

La Grammaire dialectale de l'espagnol à travers le Corpus oral et sonore de l'espagnol rural (COSER, Corpus oral y sonoro del español rural)

Inés Fernández-Ordóñez

**Édition électronique**

URL : <http://corpus.revues.org/1962>

ISSN : 1765-3126

Éditeur

Bases ; corpus et langage - UMR 6039

Édition imprimée

Date de publication : 1 novembre 2010

Pagination : 81-114

ISSN : 1638-9808

Référence électronique

Inés Fernández-Ordóñez, « La Grammaire dialectale de l'espagnol à travers le Corpus oral et sonore de l'espagnol rural (*COSER, Corpus oral y sonoro del español rural*) », *Corpus* [En ligne], 9 | 2010, mis en ligne le 04 juillet 2011, consulté le 02 octobre 2016. URL : <http://corpus.revues.org/1962>

Ce document est un fac-similé de l'édition imprimée.

© Tous droits réservés

**La Grammaire dialectale de l'espagnol
à travers le Corpus oral et sonore de l'espagnol rural
(*COSER, Corpus oral y sonoro del español rural*)**

Inés FERNÁNDEZ-ORDÓÑEZ
Universidad Autónoma de Madrid

1. Introduction

Jusqu'à tout récemment, l'étude de la variation dialectale de l'espagnol dans la péninsule ibérique était basée sur divers atlas régionaux et quelques rares monographies dialectales qui vouaient une attention particulière au castillan (par contraste avec celles, beaucoup plus nombreuses, consacrées aux domaines linguistiques de l'astur-léonais et de l'aragonais). Dans les atlas comme dans les monographies, les dialectologues prêtent davantage attention à la variation phonétique et lexicale qu'à la variation grammaticale. Les données, quant à elles, sont généralement recueillies par le biais d'un questionnaire. Le *Corpus oral et sonore de l'espagnol rural* (dorénavant *COSER*) est un corpus constitué d'enregistrements du langage oral rural dont la collecte a débuté en 1990, en complément aux sources traditionnelles, et est en expansion depuis lors.

2. Description et méthodologie

Le corpus *COSER* est restreint au langage parlé par des informateurs répondant aux critères de la dialectologie traditionnelle : des locuteurs ruraux, de préférence âgés, peu scolarisés, et natifs de l'endroit où ils ont été interrogés. En fait, le *COSER* a le même type d'informateurs que les atlas linguistiques et de nombreuses monographies dialectales, bien que sa méthodologie et ses objectifs soient différents. En 2009, 1 408 informateurs avaient déjà été enregistrés, parmi lesquels on compte 44 % d'hommes et 55 % de femmes. La moyenne

d'âge des informateurs est de 72,9 ans, celle des hommes étant légèrement supérieure à celle des femmes (73,8 contre 72).



Carte 1. Répartition géographique des localités du *COSER* (2009)

A ce jour, les enquêtes ont eu lieu dans 754 localités rurales de la péninsule ibérique, surtout au centre et au nord. Comme le montre la carte 1, la densité des points est tout à fait comparable à celle des atlas régionaux, voire même encore plus dense. Le *COSER* compte actuellement 940 heures d'enregistrements, mais ce nombre augmente chaque année grâce à de nouvelles campagnes d'enquêtes. L'objectif final est d'obtenir des enregistrements de l'espagnol parlé des régions rurales de toute la péninsule ibérique.

La méthodologie utilisée par *COSER* consiste en enquêtes semi-dirigées portant sur des sujets concernant la vie traditionnelle à la campagne. Le fait que l'entretien soit centré sur ces thèmes spécifiques n'empêche aucunement l'enquêteur, après un certain temps et après avoir gagné la confiance de l'informateur, d'aborder des sujets tels que l'éducation, les espoirs et les expériences personnelles, la vie ou la famille, tout dépendant du niveau d'aisance et de spontanéité montré par le témoin. La décision de centrer l'enquête tout particulièrement sur des sujets concernant la vie rurale « d'autrefois » vient du fait que, pour accepter d'être interrogé, l'informateur potentiel

doit reconnaître être en possession d'un savoir sur un mode de vie en voie de disparition. Ces connaissances, dues à leur propre expérience et à leur âge, confèrent ainsi à l'informateur une « autorité » face à l'enquêteur urbain. L'informateur accepte l'entretien lorsqu'il comprend que nous nous intéressons à son témoignage sur un mode de vie en déclin dont très peu se souviennent encore et sur lequel il se sait expert. Nous pensons que la coopération spontanée des informateurs serait beaucoup plus difficile s'ils étaient d'abord interrogés sur leurs opinions ou expériences personnelles, sur des questions linguistiques ou sur d'autres sujets étrangers à la vie rurale. Ce qui a souvent été déterminant dans leur décision d'accepter l'enquête, c'est le fait que l'équipe d'enquêteurs ait insisté sur son intérêt pour les traditions locales spécifiques, qui se distinguent de celles des autres localités rurales, et sur la position privilégiée du témoin en tant qu'héritier de cette tradition.

Les informateurs sont toujours contactés au hasard, sans contacts préliminaires, parmi les habitants remplissant les conditions mentionnées précédemment. L'expérience plutôt frustrante de certaines enquêtes qui ont montré une faible productivité (certains témoins parlaient peu, répondaient par des phrases très courtes ou par monosyllabes) nous a amenés à ajouter, par la suite, une condition de loquacité (« que les informateurs aiment parler ») dans le protocole de sélection des participants. Il est bien connu, pour quiconque a déjà fait du travail de terrain, que le succès n'est jamais assuré ; en commençant dans les mêmes conditions, des enquêtes peuvent être réussies ou désastreuses. Ainsi, elles ne sont pas toutes aussi productives, et le résultat dépend de la bonne volonté des informateurs, de l'habileté des enquêteurs ainsi que de l'interaction entre eux ; cependant, aucun témoignage ne doit être écarté *a priori* car on ne sait pas à l'avance s'il sera intéressant et, de plus, tous comportent des informations utiles.

Cette méthodologie ne peut certes éviter les problèmes posés par la relation à établir entre l'informateur et l'enquêteur, ou par la représentativité supposée de l'informateur choisi au hasard. Néanmoins, nous pensons que la quantité de data peut pallier ces problèmes potentiels puisque les données révèlent

toujours une cohérence géographique, et permet ainsi d'identifier les informateurs dont le parler seraient considéré comme « anormal » dans leur région. De fait, les données recueillies avec le COSER à partir de témoins pris au hasard confirment ou améliorent notre connaissance de la distribution géographique des phénomènes dialectaux présentés dans les atlas qui ont été établis à partir d'enquêtes menées auprès d'informateurs sélectionnés à l'avance.

En ce qui concerne le nombre d'informateurs de chaque point d'enquête, le *COSER* a généralement préféré interroger une seule personne, homme ou femme, mais de manière très fouillée. Toutefois, les conditions d'enregistrements n'ont pu éviter l'intervention occasionnelles d'autres personnes (généralement des membres de la famille ou des connaissances qui, attirés par l'événement que représente l'enquête, n'ont pu résister à la tentation de prendre part à celle-ci). Par conséquent, bien que 1 408 informateurs aient été enregistrés par le *COSER*, la plupart du temps conformément aux principes établis, un seul informateur par localité a été interrogé complètement, soit presque la moitié.

La durée moyenne des enregistrements est d'une heure et quinze minutes (75 minutes) par localité, mais elle varie de seulement une demi-heure à plus de deux heures et demie. La qualité des données enregistrées n'est pas directement proportionnelle à la durée et il y a d'excellents enregistrements très instructifs d'une demi-heure dont les résultats sont comparables à ceux obtenus lors de séances plus longues.

3. Utilisation et apport du *COSER*

3.1 Les distances linguistiques peuvent être mesurées dans le *COSER*

Le *COSER* est un corpus qui a pour but de mesurer les différences relevées dans le langage de groupes socioculturels ruraux peu scolarisés. Il est ainsi complémentaire à la fois des atlas linguistiques et des différents corpus consacrés aux parlers de la population urbaine et cultivée qui ont déjà été constitués, ou sont en voie de l'être, dans le monde hispanophone. Grâce à l'homogénéité de la méthodologie utilisée, on peut mesurer tant

*La Grammaire dialectale de l'espagnol à travers le Corpus
oral et sonore de l'espagnol rural*

la distance linguistique qui sépare différentes aires (distance géographique) que la distance linguistique qui sépare ce groupe social d'autres, comme par exemple, celui de locuteurs ayant un niveau socioculturel plus élevé, ou celui de plus jeunes locuteurs (distance sociale). Bien que la proportion d'hommes et de femmes interrogés ne soit pas identique (55,9 % de femmes contre 44 % d'hommes), le nombre de locuteurs de chacun des sexes est statistiquement représentatif et permet aussi la recherche des différences linguistiques associées à ce paramètre.

