *Grammatica milanese*. Milano: Bramante.

- Nyrop, Kristoffer. 1925. *Grammaire historique de la langue française*. Copenhague: Gyldendalske Boghandel. Tome V.
- Paris, Gaston. 1888. Discours publié dans le *Bulletin historique et philologique du Comité des travaux historiques et scientifiques*. 131-147.
- Parkinson, Stephen. 1988. «Portuguese.» *The Romance Languages*. M. Harris et N. Vincent (réds.). New York: Oxford University Press. 131-169.
- Perlmutter, David. 1971. *Deep and Surface Structure Constraints in Syntax*. New York: Holt, Rhinehart & Winston.
- Poletto, Cecilia. 1993. *La sintassi del soggetto nei dialetti italiani settentrionali*. Padova: Unipress.
- Pollock, Jean-Yves. 1989. «Verb Movement, Universal Grammar and the Structure of IP.» *Linguistic Inquiry* 20,3:365-424.
- Pop, Sever. 1950. *La dialectologie: aperçu historique et méthodes d'enquêtes linguistiques*. Volume I: *Dialectologie romane*. Louvain: Centre international de dialectologie générale.
- Price, Glanville. 1984. *The French Language: Present and Past*. London: Grant & Cutler.
- Prince, Alan et Paul Smolensky. 1993. «Optimality Theory: Constraint interaction in generative grammar.» MS, Rutgers University et University of Colorado.
- Renzi, Lorenzo et Laura Vanelli. 1983. «I pronomi soggetto in alcune varietà romanze.» *Scritti linguistici in onore di G.B. Pellegrini*, Padova: Pacini. 121-45.

- Revzin, Isaak I. et G.D. Juldaševa. 1969. «Grammatika porjadok i ee ispol'zovanija.»
(«Grammaire positionnelle et son utilisation.»). *Voprosy Jazykoznanija*, 42-56.
- Ritter, Elizabeth. 1996. «Pronouns from the inside out.» Conférence présentée au
Département de Linguistique, Université de Toronto.
- Rizzi, Luigi. 1986. «On the Status of Subject Clitics in Romance.» *Studies in Romance
Linguistics*. O.A. Jaeggli et C. Silva-Corvalán (réds.). Dordrecht: Foris. 391-419.
- Rizzi, Luigi. 1989. *Préface* dans Y. Roberge et M.-T. Vinet 1989.
- Roberge, Yves. 1989. «Les paramètres de la variation.» *Journal of the Atlantic Provinces
Linguistics Association* 11:94-114.
- Roberge, Yves. 1990. *The Syntactic Recoverability of Null Arguments*. Montréal: McGill-
Queen's University Press.
- Roberge, Yves et Marie-Thérèse Vinet. 1989. *La variation dialectale en grammaire
universelle*. Montréal: Presses universitaires de Montréal.
- Rohlf's, Gerhard. 1966-1969. *Grammatica storica della lingua italiana e dei suoi dialetti*.
Turin: Einaudi. Vol. II: *Morfologia*. (1968).
- Rosenqvist, Arvid. 1919. «Limites administratives et division dialectale de la France.»
Neuphilologische Mitteilungen XX:109-110.
- Rossini, Giorgio. 1975. *Capitoli di morfologia e sintassi del dialetto cremonese*. Firenze:
La nuova Italia editrice.
- Safir, Ken. 1985. *Syntactic Chains*. Cambridge: Cambridge University Press.

Safir, Ken et Osvaldo Jaeggli. 1985. «The Null Subject Parameter and Parametric Theory.»

The Null Subject Parameter. Dordrecht: Kluwer. 1-43.

Sandfeld, Kristian. 1965. *Syntaxe du français contemporain: les pronoms*. Paris:

Champion.

Sankoff, Gillian. 1982. «Usage linguistique et grammaticalisation: les clitiques sujets en

français.» *La sociolinguistique dans les pays de langue romane*. N. Dittmar et B.

Schlieben-Lange (réds.). Tübingen: Gunter Varr. 81-85.

Sapir, Edward. 1921. *Language*. New York: Harcourt, Brace & World.

Sapir, Edward. 1931. «Dialect.» *Encyclopedia of the Social Sciences*. New York:

MacMillan. 5:123-126. Nous citons du recueil de 1963: *Selected writings of*

Edwar Sapir on Language, Culture and Personality. D.G. Mandelbaum (éd.).

Berkeley: University of California Press. 83-88.

Saporta, Sol. 1965. «Ordered rules, dialect differences, and historical processes.»

Language 41:218-224.

