

L'édition et le marquage de textes de l'oralité asturienne : Une fenêtre sonore sur l'asturianité négociée.

Pour Paul Rivenc

Résumé: *compte tenu d'une structuration dialectale normale de l'asturien courant, les locuteurs asturiens font preuve d'une relative flexibilité adaptative dans le traitement de l'information: d'une part ils parlent quasi automatiquement leur propre variété d'asturien marqué, central, occidental et oriental (las serdas, mataban); et d'autre part maîtrisent une normalité centrale négociée (les serdes, mataban): c'est ce type de problématique de la variation inter- et intra-individuelle qui est ici abordé en linguistique de corpus asturien parlé transcrit, actuel et contemporain. La variation réglée de la sorte se répartit sur l'ensemble des catégories du lexique oral classes ouvertes.vox et classes fermées.vox (confondues ou non). On fait ici porter l'attention plus sur l'organisation normée de la prédication verbale orale, restituée d'après une annotation transcrite de l'oralité et sur l'approche pronominale que sur les données lexicales situationnelles, absolument triviales. La variation dialectale étant donnée au plan du corrélat phonétique du plan phonologique, seule est mise ici au coeur même de l'importance l'acquisition des structures prédictives; le tout traité par un système flexible qui facilite l'acquisition et les apprentissages du savoir dire. Autant qu'une langue, l'asturien est alors considéré comme univers, autrement dit totalité organisée par la diversité même, architecturée.*

Cette brève et modeste réflexion autour des classes et catégories de l'asturien envisagées en rapport avec son histoire sociale est conduite à partir de corpus oraux rétablis à partir de leur présentation écrite; elle porte sur le rapport entre données de terrain et théories énonciatives des analyses de discours, plus que sur les aspects proprement technologiques, qui toutefois ne sont pas négligeables, à condition que les méthodes de traitement sélectionnées soient adéquates aux objectifs; elle porte globalement et de manière synthétique sur les recherches possibles à partir du «*material de campu*» constituant un corpus de textes sonores traités en termes de catégories de langue et de catégories discursives.

Il s'agit de données audibles annotées, vérifiables dans leur homogénéité/hétérogénéité renvoyant aux conditions et contraintes d'usage courant; oraux dans leur recueil par des linguistes et ethnographes asturiens, ces textes ont été dotés d'une face écrite constituant la face visible à analyser, trace d'un projet politique de normativité; la préoccupation de base d'un linguiste romaniste applicateur, que le matériel de base soit occitan, asturien, aragonais, aranais, français ou castillan porte essentiellement sur les rapports entre normalité langagière (systématicité, règles) et diversité des réalisations de surface, aux divers plans qui impliquent la phonéticité en morphologie lexicale et syntaxe; certes cette réflexion ouvre apparemment une fenêtre sur des mondes lexicaux renvoyant à des contextes situationnels picaresques explicitement décrits dans le détail ; cette fenêtre ouvre sur les réalités sociales qui ont constitué historiquement en Europe, des sociétés rurales traditionnelles en mouvement; ce mode d'approche pose plus particulièrement la question de l'organisation d'un travail programmatique interdisciplinaire immense d'analyse des rapports entre normes (contraintes, règles) et plan d'analyse de la variation géographique inter et intra-individuelle qui s'ensuit; en effet il existe au moins deux sortes

de linguistes; les uns pour qui les règles sont, indépendamment de tout calcul; les autres pour qui les règles voient leur statut et leurs corrélats catégoriels éclairés par le calcul des poids, les densités; dans les deux cas les choses, à l'arrivée, ou tout au moins en cours de cheminement depuis les formes initiales de la matière, sont de l'ordre du calcul; il faut ajouter que cette réflexion est née en grande part de l'émerveillement toujours aussi intense qu'a su provoquer en moi Paul Rivenc, lorsque dès les années 60, avec Marie-Madeleine Rivenc-Chiclet, il ouvrit mes jeux de dialectologue et de géolinguiste aux méthodes des linguistiques travaillant à partir de corpus de données orales; depuis, cette merveilleuse «fenêtre» sur les mondes dialectaux discursifs réels, jamais opposés aux mondes formels et encore moins aux conditions institutionnelles, ne s'est refermée; nous avons appris ici qu'une fois classés correctement les faits d'oralité calculables à partir du signal et des traces acoustiques décelables, l'essentiel de l'analyse linguistique s'en trouvait largement préparé pour la suite des opérations nécessaires, à condition que le lexique soit traité en lexique phonologique sur base phonétique, dans son environnement syntaxique. Au départ une sorte d'insurrection contre l'écrit et les représentations de l'écrit achevé, éditorial; à l'arrivée, une réflexion critique sur les rapports dialectiques entre oralité (nécessairement inachevée) et écrit édité: ce qui est recherché est la mise en place d'une solution alternative à une telle conception binariste tranchée.

Autour du verbe et du pronom

L'hypothèse de base que nous développerons ici est que le lexique thématique italicisé dans les corpus asturiens écrits soumis à diverses méthodes de traitement de données, tant nominal que verbal dans la base de données textuelles indexée, constitue le donné préférentiel de chaque groupe, non de chaque individu, d'une part; et, second point, que le matériel lexical initial n'est pas comptabilisé en tant que tel dans le calcul d'une langue a minima, qui ne prend en compte que les faits de construction, et de structuration des plans (genèse et développement de structures); c'est la phonétisation systématique de l'asturien syntaxiquement et sémantiquement organisé qui constitue le premier plan visible à partir duquel se constituent les faits de démarquage visibles dans le donné morpholexical et morphonologique : *llantar/chantar / l.lantar*, pour nous limiter de manière exemplariste à un seul objet donnant statut à la langue, par la catégorisation grammaticale. Le lexique accédé, pour la psycholinguistique contemporaine est organisé catégoriellement en classes ouvertes (lexique thématique) et classes fermées (classes grammaticales); le lexique nominal thématique est forcément contraint par la thème et la situation: *gocho; capadura / capia*; quant au lexique verbal thématique de la catégorie lexicale ouverte du verbe il est également contraint par la consigne situationnelle; *criar; matar; capar; salar; cortar; afumar; embutir* etc.; c'est pourquoi il est constitué en index séparé, donné à l'accès; les objets de ce lexique ne sont constitués en objets linguistiques que lorsqu'ils ont été correctement subcatégorisés (verbes de mouvement, verbes de dire, verbes causatifs et anticausatifs; voix pronominale; d'une part; et d'autre part la reconnaissance du statut verbal implique une approche pronominale syntaxique et prosodique absolument cruciale pour la détermination du statut d'asturianité; une hypothèse de départ complémentaire est que le lexique grammatical complexe accédé est accédé en bloc de lexies et non par unités lexicales séparées, autrement dit par blocs subcatégorisés (prépositions, conjonctions, adverbes etc.). Si l'on privilégie une approche verbale de l'oralité des langues romanes, c'est en corrélation étroite avec une approche de caractère pronominal.

C'est Paul Rivenc, le linguiste applicateur, le didacticien, et l'hispaniste, qui, un des premiers, dès les années 60, m'a donné le goût de l'analyse professionnelle telle que pratiquée depuis par les linguistiques de corpus; et singulièrement de corpus oraux de langues romanes, recueillis sur base énonciativiste, autant sinon plus plus que lexicaliste; P. Rivenc a un temps cheminé aux côtés d'une équipe pédagogique de Romanistique à l'Université de Toulouse rompue à l'analyse des sources de la tradition orale; puis s'en est détaché; pensant sans doute que cette équipe se débrouillait fort bien sans lui

et sans les outils et concepts de la linguistique appliquée aux sciences des textes, oraux envisagés dans leur complexité, leur genèse, leurs représentations véhiculées, leurs formes, leurs propriétés syntaxiques et sémantiques; je ne l'ai jamais pensé et cette modeste analyse entend démontrer que je persiste et signe: il y a quelque nécessité à réviser les modèles d'une géolinguistique «fermée» appliquée élaborée à partir de corpus écrits asturiens produits par transcription de témoignages (je ne retiendrai pas *bable*, trop près de babillage); la question est soit de décrire le parler de locuteurs; soit de répondre à la question de savoir de quel parler moyen se servent les locuteurs auditeurs, lorsqu'ils se parlent en utilisant les moyens asturiens; question de haute utilité, qui repose sur le traitement de la parole lorsqu'elle est associée à un signal sonore dont les traces acoustiques sont facilement décelables et mesurables; mais les textes transcrits à partir de la perception du signal ne nous permettent pas de manière égale de remonter en tant que lecteurs au signal vrai; ils sont par ailleurs en apparence hétérogènes même si une contrainte de description encyclopédique leur confère une apparente homogénéité lexicale et terminologique: ce sont des récits de pratiques structurées dans le temps, l'espace et le corps social, ils sont caractérisés par des propriétés différentielles dont la diversité n'est ni illimitée, ni incommensurable. Toutefois cette diversité culturelle, plus ou moins nettement territorialisée, ne se mesure pas seulement en un instant (t) de l'actualité, parce que la diversité est produite dans le temps historique des interactions entre langues devenues maternelles, à des degrés divers et à des époques et moments divers de l'évolution des langues pour un Etat donné, caractérisé par sa politique linguistique des origines à nos jours. Si les noms donnés par les hommes à ce qui constitue leur milieu, aux prés, aux champs, aux vignes, au bétail, au porc en asturien ont une importance, ce n'est pas en tant que tels, mais en tant qu'ils sont porteurs d'indices d'asturianité fragmentée ou éclairent le processus de désasturianisation des sujets adultes, comme des sujets jeunes, la désasturianisation frappant d'abord les classes de sujets jeunes, du fait des modifications des conditions de transmission sociale de l'asturianité en tant que qualité de parler en langue culturelle stable. Quelques remarques préalables ne seront pas inutiles avant la mouture et la mise en farine, comme on dit en asturien (*mete_se en farina; fariña blanca / blanco...*)

