Présentation

Rita Caprini

Electronic version

URL: <http://journals.openedition.org/corpus/2289>

DOI: 10.4000/corpus.2289

ISSN: 1765-3126

Publisher

Bases ; corpus et langage - UMR 6039

Printed version

Date of publication: 1 January 2013

Number of pages: 7-13

ISSN: 1638-9808

Electronic reference

Rita Caprini, « Présentation », *Corpus* [Online], 12 | 2013, Online since 12 May 2014, connection on 24 September 2020. URL : <http://journals.openedition.org/corpus/2289> ; DOI : <https://doi.org/10.4000/corpus.2289>

Présentation

Rita CAPRINI
Università di Genova

L'édition scientifique de la revue CORPUS, que le comité de rédaction m'a confiée (et je le remercie de sa confiance), porte le titre de « Dialectologie : corpus, atlas, analyses » qui exprime bien la complexité de la tâche que certains linguistes (mes linguistes préférés...) abordent, non seulement pour composer ce numéro de la revue, mais aussi dans leur travail quotidien. Comme il fallait s'y attendre, mon appel, diffusé il y a une année, a obtenu un bon pourcentage de soumissions qui regardent les atlas linguistiques, étant donné que mon nom est désormais connu surtout pour mon travail dans le domaine de la géographie linguistique.

Ce numéro contient donc dix interventions, parmi lesquelles trois (Baiwir & Renders, Canobbio et Goebel) concernent strictement les atlas linguistiques (considérés comme *corpora*) ; un (Le Bris) est dédié à l'onomastique (autre discipline que j'aime beaucoup) ; Sauzet & Brun-Trigaud participent au débat sur la codification graphique de l'occitan, à partir des données du *Thesaurus Occitan*, base de données qui est l'objet aussi d'autres articles de ce numéro de CORPUS ; la contribution de Malfatto porte sur le recueil et l'analyse des données en dialectologie, à partir d'exemples lexicaux précis empruntés justement au *Thesaurus Occitan* ; Dalbera pose le problème général de la place de la dialectologie au sein de la linguistique générale et analyse la méthode de reconstruction qui a permis de rénover les études lexicales. Enfin, ce numéro se termine par trois études linguistiques appliquées à des données dialectales : l'une (Bucci) s'intéresse à deux processus phonologiques de certains dialectes d'Italie ; les deux autres sont des études de syntaxe, avec un article (Faure & Oliviéri) consacré à la topicalisation en occitan et un article (Ledgeway) qui montre l'importance du rôle des

données romanes, et surtout de celles tirées des variétés non-standard, dans l'évaluation et dans l'enrichissement des théories syntaxiques actuelles.

Je vais à présent discuter brièvement les résultats du travail des auteurs.

Esther **Baiwir** et Pascale **Renders**, chargées de recherches FNRS à l'Université de Liège, présentent un projet d'informatisation des atlas linguistiques pour les rendre plus accessibles à la communauté scientifique. Un tel projet est à l'étude pour l'Atlas linguistique de la Wallonie (ALW), un atlas interprétatif, dans lequel les matériaux bruts ont été enrichis par l'apport de diverses informations, de types phonétique, étymologique, historique et ethnographique. Pour les auteurs, la même chose vaut pour les atlas basés sur des sources secondaires, tels que *l'Atlas Linguistique Roman*. L'atlas informatisé deviendrait ainsi un *corpus* pour des études portant sur le(s) dialecte(s) considéré(s). Un des problèmes du projet est constitué par la transcription phonétique des formes, étant donné que chaque système de transcription est en lui-même une interprétation. Le problème de la transcription est affronté par d'autres auteurs de ce numéro de CORPUS, et les solutions proposées sont différentes. A propos de l'informatisation des atlas linguistiques, je dois ajouter, et je parle par expérience, que les atlas linguistiques sont en général des objets peu maniables et très coûteux. Il faut y penser.

