HAL
open science

L'informatique dans la recherche hispanique en France

René Pellen

To cite this version:

René Pellen. L'informatique dans la recherche hispanique en France. Bulletin de l'EPI (Enseignement Public et Informatique), 1987, 48, pp.180-184. edutice-00000997

HAL Id: edutice-00000997

https://edutice.hal.science/edutice-00000997v1

Submitted on 21 Oct 2005

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

L'INFORMATIQUE DANS LA RECHERCHE HISPANIQUE EN FRANCE

René PELLEN

l'informatique n'occupe pas encore une très large place dans la recherche hispanique en France. Mais la situation évolue rapidement. A ce jour, même si sa place reste modeste, elle est représentée dans des domaines aussi variés que l'apprentissage des langues, la documentation et la recherche linguistique et littéraire.

Dans le premier domaine, Marie-Madeleine Segonnes, par exemple (Rochefort s/Mer - Poitiers) a mis au point divers logiciels d'aide à l'apprentissage de la conjugaison, à la révision des normes orthographiques... Son travail est un travail de pionnière, car peu d'hispanistes semblent préoccupés d'élaborer des didacticiels de qualité ; ce qui limite souvent l'initiation à l'informatique, dans les universités où elle existe (ex : Dijon, responsable : Jean-Claude Villegas), à une présentation générale de pratiques et de techniques (traitement de texte, gestion des fichiers, interrogation de base de données, application de programmes statistiques). Une initiation aussi banalisée a, certes, l'avantage de valoir pour toutes les disciplines de Lettres et Sciences Humaines, mais n'apporte rien de spécifique aux étudiants et aux futurs chercheurs hispanisants.

Dans le domaine de la documentation, le G.I.S. 15 "Maison des Pays Ibériques" (C.N.R.S. et Université de Bordeaux III ; responsable : Joseph Perez) a créé une base de données (HISPABIB) qui intègre plusieurs fichiers thématiques :

- ARCHIBER (histoire ancienne et archéologie de la Péninsule Ibérique) ;
- HISPADOC (histoire, civilisation, langue et littérature hispaniques) ;
- IBERGEO (géographie des pays ibériques) ;
- BIBLIBER (catalogue collectif de périodiques couvrant le domaine ibérique).

Ces fichiers s'étendent peu à peu, au rythme de 7 000 références environ par an. Ils contiennent des renseignements sur la documentation primaire (localisation...) et une indexation des données. Implantée à Montpellier (CNUSC), HISPABIB peut être interrogée par Transpac - en ligne - ou par courrier ou téléphone.

Les hispanistes, au niveau individuel, s'intéressent également à la documentation, comme en témoigne l'ouvrage récent de J.-C.. Villegas, *La littérature hispano-américaine publiée en France 1900-1984*, Paris, Bibliothèque Nationale, 1986, 260 p. (Études, Guides et Inventaires : 4). Ce répertoire fournit la liste des notices bibliographiques correspondant aux ouvrages de littérature latino-américaine publiés en France entre 1900 et 1984. Chaque notice est accompagnée de la cote B.N. de l'ouvrage. 3 index complètent l'inventaire : index alphabétique par pays et noms d'auteurs, index des traducteurs, index des préfaciers. L'essentiel de ce travail a été réalisé sur un Apple II de 128 Ko.

Dans la recherche proprement dite, seuls quelques noms peuvent être cités, lors même que les travaux réalisés sont parfois importants. Véronique Huyhn et son équipe ("Analyse Relationnelle des Textes", rattachée au Laboratoire de Statistiques de Paris VI) ont étudié divers corpus d'espagnol contemporain, notamment à l'aide de l'Analyse des Données. Cf. l'ouvrage publié en 1977 par V. Huyhn : *Des structures syntaxiques de l'espagnol à l'analyse relationnelle des textes*, Paris, Champion, 167 p.. Carmen Pineira Tresmontant a récemment analysé la constitution espagnole, en mettant à profit l'expérience de l'U.R.L. III de l'Institut de la Langue Française (responsable : Maurice Tournier- - cf. sa thèse de 3° cycle La Constitution espagnole de 1978. Etude linguistique informatisée, Paris III, 1984. Carmen Salaun a, de son côté, effectué diverses études en recourant à l'Analyse des Données - cf. "L'Usage de l'opposition modale en espagnol : analyse d'une enquête", in *Pratique de l'Analyse des Données*, t. III, *Linguistique et Lexicologie*, Paris, Dunod, 1980.