Éléments pour une histoire sociale de l'asturianité

Si j'en suis venu, de mes positions de retranchement sur la gasconité en domaine occitan, à m'interroger sur l'asturianité en domaine ibéro-roman, il y a cela une bonne raison: celle de trouver dans ces deux grandes réussites en monde culturel ibéro-roman et gallo-roman, à la charnière même de ces deux mondes culturels caractérisables par leurs propriétés et caractéristiques territoriales, de quoi renforcer l'hypothèse selon laquelle il faudrait s'attendre, selon la formule de Paul Valéry, à ce que les (ces) mondes culturels futurs formés par les sociétés rurales soient, ici et là, de manière non totalement imprévisible, ce qu'ils sont été hier, en Aquitaine, en Septimanie, en Asturies, en Cantabrie etc. Travailler à partir de données dialectales en termes de linguistique de corpus c'est ouvrir une «fenêtre» sur les mondes possibles structurant dans sa base et son devenir la territorialité d'une langue à variétés normées; des changements importants sont intervenus dans les structures productives de la société rurale: ils constituent des mutations dans la base même, politique et économique (élevage et abattage); ces changements ont frappé de plein fouet les hommes en milieu rural dans les bases mêmes d'une économie de subsistance; si l'on peut déceler dans les témoignages les traces décelables de telles mutations, ils apportent en revanche la preuve de l'enracinement de ces pratiques en rapport avec le plan de la politique linguistique, en milieu asturien

Ce qui, pour cette analyse, prime à mes yeux n'est pas tant le coefficient optimal d'asturianité (95 %) dans les familles nucléaires traditionnelles, opposé à un indicateur minimal de castillanité (< 5%) que le rapport variable entre les deux, constituant un différentiel gradué (gradient d'asturianité langagière). Aussi la question cruciale, bien difficile, et encore plus complexe, se ramène-t-elle à une question d'apparence simple

et toute nue: comment apprend-on (qui, quels hommes, dans quel milieu, à parler une langue particulière, ici l'asturien, langue territoriale de société minière, rurale, *marinera*, d'abord dans une famille nucléaire traditionnelle asturienne, puis quand elle est, sous le poids de facteurs innovants, en interaction avec une autre langue particulière, ici l'espagnol ? sachant que cette question relative à l'acquisition des structures du parler, se distingue de la question complémentaire des structures de la langue écrite: on sait parler gascon, asturien, etc. mais on ne sait pas l'écrire, parce qu'on n'a pas appris à l'écrire, ce qui relève d'une autre discipline; celle de l'analyse des rapports entre représentations, langages et modes de communication.

C'est dans cet état d'esprit que j'aborde ici le cas du traitement linguistique robuste, phonologique et morphosyntaxique, de textes culturels asturiens plus ou moins hétérogènes, recueillis par l'ethnographie asturienne, à propos de la matière de la *matanza del gocho* en territoire asturien: ces voix asturiennes ordinaires plus ou moins localisées, si elles témoignent de profondes mutations intervenues dans les dernières décades dans la société asturienne, sous ses formes et contraintes économiques, témoignent en revanche de la permanence de la langue asturiennes; les témoignages asturiens recueillis sont, il est vrai, rapportés, indirectement, et commentés en écho par des enquêteurs collecteurs; ceci pose à l'analyste soucieux d'approcher une représentation optimale de la situation actuelle de l'asturien de réelles difficultés sous l'aspect de la problématique du traitement en géographie linguistique appliquée de ce type de données; ces données sont au moins complexes, sinon hétérogènes selon les apparences, en raison des rapports entre normalité et dialectalité; on accordera la toute première priorité de méthode à la discrimination de ce qui relève du donné d'origine (la matière asturienne) et de ce qui est du ressort du commentaire savant ou semi-savant des «prédicats ontologiques d'origine». Disons d'entrée de jeu qu'une telle approche sociolinguistique des formes prises par l'interaction dans la société hispanique en territoire asturien est, en apparence tout au moins de nature à minimiser les préoccupations d'ordre ethnographique, folklorique; certes il faut ne pas ignorer la problématique des interactions entre hommes, milieux, économie et société si on entend éclairer et mieux comprendre la problème de l'évolution de l'asturianité vers la castillanité.

Autour du nom asturien

Les recherches appliquées au domaine du nom incluent une approche lexicographique et lexicologique mais ne se limitent pas à ces deux modes d'approche du mot, conduite en termes d'histoire sociale de la langue.

Il est une première étape de l'investigation en linguistique conduite à partir de corpus, qui consiste à isoler le sous-ensemble nominal (nom, adjectif, syntagme nominal sujet et objet); ce lexique est d'une part contraint pas la contrainte situationnelle; d'autre part, indépendamment de cette contrainte il est porteur des indices de variation structurante du plan phonétique (*pellejos*), corrélat d'un plan phonologique; il faut savoir aussi comment et où, dans quel module, à quelle phase de l'accès lexical se discriminent les sub-catégories sémantiques: *adobo / adobar / matar/mata*): ces dénominations d'usage de très haute fréquence situationnelle, n'est pas un nom pour un linguiste, mais un nom verbal (*adobar*); et revanche le champ des désignations de haute fréquence utilisant le nom fait apparaître des corrélats sémantiques intéressants dans la catégorie nominale *Sw'anes*: *Sw'an El r: 'iku / Sw'an El proBe*; et l'hyperonyme qui englobe facétieusement les deux classes confondues : *los koj'ones*. Mais ce n'est pas tant la question lexicale nominale qui est ici en question, - le lexique sélectionné est trivialement propre au domaine ou champ d'expérience -, que les phénomènes plus généraux de représentation, de généralisation, d'acquisition et de transmission des contraintes spécifiques aux deux langues en interaction. ce qui intéresse l'analyse c'est la proportion de contraintes spécifiques observées, leur rapport: la grammaire du SN *farina blanco* enfreint en principe et en surface les contraintes d'accord en genre dans le GN.

En soi l'information associée à l'unité lexicale nominale d'usage *rabadal*, *tripa* n'intéresse pas l'analyse à proprement parlée; pas davantage *untu* «grasa» mais non toutes «les sortes de»); cette information est constituée par la masse des données lexicales déterminées sur le terrain, en rapport avec un ensemble de contraintes morphologiques; l'information fait partie du donné trivial qui doit avoir été au préalable déterminé en méthode lexicographique. De même, fait partie du donné brut initial le polymorphisme phonique censé connu depuis la description des mots phonologiques de la matière initiale donnée (réalisations de la phonéticité asturienne); (mais *venga/venca*; *tenga/tenca* relève-t-il de la seule variation phonétique ?) de telles interrogations montrent que la discrimination des divers plans ne va pas sans faire problème en termes de normalisation et stabilisation des textes de l'oralité qui se laisse ramener à l'existence (hypothétique) de plans alternatifs de surface: (a) *les oreyes* / (b) *les orielles* (c) *las oureichas* .; ce plan de la variation régionale, au sein même de la catégorisation, n'affecte nullement, heureusement, la catégorisation elle-même. La question posée initialement peut alors être complétée par une importante question subsidiaire; quel degré de rigidité, caractérise l'acquisition, quel degré de flexibilité, non de la langue, mais des sujets qui l'apprennent dans des conditions historiques très particulières, par étapes, progressivement, jusqu'à un degré optimal, minimal, de savoir dire? Comment un système à règles (*rabo*: *rabadal*: *rabadiel.lu* ; *tripa* : *tripu*, *untu* / *untazas*, par exemple; ou le paradigme *torga*, *torgar*, *torgada* / *torgáu*) intègre-t-il la variation - qu'elle soit non aléatoire [o/u] prétonique, traitement par réduction syllabique, allongement de syllabe tonique, changement prosodique trochée / schéma iambique - ou qu'elle comporte une part d'aléatoire (distribution apparemment aléatoire du polymorphisme roman [o/u] en position prétonique (occitan, asturien). on aura vite fait de débusquer les comportements phonétiques variables de surface les plus stables, les plus caractéristiques. Parmi ces caractéristiques, comptent les caractéristiques prosodiques segmentales; *fegadu* (astur) est de même statut prosodique que cast. *hígado*: propatoxyton, et ceci est une constante prosodique. Voilà de quoi donner du grain à moudre à l'intersection des mathématiques et des sciences humaines, loin de tout grand spectacle médiatique. En fin de compte que faut-il sauver ? la variation, la diversité culturelle, ou les Sciences Humaines qui la chassent et la débusquent pour mieux la comprendre ? quelle meilleure forme de compréhension substituer à une conception de patoisement du castillan: un castillan déformé, un patois du castillan, conception longtemps propagée, on le sait, en Espagne même, comme en France? ou une langue définie à l'intersection des axes de la culture et de l'histoire des civilisations ibériques ?

La question se complique du fait de la volonté d'écrire selon une normalité de correction l'asturien minimal, ce qui est du fait des décisions prises par les rédacteurs; qui dit volonté et décision se situe entre plan de la description des réels et politique linguistique.