Sabina **Canobbio** (Università di Torino) nous parle de son travail à l'ALEPO (*Atlante linguistico ed etnografico del Piemonte occidentale*, trois volumes déjà parus chez les Edizioni dell'Orso, Alessandria : I. La fauna, II. Caccia e pesca, III. Il mondo animale). Canobbio esquisse un bilan de son expérience, du recueil des données à la constitution d'un *corpus* « imposant et complexe », dans lequel les ethnotextes jouent un rôle particulier. Ces textes sont des chansonnettes d'enfants, des devinettes etc., qui ont fait la différence par exemple entre des grandes œuvres de la géographie linguistique comme *l'Atlas Linguistique de France* et *l'Atlante Italo Svizzero*. Dans l'AIS, la partie ethnographique est souvent imposante (il suffit de

Présentation

revoir la carte des noms de *coccinella septempunctata* littéralement couverte par des ethnotextes qui ont déjà fait l'objet de nombreuses études).

Le but de l'article de Hans **Goebel** (Université de Salzburg) est d'analyser le substrat méthodique et idéologique de l'ALF et de son créateur Jules Gilliéron, mort en 1926. L'auteur se demande pourquoi les atlas linguistiques ont vu la lumière en France et pas ailleurs. La réponse se trouve dans la départementalisation de la France survenue après la Révolution Française, qui a créé une nouvelle conscience de l'espace national basée sur l'idée de quadrillage, de l'équidistance et de la parité fondamentale des nouvelles subdivisions territoriales et a offert des nouvelles possibilités pour les intellectuels œuvrant au service de l'administration nationale. La différence avec d'autres pays, comme l'Italie, et d'autres atlas, comme l'AIS, est bien évidente. Il suffit de regarder, dans les atlas comme l'ALiR ou l'ALE auxquels nous travaillons aujourd'hui, la différence entre le quadrillage français et celui des autres pays... Pour Goebel, la bipartition linguistique de la France en Oc et Oïl a fini par devenir une pomme de discorde scientifique et presque un risque pour l'unité de la nation. En Allemagne, la même question s'est posée pendant le XIX^e siècle pour la frontière linguistique entre Hoch- et Niederdeutsch, question qui a aussi aidé la géographie linguistique à se développer comme science autonome au sein de la linguistique. Goebel souligne que pour faire d'un atlas un *corpus* utilisable par tous il faut s'en tenir aux principes de la collecte sélective de données, en laissant de côté les réponses doubles trop nombreuses (« Qui trop embrasse, peu étreint »). De ce principe s'est inspiré l'ALD (*Atlante linguistico del ladino dolomitico e dei dialetti limitrofi*) de Goebel.

Daniel **Le Bris** (CRBC Brest, UBO-UEB) prend en compte les patronymes des personnes nées entre 1891 et 1990 dans les cinq départements de la Bretagne historique, à partir de l'ensemble des fichiers patronymiques conservés par l'INSEE (Institut National de la Statistique et des Etudes Economiques). Un nouveau logiciel permet maintenant de cartographier tous ces fichiers. Le territoire considéré par Le Bris constitue la zone celtique de la péninsule armoricaine, et les cartes produites par

l’auteur nous permettent de constater la longue durée des frontières linguistiques : voir par exemple la carte 2 « Anciens territoires gaulois en Armorique ». Le Bris constate en conclusion la relative inertie de la population bretonne sur plusieurs siècles, et la concordance entre aires linguistiques et aires culturelles dans la péninsule armoricaine. Cette concordance pourrait bien s’inscrire dans le paysage depuis l’Age des Métaux. Je souligne que le travail de Le Bris nous aide encore une fois à hausser les dates de l’histoire linguistique d’Europe. En ce qui concerne l’étude des noms de famille, il faut noter que ce champ de travail n’est fouillé, avec une certaine régularité, que depuis quelques dizaines d’années. Nous disposons certainement d’un travail séculaire des linguistes surtout sur l’étymologie des noms de famille en Europe, mais depuis peu seulement les données onomastiques sont abondamment mises à la disposition des chercheurs, grâce surtout au traitement informatique des données, qui demandait jadis une grande fatigue manuelle.