Le fait que les médias soient la source de la plupart des corpus oraux de l'espagnol confère une certaine singularité au *COSE*R, dont les informateurs sont rarement enregistrés dans ce domaine. Les données obtenues par le *COSE*R, comparées avec celles des autres corpus de l'espagnol oral, mettent clairement en évidence les différences socioculturelles¹. A cet égard, le *COSE*R se révèle particulièrement utile puisqu'il permet d'étudier des productions grammaticales d'un type non standard qui sont généralement évitées dans la langue écrite et dans le parler des groupes socioculturels plus cultivés. Pour cette raison, Chambers (1995) a proposé, comme universel sociolinguistique, le caractère qualitatif (la présence / l'absence) de variables grammaticales dans l'échelle sociale, par opposition au caractère quantitatif des variables phonétiques.

Les langues standard reflètent moins facilement la variation grammaticale. Par conséquent, les variables de ce type sont fréquemment soumises à un filtre sociolinguistique pouvant altérer les principes linguistiques qui expliquent leur fonction d'origine. C'est le cas, par exemple, des emplois – considérés comme anormaux – des pronoms atones qu'on appelle *leísmo*, *laísmo*, y *loísmo*². Grâce aux enquêtes sociolinguistiques de

1 Les conclusions tirées de ce contraste entre les groupes linguistiques sont méthodologiquement correctes puisque le type d'entretien utilisé par le *COSE*R et les autres corpus est identique ; l'interview (type de conversation de type question-réponse) est toujours le cadre qui génère les données du *COSE*R, et c'est aussi souvent le cas dans d'autres corpus oraux de l'espagnol.

2 Comme nous le verrons plus loin (cf. 2.2.1), le *leísmo* est l'emploi du pronom datif *le* à la place des pronoms accusatifs *lo* et *la* en tant

Klein-Andreu (1979, 1981, 2000) et du *COSER* (cf. Fernández-Ordóñez 1994, 1999), nous savons désormais que ce que les grammairiens considéraient comme des emplois déviants de l'usage normal des pronoms sont en fait des manifestations partielles d'autres paradigmes pronominaux dans lesquels les pronoms sont choisis selon des principes linguistiques différents de ceux à l'œuvre en espagnol standard. Certains de ces paradigmes, comme le paradigme référentiel du castillan, ne figurent en totalité que dans le parler de groupes sociolinguistiques de statut moins élevé. Au fur et à mesure que le statut social s'élève, la plupart des emplois caractéristiques de ces paradigmes (*leísmo* utilisé pour les objets inanimés, *laísmo* et *loísmo*) disparaissent. La répartition sociolinguistique a mené à des confusions dans l'interprétation traditionnelle des faits, puisque la plupart des chercheurs ont fondé leurs hypothèses sur ce sujet en ne prenant en compte que les données partielles fournies par la langue écrite et soutenue (dans laquelle le *leísmo* pour une personne masculine est accepté tandis que les autres *-ismos* sont normalement rejetés). Ainsi, les données du *COSER* ont permis de comprendre des variables grammaticales dont les règles linguistiques devenaient confuses puisqu'elles ne faisaient pas partie, ou très peu, de la langue standard.

Le *COSER* offre par ailleurs un autre aspect particulièrement intéressant : la possibilité d'étudier les changements linguistiques en temps réel. Etant donné que les informateurs du *COSER* appartiennent au même groupe social que ceux des atlas linguistiques, il devient possible de comparer le langage de plusieurs générations consécutives de locuteurs appartenant à ce groupe. Malgré les différences entre la méthodologie des questionnaires des atlas et le type d'enquête utilisé par le *COSER*, le contraste entre les données obtenues par les atlas linguistiques de l'espagnol péninsulaire et celles des enquêtes du *COSER* permet notamment d'étudier les changements en temps réel : le temps écoulé depuis le début du XX^e siècle

qu'objets directs. Le *laísmo* est l'emploi des pronoms accusatifs *la* et *las* à la place des pronoms datifs *le* et *les* en tant qu'objets indirects, et le *loísmo* est l'emploi des pronoms accusatifs *lo* et *los* à la place des pronoms datifs *le* et *les* en tant qu'objets indirects.

jusqu'au début du XXI^e siècle (distance chronologique). Il a été ainsi possible de remarquer le déclin de certains aspects de la grammaire rurale, tel que les séquences *me se, te se* (Heap 2006), ou l'utilisation de *yo* et *tú* comme pronoms régis par une préposition (Pato 2009), tout en montrant la relative stabilité d'autres mécanismes, tel que l'emploi du conditionnel plutôt que de l'imparfait du subjonctif (Pato 2004) ou l'accord avec le sujet des verbes pronominaux à l'infinitif (Pato & Heap 2009).

3.2 La variation grammaticale est mieux expliquée dans le COSER

Il est important de souligner que les enquêtes du *COSER* se sont montrées particulièrement utiles pour enregistrer des phénomènes dialectaux d'ordre grammatical, un aspect traditionnellement peu représenté dans les monographies dialectales et dans les questionnaires des atlas linguistiques. En effet, le déroulement de l'enquête donne la possibilité de faire des recherches sur l'emploi de n'importe quel phénomène grammatical dans un contexte réel : au lieu des phrases isolées, hors contexte et peu naturelles qui sont typiques des questionnaires, l'enquête recueille des phrases émises lors d'un discours réel, dans lequel il est possible d'étudier les valeurs contrastives, les motivations affectives et les conclusions pragmatiques liées à une structure donnée. Par conséquent, les données du *COSER* permettent de mieux comprendre une structure qui existait en ancien espagnol mais qui, de nos jours, se rencontre uniquement dans certaines variétés rurales spécifiques, et qui a clairement une valeur focale (de focus informatif ou neutre et, parfois aussi, de focus contrastif, cf. Zubizarreta 1999) : l'emploi de l'article suivi de l'adjectif possessif (*el mi hijo* « le mon fils »), qui, dans ces variétés, est employé à la place de la structure possessive emphatique régulière en espagnol (*MI hijo, el hijo mío* « MON fils, le fils à moi »), et contraste avec les constructions non emphatiques de la possession inaliénable (*el hijo*). Le caractère focal de la forme explique le fait que ces structures s'emploient plutôt avec les possessifs de la première et de la deuxième personnes, concernant le locuteur et le récepteur, et avec des compléments mettant en évidence la

relation entre le possesseur et le possédé³, aspects qui peuvent être difficiles à repérer dans des phrases isolées comme dans les questionnaires des atlas ou celles qui sont sporadiquement citées dans les monographies dialectales⁴.

3.2.1 Des paradigmes pronominaux au lieu de « déviances »

Comme on l'a annoncé plus haut, les apports de la méthodologie de l'enquête se sont révélés d'une importance primordiale pour expliquer un aspect de la syntaxe espagnole traditionnellement mal compris : le *leísmo*, le *laísmo* et le *loísmo*.

La grammaire normative définit le *leísmo* comme l'emploi de la forme *le* à la place du pronom accusatif *lo* (ou exceptionnellement *la*) pour référer à l'objet direct : une extension du datif aux dépens de l'accusatif. Les auteurs ont distingué plusieurs types de *leísmo* qui ont divers degrés d'incidence : 1) le *leísmo* le plus fréquent et le plus répandu est celui qui réfère à un objet direct masculin, singulier et humain (1a) ; 2) le *leísmo* masculin singulier qui réfère à un objet direct non humain présente une diffusion plus limitée (1b) ; 3) le *leísmo* pluriel apparaît moins souvent que le singulier, bien qu'il soit plus fréquent s'il concerne une personne

3 Cela concorde aussi avec le fait que, en ancien espagnol, la structure est beaucoup plus fréquente dans le discours direct ou rapporté que dans les autres types de textes (cf. Company 2009).

4 En outre, l'*Atlas linguistique de la péninsule Ibérique (ALPI)*, que l'on avait perdu jusqu'à récemment, est le seul atlas consacré à l'espagnol qui inclut deux questions permettant de documenter cet emploi : No. 260, *Mis cuñados y mis primos* « Mes beaux-frères et mes cousins », No. 261, *Sus corderos están en nuestro prado* « Leurs moutons sont dans notre champ ». Heureusement, le matériel d'*ALPI*, dont seulement un volume a été publié (cf. Navarro Tomás 1962), peut désormais être consulté en ligne (cf. Heap 2002, 2003-). Bien que certains atlas régionaux incluent la question *Mi mujer va a menudo al médico* « Ma femme va souvent chez le médecin » (*Atlas linguistique et ethnographique de Navarre, Aragon et La Rioja*, ou *ALEANR*, carte 1743, *Atlas linguistique de Castille et Leon*, ou *ALCyL*, carte 165, *ALECMAN*), seule la réponse relative à la locution adverbiale est cartographiée, et non celle relative au possessif. *El ALECMAN* est le seul à reprendre la question de l'*ALPI Mis cuñados y mis primos*. Étonnamment, l'*Atlas linguistique et ethnographique cantabrique* (ou *ALECCANT*) n'a inclus aucune question concernant cet usage en dépit du fait qu'il soit très répandu dans la région, comme le montrent les notes de caractérisation linguistique introductives de chaque localité.