Schogt, Henry G. 1966. «Baudoin de Courtenay and Phonological Analysis.» *La*

Linguistique 2:15-29.

Schogt, Henry G. 1968a. «La dynamique du langage.» *Le langage*. A. Martinet (éd.)

Paris: Gallimard (Encyclopédie de La Pléiade). 773-813.

Schogt, Henry G. 1968b. *Le système verbale du français contemporain*. Paris: Mouton.

Simpson, Jane, et Meg Withgott. 1986. «Pronominal clitic clusters and templates.» *Syntax*

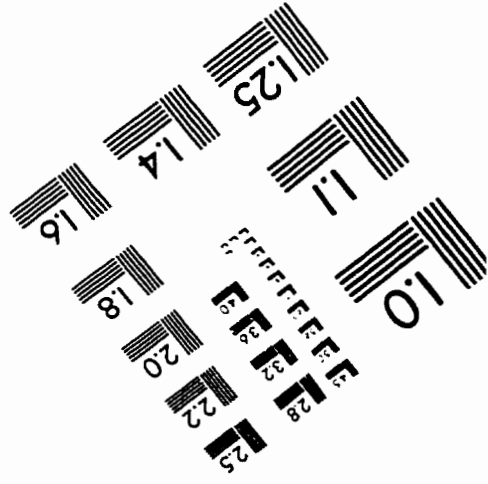
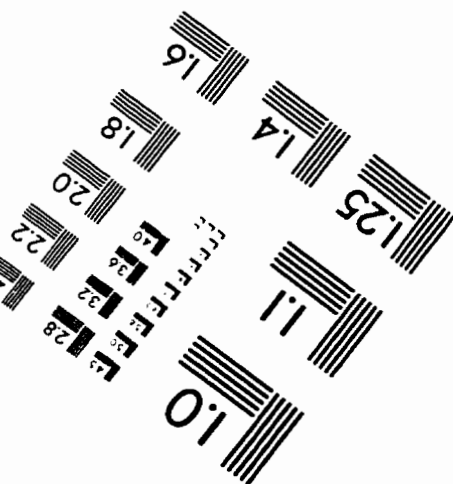
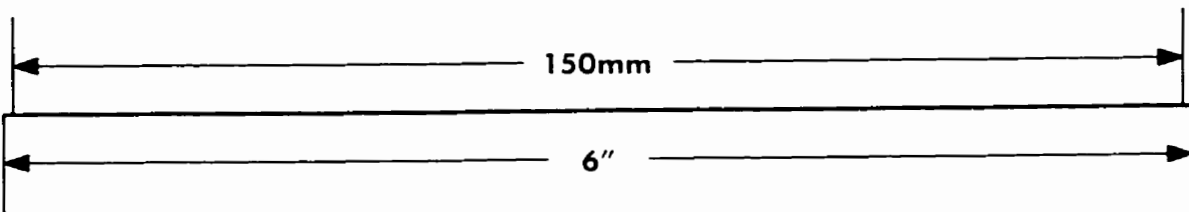
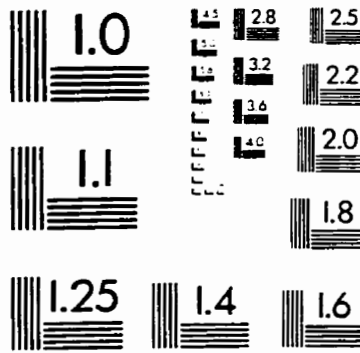
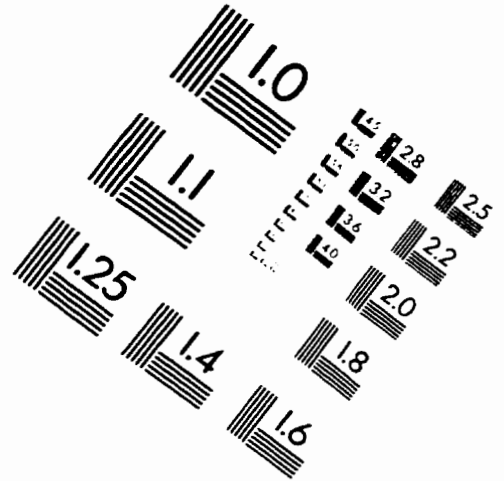
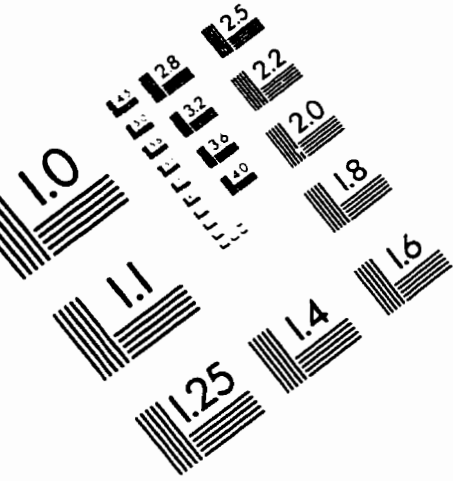
and semantics 19: *The syntax of pronominal clitics*. H. Borer (éd.). 149-174.