Il y a en effet de fortes chances pour que les outils de comparaison actuellement utilisés, comme les linguistes comparatistes eux-mêmes, comparent ici des données fort hétérogènes, s'ils ne prennent pas en compte les éléments qui constituent le paramètre énonciatif: l'informateur dit (a); le philologue normateur écrit (b) là où il a recueilli un donné (a) qu'il traite en optimisant et réduisant le caractère aléatoire de la variation dialectale. Cet ordre de choses minimales relève de la relation entre peuples et cultures, donc, à ce titre, de la politique linguistique des Etats. En revanche des suites grammaticales inventoriées comme blocs *da se_jos*; *a les mujeres, tocava_yos traballar_lo*; *untando_j_lu* ; *carne ensangrientado / otra amasadura* etc. relèvent strictement de la description et de l'analyse linguistique de l'organisation spécifique de la matière asturienne de base en termes catégoriels (classes ouvertes et classes grammaticales fermées du lexique asturien) et règles. Loin de prétendre ici à renouveler en profondeur la description ou l'analyse, je voudrais flécher certaines orientations de l'analyse pour servir à une linguistique dialectale appliquée des langues romanes dans leur stabilité: on accordera la priorité à la hiérarchisation des préoccupations (a) données externes: *sopa de f'egadu* (b) données

internes construites, soumises à analyse ; par exemple les expressions normées à infinitif prépositionnel construit avec sujet exprimé et postposé: *pa cenar los homes; syntaxe d'accord dans le SN carne insangrentado / otra amasadura*

D'emblée disons que le linguiste applicateur n'a pas en tant que tel vocation à formuler des propositions relatives à la crise de la filière porcine ici ou là, en France, dans les Asturies; et pas davantage à philosopher sur les désignations traditionnelles données du *cabu mundu* (astur.), forme traduite retenue pour *estantin ciegu*, *tripa ciego*, *cap del monde* (occitan), qui constituent une partie de la matière initiale. Les unités de ce répertoire spécialisé ont été italiquées, pour être extraites du corpus d'observation de la stabilité du construit syntaxique, prosodique, sémantique. Ce qui compte pour l'analyse c'est la comparabilité de données phonologiques et morphologiques répétées jusqu'au rassasement à satiété. une fois donné unu, vite on est conduit à comparer des gabarits morphosyntaxiques comparables sous l'aspect du gabarit même: les trois suites syntaxiques *l'unu, n'unu, et xuntu*, et de même *unu*: *l'unu / n'unu, d'unu* sont des entrées de même gabarit pour deux (n) données lexicales discriminées: *untu, uno* et *xuntu*, soit dans les classes lexicales ouvertes, soit dans les classes grammaticales fermées du lexique (*xuntu, d'unu, n'unu*).

Material de campu : terrains, théories et technologies de manipulation

La matière initiale du corpus constitué et mis par écrit concerne jusqu'à satiété et écoeurement, selon les points de vue, la place du porc (élevage et fabrication) dans l'économie familiale de la société rurale asturienne jusque dans les années 60-80 du 20ème siècle; c'est un corpus «pur porc», comme l'Antigone de Sophocle un corpus pur sang; l'homologie entre les deux plans n'est pas du tout fortuite; les deux corpus sont joués, ils se jouent dans des contraintes institutionnelles tragi-comiques homologues et caractérisent ici la culture asturienne comme Antigone caractérisait la culture populaire athénienne. Sans doute pour un linguiste, ces contraintes situationnelles rendent-elles le corpus quelque peu lassant; mais cette répétition navrante à pour corrélat une propriété intéressante: elle facilite la comparaison des observables aisément étalonnés au plan énonciatif à dominante descriptive; et par ailleurs tant au plan des réalités qu'au plan de l'expression, la collection de témoignages rassemblés (témoignages rapportée, plus ou moins directement, de sujets adultes, des classes d'âge, - n'excluant pas les productions de jeunes interprètes - *capar los gochos* dans les jeux !) met en scène les facteurs de cohésion d'une société soudée, qui se fait et se défait avec les mouvements sociaux; on y entrevoit souvent ce que pouvait être la réalité sociale et les modes d'expression dans les sociétés minières asturiennes jusqu'aux portes du 21ème siècle. Il sera par conséquent aisé aux linguistes des linguistiques de corpus, aux sémanticiens de toutes sortes, d'ironiser à bon marché sur la portée patrimoniale des collections que j'interroge non pour leurs contenus, mais pour leurs formes mêmes (plan de l'expression) qui organisent les représentations; ce sont ces représentations, y compris les représentations des faits de langue qui assurent la cohésion d'une forme de société.

Voici offert à Paul Rivenc en forme de muestra et de prueba _ pa probar, un bref échantillon de la base manipulée :

la S'ente ke tj'eù g'otSa kr'ija kujn'inos / pOr_ke soluTj'ona la papel'ata as'ina / i la ke n'on / k'Ompra_lo < de tS'ikos > .

les k'otSes paridj'eGes tj'eneù_ke []tar muj bj'en TeB'a[]es / i kw'enta mio Gw'ela De bOrBud'iu ke

< paiT'ija ke nun s enLen'aBeù kOn n'a: > / un eSemplu De < lo pereT'i[]es > ke [] t'aBen je [] ke s'iGe D_ar:ew / em bOrbud'in / ten'ijen una d estes g'otSes n_uù kuB'il ke ten'ija una r:endiS'ina pEr_de_b'aSo la pw'Erta / i pEr eL'i metjo_se una t'arde ke LuB'ija i faT'ija fr'iw / uù kOrder'in:r:ujn'in ke []t'aBen krij'ando a_mamona eù_k'asa

/ kwando f'oren a TeB'ar la gotS'ona / atop'aron los GweT'inos de < SakoB'in > / _ke as'ina_j Lam'aBen al iMfel'iT kOrd'eru / les g'otSes de kr'ija k'omen ma'iT / LaB'aTa kOM far'ina De ma'iT / pan'ojes um_poko B'erdes / pat'ates koT'i[]es i lo ke p'iLen / mjo m'aDre kw'enta_me k en t'araNu tamj'en_jos 'etSen r:emol'atSa / kalaB'aTa / kast'aNes i b'erTes / D_eLes_beg'a[]es iMf'esten_se les k'ases de r:atos / pOr_ke nun se_jos pwe[] [e]tS'ar bel'enu si Ban al kuB'il de los g'otSes / i 'estos los k'omen / i embel'enen_se depw'es /

kwando una g'otSa p'are / aj_ke kudj'a_la unes nw'etSes pa_ke num m'ate los f'ijos / etS'ando_se en_r:iBa_D 'eLos / sin da_se_kw'enta / [pOrke los gotS'inos 'anden komo fOrm'iGes rebOlt'oses per ente la m'aDre / i kw'er:eù pel'iGru as'ina /]

los kujn'onos m'amen asta []l mes i m'ediu i los dos m'eses / i je ent'os kwando empj'eTen a komer da_ke al mezmu tj'empu_ke m'amen d_algo tamj'en / k'omen pat'ates koT'i[]es kOn L'etSe / i um_p'oko_De far'iNa / a eso_De los kw'atro m'eses L'amen_se LeB'aikos / i sw'eLeù kap'a_se n esa eD'a[] / si nun se D'eSeù pa ber:'ones / as'ina la k'arne nun s'abe a k'otSu Ber:'i[]u /

los g'otSes ke Baù mat'a_se tj'eneù_ke []tar bj'en TeB'a[]os / si []tan kOn 'otros ke nun son pal sa_martin / ak_ke kambj'a_los pa 'otru kuB'il / pa Da_jos de kom'er / i eùgOrd'a_los da_f'etSu / aj kjen_jos 'etSa L'etSe / pero lo nOrm'al je TeB'a_los kOm mas LaB'aTa / i mas far'iNa de ma'iT / i ma'iT eù gr'anu / mjo m'aDre kw'enta k en t'araNu ab'ija S'ente

ke los etS'aBa pal pw'Ertu pa_ke pastj'aren i komj'eren aBej'otes / i as'ina D'iT ke sal'ijen pEr_bw'enos los Sam'ones / esta T'eba d'eBe aniTja_se dos meses enantes_de la mat'anTa /

Bibliographie de la matière asturienne traitée

- Gonzalez-Quevedo (Roberto).- Economía y cultura del gochu.- dans *Cultures 9, 1999*, pp. 11-25 (introducción)
- Gonzalez-Quevedo (Roberto).- Agricultura y Ganadaria en Asturias.- dans *Enciclopedia Temática de Asturias IX, 1987*, pp. 127-190
- Gonzalez-Quevedo (Roberto).- La cultura del gochu.- dans *Cultures 2, 1992*, pp. 191-209

Les erreurs contenues dans le corpus transcrit sont de mon fait. Sans cette enquête ethnographique, le matériau empirique n'existerait pas.

Préparation du corpus asturien en fonction des objectifs de l'analyse linguistique : vocés asturienes et asturiano a mínimo

En préparant le corpus asturien de textes couplés à un dictionnaire de base de l'asturianité dialectale (DALLA), je me suis mis en position d'acquisition d'une langue particulière, d'une part dans son homogénéité (contraintes, structures) et d'autre part dans sa diversité de surface telle que je pouvais la percevoir de mon poste d'observation empirique.