*La Grammaire dialectale de l'espagnol à travers le Corpus
oral et sonore de l'espagnol rural*

(1c) qu'une « chose » (1d) ; 4) le *leísmo* le plus rare de tous est celui qui réfère à un objet direct féminin, normalement humain, qu'il soit singulier (1e) ou pluriel (1f) ; 5) on ne trouve jamais *leísmo* avec référence à un neutre.

- (1a) ¿Conoces a Juan? Sí, *le* conozco hace tiempo.
Tu connais Juan ? Oui, je le connais depuis longtemps.
- (1b) ¿Sabes dónde está mi libro? No, no *le* he visto por aquí.
Tu sais où est mon livre ? Non, je ne l'ai pas vu ici.
- (1c) Esta tarde voy a recoger a los niños del colegio y *les* llevaré al parque.
Ce soir, je vais prendre les enfants au collège et je les emmènerai au parc.
- (1d) Fui a buscar los discos que querías y *les* encontré en la tienda de abajo.
Je suis aller chercher les disques que tu voulais et je les ai trouvés dans la boutique d'en bas.
- (1e) A María hace tiempo que no *le* veo.
Marie, ça fait longtemps que je ne l'ai pas vue.
- (1f) Aquí no hay monjas. En la guerra *les* mataron a todas.
Ici, il n'y a pas de nonnes. Pendant la guerre, ils les ont toutes tuées.

Le *laísmo* concerne l'emploi des pronoms accusatifs *la* à la place du pronom datif *le* : une extension de l'accusatif au dépens du datif. Le *laísmo* est fondamentalement réservé à l'humain (2a), bien qu'il existe aussi des exemples avec des antécédents non humains (2b), aussi bien singuliers que pluriels (2c, d)⁵. Le *laísmo* est d'un usage moins répandu que le *leísmo* humain et on a observé qu'il a une fréquence légèrement plus élevée au singulier qu'au pluriel.

- (2a) Cuando vi a Pepa, *la* di su regalo.
Quand j'ai vu Pepa, je lui ai donné son cadeau.
- (2b) Coges la sartén, *la* das la vuelta y ya tienes lista la tortilla.
Tu prends la poêle, tu lui fais faire un tour et l'omelette est prête.

5 Cela est attribué au fait que la plupart des antécédents des datifs sont animés.

- (2c) A las niñas de hoy ya no *las* gusta coser.
Les filles d'aujourd'hui n'aiment pas coudre.
- (2d) A esas rosas hay que cortar*las* los tallos secos.
Ces roses, il faut leur tailler les tiges mortes.

Le *loísmo* est le détournement le moins fréquent des emplois des pronoms. Il s'agit de l'utilisation de *lo* au lieu de *le* pour le datif avec un antécédent masculin ou neutre. Dans tous les exemples traditionnellement cités, la plupart ont un antécédent masculin pluriel, le plus souvent humain (3a), bien que les antécédents non humains ne soient pas complètement absents (3b)⁶. Au contraire, au singulier, on a observé une plus grande fréquence de *loísmo* référent à des antécédents non humains (3c) et neutres (3d, e), les exemples avec un antécédent humain restant marginaux (3f).

- (3a) Cuando recojo a los niños del colegio, *los* llevo la merienda.
Quand je vais chercher les enfants au collège, je leur apporte le goûter.
- (3b) Para arreglar esos trajes, hay que sacar*los* el bajo.
Pour arranger ces vêtements, il faut leur raccourcir le bas.
- (3c) Cuando el arroz está cocido, *lo* echas la sal.
Quand le riz est cuit, tu lui ajoutes le sel.
- (3d) Yo no *lo* doy ninguna importancia a eso.
Moi, je ne lui accorde aucune importance à cela.
- (3e) Antes iba a esquiar, pero luego *lo* cogí miedo y lo dejé.
Avant j'allais skier mais après j'ai eu peur et j'ai abandonné.
- (3f) Cuando vi que el ladrón me iba a asaltar, *lo* pegué un empujón y salí corriendo.
Quand j'ai vu que le voleur allait m'attaquer, je lui ai donné un coup et je suis parti en courant.

Bien que le lien entre ces emplois ait été observé depuis longtemps, les chercheurs n'avaient pu les expliquer comme relevant d'un principe linguistique cohérent et, afin d'en expliquer l'origine, ils devaient avoir recours à la combinaison

6 Pour les mêmes raisons exposées à la note 5.

de deux tendances contradictoires⁷. L'une est la tendance à distinguer les compléments d'objet direct humains (avec *le* et le *leísmo*) des non humains (en employant *lo*, *la*, sans le *leísmo*) ; cette hypothèse expliquait le *leísmo* humain, considéré comme une façon de marquer morphologiquement les compléments animés mais ne rend pas compte du fait que cette extension de la morphologie dative concerne surtout les compléments masculins, et non les féminins, et n'éclaire en rien les occurrences de *leísmo* non animé, de *laísmo* et de *loísmo*. Il fallait donc recourir à une autre tendance, qui aurait pour but de sélectionner les pronoms uniquement selon le genre de leur antécédent, sans considérer leur fonction syntaxique dans la phrase. Selon cette perspective, le *leísmo*, le *laísmo* et le *loísmo* élimineraient les distinctions de cas au profit des distinctions de genre avec *le* comme pronom masculin, *la* comme pronom féminin et *lo* comme pronom neutre. Cette tendance expliquait alors le *leísmo* masculin (humain et non humain) et le *laísmo* ; cependant, cela n'éclairait pas le fait que l'emploi du *leísmo* ne peut être généralisé pour tous les objets masculins – étant toujours plus fréquent avec un antécédent humain que non humain – ni le fait que l'emploi du *leísmo* reste plus rare au pluriel, où il est en compétition avec le *loísmo*.

A la suite de ces observations classiques, les auteurs des atlas linguistiques ont élaboré des questionnaires qui avaient pour but d'enregistrer surtout le *leísmo* humain masculin, le *laísmo* et le *loísmo*, c'est-à-dire qu'ils voulaient recueillir les manifestations fondamentales de la première et de la deuxième tendance, respectivement. Mais les autres cas ont été négligés, comme l'emploi du *leísmo* non humain, ou celui du pronom *lo* référant à des entités non dénombrables (masculines et féminines), bien qu'ils soient indissociablement liés, comme nous allons le montrer⁸.

7 Les défenseurs les plus connus de cette interprétation sont Cuervo (1895), Fernández Ramírez (1987) et Lapesa (1968). Des hypothèses semblables ont été émises par García (1975) et Flores Cervantes (1997, 2002, 2006). Un compte rendu critique de ces analyses se trouve dans Fernández-Ordóñez (1993) et (2001), avec des arguments prenant en compte l'ensemble du contexte roman.

8 Ainsi, l'ALPI consacre cinq questions au *leísmo* humain (350 *A Miguel le cogieron preso* « Miguel a été fait prisonnier », 351 *Le llevaron a la cárcel*

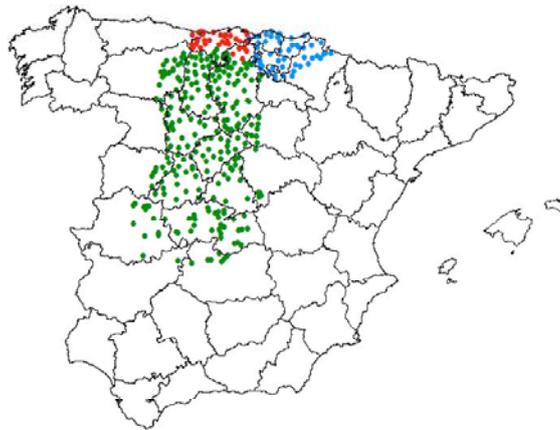
L'analyse de l'enregistrement des données du langage rural a permis d'établir la délimitation géographique exacte des régions dans lesquelles on retrouve chacun de ces emplois et de démontrer que l'apparent manque de cohérence des phénomènes