Patrick **Sauzet** (Université Toulouse-Le Mirail) et Guylaine **Brun-Trigaud** de l’UMR 7320 « Bases, Corpus, Langage » déjà citée ci-dessus, se proposent de présenter la logique de la codification graphique « classique » de l’occitan, en partant des données du THESOC. La graphie « classique » et celle « mistralienne » ont des traits en commun et des différenciations que l’article se propose d’illustrer.

Albert **Malfatto** de l’Université Nice Sophia Antipolis travaille sur des *corpora* lexicaux dialectaux et participe à des enquêtes de terrain. Le *corpus* utilisé pour cette étude est issu de la base dialectale occitane *Thesaurus Occitan* (THESOC) qui regroupe les données dialectales occitanes qui se trouvent dans les *Atlas linguistiques et ethnographiques de France par régions*, avec les données supplémentaires recueillies par l’équipe de dialectologie du laboratoire dans lequel Malfatto travaille (UMR 7320 « Bases, Corpus, Langage », CNRS / Université Nice Sophia Antipolis). Les localités présentes dans la base sont au nombre de 845. Malfatto nous présente une réflexion sur le thème de l’élevage, il examine en particulier les dénominations dialectales du système notionnel des ânes, ce qui lui permet de conclure que les catégorisations établies par les

Présentation

communautés des locuteurs sont souvent plus complexes que le niveau de détail sollicité par le questionnaire lexical.

Jean-Philippe **Dalbera**, de l'Université Nice Sophia Antipolis, nous offre une discussion sur la dialectologie comme science. Il se demande si elle est devenue « une discipline vieillotte, passéiste, folkloriste, un savoir superflu », une discipline mineure pour les linguistes, l'étude des *parlers locaux* moribonds, qui semble s'occuper parfois seulement de favoriser la diffusion et promouvoir l'enseignement de telles langues à des générations qui ne la connaissent plus. En réponse à ces questions, Jean-Philippe Dalbera souligne le maître-mot-clé définitoire de la dialectologie qui est *reconstruction*. En effet, ce qui caractérise la dialectologie c'est la prise en compte conjointe du temps et de l'espace.

Le principe de la reconstruction a été bien appliqué aux domaines de la phonologie et de la morphologie. Le secteur du lexique l'a beaucoup moins sollicité. Nous avons essayé, Dalbera, moi et quelques autres, de remédier à ce manque, suivant surtout le sillage de Mario Alinei. La grande difficulté de cette démarche est que les chercheurs doivent travailler avec leur esprit et leurs représentations modernes, c'est-à-dire exactement avec l'outil qui a le plus changé au cours des derniers siècles... Il est difficile pour le chercheur qui se penche aujourd'hui sur la reconstruction sémantique dans le domaine des parlers locaux de comprendre comment les locuteurs pouvaient autrefois croire qu'un oiseau sauvage soit l'ancêtre ou bien le père nourricier des hommes et des femmes de son village, mais c'est pourtant bien cela dont on a la trace. Et c'est ainsi que l'on trouve un *père-loriot* et une « tante renarde » etc. un peu partout dans la Romania, comme on peut le voir dans les données des atlas linguistiques. Et qui se souvient encore qu'autrefois on offrait sur une petite table du pain et du fromage à la belette, comme à un dieu, pour l'adorer ou bien pour détourner son inimitié ? Qui pourrait aujourd'hui craindre qu'une chenille puisse sucer le sang d'un homme jusqu'à sa mort ? Je cite ici des croyances qui ont donné origine à plusieurs noms romans des petites bêtes citées. Autant de dénominations qui ont constitué de notoires *cruces* de la recherche étymologique.