La phase de translittération par les enquêteurs asturiens a donné une forme figée aux témoignages recueillis, dont on ignore s'ils ont été enregistrés en tout ou partie, puis transcrits, et selon quelle méthode, avec quelles priorités; mon travail, très risqué, a ici consisté à retourner à une source première orale du témoignage en donnant une idée aussi précise que possible de ce qui est perçu par l'oreille du témoignage: je transcris ce que je me suis mis en tête dans la tête, et par conséquent non une réalité articulatoire ou acoustique mais une de mes représentations de l'asturianité, qui seule existe; que de fois ai-je hésité sur l'aperture d'une voyelle prétonique ou post-tonique, sur les liaisons

consonantiques; autant que possible, j'ai tenté de respecter - principe de précaution et éthique de transcription - , les décisions graphiques des éditeurs, quand leurs stratégies sont explicites: *gocho* / *guchu* / *gochu* se partagent le terrain, dans quelles proportions? Les décisions prises modifient la représentation du mot phonologique, du syntagme phonologique, de la phrase asturienne sonore, forme cruciale d'une hispanité asturienne.

Autour du mot roman mélodie et ses « charges mémorielles » :

(1) Entre tradition et innovation

Il ne s'agit pas ici d'innovation au plan des techniques et des mouvements sociaux, mais au plan de l'expression, à la mesure du poids de la castillanité et du poids de l'asturianité, leune comme l'autres représentées en contraste, sinon en opposition, voire en conjonction, si les deux langues sont en situation d'adstrat.

(2) Entre dissémination et polarisation

Très souvent, les témoignages permettent à l'auditeur fictif de se forger une représentation de la différenciation, autant que des caractéristiques répétées de la dialectalité asturienne; c'est la répétition qui finit par construire une représentation de la polarisation: zone côtière contre sone intérieure de montagne, à proximité, ici, comme en domaine cantabrique, comme en domaine basque aquitain et ibérique. Ce n'est pas ici le lieu de développer la problématique de l'établissement de la carte lexicale de densité.

(3) Industrie équitable des langues au service d'une grammaire énonciative fonctionnelle

Les méthodes et techniques d'industrie des langues ici utilisées ont un caractère robuste, sinon fruste, à la hauteur des ressources limitées dont je dispose: un indexeur, un compteur, un classeur, un comparateur, un cartographeur. Les petites langues pour leur traitement n'ont pas constamment bénéficié des méthodes lourdes déployés pour le traitement industriel des grandes langues: ceci est vrai dans les Espagnes comme cela l'a été pour les langues de France et constitue un point de leur histoire sociale.

On va a présent entrer plus en détail dans la problématique de la validité des représentations construites avec les ressources et méthodes utilisées: la carte, l'indexation, la lemmatisation des formes, la catégorisation, la représentation prosodique du mot, du syntagme, de la phrase: rien, disons-le ne remplacera jamais le témoignage oral fixé, numérisé, travaillé; on opère ici par défaut.

La carte synthétique, outil de la géographie linguistique des langues romanes : l'apprentissage des savoir parler dans les communautés asturiennes

Pour connaître le plan de la variation, l'expert en géolinguistique utilise soit la cartedonnées du dictionnaire de prononciation, soit la carte_données d'un atlas, s'il existe.

Le dictionnaire asturien DGLA de Xosé Lluis Árias García permet une représentation géolinguistique suffisante des aires de variation, comme le permettrait un atlas linguistique asturien s'il existait, permettant la comparaispon des caractéristiques langagières sur la totalité de la cordillère cantabrique. Il réussit ce tour de force d'être à la fois une organisation de dictionnaire et d'atlas. La source est établie au plan de surface phonétique); la cible est établie au plan de la phonologie lexicale.

asitiáu / asitia[]o (source localisée)/ ==> *asitiado* (cible de système)

[]t'aben / *estaben* ==> *estaban_estar* (cible de système)

Je me limiterai ici à ces deux exemples, encourageant ici l'objection fort légitime d'exemplarisme, souvent formulée à l'encontre des approches dialectales. Pour faciliter la formulation des questions auxquelles on tente d'apporter une des réponses optimales.

Première question: la carte des variantes de système nous renvoie-t-elle au concept de polymorphisme linguistique asturien organisé géographiquement en variantes centrales, orientales, occidentales?

Seconde question: à cette première représentation cruciale pour l'histoire sociale de l'asturianité, se superpose une seconde question: quelles variantes et à quel niveau, relèvent de la mise en oeuvre du concept de polynomie au sens qu'a donné à ce terme Jean-Baptiste Marcellesi ? Les deux questions, qui concernent de plein fouet l'histoire sociale des langues régionales de l'Europe, méritent d'être creusées, à partir d'une des failles multiples que la linguistique d'observation permet de déceler dans les surfaces apparentes et audibles des parlers, pour les langues du monde.

Comme on ne peut en même temps voyager sauf au prix d'un comparatisme de grande envergure dans les mondes indo-iraniens, indo-aryens et indo-européens, on commence par observer ici ou là entre Clermont Ferrand et Pontevedra, en domaine celtique romanisé, une faille infime dans un des mondes romans de l'Europe Méridionale, au coeur du domaine asturien, entre Cantabrie (Santander, Torrelavega) et Galice (Pontevedra).

Posons le principe (hypothèse nulle de départ) que dans ce secteur géographique des Espagnes, on parle actuellement et nécessairement - au temps *t* résultant de l'actualité ibérique l'espagnol, comme en Belgique on parle le belge et en Suisse le suisse; comme à Toulouse le français. Cette hypothèse constitue de fait une partie de la question de départ à laquelle nous entendons répondre, sachant que nous ne détenons pas en tant que géolinguistes les solutions au problème posé par le traitement de l'oralité asturienne dialectale, à partir de textes oraux transcrits.

La première solution proposée est une solution cartographique, qui a été largement explorée par la géolinguistique romane (ALAR).

On peut voyager à travers les cartes dessinées d'atlas linguistiques (*el gocho; el matanzu; el pintoriu/pintón; el mercáu* []l *gochu* etc.); ce sont des cartes initiales de matière lexicale; si l'on a pour objectif de dresser l'atlas linguistique d'un domaine linguistique roman, européen etc., on entreprend alors de dessiner, selon différentes méthodes la relation des lieux à l'espace de l'asturianité; tel n'est pas mon objectif à court terme; toutefois une représentation dessinée des configurations de caractéristiques systématiques n'est pas inutile, il s'en faut, à une meilleure représentation de ce qu'est dans l'espace et le temps l'asturianité, définie en termes de représentation, à partir d'observables localisés. Le rapport, de la normalité à la « normaison » pose d'énormes problèmes à la préparation des corpus entre linguistique fondamentale et appliquée (didactique des langues vivantes). L'approche dialectale par la carte pose elle-même de très importants problèmes du fait même de ses conditions d'établissement; dans la mesure où elle peut faciliter sinon créer une illusion d'optique chez des esprits faibles qui pourraient croire qu'on parle en Asturie des années 2004 comme on y parlait, sur les marchés de La Pola ou dans les *mantequerías* de La Pola et Marcenao dans les années 50, les années d'avant le *mercau* industriel.

Ontologies de domaines et sémantique : histoire du mouvement et organisation modulaire du dictionnaire d'apprentissage

Même si aucune analyse linguistique ne peut faire l'impasse des répertoires lexicaux par indexation des contextes et concordances: il ne s'agit en aucun cas d'un lexique de spécialité, mais d'un lexique général (classes ouvertes) comportant un secteur terminologique dialectal polarisé et disséminé à la fois, rapporté aux catégories de pensée, d'action et d'expression (mythe del gotcho; merde des porcheries; fumier: catégories des pratiques dans une société rurale traditionnelle d'avant la crise).

Une observation absolument triviale, bien que non dénuée d'intérêt est que l'ensemble du processus énonciatif s'organise, au plan de l'expression autour des fonctions (causativité) et des «voix» (transitivité, intransitivité, voix moyenne, passivité, changement d'état) ; donc autour d'un nombre restreint de verbes d'action, d'état, de mouvement, de dire, causatifs, anti-causatifs, considérés comme pivots et formant le noyau verbal d'une hispanité fondamentale, avec caractéristiques d'asturianité au plan phonique. Cette observation est de taille, au regard de l'identification de la nature d'un objet désigné métaphoriquement *paxarina*, *paxarineta* (par convention largement polarisée: sans doute le diaphragme du porc, très plausiblement, si ce n'est la rate, le pancréas (pourquoi non ?) etc. pour la petite histoire, anecdotique). Toutefois une longue habitude m'a appris à observer jusque dans le menu détail et toujours, avec l'idée qu'il en restera toujours quelque trace en description des faits de langage (communication, représentation rapportées aux hommes, à leur milieu, à leur histoire, à leur organisation territoriale).

Cette opération de préparation lexicale du corpus par catégorisation et subcatégorisation, absolument triviale, n'en demeure pas moins une nécessité d'ordre absolue, pour qui veut poursuivre l'analyse au-delà des limites et finalités de la lexicologie quantitative.

On constitue ici une sorte de réservoir lexical, de répertoire lexical qui n'est pas un registre individuel, mais le registre collectif ou registre fondamental pour la classe ouverte constituant le lexique verbal accédé par tous, ouvert à tous, clé de l'asturianité de base: on a souvent parlé d'un asturien minimal.