« Il a été envoyé en prison », 352 *Al padre le vieron llorando* « On a vu le père pleurer », 353 *A los niños les socorrieron los vecinos* « Les enfants ont été secourus par les voisins », 355 *Al enfermo hay que cuidarle* « Le malade doit être soigné ») : outre le grand nombre de questions destinées à enregistrer le même phénomène, le caractère standard du *leísmo* de personne masculin est montré par le fait que les questions sont formulées avec le *leísmo*. Par opposition, celles consacrées au *loísmo* (356 *Al niño le pusieron un vestido* « L'enfant portait une robe », 357 *Tráete los candiles para echarles aceite* « Apporte les lampes à l'huile pour les remplir ») et au *laísmo* (359 *A la madre no le dieron la limosna* « On n'a pas fait d'offrir à la mère », 360 *Aquella desgracia le costó a ella la vida* « Ce malheur lui a coûté la vie », 361 *A las hermanas les enviaron unas cartas* « On a envoyé des lettres aux sœurs », 362 *A la yegua le cansa el trabajo* « Le travail fatigue la jument ») sont exprimées avec les formes pronominales normales. Aucune question concernant le *leísmo* masculin non humain n'a été prévue. Néanmoins, les questions 312 et 313, destinées à la conjugaison du verbe *vaciar* « vider », peuvent permettre l'étude du *leísmo* non humain (312 *¿Dónde vacían el cántaro?* « Où est vidée la cruche ? », 313 *No lo vacies en la calle* « Ne la vide pas dans la rue »). L'ALEANR consacre moins de questions à cet emploi, et de plus, la plupart d'entre elles sont calquées sur celles du questionnaire de l'ALPI (elles reproduisent les entrées 350-351, 353, 356, 359, 362 correspondant aux cartes 1708-1711). Bien qu'il n'y ait aucune question pour le *leísmo* masculin non humain, il en est une qui permet l'enregistrement du *leísmo* féminin humain (*A la madre la vio en la calle* « La mère a été vue dans la rue », carte 1713). Seuls l'ALECant et l'ALCyL incluent de nouvelles questions visant le *leísmo* non humain (avec un antécédent animé, *Al lobo lo vimos* « Nous avons vu le loup », cartes 1194 et 118 respectivement, et avec un antécédent inanimé, *El libro lo olvidé en casa* « J'ai oublié le livre à la maison », ALECant 1195, *El paquete lo olvidé* « J'ai oublié le colis », ALCyL 116). Ces deux atlas régionaux reproduisent également les questions 350, 352-353, 356, 359 et 362 de l'ALPI (ALECant, 1243, 1245-1247, 1192, 1197; ALCyL, 111-114, 117, 120) et 1713 d'ALEANR. Dans l'ALECMan, on retrouve également les questions 350-353, 356, 359 et 362 de l'ALPI, auxquelles ont été ajoutées de nouvelles questions pertinentes pour le *laísmo*, *A las niñas no (les/las) gusta estudiar* « Les filles n'aiment pas étudier », *La torre desde aquí se (le/la) ve* « La tour qu'on voit d'ici ». Aucun de ces atlas ne permet d'observer l'absence de *leísmo* quand l'antécédent est un complément masculin ou féminin non comptable (comme *pan* (pain), *vino* (vin), *trigo* (blé), *agua* (eau), etc.), pas même l'ALECant, alors que la Cantabrie est une région où l'existence du neutre de matière a été bien décrite.

*La Grammaire dialectale de l'espagnol à travers le Corpus
oral et sonore de l'espagnol rural*

est en fait dû à l'existence de plusieurs paradigmes pronominaux, différents du système de l'espagnol standard, qui étaient confondus dans les études antérieures, faussant ainsi les analyses (Fernández-Ordóñez 1994, 1999, 2001). Globalement, il y a trois paradigmes principaux : le premier est le paradigme propre au roman parlé au contact du basque (Tableau I), le deuxième est le paradigme cantabrique (Tableau II) et enfin, le troisième, le castillan, se situe dans la Castille ouest et est parfois appelé système référentiel (Tableau III). Le paradigme castillan vient de l'évolution du cantabrique, après l'élimination de la catégorie du cas. Seul le paradigme castillan présente des cas de *laísmo* et de *loísmo*, alors que les trois paradigmes présentent des cas de *leísmo* masculin humain, ce qui explique que le *leísmo* s'est trouvé être le phénomène le plus étudié, d'autant qu'il est aussi le seul à être admis dans la langue soutenue et à l'écrit.

La distribution géographique de ces paradigmes est illustrée par la carte II.



Carte II. Distribution géographique des paradigmes cantabrique (en rouge), castillan (en vert) et roman en contact avec le basque (en bleu)

Le paradigme basque (en bleu sur la carte II) montre la concrétisation de l'hypothèse du *leísmo* comme étant une extension du datif aux compléments humains masculins et féminins. Par contre, les paradigmes cantabrique (en rouge) et castillan (en vert) se caractérisent par le fait que la sélection

pronominale est fondée sur la catégorisation sémantique de l'antécédent comme dénombrable ou non dénombrable, une catégorie linguistique qui n'avait pas été prise en considération par le passé, et qui explique le fait que le *leísmo* soit généralisé avec les antécédents humains masculins (toujours comptables et donc représentés par *le*), mais pas systématique avec les antécédents non humains, puisque ceux-ci peuvent être comptables (donc *le*) ou non comptables (donc *lo*). Le paradigme castillan, pour sa part, se distingue du paradigme cantabrique par l'absence de cas, généralisant ainsi *la(s)* et *lo* en tant que pronoms datifs. Pour compliquer les choses davantage, au masculin pluriel, le système castillan présente au moins deux issues différentes selon le pronom préféré : (a) *les*, employé dans le nord (nord-ouest de Burgos, Palence et Valladolid) ; (b) *los*, caractéristique du sud (est de Salamanca et Cáceres, Ávila, ouest de Tolède et Madrid). Par conséquent, l'aire du système castillan comprend le centre et l'ouest de la Castille, du sud des monts Cantabriques jusqu'à La Manche.

Pour faciliter la compréhension, les formes de ces trois paradigmes qui divergent partiellement de celui de l'espagnol standard (Tableau IV) sont en italiques.

Tableau I. Paradigme du roman en contact avec le basque

	ANIMÉS		INANIMÉS			
	MASCULIN / FÉMININ		MASCULIN		FÉMININ	
	SINGULIER	PLURIEL	SINGULIER	PLURIEL	SINGULIER	PLURIEL
ACCUSATIF	<i>le</i>	<i>les</i>	ø / lo	ø / los	ø / la	ø / las
DATIF	<i>le</i>	<i>les</i>	le	les	le	les

Tableau II. Paradigme cantabrique

	COMPTABLES				NON COMPTABLES		NEUTRE
	SINGULIER		PLURIEL		MASC.	FÉM.	
	MASC.	FÉM.	MASC.	FÉM.			
ACCUSATIF	<i>le</i>	la	los	las	lo	<i>lo</i>	lo
DATIF	le		les		le		le

*La Grammaire dialectale de l'espagnol à travers le Corpus
oral et sonore de l'espagnol rural*

Tableau III. Paradigme référentiel castillan

	COMPTABLES				NON COMPTABLES		NEUTRE
	SINGULIER		PLURIEL				
	MASC.	FÉM.	MASC.	FÉM.	MASC.	FÉM.	
ACCUSATIF	<i>le</i>	<i>la</i>	<i>les</i> (a) <i>los</i> (b)	<i>las</i>	<i>lo</i>	<i>lo</i>	<i>lo</i>
DATIF	<i>le</i>	<i>la</i>	<i>les</i> (a) <i>los</i> (b)	<i>las</i>	<i>lo</i>	<i>lo</i>	<i>lo</i>

Tableau IV. Paradigme de l'espagnol standard

	MASCULIN	FÉMININ	NEUTRE
ACCUSATIF	<i>lo(s)</i>	<i>la(s)</i>	<i>lo</i>
DATIF	<i>le(s)</i>		<i>le</i>

Le fait que les grammairiens et les dialectologues de l'espagnol n'aient pris en compte qu'une partie des emplois différenciés du système standard et ne les aient pas intégrés dans des paradigmes pronominaux soumis à des principes linguistiques différents explique leur incompréhension de ces mécanismes. De plus, ils n'en ont étudié que les emplois les plus fréquents, sans tenir compte des autres, minoritaires d'un point de vue global, tels que le *leísmo* référant à des compléments humains féminins (*A María le vi ayer* « J'ai vu Marie hier ») ou à des pronoms nuls (*Los libros te Ø he dado* « Je t'ai donné les livres »), particularité du roman en contact avec le basque, ou encore le *lo* référant à des compléments de matière féminins (*La lana lo venden* « La laine est vendue / Ils vendent de la laine »), typique des aires cantabrique et castillane.

3.2.2 *Substitutions de mode, variantes minoritaires et quantification des données*

Ainsi, comme nous venons de le voir, le *COSER* permet de meilleures interprétations des principes linguistiques opérant dans les variétés orales. Et cet avantage est indubitablement lié à la possibilité de quantifier les data : pour une variable linguistique donnée, l'enquête permet de quantifier les variantes, tant dans une localité en particulier que dans les

différents contextes dans lesquels elles apparaissent, alors que dans les atlas, cette quantification est généralement impossible puisqu'une seule réponse n'est enregistrée pour chaque localité, et parce qu'il n'y a que très peu de questions concernant chaque variable. Par conséquent, les variantes minoritaires d'une variable donnée apparaissent rarement dans les atlas.

C'est la conclusion qui est tirée, par exemple, de l'étude d'un emploi grammatical original observé dans les régions centrale et septentrionale de la péninsule ibérique (cf. carte III) : l'emploi du conditionnel simple (-ría) au lieu de l'imparfait du subjonctif (-ra / -se), et ce dans tous les contextes syntaxiques de l'imparfait du subjonctif en espagnol (Pato 2004).



