

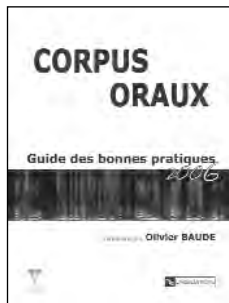
Corpus oraux

Un guide des bonnes pratiques

La Délégation générale à la langue française et aux langues de France (DGLFLF) du ministère de la culture et de la communication vient de publier, dans le cadre de son programme « Corpus de la parole », un Guide des bonnes pratiques destiné aux chercheurs collectant et diffusant des archives sonores.

Le programme « Corpus de la parole »

Depuis 2004, la DGLFLF a entrepris un grand programme pour la numérisation et la valorisation des corpus oraux. Un corpus oral est une collection ordonnée d'enregistrements de productions linguistiques orales et multimodales réalisées par des chercheurs. La France dispose d'une grande richesse linguistique fondée sur la diversité. À côté du français, langue nationale, présent sur les cinq continents, les langues de France constituent un patrimoine culturel unique. Or ce patrimoine est méconnu, et si des archives sonores existent pour la quasi-totalité de ces langues, force est de reconnaître que cette richesse, constituée du français parlé et de la diversité des langues de France, n'est accessible ni à l'ensemble de la communauté scientifique, ni au grand public. Plus grave encore, de nombreux documents sonores uniques, conservés sur des supports physiques en fin de vie (bandes magnétiques), sont voués à disparaître à tout jamais dans un délai très bref. Il s'agit souvent des derniers et seuls documents sur des langues de France et sur le français. Le ministère de la culture (DGLFLF) a souhaité initier un programme de numérisation qui offre non seulement la possibilité de sauver ces documents, mais aussi de les valoriser en les transformant en



Corpus oraux, Guide des bonnes pratiques, 2006
Coordonné par Olivier Baude
Paris : CNRS-Éditions et Presses universitaires d'Orléans, 2006.
192 p., 14 €.

de véritables ressources linguistiques numériques, pour la recherche en sciences humaines, l'enseignement et l'ingénierie des langues.

Le guide des bonnes pratiques

Cependant, la numérisation de ces corpus et leur valorisation par une accessibilité à un très large public ne vont pas sans poser de nombreux problèmes juridiques et éthiques. La DGLFLF a donc commencé par créer un groupe de travail chargé de rédiger un guide de bonnes pratiques. La rédaction de ce guide représente une expérience unique. Elle réunit différentes communautés : chercheurs universitaires et du CNRS, conservateurs, juristes et informaticiens pour élaborer une démarche commune dans le respect de la loi et de l'éthique.

Cette démarche a comporté plusieurs étapes. Une lecture croisée des textes juridiques par les linguistes et les juristes a permis de repérer les problèmes. Les chercheurs ont ensuite accepté d'explicitier leurs pratiques au regard de la législation. Cette étape fondée sur la réflexivité a permis d'élaborer des propositions pour de *bonnes pratiques* partagées par la communauté scientifique et de repérer des aspects juridiques qui posent des difficultés dans l'état actuel du droit.

Le guide publié en mai 2006 est l'aboutissement de deux années de travail, ponctuées par l'organisation d'une journée d'étude tenue à la BnF et consacrée à une version provisoire. Une version anglaise est en cours, et sera suivie d'une version espagnole. Il s'agit bel et bien de la première étape d'un programme qui doit permettre de pallier le retard de la France dans la diffusion du patrimoine linguistique numérique national, reposant sur une initiative, unique en Europe et au niveau international, de numérisation du patrimoine dans le respect des nouvelles technologies de conservation.

Olivier Baude

Ministère de la culture et de la communication
Délégation générale à la langue française et aux langues de France
olivier.baude@culture.gouv.fr

Corpus de la parole Langues et cité n° 6, mai 2006. 12 p.

Ce numéro présente un état de la recherche sur les corpus oraux, ainsi que les actions du programme « Corpus de la parole » de la Délégation générale à la langue française et aux langues de France (DGLFLF) du ministère de la culture. *Langues et cité*

est le bulletin de l'Observatoire des pratiques linguistiques de la DGLFLF.

À télécharger :

http://www.culture.gouv.fr/culture/dglf/Langues_et_cite/Langues_cite6.pdf



26 septembre 2006

Journée européenne des langues

Depuis 2001, Année européenne des langues, le Conseil de l'Europe a fait du 26 septembre la Journée européenne des langues, pour célébrer la diversité linguistique et informer les Européens sur le plurilinguisme. Organisée autour de trois thèmes : parler, comprendre, traduire, cette journée mettra notamment en avant :

- le principe de l'intercompréhension entre langues apparentées ;
- les technologies les plus avancées de traitement automatique du langage (traduction assistée par ordinateur, logiciels de gestion des mémoires de traduction...).

La DGLFLF propose une sélection d'études et rapports sur le plurilinguisme :

<http://www.dglff.culture.gouv.fr> (rubrique « études et recherches »)

À télécharger, la brochure *L'intercompréhension entre langues apparentées. Références 2006* (12 p.) : <http://www.culture.gouv.fr/culture/dglf/publications/intercomprehension.pdf>

Le n° 1 des Cahiers de l'Observatoire des pratiques linguistiques

Les rectifications orthographiques de 1990/ Analyses des pratiques réelles (Belgique, France, Québec, Suisse, 2002-2004)

L. Biedermann-Pasques, F. Jecic éd.

Orléans : Presses universitaires d'Orléans, 2006. 160 p.

Cette nouvelle collection propose des « états des lieux » des pratiques langagières en France. Le premier volume est une analyse des pratiques orthographiques réelles des francophones aujourd'hui, à partir d'enquêtes sur les dictionnaires et auprès d'étudiants belges, français, québécois et suisses, quinze ans après les rectifications proposées par le Conseil supérieur de la langue française et acceptées par l'Académie française.

À télécharger : http://www.culture.gouv.fr/culture/dglf/publications/rectifications_ortho.pdf

Document imprimé disponible sur demande à la DGLFLF.

À voir sur ce thème le numéro 7 de *Langues et cité*.