Module du répertoire collectif initial en approche verbale

C'est le module d'origine, à cette étape non subcatégorisé, uniquement marqué verbal, prétraité par indexation brute. C'est, dans les classes ouvertes du lexique, ce module qui est au cœur de notre observation et organise le plan de l'expression: dans mon hypothèse de travail, ce sont, parmi les classes ouvertes du lexique, ces unités du lexique verbal qui sont les premières accédées (*afumar*, *curar*, ..., n), avec quelque décalage par rapport aux catégories nominales; le module nominal intervient en second lieu, par incorporation: il forme ainsi une sorte de réservoir terminologique collectif de domaine d'expérience (*san_martin*): mais l'opération de traitement peut créer l'illusion d'un répertoire individuel par informateur; les enquêteurs, toutefois ne s'y sont pas trompés, et ne se sont pas laissés endormir par le mythe du répertoire patrimonial individuel.

Au préalable on se sera débarrassé par décision des lexies grammaticales temporelles, spatiales, versées à un répertoire spécifique: *un pouco más alantre*; cette décision ne sera pas mise en discussion ici. Cet index ne produit que la nomenclature d'entrée du dictionnaire d'apprentissage, tel qu'il est inscrit dans le carnet de l'élève, à la première étape du repérage brut. On y notera toutefois entre parenthèses, les éléments d'un premier repérage des relations; d'entrée de jeu, disons que ce n'est pas tant sur l'entrée de donnée - *botar*, par exemple, que portera notre attention que sur les contraintes phonologiques et morphologiques qui régissent une réussite asturiana et galego-asturiana de gabarit, taille et rythme des distributions verbo-pronominales *botábamos jes/you*, combinaison dont la solution phonologique repose sur une trace sonore organisée à la façon galega-asturiana, par ailleurs bien explorée par les phonologues et la morphosyntaxe des langues romanes,

en rapport avec la description des faits de clise pronominale, qui toutefois n'expliquent pas tout. Le seul index verbal permet de parcourir deux échantillons de taille e 50.000 mots phonologiques (soit un total de plus de 100.000 mots phonologiques des corpus A et B comparables et d'identifier en approche pronominale la totalité des structures syntaxiques et sémantiques de représentation.

aballar
abarruntar
ablandar (blando)
ablandiar (blando/blandio/blandecer)
abrir
 abrir a canal
 abrir en canal
aburacar (buracu)
aburrueirar
acebonar (cebo, cebón)
acertar (cierto)
achegar
achoquinar
achorizar (chorizu)
aclarar (claro)
acocorar (cocorar)
acogomellar (de)
acolorazar (se) (color)
acompañar (compangu)
acoraxar (cor, corón)
acorralar (corral)
acostumar (costume)
actuar
acuriosar (curioso)
adelantar (delante)
adelgazar (delgado)
afaitar (afeitar)
afallecer
aferrar (ferro)
afilar (filo)
afincar (finca)
afitar (fita)
afogar (fueu)
afondare
afroxar
afumar (fumo: afumiar)
afuracar (furaco / furar)
aganchar (gancho)
agarrar (garra)
agasallar (gasalla)
agolpar (golpe)
agradecer
agriar (agro)
aguantar (agua)
agurrinar (gorro, gorrin, güerre)
agueyar
aisllar

alambrar
alcontrar (contra)
algamar
alimentar
aliviar
allargar (llargo)
allugar
allumar (llum)
almorzar
amanecer
amañar
amarar (amaro)
amargar (amargo)
amarrar
amasar (masa)
amastraconar (se)
amburar (se)
amestar
amesturar (mestura; mesturar)
amolar (mola)
amollecet (molle, amolecet)
amontonar (montón)
andar
 andar a la yerba
 andar a la pia
 andar de mondongu
 andar al buey
 andar al burrón (varrón)
andechar (facer andecha)
anillar (anillo)
antroír (antroxo)
añadir
apalabrar (se)
apañar (pañar)
aparar
apareyar
apartar (se)
apegar (se) (pega, pegar)
apelmazar (se)
apellidar
apertar
aplastar
aporondar
aportar
apreciar (precio)
apretar (prieto)
aprovechar
apuntar (que p / punto)
apurrar (se)
apurriar
apurrir
arimar (se)
arrallar

arrancar
arrañar
arrastrar (rastros, rastrillo)
arredondiar (se) (redondo)
arreglar (regla)
arremangar (se)
arregar
arresistir
arrimar (se)
arrodiar (se)
arrojar (arrojar)
asar (asadura)
ascotrollar
asegurar (seguro)
asemeyar
aserenar (sereno)
asitiar (se) (sitiu)
asonseñar
asoriar
asustar (se)
asuxetar (sujeto)
atapecer
atar
atardecer (tarde)
atender
atopar
atorrollar (torrollo)
atravesar (travieso)
atropar
averar (vero)
avezar (vezu)
aviar
 aviar el pote
avisar
avinagrar (se) (vino)
ayudar (axudar)
axeitar (se)
axuntar
baldar (se)
baldiar (balde)
bañar (baño)
barruzar
baxar (baxo)
beilar
berrir (verrir, verrón)
berzar (berzes)
bixilar
blanquecer (blanco, branquecer)
blanquear (blanco, branco, branquear)
botar
 botar a culpa
bromar (bromada)
buscar (rebuscar)

camudar
caber
caer
cagar
calcar (se)
calentar
caltener (se)
caminar (encaminar)
cansar
capar
carecer
cargar
casar (se) / casa
catar
cavilar
cayer (caer)
 cayer en desusu
cebar (cebo)
celebrar (celebre)
cenar (cena)
chamar (llamar)
charrar
chegar (llegar)
chillir (chillar, perchillar)
chiflar (chifrar)
chinchurrar (chinchorro)
chucar (se) Va ensugar (se)
chupar
cimblar
clavar (clavo)
cobrar
cocer
cocinar (cocina)
cofar (cofa)
coger (recoger)
colar
colingar (se)
comentar
comer
comerciar (comercio)
comprar
comprometer (se)
condimentar (condimento)
conservar
consistir (en)
contaxar
conxelar
corar (gorar / corón)
corcollar (corcol.lar)
correxir
cortar
costar
 costar trabayu

coyer (recoyer, encoyer)
coyer el relevo
coser
crecer
criar
cuadrar (se) / cuadro
cuayar (cuayo, cuacho, cuachar)
cubrir (/ descubrir)
cuchar (cuitu, cuchu)
cuelgar (colgar/descolgar)
cuidar (cudiar, curiar)
cumplir
curar
curiar
curtir (curte)
dar
dar la prueba
dar anuncia
dar vuelta (vueltas)
dar mala suerte
dar puxu a
dar gustu
dar sabor (gustu, golor, ..., n)
dar sa_martinu (samartin)
decir
deitar
depende (de)
derrengar
derretir
desangrar (sangre, sangrar, sangradie.las)
desapegar (se)
desatetar (teta)
desayunar
desenbalustrar (balustra)
desenredar (red)
desentrellizar (entrelliz)
desfacer (facer)
desgollar (desguelle)
desgüesar (hueso, desosar/oso)
desingrir (sangre)
desmangar (mango)
despedir (despidir)
despegar (pega)
despostar (poste, postar, postrar)
destorbar (estorbar)
destrozar (trozo)
desurdir
desvestir (veste)
devolver (volver)
dexar VA echar
dir (ir)
disluir (diliir,se)
dividir (se)

dixebrar (se)
doler
echar
 echar l'agua
 echar en pote (al pote)
 echar a sal
 echar n'untu
 echar a vultu
 echar al buey
ellaborar (llabor)
embafarar (se)
emberzar (berza)
embizcar
emborcar
emborrachar (borracho)
embotar (botiellu, buchello) emboutar
embozar
embruxar (bruxa, bruxu)
embuchar (buche)
emburriar
embutellar (botello, buchello)
embutir (botello)
empanar
empedrar (piedra)
emplegar (se) (emplear)
emporcar (porcu)
empreñar
emprimir (primo)
enanchar (ancho)
enaugar (augua, agua)
encaminar (camino)
encargar (carga)
encaxar (caxa)
encebollar (cebolla)
encender
enchorizar (chorizo)
enchumazar (chumo)
encolingar (se)
encomenciar
encomendar
encostillar (costilla, costillar)
encuruxar (se)
endolcar (desendolcar)
endurecer (duro)
enforfugar
enfriar (frio)
engalar
enganchar (se) (gancho)
engañar
engarabir (se)
engayar
engolar (gola)
engordar (gordu)

engoruñar (se)
engrasar
engrispiñar (se)
enlatar (lata)
enmangar (mango)
enprencipiar
enredar
enrollar (se) (enrolar / desenrolar)
enroscar (rosca, se)
ensangrentar (sangre, se)
ensugar (se)
entablar (tabla, tablero)
entainar (a)
entamar
entarrar (tarro)
entender
entevenar(se)
entisnar
entoirar
entouriar (se)
entrar
 entrar en aición
entrecallar
entregar (se) (entriar_se)
entrellizar (entrelliz)
entrepernar (perna)
envasar
esbardiar
esbastardir (bastard)
escaldar (caldo)
escamar
escapar (s')
escaer (se)
escarnar (carne)
escarniar (carne)
escarnizar (carne)
escatanar (se)
esclarar (claro)
esclarir (se)
escocer (cocer)
escostiellar (costiella)
escuachar
escuadrar (cuadro)
escuartizar (cuarto)
escucarazar (se)
escuerriar (cuerria)
escuitar
escurecer (al --)
escurrir
esfaroliar
esfocicar (focico)
esfoyar (foya, esfollar, desfuelle)
esfrecer

