



Une base de données



par et pour

l'ED268



-
-
-
-
-
-
-
-
-
-
-

“ Propositions de Normalisation pour

une Base de Corpus Multimédia ”



à l'ED268

Sonia Branca, Serge Fleury, Cédric Gendrot, Michel Jacobson, Thierry Pagnier,
Patrick Renaud, André Salem, Pollet Samvelian, ... • • • • •

-
-
-

Le cœur du projet

- 1 L'élaboration d'une base de corpus (principalement oraux) regroupant des données de langues de différentes natures
- 2 Des propositions de normalisation pour l'encodage de corpus de langue.

•
•
•

Objectifs ...

- Une approche interdisciplinaire, et dont l'objectif évident est une contribution à la linguistique de corpus.
- **Parmi les membres du projet :**
 - Sociolinguistique : P. Renaud; S. Branca; T. Pagnier
 - Phonétique : C. Gendrot
 - TAL / Informatique : S. Fleury; M. Jacobson; A. Salem
 - Syntaxe, Sémantique : P. Samvelian

-
-
-

Utilités d'une base de corpus

- Archivage de langues. (site LACITO : <http://lacito.vjf.cnrs.fr/archivage/>)
- Analyse / synthèse de la parole.
- Favoriser les collaborations pluri-disciplinaires.
- Aider les jeunes chercheurs pour la constitution de leurs corpus

•
•
•

Offrir aux chercheurs la possibilité de :

- Mettre à disposition leurs corpus.
- Utiliser pour leurs recherches les données stockées dans cette base de données

•
•
•

Contacts et Liens ...

- Site Web :

<http://www.cavi.univ-paris3.fr/ilpga/ilpga/sfleury/pi-ED268-0406/index.html>

- Serge Fleury :
fleury@noos.fr

- Cédric Gendrot :
cgendrot@univ-paris3.fr

• • • • • • • •