Carte III. Zone de substitution du subjonctif dans la péninsule ibérique

Cet emploi a été consigné dans les atlas, mais de manière bien insuffisante, on y omet le fait que l'imparfait du subjonctif est non seulement remplacé par le conditionnel -ría (la variante majoritaire), mais aussi par l'imparfait de l'indicatif -ba (la variante minoritaire). Les exemples (4) et (5) du *COSE*R montrent les deux variantes de cet emploi chez le même informateur de Santervás de la Vega (Palence) :

- (4) Las costillas y todas esas cosas se metían en ollas para que se *conservarían*.
Les côtelettes et toutes ces choses étaient mises dans des marmites afin qu'elles se conserveraient.
Les côtelettes et toutes ces choses étaient mises dans des marmites afin qu'elles se conservent.

*La Grammaire dialectale de l'espagnol à travers le Corpus
oral et sonore de l'espagnol rural*

- (5) Se las colgaba en la cocina o en una habitación, o como fuera... que las diera un poco el sol, para que *estaban* más buenas [las morcillas].

Ils étaient suspendus dans la cuisine ou dans une pièce ou quelque part... pourvu qu'ils reçoivent un peu de soleil afin qu'ils étaient meilleurs [les boudins].

Ils étaient suspendus dans la cuisine ou dans une pièce ou quelque part... pourvu qu'ils reçoivent un peu de soleil afin qu'ils soient meilleurs [les boudins].

En comparant les cartes des atlas régionaux (*ALCyL*, *ALEANR*, *ALECAnt*) avec la carte faite à partir du matériel du *COSER*, il est possible de confirmer, d'une part, la concordance générale de la zone linguistique dans laquelle ce phénomène se produit (cartes IV et V)⁹ :



Carte IV. L'emploi de *-ría* (en rouge) au lieu de *-ra* / *-se* (en bleu), selon les atlas régionaux (Pato 2004)

⁹ Toutefois, on note une différence entre les données d'*ALCyL* et celles du *COSER* concernant la zone affectée par le phénomène : l'atlas le situe plutôt au centre et au sud de la province de Soria mais ni dans la province de León, ni dans celle de Valladolid. Dans ces deux derniers cas, l'absence du phénomène s'explique par sa rareté. En revanche, la divergence des données pour Soria reste incompréhensible, puisque le *COSER* et l'*ALCyL* ont tous les deux mené des études récentes dans cette province.

I. FERNANDEZ-ORDOÑEZ



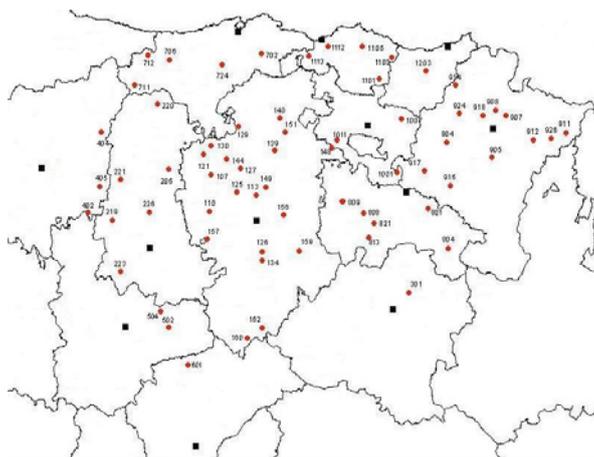
Carte V. L'emploi de *-ría* (en rouge) au lieu de *-ra* / *-se*, selon *COSER* (Pato 2004)

D'autre part, il est clair que les atlas sont incapables de refléter la variante minoritaire *-ba*, qui est cependant régulièrement enregistrée dans tout le territoire par le *COSER* (Cartes VI et VII).



Carte VI. L'emploi de *-ba* (en rouge) au lieu de *-ra* / *-se* (en bleu), selon les atlas régionaux (Pato 2004)

*La Grammaire dialectale de l'espagnol à travers le Corpus
oral et sonore de l'espagnol rural*



Carte VII. L'emploi de *-ba* (en rouge) au lieu de *-ra / -se*,
selon *COSER* (Pato 2004)

Comme nous l'avons vu, bien que les deux variantes existent, elles n'apparaissent pas avec la même fréquence, ce qui explique le fait que la variante minoritaire ne figure pratiquement pas dans les atlas : quand l'imparfait du subjonctif est remplacé par ces formes de l'indicatif, *-ría* domine à 96 %, tandis que *-ba* n'apparaît que dans 4 % des cas.

La quantification des données n'est pas impossible avec les données des atlas, mais statistiquement, elle est beaucoup plus juste si les données proviennent d'un corpus comme le *COSER*. Une première raison en est que le phénomène apparaît dans des contextes qui n'étaient pas forcément prévus lors de l'élaboration des questionnaires des atlas. Ce fut justement le cas pour les *leísmo*, *laísmo* et *loísmo* que nous avons décrits plus haut. C'est également ce qui se produit avec l'emploi de *-ría / -ba* au lieu de *-ra / -se*. En effet, les atlas avaient prévu d'enregistrer cet emploi plutôt dans la protase des phrases conditionnelles et dans les phrases désidératives employant *ojalá* « Je souhaite / J'espère »¹⁰ alors que le phénomène se

¹⁰ L'*ALPI* inclut quatre questions pertinentes (386 *Si tuviera dinero lo compraría* « Si j'avais de l'argent, je l'achèterais », 387 *Si estudiase aprendería* « Si j'étudiais, j'apprendrais », 388 *Si pudiera la mataría*

manifeste dans les propositions complétives, relatives, finales, causales, etc., soit toute sorte de proposition subordonnée dans laquelle on trouvera vraisemblablement l'imparfait du subjonctif en espagnol (comme l'ont observé Ridruejo (1975), Silva Corvalán (1985) ou Martínez Martín (1983) dans des études portant sur les territoires de La Rioja et de Burgos). Que cela concerne l'utilisation des formes pronominales ou celle des formes verbales, le questionnaire d'atlas présente comme des déviations partielles de l'usage standard ce qui est en fait un autre système régi par des principes linguistiques différents et dont les manifestations apparaissent dans une gamme de contextes beaucoup plus vaste.

En second lieu, le nombre d'occurrences d'un phénomène donné est toujours plus grand dans un entretien que dans les réponses à un questionnaire d'atlas, même dans le cas hypothétique où on aurait inclus dans le questionnaire tous les contextes syntaxiques où le phénomène est susceptible d'apparaître. C'est ce nombre significatif d'enregistrements qui nous permet de détecter la présence des variantes minoritaires, qui sont en fait dissimulées dans les atlas. Par conséquent, du point de vue statistique, la quantification de données d'un corpus comme le *COSE*R permet de tirer des conclusions beaucoup plus proches de la réalité des usages linguistiques. Cette quantification permet, par exemple, de nuancer les cartes précédentes (cartes IV, V, VI et VII), et de préciser où se trouve le centre de rayonnement de l'emploi de *-ría* / *-ba* au lieu de *-ra* / *-se*, et quelles sont les aires de transition. La carte VIII suivante montre que le centre de rayonnement est clairement

« Si je pouvais, je la tuerais », 390 (*Ojalá lloviese* « Si seulement il pouvait pleuvoir »), parmi lesquelles la première et la dernière ont été reprises dans l'*ALEANR* (cartes 1704, 1706), dans l'*ALECant* (cartes 1216, 1220) et dans l'*ALCyL* (148, 152). L'*ALEANR* a enrichi les contextes syntaxiques en ajoutant une entrée avec une complétive (1705 *Le dijo que trajera un pan* « Il lui a dit d'apporter un pain »), de laquelle ont hérité l'*ALECant* et l'*ALCyL* (cartes 1218 et 150, respectivement). L'*ALECant* a ensuite ajouté une proposition concessive à la liste (1217 *Aunque pudiera no lo haría* « Même si je le pouvais, je ne le ferais pas »), reproduite dans l'*ALCyL* (carte 149). Enfin, seul le questionnaire de l'*ALCyL* inclut une proposition finale (151 *Esto te lo dije para que fueras bueno* « Je te l'ai dit pour que tu sois un bon garçon »).

*La Grammaire dialectale de l'espagnol à travers le Corpus
oral et sonore de l'espagnol rural*

délimité au nord et à l'est de Burgos, et dans les régions
voisines de la Cantabrie, la Biscaye, l'Alava et la Haute Rioja.



Carte VIII. L'emploi de *-ría / -ba* au lieu de *-ra / -se*, selon *COSER*,
incluant la quantification du phénomène ● 75-100 % / ● 50-75 % / ● 25-50 %
(Pato 2004)

D'autre part, les données du *COSER* ont permis d'identifier une autre propriété générale de ce changement de mode : il est caractéristique des temps simples, qui atteignent une fréquence moyenne de 61,9 % dans la zone, tandis qu'il se produit rarement avec les temps composés (21,6 %).