esgarar
esgargachar
esgüazar
esixir
esllagullar
esllavar (llavar)
eslleyir (lleyir)
esmocar
esmondar
esmondongar (mondongo/mendongo, esmondongar))
esmourecer (se) (mouro)
espartir
espatalexar (pata)
espedazar (pedazo)
esperar (asperar)
espernar (esparnar / perna)
espetar
espiezar (pieza)
espolveirar (polvo, polveira)
espoixigar
espuchar (se)
espurrrir
esqueicer
estacar (estaca)
estaferiar
estañar
estayar (tayar/tayu)
estar
 tar []tar
 estar a remueyu
 estar de matanza
 estar de mondongo
 estar en celo
 estar al gustu de
 estar al tanto que (p)
 estar a puntu
 estar al tanto
 estar al tiempu de
estazar
estirar
estoucinar (estocinar)
estragar
estrañar (se)
estrapellar
estrar
estremar
estripar (tripa)
estroldar
estropiar
estrozar (trozo)
estruñar (estruñer, estruñir)
estruyer
esvaciar (vaciu, bacía)

existir
facer (fer, faer)
 facer el sa_martin (matar)
 facer el matanzu (matar)
 facer un corte (cortar)
 facer_falta
 fer el quite
 fer el óso
falar
fartucar (farto)
fender
ferventar
ferver (ferviendo)
filar (filo)
fixar
finar
fincar
forrar
fozar (hozar) / focicu
fregar (esfregar)
freír (freyir)
furar (furacu)
garantizar
garrar (agarrar)
gastar
glayar
gotiar (gota)
gramar
grillar
gruñir (gruño)
gruxir
guetar (gueta, a la gueta de)
guisar (guisado)
gurniar (urniar)
gustar
haber
haber_que
hinchar (se)
iguar
impedir
inflar
intentar
intercambiar
llamber (rellamber)
llambiotar (llambión)
llantar (chantar)
llapar (chapar)
llegar (chegar)
llemer
llenar (chenar/lleño/cheno)
llendar (llende)
llevar (levantar)
llevar (levar, chevar)

llevar la prueba
llevar l'antroxu
llevar trato (con)
llevar peso
llevar golpes
llibrar (librar) (se)
lligar
llimpiar (limpiar)
llindar
llindiar
llograr (se) / l.lueu
machacar
macizar
madreñar
mamar
manchar
mandar
manexar (mano)
mangar
mantecar (mantegar, manteigar, manteca, manteiga)
mantener (manter)
marcar (marca)
matar
 estar de matanza
mayar
mazar (maza, mazón, mazona)
medir (midir)
medrar
mercar (mercáu)
mesturar (se)
meter
 meter a remueyo (a moyo)
 meter en sal
 meter n'untu
mezclar
mexar (mexo, mexadeiro)
moler (mola)
mondonguiar
 estar de mondongo
moyar (remoyar, moyo)
mullir (muller)
nacer
negar (se)
obrar (obra)
ocultar (se)
olecer
oler (goler)
ortigar (ortiga)
oyer
parar (emparar)
parir (paridiega, paririega)
participar (a)
partoriar (parturiar, parto)

pasar
 pasar fame
 pasar miedu
 pasar noite
pastiar
pegar
pelar (pelo)
pelear
pender (pintor)
peneirar
pesar (peso)
pezquizar (pezcuezu)
picar (picadillu)
pillar (piyar)
pinchar
pingar (pinga)
podreecer
poner (poñer)
 poner_se
 poner_se_d'acuerdu (d'acordo)
postrar
preguntar (perguntar)
prestar
principiar
producir (productu)
pruebar (prebar)
 facer la prueba (preba)
pulgar
puxar (puxu)
punxar
 punxar puntu a
queixar (se)
quemar
quitar (se)
rabilar (rabil)
rancear (rancio)
rascar
raspar (rapar)
rayar (se) (raya)
raxar (se)
recochar
recordar (se)
recortar (cortar, corte, curtiu)
recluir
refregar
regalar
rellenar (llenar / lleno ; chenar/cheno)
remasar
rematar (matar)
remolar (mola)
repetir
reponer (se)
reposar (posar, se)

requerir (querir)
resbalar
restorcer
reteyar (teya)
reventar
ripiar
riscar
rodalar (roda)
romper
roñar
rozar (rozu)
rucar
rustir
ruxir
saber
saber (a)
saborear (sabor)
sacar
 sacar la gota
 saca_se_y l'untu
sacionar (sazonar)
salar (sal)
salsiar
salcochar
salir
 salir n'est'au
sallar
saltar
salvar
samartinar (samartin)
sanar
sangrar (sangre)
secar (se)
segar (sega)
segotar (sega)
seguir (seguida)
semar
señalar
ser
 ser de confianza
serenar (sereno)
servir
sobrar
solmenar
soltar
sostener
subastar
subir
sudar (sudor)
suspender
sustituir (se)
tachar(tayar)
tardar (tarde)

taurir (taurida)
tayar (estayar, tachar)
tener (ter)
tener aveiro
 tener calor (ter calor)
 tener cuenta (en cuenta)
 tener cuidao
 tener fame
 tener fama (ter fama)
 tener procuru
 tener tinu
 tener a puntu
tener_que
tirar (tira)
tizar
triyar (trigar)
tomar (se)
 tomar a parva
torgar (torga)
tostar
trabayar (trabayu, trabachu, trabachar)
trancar
tratar (se)
trayer (traer)
tresformar (se)
trevolgar (se)
tumar
untar (unto)
vaciar (baciú)
valer
vender
venir
 venir bien
ver
verrir (verrida)
vezar (vezu)
vivir
volver (revolver, devolver)
xamuscar
xarriar
xergar
xibrar (xibre)
xintar (xinta)
xorrascar
xugar (xuegu)
xunir (xuncir)
zampar (zampa)
zarrar (cerrar)
etc.

Soit un bilan de plus de 650 unités lexicales verbales dessinant le profil du lexique asturien minimal (fondamental, si on préfère): ceci n'a d'importance que dans la mesure où ce lexique permet d'observer comment s'opère en asturien la prédication,

du point de vue de la description, comme du point de vue des apprentissages majeurs; la morphologie verbale est supposée connue de par ailleurs..

Voilà donc offert à nos yeux et ouvert à nos oreilles, pour faciliter la compréhension, un premier pré-répertoire donné à parcourir en ballade aléatoire ou non aléatoire à travers le corpus oral segmenté et balisé, quelle que soit la méthode de balisage retenue.

Cet index permet un balayage systématique, en mode aléatoire ou non, selon les méthodes et objectifs, de l'ensemble du corpus; du même coup se trouve minimisée, optimisée, la place de la variation régionale de l'asturien (central, occidental et oriental) sans que le débat soit fermé pour autant.

Non seulement on peut considérer le rapport de cet index asturien de classe fondamental avec le répondant castillan, mais encore à partir de cet index on peut balayer la somme théorique de tous les prédicats produits et prévisibles pour la situation et les contraintes de tâche liées à la description ethnographique. De plus le balayage par concordances rend ces corpus produits totalement comparables, donc manipulables par un comparateur assistant le descripteur artisanal, qui ne travaille plus exactement comme un ouvrier d'atelier (scriptorium) du 12ème ou du 14ème siècle.

L'annuaire de la classe ouverte verbale est automatiquement dressé par un algorithme de traitement du donné grammatical qui ne laisse aucune place à l'aléatoire. Une fois constitué l'inventaire et ordonné en subcatégories sémantiques, une grande partie des tâches préliminaires de base, préalables au traitement linguistique proprement dit, est achevée. La priorité méthodologique accordée à la réalisation de ce pré-répertoire est justifiée par la priorité à accorder à l'acquisition du module syntaxique et sémantique organisé par le verbe, indépendamment de la différenciation phonétique de surface, ici totalement neutralisée.

Chaque unité en termes de linguistique de corpus, est de manière triviale, associée à son indicateur de densité lexicale, en même temps qu'à sa catégorie sémantique: bien des outils de lexicologie en linguistique de corpus oublient l'un des deux indicateurs, ce qui invalide l'opération.

Une des tendances premières de la linguistique appliquée à conduit dans une perspective didactique à appliquer ici les principes de base d'une métrique lexicale, permettant de mettre en évidence le noyau des lexiques fondamentaux., assez loin, parfois, du front de la théorie de la constitution du noyau phonologique lexical. Ce qui fait que de Madrid ou de Paris, on a parfois accordé une importance très insuffisante à la constitution de l'archive sonore des voix. D'autre part il était tentant de rester entre soi et de collectionner les voix d'étudiants des Universités, directement en mains et sous la main; c'était moins coûteux en termes de travail que d'aller sur le tas entre Réquista et Lodève, recueillir la matière même, ou entrer Santander et Ourense (domaine cantabrique et asturien, aux portes de la Galicia et du León). On obtenait ainsi des voix épurées, aussi dépourvues que possible des caractéristiques anciennes, appelées dialectales, parfois incorporées ou encapsulées au lexique verbal asturien basique, dans des conditions contestables.

Il est très aisé alors, selon cette méthode, de ramener à un seuil inférieur à 5% le taux d'asturianité dialectale, pratique courante en linguistique de corpus se fondant sur des artefacts écrits déjà fortement influencés par la normalité espagnole. Ce seuil détermine alors le «ratio» de régionalisme acceptable, par tolérance, selon les méthodes classiques: moyennes: moyennes des moyennes; distances; indicateur de similarité ou voisinage; écarts, écarts types; indicateur V de variance etc. ; mais la question posée initialement ne trouve pas ici une solution, tant s'en faut. Le problème posé, voire mal posé, reste entier, précisément parce que non vraiment posé.