Jean Roudil (directeur du Séminaire d'Études Médiévales Hispaniques, Paris XIII) a fait dépouiller sur ordinateur plusieurs grands textes médiévaux ; notamment la *Primera Cronica General* (XIIIe siècle), dont il a publié en 1979 l'Index alphabétique des formes (in *Cahiers de Linguistique Hispanique Médiévale*, n° 4 bis, p. 205-365). D'autres index sont en préparation - notamment l'index des adjectifs - ainsi que diverses études.

René Pellen (Université de Poitiers/C.N.R.S.), poursuit depuis 1967 des recherches informatisées sur l'épopée, en particulier sur la *Geste du Cid*. Cf. *Poema de Mio Cid : dictionnaire lemmatisé*..., Paris, Université de Villetaneuse, 1979, 307 p. (Annexes des *Cahiers de Linguistique Hispanique Médiévale* ; 1). Il a publié les résultats de ses recherches dans les CLHM (1976-1985) et dans d'autres périodiques. Ses études concernent la langue du Cid, la datation et la structure du poème, la prosodie épique, l'évolution de l'espagnol, la quantification en linguistique, la méthodologie de la recherche informatisée (cf. "PECI : un système de dépouillement de textes", en collaboration avec Claude Siredey, in *La Licorne*, publication de la Faculté des Lettres et des Langues de l'Université de Poitiers, n° 4, 1980, pp. 122-141 ; "DILEM : construire un dictionnaire lemmatisé avec l'informatique...", ibid., n° 7, 1983, pp. 197-231).

C'est maintenant dans le cadre de l'U.A. 998 du C.N.R.S. (Laboratoire d'Études et de Recherche sur le Moyen Age) et de l'une des principales composantes, l'ancienne R.C.P. "Sources médiévales antérieures à 1300", qu'il prépare le dépouillement de nombreux textes antérieurs à 1250 : *Liber Regum*, *Fazienda de Ultramar*, *Roncesvalles*, *Milagros de Nuestra Senora*, *Documentos linguisticos* publiés par R. Menéndez Pidal, etc. Tous ces textes ont été saisis : certains ont déjà été dépouillés et sont en cours d'étude ; le premier volume du *Dictionnaire Glossométrique des "Milagros"* est presque achevé. L'Index Lemmatisé du *"Liber Regum"*, lui aussi, devrait être publié prochainement.

L'un des projets de l'U.A. est, en effet, de publier des instruments de travail (index, dictionnaires, concordances...) qui devraient renouveler la connaissance des langues médiévales et les conditions de la recherche diachronique. L'U.A. a créé de nombreux logiciels propres pour dépouiller les textes (PECI) et pour lemmatiser leur vocabulaire (DILEM).

Elle assure également, en collaboration avec la Bibliothèque Universitaire de Poitiers, le dépouillement de dix périodiques spécialisés. À partir du fichier ainsi créé (qui va se convertir, avec l'informatisation de la Bibliothèque Universitaire, en une base de données interrogeable à distance), sont publiées des bibliographies récapitulatives : cf. R. Pellen et Jacques Pradines :

-"L'Informatique et les Humanités d'après quelques périodiques spécialisés (1978-1981)", in *La Licorne*, n° 7, 1983, p. 323-356 ;

- "L'Informatique et l'étude des langues... Bibliographie 1979-1983. I. Langues 'rares'", ibid., n° 11, 1986, p. 139-161 ;

- "L'Informatique et l'étude des langues... Bibliographie 1979-1983. II. Langues les plus étudiées", ibid., sous presse ;

- "Méthodologie informatique dans les Humanités. I. Logiciels. Bibliographie... 1979-1984", in *Linguistic and Literary Computing*, 1986, vol I, n° 2, p. 86-92 ;

- "Méthodologie informatique dans les Humanités. II. Statistique... Bibliographie... 1979-1984", ibid., 1986, vol I, n° 3, p. 173-175 (A-J), et n° 4, p. 216-220 (K-Z) ;

- "L'informatique et les Humanités. Bibliographie 1984-1986...", ibid., sous presse.