La quantité de données rend aussi possible l'application de tests statistiques comme la régression logistique, qui permet de déterminer l'influence simultanée de plusieurs variables sur la manifestation du phénomène. En ce qui concerne l'emploi de *-ría / -ba* au lieu de *-ra / -se* dans les variétés castillanes, on a pu démontrer que l'idée la plus répandue à ce sujet – qui situait l'origine du phénomène dans la protase des phrases conditionnelles – était erronée. C'est plutôt la prévalence de *-ría* et *-ba* aux dépens des formes du subjonctif *-ra / -se* qui se serait manifestée d'abord dans les complétives, pour s'étendre ensuite aux relatives, puis aux conditionnelles, et finalement aux autres contextes syntaxiques (Pato 2003, 2004).

Fréquence de la substitution du subjonctif selon le type de proposition :

Complétives (72,1 %) > Propositions relatives et adverbiales
(manière, lieu et temps) (61,7 %) > Conditionnelles /
Finales (57,5 %)

L'ordre *complétives* > *relatives* > *conditionnelles* > *reste* est peut-être passé inaperçu parce que dans les propositions complétives et relatives de l'espagnol standard, les deux modes peuvent alterner sans que l'on perçoive nécessairement une différence très claire dans l'interprétation du contenu modal. Dans les propositions complétives, on explique habituellement le choix de l'un ou l'autre mode par la valeur de vérité présupposée de ce qui est exprimé dans la subordonnée, si c'est vrai, sûr (ce qui implique le choix de l'indicatif) ou non (avec le subjonctif) :

(6a) María sabía que Jaime vendría (*viniera) a visitarla
*María savait que Jaime viendrait (*vînt/*vienne) la voir*

[+ certitude]

(6b) María esperaba / no creía que Jaime vendría / viniera a visitarla.
*María espérait / ne pensait pas que Jaime viendrait / vînt /
viene la voir*

[+/- certitude]

(6c) María deseaba que Jaime viniera (*vendría) a visitarla
*María souhaitait que Jaime vînt/vienne (*viendrait) la voir*

[- certitude]

Dans l'exemple (6a), on sait que Jaime va venir, c'est pourquoi l'indicatif est requis. En (6c), il est incertain que Jaime vienne, le subjonctif est donc obligatoire. Et en (6b), on ne sait si Jaime viendra ou non, c'est pourquoi l'un comme l'autre sont possibles.

On retrouve des cas similaires avec les relatives et les adverbiales. Comme on le sait, l'alternance des modes s'explique en espagnol par le caractère plus ou moins spécifique de l'antécédent : s'il est spécifique, existant ou identifiable on utilise l'indicatif, et s'il est non spécifique, c'est-à-dire s'il n'est

*La Grammaire dialectale de l'espagnol à travers le Corpus
oral et sonore de l'espagnol rural*

pas sûr qu'il existe, le mode subjonctif est alors requis. Comme on le voit dans les phrases suivantes, le conditionnel et l'imparfait du subjonctif peuvent aussi apparaître tous deux dans des contextes d'interprétation incertaine (cf. 7b vs 7a,c) :

(7a) El hombre que sabría (*supiera) aquel misterio hacía tiempo, había desaparecido
*L'homme, qui connaîtrait (*connût) ce mystère il y a longtemps, avait disparu*
L'homme, qui connaissait ce mystère il y a longtemps, avait disparu

[+ spécifique]

(7b) El hombre que sabría / supiera aquel misterio hacía tiempo había desaparecido
L'homme qui connaîtrait / connût ce mystère il y a longtemps avait disparu
L'homme, qui connaissait ce mystère il y a longtemps, avait disparu

[+/- spécifique]

(7c) No hubo nadie que supiera (*sabría) aquel misterio
*Il n'y avait personne qui connût (*connaîtrait) ce mystère*
Il n'y avait personne qui connaissait ce mystère

[- spécifique]

En (7a) la relative est explicative, d'où le caractère spécifique de l'antécédent : par conséquent, le subjonctif est impossible. Dans la phrase (7c), l'antécédent est non spécifique, *nadie* « personne » suppose l'inexistence du référent : l'indicatif est ici agrammatical. La phrase (7b) est ouverte à une interprétation +/- spécifique de l'antécédent, ce qui permet par conséquent l'emploi de l'un ou l'autre mode.

Conformément à tout cela, on observe que, en castillan septentrional, plus l'antécédent est spécifique (défini et exprimé), plus l'indicatif est employé à la place du subjonctif, comme l'illustre l'échelle suivante :

Fréquence de la substitution du subjonctif selon le type d'antécédent dans les relatives

Antécédent défini explicite (*el N que*, 70,2 %) > Antécédent indéfini explicite (*un N que*, 60,3 %) > Antécédent défini non explicite (*el que*, 58,8 %) > Antécédent indéfini non explicite (*uno que*, 56,4 %) > *nadie que* (0 %)

Face à ces alternances de mode qui paraissent banales en espagnol, la présence du conditionnel à la place du subjonctif dans la protase de la proposition conditionnelle (*Si tuviera / tendría dinero, lo compraría* « Si j'eusse / j'aurais de l'argent, je l'achèterais ») a toujours été considérée comme un usage limité à certains dialectes. Dans la péninsule ibérique, ce tour a été et est encore aujourd'hui un stéréotype du parler des Basques, ce qui a peut-être conduit à situer l'origine du phénomène dans ce contexte syntaxique. Mais en réalité, en castillan septentrional, la protase conditionnelle n'est pas le contexte à l'origine du phénomène, mais plutôt la partie qui semble avoir été traditionnellement la plus « visible » pour les grammairiens et les dialectologues.

La perte du subjonctif selon l'échelle *complétives > relatives > conditionnelles > reste* pourrait probablement s'appliquer à d'autres variétés de l'espagnol, comme l'espagnol d'Amérique, voire à d'autres langues romanes. D'un point de vue typologique, il s'agit d'un changement prévisible, tant pour le mode que les temps concernés, le subjonctif passé et non le présent, puisque l'indicatif est moins marqué que le subjonctif et le présent moins marqué que le passé.

3.2.3 Accord neutre de matière, variables non existantes et implications typologiques

Dans les cas que nous venons de présenter, les enregistrements du *COSEER* ont permis d'étudier et d'élucider des phénomènes dialectaux connus, mais parfois mal compris. Mais le *COSEER* est encore plus intéressant puisqu'il a aussi conduit à découvrir des phénomènes dialectaux complètement ignorés jusqu'ici par les grammairiens et les dialectologues. Le meilleur exemple en est ce qu'on appelle traditionnellement l'accord neutre de matière. Cet accord était bien connu dans les parties centrale et orientale des Asturies et de la Cantabrie, mais était passé

*La Grammaire dialectale de l'espagnol à travers le Corpus
oral et sonore de l'espagnol rural*

inaperçu en Castille. Grâce aux enregistrements du *COSER*, la zone géographique de l'accord neutre de matière s'est considérablement étendue au sud (Fernández-Ordóñez 2007, 2006-2007) comme le montre la carte IX.



Carte IX. La zone géographique de l'accord neutre de matière

En quoi consiste l'accord neutre de matière ? Le « neutre de matière » est l'expression morphologique, pour les pronoms et les adjectifs qui s'accordent, de la catégorisation du nom en dénombrable (comptable) ou non dénombrable (non comptable). Il convient de ne pas confondre l'ensemble des accords appelé « neutre de matière » avec l'existence d'un troisième genre « neutre », puisque les nombres qui ont une interprétation non dénombrable sont, du point de vue du lexique, masculins ou féminins. Il s'agit d'un accord en termes de traits sémantiques, et non lexicaux. Ainsi, dans ces dialectes, le pronom neutre apparaît non seulement avec des antécédents non lexicaux (ce qui constitue la norme en espagnol), mais aussi avec des antécédents lexicaux dont le substantif masculin ou féminin (singulier ou pluriel) est interprété comme une matière, une entité non dénombrable. Cet « accord neutre » ne se limite pas aux pronoms, il s'étend aussi aux adjectifs post-nominaux (épithète ou attribut) et prédicatifs. Néanmoins, l'accord neutre de matière ne se produit jamais avec des éléments pré-nominaux, comme les articles ou les adjectifs, et rarement avec le substantif lui-même, comme le résume le tableau V.

Tableau V. Distinction des genres selon la nature du mot

Déterminants (articles et démonstratifs)	Nom	Adjectifs	Pronoms personnels et démonstratifs
MASC./FÉM.	MASC./FÉM.	MASC./FÉM./NEUTRE	MASC./FÉM./NEUTRE

- (8a) El buen vinu blanc-o se toma frí-o. Pruéba-lo
le.M bon.M vin.M blanc-N se boit froid-N. goûtes-y.N
Le bon vin blanc se boit froid. Goûtes-y
- (8b) La buen-a leche fresc-o se toma templad-o. Pruéba-lo
la.F bonne-F lait.F frais-N se boit tiède-N. goûtes-y.N
Le bon lait frais se boit tiède. Goûtes-y
- (8c) La-s medicina-s es car-o. Lo compramos en la farmacia
Les.F-PL médicaments.F-PL est cher-N nous l'.N achetons à
la pharmacie
Les médicaments sont chers. Nous les achetons à la pharmacie

En (8) nous voyons que l'accord « neutre » se produit avec les noms interprétés comme « matière », qu'ils soient au masculin ou au féminin, et au singulier comme au pluriel. On observe également que les adjectifs attributs et les adjectifs épithètes prennent l'accord neutre, mais cette situation maximale pour l'accord neutre de matière se limite aux Asturies. En Cantabrie et en Castille, il est pratiquement inexistant avec les adjectifs épithètes ; il est limité aux adjectifs et aux pronoms attributs.