- Spencer, Andrew. 1991. *Morphological Theory: An Introduction to Word Structure in Generative Grammar*. Oxford: Blackwell.
- Tennant, Jeff. 1995. *Variation morphophonologique dans le français parlé des adolescents de North Bay*. Thèse de doctorat, Université de Toronto.
- Todolí, Júlia. 1995. *Aproximació a la sintaxi del pronoms clítics catalans*. Thèse de doctorat, Université de València.
- Torrego, Esther. 1984. «On Inversion in Spanish and Some of its Effects.» *Linguistic Inquiry* 15,1:103-129.
- de Tourtoulon, Charles, et Olivier Bringuier. 1876. «Rapport sur la limite géographique de la langue d'oc et de la langue d'oïl.» *Archives des missions scientifiques et littéraires*. Paris: Imprimerie Nationale (Ministère de l'Instruction Public et des Beaux Arts).
- Vance, Barbara. 1988. «L'évolution de Pro-drop en français médiéval.» *Revue québécoise de linguistique théorique et appliquée* 7,3:85-112.
- Vanelli, Laura, Lorenzo Renzi et Paola Beninca. 1985. «Typologie des pronoms sujets dans les langues romanes» *Actes du XVIIeme Congres International de Linguistique et Philologie Romanes* (Aix-en-Provence. 1983), III:163-76.
- Vincent, Nigel. 1988. «Latin.» *The Romance Languages*. M. Harris et N. Vicent (réds.). New York: Oxford Universtiy Press. 26-78.

- Wanner, Dieter. 1974. «The evolution of Romance clitic order.» *Linguistic Studies in Romance Languages*. R. J. Campbell, M. Clayton et M. Goldin (réds.) 158-177.
- Wanner, Dieter. 1977. «On the order of clitics in Italian.» *Lingua* 42.101-128.
- Wanner, Dieter. 1987. *The Development of Romance Clitic Pronouns: from Latin to Old Romance*. Berlin: Mouton de Gruyter.
- Wanner, Dieter. 1991. «The Tobler-Mussafia law in Old Spanish.» *Current Studies in Spanish Linguistics*. F. Martínez-Gil et H. Campos (réds.). Washington: Georgetown University Press. 313-378.
- Wanner, Dieter. 1993. «L'expression du sujet dans les langues romanes.» *Actes du XXe Congrès International de Linguistique et Philologie Romanes*. W. Raible et W. Oesterreicher (réds.). Tome III:IV (Typologie des langues romanes). 449-460.
- Wanner, Dieter. 1995. «Dialect Variation as a Consequence of Standardization.» *International Colloquium on Dialects and Syntax*. University of Düsseldorf.
- Wanner, Dieter. À paraître. «El orden de los clíticos agrupados en castellano.» *Thesaurus*.19
- von Wartburg, Walther. 1950. *Die Ausgliederung der romanischen Sprachräume*. Bern: Franck. Nous citons de la traduction française: Jacques Allières et Georges Straka. 1967. *La fragmentation linguistique de la Romania*. Paris: Klincksieck.
- Weinreich, Uriel. 1954. «Is a structural dialectology possible?» *Word* 10:388-400.

- Weinreich, Uriel, William Labov, et Marvin Herzog. 1968. «Empirical foundations for a theory of language change.» *Directions for Historical Linguistics*, W.P. Lehmann et Y. Malkiel (réds.). Austin: University of Texas Press. 95-195.
- Wexler, Kenneth et Maria R. Manzini. 1987. «Parameters and learnability in Binding Theory», dans T. Roeper et E. Williams (réds.) *Parameter Setting*. Dordrecht: Reidel.
- Zwicky, Arnold. 1977. *On Clitics*. Bloomington: Indiana University Linguistics Club.
- Zwicky, Arnold, et Geoffrey Pullum. 1986. «The Principle of Phonology-free Syntax: Introductory Remarks.» *Working Papers in Linguistics*. Columbus: The Ohio State University. 63-91.

IMAGE EVALUATION TEST TARGET (QA-3)



APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989

© 1993, Applied Image, Inc., All Rights Reserved