Pour en revenir au domaine de l'hispanisme, un autre chercheur, Michel Garcia (professeur à Paris III) a rejoint l'U.A. en 1983. Il a commencé avec l'informatique diverses études, notamment sur les *Lettres du Maure de Grenade* (cf. communication au 1er Congrès de Linguistique Hispanique - Rouen, 1985 - : "Les Lettres du Maure de Grenade", Actes..., *Cahiers du C.R.I.A.R.*, n° 6, 1986, pp. 175-178. Le texte des mss. E et N du *Rimado de Palacio*, du Canciller Ayala (fin du XIVe siècle) a été ensuite mémorisé : lorsqu'il sera au point, il fera l'objet d'une recherche conjointe de R. Pellen et de M. Garcia sur la langue du Canciller. La publication du vocabulaire lemmatisé est prévue pour une étape ultérieure.

Avec des préoccupations différentes, mais selon une démarche qui les confronte aux mêmes problèmes que la plupart des chercheurs précédents (lemmatisation, étude quantitative du vocabulaire et des structures linguistiques...), Colette Noyau et Jorge Giacobbe s'intéressent, quant à eux, à l'acquisition d'une langue par des adultes immigrés. Leurs recherches relèvent d'un programme européen qui porte sur dix couples de langues. Le G.R.A.L. (Groupe de Recherche sur l'Acquisition des Langues Paris VIII), dont ils font partie, se consacre au couple espagnol langue source-français langue cible. Les recherches actuelles concernent plus spécialement le lexique des apprenants, son évolution, les méthodes à mettre au point pour l'analyser et pour comparer les résultats obtenus par le G.R.A.L. avec les résultats des autres équipes européennes. C. Noyau a exposé l'état de ces recherches dans plusieurs communications récentes : cf., par exemple, l'exposé fait au Colloque de Brest (IIe Colloque de Linguistique Hispanique, avril 1986), "Le Développement du lexique en langue étrangère : lemmatisation des données orales d'acquisition" (à paraître dans les Actes du Colloque).

Si, comme on le voit (mais ce bref panorama n'est sûrement pas exhaustif), le nombre des hispanistes engagés dans l'informatisation reste encore peu élevé, le nombre des chercheurs qui ressentent le besoin de recourir à l'informatique ne cesse de croître. Ce besoin s'est manifesté avec vigueur au Colloque de Brest. Une table ronde consacrée au thème "Linguistique et Informatique" et diverses communications ont abouti à des projets qui sont en cours d'étude. Certaines réunions ont eu lieu (Paris, octobre 1986). D'autres s'annoncent (Bordeaux ou Grenoble, 1978). Une collaboration s'est établie entre l'Université de Rouen (Jack Schmidely) et l'Université de Poitiers (R. Pellen) pour le dépouillement et l'étude de textes contemporains, sur lesquels travailleront, entre autres, les étudiants de maîtrise des deux universités.

Sur le plan documentaire, à la suite des demandes formulées au Colloque, le G.I.S. de Bordeaux a fait des propositions :

- réalisation d'une base de données de linguistique hispanique à partir du dépouillement régulier de quelques revues spécialisées ;
- réunions thématiques sur les problèmes concernant l'informatisation de la documentation.

Un autre projet est en train de prendre corps : il consisterait à créer, en France, un Observatoire de la langue espagnole. La plupart des participants au Colloque de Brest ont manifesté leur intérêt pour la création d'un tel Observatoire.

Enfin, les hispanistes français ont décidé de prendre contact avec leurs collègues espagnols et, d'une manière plus générale, avec leurs collègues étrangers, pour harmoniser d'un pays à l'autre l'informatisation de la recherche. Étant donné que dans tous les pays l'informatique s'implante désormais dans la recherche au niveau de la maîtrise - ou de son équivalent étranger - (quand ce n'est pas plus tôt), il est probable que dans les années qui viennent le nombre des hispanistes informatisants va croître rapidement. Il est probable aussi que le contexte général de la formation des étudiants (initiation à l'informatique ou recyclage) et de développement des infrastructures de recherche (informatique documentaire : réseaux REUNIR, EARN...), de même que l'extension de la micro-informatique vont modifier profondément les méthodes et les comportements. De toute façon, les années 1986-87 auront marqué un tournant pour l'hispanisme sous la forme d'une prise de conscience collective de l'impossibilité, à l'heure qu'il est, de se passer de l'informatique

René PELLEN,
Université de Poitiers