D'autre part, la répartition statistique et géographique de cet accord en Asturies, Cantabrie et Castille (cf. Tableau VI¹¹) semble refléter le chemin de l'expansion de cet accord.

Tableau VI. Pourcentage de l'accord neutre de matière selon la position syntaxique et la nature des mots

Accord de matière (N fém.)	Déter- minants	Nom	Epithète	Attribut (<i>ser</i>)	Attribut (<i>estar</i>)	Prédica- tifs	Pronoms objets directs
Asturies	-	-	+ 29 %	+ 40,5 %	+ 58,7 %	+ 65 %	+ 86,4 %
Cantabrie	-	-	- 10 %	+ 35 %	+ 55 %	+ 59,3 %	+ 81,5 %
Castille	-	-	-	- 18,5 %	+ 51,2 %	+ 53,3 %	+ 76,5 %

¹¹ Dans ce tableau, le signe + indique 30 % ou plus de concordance neutre de matière et le signe - indique moins de 30 %.

*La Grammaire dialectale de l'espagnol à travers le Corpus
oral et sonore de l'espagnol rural*

Il s'est produit d'abord avec les pronoms (nécessairement situés dans une autre proposition que l'antécédent nominal) et il s'est graduellement étendu : tout d'abord aux adjectifs prédicatifs (qui sont justement structurellement dans une moindre cohésion avec le nom), puis aux attributs (situés eux aussi dans le prédicat mais à travers un verbe copule, sans dénotation lexicale). Enfin seulement, sont concernés les épithètes, les plus proches du nom. Ce parcours, de droite à gauche, explique que l'accord ne se manifeste jamais avec les éléments qui précèdent le nom : déterminants, quantificateurs et adjectifs prénominaux. Cet itinéraire semble prédire que l'accord neutre n'atteindra ces éléments que lorsqu'il sera pleinement établi sur la catégorie du nom, ce qui n'est pas encore le cas.

Ces faits sont particulièrement intéressants non seulement parce qu'ils permettent une meilleure connaissance de la grammaire dialectale, mais aussi parce que ce modèle de diffusion coïncide avec la Hiérarchie de l'Accord proposée par Corbett (1991, 2006) :

Agreement Hierarchy

attributive > predicate > relative pronoun > personal pronoun

For any controller that permits alternative agreements, as we move rightwards along the Agreement Hierarchy, the likelihood of agreement with greater semantic justification will increase monotonically (that is, with no intervening decrease). (2006 : 207).

Modèle de concordance neutre de matière dans les dialectes romans ibériques

épithète > attribut > prédicatif > pronom personnel

La recherche sur l'accord neutre de matière a ainsi révélé avoir des conséquences typologiques. Grâce à l'analyse détaillée et progressive des données que permettent les dialectes, il est maintenant possible de proposer un raffinement à l'« *Agreement Hierarchy* ».

En effet, on note une différence intéressante entre le pourcentage d'occurrences des adjectifs attributs avec la copule *ser* (attributs de propriété individuelle) et celui obtenu avec la

copule *estar* (attributs épisodiques). L'accord neutre de matière est définitivement plus fréquent lorsque les adjectifs dénotent un attribut épisodique. Ainsi, la position « attribut » de la hiérarchie pourrait être subdivisée en attributs primaires (attributs de propriété individuelle) et secondaires (attributs épisodiques), le deuxième étant le plus susceptible d'être soumis à l'accord sémantique. Par conséquent, les données romanes ibériques suggèrent de réviser comme suit l'« *Agreement Hierarchy* » afin de tenir compte de cette distinction :

Proposition de précision à l'« Agreement Hierarchy »

épithète > attribut de propriétés individuelles > attribut
épisodique > relatif > pronom

Quoi qu'il en soit, la grammaire dialectale est une mine, une source essentielle afin d'améliorer notre compréhension de nombreux principes interlinguistiques universels et ouvre de nouvelles voies pour tester leur validité.

3.2.4 Les atlas linguistiques et les corpus oraux sont complémentaires

Les corpus comme le *COSER* contribuent à l'approfondissement de la connaissance de la grammaire dialectale. Toutefois, bien que les atlas régionaux et les monographies dialectales montrent certaines faiblesses concernant l'étude de la grammaire, il ne faut pas négliger leur apport, même en ce qui concerne la morphosyntaxe. Lorsque la plupart de ces outils furent élaborés, ni la syntaxe ni la sociolinguistique n'avaient atteint le développement théorique qu'ils ont connu ces cinquante dernières années. Bien que la méthodologie du questionnaire employée dans l'*ALPI* et dans les atlas régionaux ultérieurs pour enregistrer la langue orale soit très différente de la méthodologie de l'enquête du *COSER*, on doit reconnaître qu'elles sont toutes deux le produit de l'état théorique de la dialectologie à leur époque. Le développement de la sociolinguistique a montré les limites de la méthodologie des atlas ; cependant, il est important de se rappeler que, puisqu'il n'y a pas d'enregistrement du langage dans le passé qui soit équivalent aux enregistrements actuels (et qu'il n'y a aucun moyen humainement possible pour les obtenir), les données des

*La Grammaire dialectale de l'espagnol à travers le Corpus
oral et sonore de l'espagnol rural*

atlas demeurent un précieux témoignage – aussi imparfait soit-il – pour l'étude du parler rural et de la grammaire, ainsi que l'ont montré certaines études (comme celles de Heap (2000) ou de Benito (2009), par exemple)¹². En outre, on ne rappellera jamais assez que la comparabilité des données d'un questionnaire est rarement obtenue par la méthodologie de l'enquête, dans laquelle les chercheurs peuvent essayer d'obtenir certaines données, sans jamais être complètement sûrs de réussir à atteindre leur but. C'est la raison pour laquelle des projets comme le *Dynamic Syntactic Atlas of the Dutch Dialects* (DynaSAND) (Barbiers 2006) utilisent d'abord des questionnaires qui sont ensuite complétés et précisés avec des enquêtes orales. D'autre part, les atlas linguistiques offrent de l'information que ne fournissent pas les corpus comme le *COSER*. L'enquête s'est révélée particulièrement utile pour enregistrer des phénomènes de nature grammaticale mais beaucoup moins pour le lexique. Etant donné qu'il s'agit d'une conversation semi-dirigée, les mots de nature dialectale recueillis ne sont pas toujours répétés ; pour ce qui concerne le vocabulaire, on ne peut tirer aucune conclusion comparable à celle d'un atlas. Le *COSER* constitue donc un complément au matériel publié dans les atlas et dans d'autres types de sources dialectales, un complément qui ouvre des perspectives d'avenir enrichissantes pour l'étude de la grammaire dialectale. Au-delà de l'intérêt en phonétique et en lexicologie, les dialectologues savent que la grammaire dialectale est une source d'information précieuse – et jusqu'à maintenant sous-estimée – non seulement pour la caractérisation d'un domaine linguistique en particulier, mais aussi pour l'étude typologique des langues (voir Kortmann 1999, 2004a, 2004b).

12 A cet égard, les matériaux de l'*ALPI* seront précieux puisque c'est le seul atlas de toutes les langues romanes de la péninsule. Fort heureusement, bien qu'un seul volume ait été publié (cf. Navarro Tomás 1962), ces données sont en partie disponibles sur Internet (cf. Heap 2002, 2003-). De plus, le projet coordonné par Pilar García Mouton, « Edición y elaboración de los materiales del *Atlas Lingüístico de la Península Ibérica* de Tomás Navarro Tomás », permettra dans quelques années de mettre à disposition de la communauté scientifique le contenu des questionnaires d'enquête, avec la possibilité de réaliser des cartes interactives sur Internet.