Mais sans subcatégorisation sémantique, sans identification des prédicats

ontologiques d'origine, une telle liste est inefficace et ne peut entrer telle quelle dans le dictionnaire didactique. Il est en effet trivial, mais crucial, de discriminer les causatifs des anticausatifs, les statifs, les attributifs, les verbes scéniques d'action, les verbes de mouvement; les lexies simples des lexies verbales figées (GMF verbaux) etc.

Dictionnaire, dissémination et polarisation : place de la dialectalité dans le dictionnaire

On peut, si l'on a pour mission de dresser le dictionnaire général d'un domaine, se donner les moyens de voyager (de naviguer en mode aléatoire ou non) à travers les entrées, contextes et structures du dictionnaire (Meyer-Lübke 1972; Fritz Krüger 1991; García de Diego 1970; DALLA de Xosé Lluis Arias García); tel n'est absolument pas mon objectif; l'excellent DALLA y suffit à lui seul. Alors ? On maintient cependant que de l'habitude de traiter systématiquement le lexique (matière et formes sonores de la matière asturienne), il reste toujours des traces professionnellement fort intéressantes. La méthode est simple; sur une surface apparemment plane, lisse, unifiée, on observe un tout petit trou; on gratte au marteau; on élargit le trou; on mesure la faille; on y entre dedans, on explore le site découvert de fond en comble, on analyse avec les outils appropriés les éléments qui font dissémination et polarisation en même temps, en rapport avec les fonctions de communication. On entend au fond, des choses inaudibles de la surface, qui sont du domaine du dicible effaçable. J'ai voulu, à partir de corpus écrits ou de pseudo-écrits tenter une exploration de ce type; si le voyage vaut le détour, que l'on n'attende pas toutefois de nous de démontrer que nous y avons trouvé une solution; toutefois nous aurons au moins contribué à mieux comprendre les éléments de la problématique.

L'opération de subcatégorisation du lexique des classes verbales ouvertes, toutes caractéristiques dialectales, à ce stade, non prises en compte, permet d'optimiser le dictionnaire d'apprentissage.

Voix asturiennes : matière, formes sonores organisées

Comme on parlait occitan au fond des mines de Carmaux, on parlait astirien au fond des mines asturiennes; à force de gratter ces deux petits trous ou gisement, on finit par mieux entendre et donc mieux comprendre ce qui fait la spécificité des voix ici entendues; il faut ici comprendre que cette métaphore du gisement peut être élargie à l'ensemble des situations souterraines où se parlent les langues vivantes réelles de caractère dialectal, en Europe, par exemple.

Voyager dans l'archive sonore des parlars romans ordinaires présente une triple dimension: charge mémorielle lexicale sonore, définie en termes de phonologie lexicale de la matière sonore; charge mémorielle syntaxique définie en termes de phonotaxe de l'oralité, de parataxe, de syntaxe; charge mémorielle sémantique (*refranes; cuentos, formularios*): mémoire des couches anciennes, des couches innovées, mémoire adaptative, mémoire discursive. Mais quels avantages présente cette stratégie de présence sur l'ensemble des fronts et en même temps? et quels en sont les inconvénients ?

Voilà donc défini, partiellement au moins, mon objectif, partiellement atteint par le biais de l'annotation des textes édités.

Mais se cantonner ainsi dans le capital langagier décrit, dans le capital lexical décrit - réservoir de prétendues capacités individuelles, représenterait une réduction abusive des dimensions du problème. Ce qui doit être mis en évidence c'est le concept même de capital social, ici asturien, tissé des interactions entre individus dans les groupes, communautés rurales /urbaines et urbaines faisant classes: ainsi, le plus souvent, contre une représentation disséminée, individu par individu, la politique éditoriale des linguistes ethnographes et sociologues asturiens les a conduits à procéder par polarisation: c'est le capital social des communautés qui est représenté en même temps que le capital

langagier de groupe; sans doute une représentation optimisée peut-elle être atteinte par calcul des proximités entre groupes polarisés: à quel groupe appartient Riosa de Llena, Pola de Llena. On a de fortes chances de mettre ainsi en évidence non les facteurs de dissémination dialectale (principe de distanciation maximale); mais au contraire un facteur de polarisation qui joue un rôle de gouvernance: groupe occidental, groupe oriental, groupe central? gouvernance s'entend ici hors acception de volonté politique, hors intervention sur la langue dans les sociétés de volontés politiques; ce qui constitue le coeur même du projet asturien, bien compris, dans les Espagnes linguistiques et culturelles.

Voyager dans des paysages sonores non fictifs, en plein coeur des Asturies, aux portes de la Galice, de la Cantabrie à la Galice, au León, comment faire, quand on n'a que de pauvres écrits dialectaux, en France on dirait patoisants? Il faut les rendre sonores, visibles, inoubliables en quelque sorte, parce qu'entendues avec assez d'insistance pour que les «charges mémorielles» puissent porter leur fruit; de cette lecture sonore des textes, il reste toujours des traces; qu'il s'agisse des monologues de l'Antigone sur ses obligations tragiques, ou de plus modestes textes encyclopédiques, sur des sujets humbles: l'essentiel est d'accéder à la sémantique par les formes sonores même de la langue à apprendre: ce peut être l'Occitan de l'Auvergne ou du Limousin à la Gascogne; ce peut être l'asturien dans ses formes centrales, occidentales ou orientales.

Ne pouvant aborder l'asturien par une description d'atlas inexistant, j'ai commencé à entrer en asturianité dialectale (langue et dialectes) par les structures du **DALLA**, puis par l'annotation, en termes de perception sensorielle ou de restitution, à partir de textes graphiques asturiens, ce qui ne va pas sans problèmes et incertitudes.

Pour moi, l'asturien c'e n'est pas du tout comme le javanais, mais c'est cependant le résultat d'une transformation mélodique ordonnée, polarisée, réarrangement stabilisé d'une langue qui pourrait être au départ une des langues d'Espagne, que l'on pourrait appeler l'espagnol enfoui: vrai trésor mélodique enfoui dans des répertoires, une archive sonore et écrite entremêlées dans le temps et une longue histoire. Mais, toutes proportions gardées, la matière gasconne, par rapport à la matière occitane, de par sa position géo-historique même, n'est-elle pas à l'occitan ce que serait la matière asturienne à la matière espagnole, matière des Espagnes?, Ce n'est pas pour cela que l'asturien serait un peu le javanais ou un créole du castillan ?

A la découverte de similarités significatives ; systématité complexe : règles et contraintes sur la grammaire asturienne

Il nous faut revenir à la méthode comparative des prédécesseurs (F. Krüger, G. Rohlfs, Schmidt): les *duernes* et *duirnis* de la mélodie asturienne, ne sont-elles pas, prosodiquement parlant, les mots mélodiques *dornas*, *dornha*, *duèrnas*, *duòrnas*, conformes aux contraintes de la mélodie des langues romanes accentuées, la mélodie occitane, dournes.R de la mélodie du lexique français (*douirnes). Ce n'est pas tant la similarité lexicale fondée sur l'existence d'un lexème commun de fonds celtique qui importe que la structure même du mot phonologique même, quelle que soit son origine, dans un système à régularité flexible. L'asturianité ici s'inscrit sur une portée de type hispanique (avec transition d'aperture: *duòrna /duèrna*; l'occitanité sur une portée de type occitan et portugais (sans transition d'aperture: *dorna*, *dornha*: ou avec étapes de différenciation: *duòrna /duèrna*).

Deux expériences majeures m'ont conforté dans cette conception: une enquête sur la prosodie de l'espagnol parlé en aire culturelle cantabrique; une enquête sur la mélodie du français parlé en Poitou méridional; et enfin l'ensemble des enquêtes dialectales conduites en France méridionale sur la prosodie des parlers appelés dialectes (Occitan parlat jos terra, par exemple, en plein pays minier de Carmaux). Il faut ajouter à ceci que j'ai gardé

le souvenir assez précis de la prosodie des mots hérités des langues indo-européennes: les mots grecs de l'Antigone, prosodie des mots du Sermo vulgaris de Plaute, Apulée; de cette habitude d'interroger de manière têtue les données actuelles rapportées aux sources indo-européennes, indo-aryennes, il reste toujours quelque trace.

Ma conception du mot roman, quelle que soit l'application est une conception héritée d'une part de la linguistique des formes sonores de la matière, définies en termes de positions, éléments, portée, mélodie, degrés de liberté, interaction de contraintes et transgression, ancrage initial, mouvement à partir de l'ancrage initial, remplissage des positions vides; multiplicité des ancrages et désancrages possibles, conservation de l'ordre des éléments: mon maître en la matière est l'énigmatique Pancho Solares Huerta, de l'Université de Xalapa, Ver. (CGR10, 1985, 177-179): l'intérêt de la doctrine est de se ramener à ces deux pages de base. On ne trouve ici que l'application du premier principe, à l'exclusion du second.

Les systèmes de règles y sont conçues en termes de degrés de libertés, corrélat de contraintes sur la langue.