Dalbera propose donc ici un outil très performant pour la reconstruction sémantique, le cycle qui va de la motivation, à la convention, à l'arbitraire, pour recommencer avec une nouvelle motivation (remotivation), et cela sans fin. La démarche étymologique classique constate que la forme phonique est opaque, qu'il n'y a pas un mot latin auquel la relier, et tourne court. Seule la reconstruction sémantique, avec son cycle, permet de trouver l'explication de nombreux mots *opaques* des langues romanes.

La dialectologie n'est donc pas la discipline « défaite, moribonde » que ses voisins rognent de toutes parts, mais elle a retrouvé une place de choix au sein des sciences du langage avec l'idée d'une double reconstruction des deux plans du signe, en proposant la reconstruction du signifié à partir du motif utilisé comme clef.

Jonathan **Bucci** de l'Université Nice Sophia Antipolis nous offre la seule intervention phonologique de ce numéro. Il se propose d'étudier deux processus phonologiques bien distincts qui sont la réduction des voyelles atones à schwa que l'on trouve dans certains dialectes italiens et le redoublement phono-syntaxique induit par l'accent (*Raddoppiamento Fonosintattico*). L'enquête est conduite sur le coratin, un dialecte de l'aire apulo-barese parlé dans la ville de Corato située dans la région des Pouilles à environ 40 km au nord de Bari. Le *raddoppiamento fonosintattico* induit par l'accent est un processus de sandhi externe qui consiste, dans une séquence [mot1+mot2], à redoubler la consonne initiale du second lorsque la voyelle finale du premier est tonique (ex : *città ppulita* vs *casa pulita*). L'objectif de l'article de Bucci est de montrer que les deux phénomènes sont incompatibles, i.e. qu'aucun système ne peut les combiner.

Richard **Faure** et Michèle **Oliviéri** (Université Nice Sophia Antipolis) placent leur intervention dans le cadre de la grammaire générative, dont les dernières avancées permettent de repenser un problème d'ordre syntaxique comme la topicalisation en occitan. Le matériel linguistique sur lequel les auteurs travaillent est fourni par les études de Lahne et Sauzet et également par la base de textes du *Thesaurus Occitan*, qui est exploité aussi par d'autres auteurs de ce numéro de CORPUS.

Présentation

Les divers types de dislocation / topicalisation dans les langues romanes ont été bien décrits depuis les années 1980 et Faure & Oliiviéri observent que les dialectes occitans modernes ont développé une stratégie particulière : dans la périphérie gauche de la proposition, on trouve une tête fonctionnelle porteuse d'un trait fort entraînant une topicalisation « suspendue » dans les subordonnées, et les dialectes occitans montrent une partition en fonction de ce paramètre. Cette « troisième topicalisation » en occitan pourrait faire de cette langue un isolat parmi les langues romanes, malgré des phénomènes ressemblants dans certains dialectes italiens.

Adam **Ledgeway** (University of Cambridge) dans son imposant article se propose de mettre en relief, en puisant dans les données des variétés dialectales romanes, l'importance de leur rôle dans l'enrichissement des théories syntaxiques actuelles. Ledgeway travaille sur des *corpora* rassemblés par lui-même pendant ses recherches surtout en Italie et discute de nombreux exemples sur le problème du sujet nul, l'ordre des mots, la position du verbe ou de l'adjectif dans la phrase, etc. Il constate que la variation typologique dans les langues romanes (et dans l'indo-européen) est plus grande que celle que la tradition admet. En conclusion, il démontre encore une fois la grande importance de la linguistique romane pour les études de linguistique générale. En effet, les *corpora* recueillis dans les dictionnaires (comparés et dialectaux) et les atlas linguistique de l'aire romane, ajoutés à la grande histoire du latin, nous aident à comprendre la variation dans le temps et l'espace qui dans d'autres traditions linguistiques (en Europe et ailleurs) est plus difficile à discerner.