Références bibliographiques

- ALCyL* : Alvar M. (1999). *Atlas lingüístico de Castilla y León*, 3 vol. Valladolid : Junta de Castilla y León.
- ALEANR* : Alvar M., *et al.* (1979-1983). *Atlas lingüístico y etnográfico de Aragón, Navarra y La Rioja*, con la colaboración de Antonio Llorente, Tomás Buesa y Elena Alvar, 12 vol. Zaragoza : Institución Fernando el Católico / CSIC.
- ALEcant* : Alvar M., *et al.* (1995). *Atlas lingüístico y etnográfico de Cantabria*, 2 vol. Madrid : Fundación Marcelo Botín.
- ALEcMan* : García Mouton P. & Moreno Fernández F. (dir.). *Atlas Lingüístico (y etnográfico) de Castilla - La Mancha*. Universidad de Alcalá : <<http://www2.uah.es/alecman>>
- ALPI* : Navarro Tomás T. (dir.) *et al.* (1962). *Atlas lingüístico de la Península Ibérica*, vol. 1, *Fonética*, con la colaboración de Francesc de Borja Moll, Aurelio M. Espinosa [junior], Luís F. Lindley Cintra, Armando Nobre de Gusmão, Aníbal Otero, Lorenzo Rodríguez-Castellano and Manuel Sanchis Guarner. Madrid : CSIC.
- Barbiers S. *et al.* (2006). *Dynamic Syntactic Atlas of the Dutch dialects (DynaSAND)*. Amsterdam : Meertens Institute <<http://www.meertens.knaw.nl/sand/>>.
- Benito C. de (2009). Descripción y análisis del pronombre concordado con el sujeto en el *Atlas Lingüístico de la Península Ibérica*. Memoria de Máster, Universidad Autónoma de Madrid.
- Chambers J. K. (1995). *Sociolinguistic theory*. Oxford : Blackwell.
- Corbett G. C. (1991). *Gender*. Cambridge : Cambridge University Press.
- Corbett G. C. (2006). *Agreement*. Cambridge : Cambridge University Press.
- COSER* : Fernández-Ordóñez I. (dir.) (2005-). *Corpus oral y sonoro del español rural*. Madrid : Universidad Autónoma de Madrid <<http://www.uam.es/coser>>.

*La Grammaire dialectale de l'espagnol à travers le Corpus
oral et sonore de l'espagnol rural*

- Cuervo R. J. (1895). « Los casos enclíticos y proclíticos del pronombre de tercera persona en castellano », *Romania* 24 : 95-113, 219-263.
- Company C. (2009). « Artículo + posesivo y estructuras afines », in C. Company (dir.), *Sintaxis histórica de la lengua española*, Segunda parte : La frase nominal. México : Universidad Nacional Autónoma de México / Fondo de Cultura Económica, vol. 1, 759-880.
- Fernández Ramírez S. (1987²). *Gramática española*. 3.2. *El pronombre*. Madrid : Arco/Libros.
- Fernández-Ordóñez I. (1993). « Leísmo, laísmo y loísmo : estado de la cuestión », in O. Fernández Soriano (éd.), *Los pronombres átonos*. Madrid : Taurus, 63-96.
- Fernández-Ordóñez I. (1994). « Isoglosas internas del castellano. El sistema referencial del pronombre átono de tercera persona », *Revista de Filología Española* LXXIV : 71-25.
- Fernández-Ordóñez I. (1999). « Leísmo, laísmo y loísmo », in Ignacio Bosque, Violeta Demonte (éd.), *Nueva gramática descriptiva de la lengua española* (3 vol.). Madrid : Espasa-Calpe, vol. 1, 1317-1397.
- Fernández-Ordóñez I. (2001). « Hacia una dialectología histórica. Reflexiones sobre la historia del leísmo, el laísmo y el loísmo », *Boletín de la Real Academia Española* LXXXI : 389-464.
- Fernández-Ordóñez I. (2007). « El neutro de materia en Asturias y Cantabria. Análisis gramatical y nuevos datos », in Inmaculada Delgado Cobos y Alicia Puigvert Ocal (éd.), *Ex admiratione et amicitia. Homenaje a Ramón Santiago*. Madrid : Ediciones del Orto, 395-434.
- Fernández-Ordóñez I. (2006-2007). « Del Cantábrico a Toledo. El “neutro de materia” hispánico en un contexto románico y tipológico », *Revista de Historia de la Lengua Española* 1 : 67-118 ; 2 : 29-81.
- Flores Cervantes M. (1997). « Individuación de la entidad en los orígenes de *leísmo*, *laísmo* y *loísmo* », in C. Company (éd.),

- Cambios diacrónicos en el español*. México : Universidad Nacional Autónoma de México, 33-63.
- Flores Cervantes M. (2002). *Leísmo, laísmo y loísmo : sus orígenes y evolución*. México : Universidad Nacional Autónoma de México.
- Flores Cervantes M. (2006). « Leísmo, laísmo y loísmo », in C. Company (dir.), *Sintaxis histórica de la lengua española*, Primera parte : La frase verbal. México : Universidad Nacional Autónoma de México / Fondo de Cultura Económica, vol. 1, 671-749.
- García E. C. (1975). *The role of theory in linguistic analysis. The Spanish pronoun system*. Amsterdam : North-Holland Publishing Company.
- Heap D. (2000). *La variation grammaticale en géolinguistique : les pronoms sujet en roman central*. München : Lincom Europa.
- Heap D. (2002). « Segunda noticia histórica del ALPI (on the 40th anniversary of the publication of its first volume) », *Revista de Filología Española* LXXXII : 5-19.
- Heap D. (2003-). *Atlas lingüístico de la Península Ibérica. ALPI searchable database*. London, Ontario : University of Western Ontario <<http://www.alpi.ca>>.
- Heap D. (2006). Secuencias « invertidas de clíticos : un cambio (¿?) en tiempo real », in José Jesús de Bustos et José Luis Girón Alconchel (éd.), *Actas del VI Congreso Internacional de Historia de la Lengua Española* (Madrid, September 29 - October 3, 2003). Madrid : Arco/Libros, I, 785-98.
- Klein-Andreu F. (1979). « Factores sociales en algunas diferencias lingüísticas en Castilla la Vieja », *Papers : Revista de Sociología* 11 : 46-67.
- Klein-Andreu F. (1981). « Distintos sistemas de empleo de ‘le, la, lo’ : perspectiva sincrónica, diacrónica y sociolingüística », *Thesaurus* XXXVI : 284-304. (Reímpreso in O. Fernández Soriano (éd.), *Los pronombres átonos*. Madrid : Taurus, 1993, 337-353).

*La Grammaire dialectale de l'espagnol à travers le Corpus
oral et sonore de l'espagnol rural*

- Klein-Andreu F. (2000). *Variación actual y evolución histórica : los clíticos le/s, la/s, lo/s*. München : Lincom Europa.
- Kortmann B. (1999). « Typology and dialectology », en B. Caron (éd.), *Proceedings of the 16th international congress of linguists*, CD-ROM. Amsterdam : Elsevier Science.
- Kortmann B. (2004a). « Why dialect grammar matters », *The European English Messenger* XIII : 24-29.
- (éd.) (2004b). *Dialectology meets typology. Dialect Grammar from a Cross-Linguistic Perspective*. Berlin / New York : Mouton de Gruyter.
- Lapesa R. (1968). « Sobre los orígenes y evolución del leísmo, laísmo y loísmo », in K. Baldinger (éd.), *Festschrift Walther von Wartburg*. Tübingen : Max Niemeyer, 523-551. (Réimprimé in R. Lapesa, *Estudios de morfosintaxis histórica del español*. Madrid : Gredos, 2000, 279-310).
- Martínez Martín M. (1983). « La sustitución de *cantara* (*cantase*) por *cantaría* en el habla de la ciudad de Burgos », *Lingüística Española Actual* V/2 : 179-204.
- Pato E. (2003). « Contextos neutralizadores de la oposición modal y relaciones de alomorfismo desde el español medieval: Las formas *cantase*, *cantara* y *cantaría* », *Moenia* 9, 223-252.
- Pato E. (2004). *La sustitución de cantara / cantase por cantaría / cantaba (en el castellano septentrional peninsular)*. Madrid : Universidad Autónoma de Madrid. <http://joule.qfa.uam.es/coser/publicaciones/enrique/2_es.pdf>
- Pato E. (2009). « Nivelación lingüística y simplificación : el uso de preposición + *tú* en la historia de la lengua », in E. Montero *et al.*, *Actas del VIII Congreso Internacional de Historia de la Lengua Española* (Santiago de Compostela, 14-18 de septiembre de 2009), sous presse.
- Pato E. & Heap D. (2009). « Plurales anómalos (el morfema verbal *-n*) en los dialectos y en la historia del español », in E. Montero *et al.*, *Actas del VIII Congreso Internacional de Historia de la Lengua Española* (Santiago de Compostela, 14-18 de septiembre de 2009), sous presse.

I. FERNANDEZ-ORDOÑEZ

- Ridruejo E. (1975). « *Cantaría* por *cantara* en la Rioja », *Berceo* 89 : 123-134.
- Silva Corvalán C. (1985). « Modality and semantic change », in Jacek Fisiak (éd.), *Historical semantics. Historical word-formation*. Berlin / New York : Mouton de Gruyter, 547-572.
- Zubizarreta M. L. (1999). « Las funciones informativas : tema y foco », in I. Bosque, V. Demonte (éd.), *Nueva gramática descriptiva de la lengua española* (3 vol.). Madrid : Espasa-Calpe, vol. 3, 4215-4244.