(1) Morphologie concaténative de l'asturien

Le premier principe de liberté (contre les linguistiques à règles) est ainsi formulable, en termes de morphologie concaténative: chaque élément sonore d'une suite (bien formée), _ ainsi le trivial asturien / r:ʃaùkotSu / le non moins trivial *vasherer* gascon / braS'E /, le banal /mut'On / du dictionnaire de prononciation française (12 Naucelles, 46, Gourdon) _ est librement uni à une position de la portée: *fai; fai-se; faga; faiga, traiga; tenga, duernu, duirnu (duerno); cuirnu (cuerno)*: cette organisation phonologique est vérifiable en rapport avec la chaîne phonétique en dispositif d'analyse phonétique instrumentale. C'est ce plan phonologique seul qui est ici ouvert.

(2) Morphologie non concaténative de l'asturien

Le second principe de liberté n'entre en jeu que dans les schémas de morphologie non concaténative: les positions restées libres sont unies soit à ces mêmes éléments, soit aux éléments d'un quasi-morphème incorporé, encapsulé, positionné; l'analyse est conduite à partir du fameux /truvydyk'ykav'y /; ainsi Josepa permet-il selon les contraintes asturiennes (voire occitano-catalanes), soit Bepa / soit Xepa: ces faits ne sont pas ici examinés.

Ainsi considérons, dans une fenêtre dialectométrique, un mot mélodique occitan de mélodie monosyllabique en pied fermé non par un quasi-morphème, mais par un morphème -s/j en coda complexe en SN pluriel: *las vacas /les b'akes /* serait la mélodie initiale, le point d'émergence de nouveaux éléments positionnés sur la chaîne, avec resyllabation : ==> leh b'akes / lej b'akes/ , avec polarisation territoriale des solutions ou événements sur axe occitan d'ouest en est, si l'on s'en tient à un comparatisme intra-occitan. Il en serait de même de manière prévisible en domaine catalan occidental, valencien, asturien, aragonais pyrénéen (Benasque)

Savoir si les textes asturiens relèvent de l'analyse des parlers régionaux d'Espagne du Nord, des langues officielles ou non officielles de l'Europe du Sud revêt à mes yeux une moindre importance: l'asturien est une des langues réelles du Principado de Asturias en Espagne du Nord; son rapport aux anciens dialectes de la Castille, par rapport à la situation des Espagnes au Nord relève du projet de la Linguistique historique et de la géopolitique des langues à caractère dialectal en Europe. Comme le Gascon est un des extrêmes de l'occitan, l'asturien est un des extrêmes du castillan, entre Cantabrie et Galice, entre le León et Miranda de Douro.

Il importe en revanche de relever que la distance entre les versions comporte un aspect géolinguistique qui justifie une mesure métrique, pour les paramètres phonétique, phonologique et morphologique :

- critères phoniques mineurs sur listes de mots phoniques

Statut diphtongal, gabarit et taille de la base observée dans une perspective typologique

duerna / dorna, par exemple illustre une de ces mesures qui accorde une place relative à l'aléatoire dans les distributions constatées. Constaté le polymorphisme asturien n'explique en rien les cas de similarité avec la phonologie métrique occitane, qui n'est pas surprenante, si l'on a présent à l'esprit le témoignage des historiens (Higounet, entre autres).

bon/bueno

duerna dorna duorna

prueba, preba, proba; probar, pruebar prebar

frio / friu ? rio /riu ?

Métaphonie asturienne

duernu

gotxu / gochu / cochu

tablero, tableiro, tabliru

- critères morphosyntaxiques

La clise pronominale, fait de phonostructure asturienne majeur

pide_se_y

Le GN et L'accord

tripa delgado, tripa gordo : quand le non-accord institue une normalité morphosyntaxique asturienne.

GN=N1deN2

cabu mundu (cabo del mundo; occitan cap de monde, cap del monde)

Edition et projet Pros_edit astur

Seuls les éditeurs et enquêteurs peuvent valider nos propositions d'annotation, au vu de la connaissance qu'ils ont du parler enquêté.

Annotation et marquage

L'annotation est opérée en apinewWord pour traitement en environnement PRAAT 4.2, après segmentation: quand les sonorités documentaires n'existent pas, il faut les créer en mettant en oeuvre des méthodes soit de reconnaissance de la parole soit de synthèse vocale.

Segmentation

Le critère de segmentation qui est suivi obéit à des discriminations prosodiques et syntaxiques

On utilise des marqueurs robustes, en mode HTML par exemple, ou toute méthode de marquage similaire.

Algorithmique de génération de la prononciation à partir des graphies

Elle est parfois soluble, parfois non: c'est selon. Et il faut s'y faire. Les recherches en cours ici et là concourent à l'établissement d'un programme d'édition des textes sonores base de recherches utilisant les technologies dites de traitement des données (Data Processing).

Indexation des unités

A l'indexation des unités lexicales qui suppose une réflexion préalable sur les suites syntagmatiques ou syntaxiques insécables: lexies composées des classes ouvertes ou fermées, on substitue le principe d'indexation du flot verbal par unités grammaticales.

Classes lexicales ouvertes sur le lexique

On établit ici les répertoires de groupes, sur un axe qui va de la dissémination maximale à la polarisation.

Classes lexicales fermées

Le lexique unifié ici établi obéit à la tendance à la normativité maximale.

Dans cette catégorie, les lexies complexes ou composées, groupes de mots figés, constituent un des prérequis obligatoires de toute opération de traitement dans la perspective de l'élaboration du dictionnaire d'apprentissage.

Conditions et contraintes : le savoir dire asturien

(1) Opérateurs phonologiques

Les critères de réalisation phonétique aboutissent à des représentations présentant sous l'angle de la dissémination un caractère maximal, qu'il convient d'optimiser.

(2) Opérateurs morphologiques

(3) Opérateurs syntaxiques

Ce sont les opérateurs phonologiques, morphologiques, syntaxiques, qui construisent la noamalité dialectale (polarisation)

d'aBa_se_jos

(4) Contours et opérateurs prosodiques

On s'assigne de repérer des indicateurs d'asturianité.

(5) Incertitudes

Elles sont dues à l'hétérogénéité des graphies et des représentations qui peuvent leur être associées.

(6) Polymorphisme : un concept douteux, mais nécessaire ?

Le polymorphisme, en rapport avec la diversité de la matière même transmise, inscrit dans la matière même, renvoie une image optimisée de la partition de l'asturien en trois sous-ensembles: central, occidental et oriental, comme on peut s'y attendre. Mais la

géographie linguistique permet de représenter, à l’opposé, les facteurs de dissémination maximale, de fragmentation maximale.

Cela suffit pour rassurer le lecteur: non, nos objets de linguistes applicateurs de sont ni des produits de l’élevage, ni des andouilles, ni des produits de charcuterie, ni des culs de vaches, chers à Jean Belgarric et Marie-Madeleine Chiclet-Rivenc - mais bien des représentations, monde de théories et modèles, en rapport indirect avec le capital tissé de dimensions sociale, culturelle et langagière.

Mais pourquoi, pourquoi ne nous a-t-on pas appris à l’école, non pas les langues des pays, mais ce qui, à partir d’elles, pouvait devenir objet de connaissance, et donner matière à reconnaissance de gigantesques avens culturels cachés, souterrains, fleurines enfouies à jamais, noyées dans le prétendu patrimoine monumental? lieu d’exploration des rapports historiques complexes entre peuples, cultures et pouvoirs? ce ne sont pas les corpus qui sont importants; pas davantage les linguistiques de corpus, les outillages, mais ce qui, avec ces instruments, avec ces outils, et à partir de la matière de base, peut permettre au plus grand nombre (au peuple, par conséquent, à Gaby au mas Raynal, par exemple, à Monicien dans la Grande Lande, comme aux campesinos des Asturies) de se retrouver dans la langue de sa culture, au coeur des avens: il y a loin de Gaby (voix des avens) à José Bové (voix du Festival de Cannes); comme il y a loin des cultures de la didactique aux voies des linguistiques fondamentales; être sur tous les fronts à la fois; et, du fonds d’un aven, ou d’une brana asturienne, à partir de données basses, - *des granhòtas, des abelhai, des ruscats* - réinventer et réactualiser quotidiennement la méthode VOX, chère à Paul Rivenc, loin des grands axes commerciaux du commerce libéral des grandes langues, loin des grands spectacles (shows), et expositions: n’est-ce pas une des voies d’une industrie plus équitable des langues pour qui le traitement est trop longtemps demeuré un luxe ?

Ce qui prime ici est une méthode qui permette en mode simultané d’ouvrir des fenêtres sur l’articulation des dimensions du fait dialectal asturien, déjà complexe en soi, quand il est confronté à la dimension de la castillanité: hommes et leur milieu; dimension variationniste territoriale; représentation des langues en contact; variation normée et fenêtrage des variétés territorialement normées. Les dialectologues empiriques passeront longtemps encore pour des collectionneurs «exemplaristes» au regard des théoriciens de la communication et des Sciences du langage; ce n’est pas ce qui nous détournera du combat qui nous mobilise. L’essentiel est que nous avons très tôt appris, grâce à la linguistique appliquée, telle que pratiquée au CREDIF, à Toulouse, Besançon, etc. à retourner la satire et l’esprit critique contre nous mêmes et appris à nous sauver des grandes machines de traitement lexicométrique en maîtrisant des outillages de traitement suffisants, robustes, appropriés à la nature des problèmes soulevés, compte tenu de la taille et de la qualité des données dialectales à mesurer, en rapport avec une systématisme complexe. Ce qui nous conduit à renier bien des pas antérieurs nécessaires : *cau denegar cada pas* !

La base de textes offerte à P. Rivenc est accessible sur notre site erss/clid/occitan.