

*Collecte et exploitation de corpus
dans le domaine berbère*

Naima Louali

UMR 5596 Dynamique du Langage

CNRS-Lyon 2

naima.louali@ish-lyon.cnrs.fr

Collecte et exploitation de corpus dans le domaine berbère

Pour le berbère, on dispose :

d'importants recueils de textes anciens datant des XIX^{ème} et XX^{ème} siècles

de plusieurs travaux récents (thèses et différents mémoires) s'appuyant sur des textes transcrits

Initier et sensibiliser à la collecte de corpus oraux

Le choix des locuteurs, des situations

La nature du corpus (type et taille)

Les techniques d'enregistrement, etc.

Le travail de terrain (1)

Vaux, Bert & Justin Cooper (1999) *Introduction to Linguistic Field Methods* : Textbooks in Linguistics 1, Munich, Lincom Europa, 196 pages.

Newman, Paul & Martha Ratliff, ed. (2001) *Linguistic Fieldwork*, Cambridge University Press, 288 pages.

Le travail de terrain (2)

La nature des textes

Le matériel et méthode d'enregistrement

Les problèmes de transcription

Le travail de terrain (2)

La nature des textes

Le matériel et méthode d'enregistrement

Les problèmes de transcription

I. La nature des textes

différents types de textes pour *chaque*
localité berbère étudiée

Un échantillon représentatif comme
champ d'application

I. La nature des textes (suite)

- Linguistique
- Traitement automatique de la parole
- Anthropologie
- Didactique (apprentissage à partir de textes), etc.

I. La nature des textes (suite)

-répertoire de contes, légendes, fables et proverbes
(littérature orale) 

étude lexicographique

-récit personnel  **étude prosodique**

 - récit sollicité (« la bise et le soleil » par exemple)

-récit ethnographique

Format de fiche

Date : 3 août 2002

Nom : F.

Prénom : Aziz

Sexe : masc.

Age : 28

Localité : Asni

Tribu : Ighaghayen

Dialecte : tachelhit

Compétence ling. : bilingue
tachelhit / arabe dialectal

Situation géo. : Sud Maroc

Type de textes : récit de vie ,,
proverbe, poème, chant, texte
sollicité, berceuse, conte

Autre :

Commentaires : 2 ans à Marrakech
(adulte)

Enquêteur : N. Louali

Format de fiche (suite)

Date : 3 août 2002

Nom : F

Prénom : Aziz

Sexe : masc.

Age : 28

Localité : Asni

Tribu : Ighaghayen

Dialecte : tachelhit

Compétence ling. : bilingue
tachelhit / arabe

Situation géo. : **Sud Maroc**

Type de textes : récit de vie ,
conte, proverbe, poème, chant,
texte sollicité

Autre :

Commentaires : 2 ans à
Marrakech

Enquêteur : N Louali

Point d'enquête



II. Le matériel et méthodes d'enregistrement

celle qui préconise l'introduction tardive de l'enregistrement tout en lui aménageant des séances spécifiques au cours d'un travail de terrain

celle qui fait de l'enregistrement un outil présent tout au long de l'enquête.

II. Le matériel et méthodes d'enregistrement (suite)

la qualité du matériel (enregistreur +micro)

la disposition du micro à la bonne distance du
locuteur

1 

2 

Pour ce qui est des dialogues : choisir
l'enregistrement sur des pistes séparées

III. La transcription

Une première transcription rapide accompagnée d'une traduction «approximative»

Affiner la transcription en s'aidant de l'enregistrement

- Quelle transcription ?
- Quel alphabet ?

III. La transcription (suite)

La transcription phonétique

La transcription phonologique

La notation ou 'la notation usuelle'

III. La transcription (suite)



1- [jusəd ... akʷbbi lmalik əttlətʃannəs]

III. La transcription (exemple)

1- [jusəd ... akʷbbi lmalik əttlətʃannəs] 

2- /jusd akʷbbi i lmalk d tltʃa-nns/

Il-est-venu enfant à roi et fille-de-lui

“L’enfant est venu auprès du roi et sa fille ”

III. La transcription (exemple ...)



1-[nətta lwahid wənn gizida amafbosnnes
enqreta]

Consonnes
emphatiques /
pharyngalisées

ḏ ḏ^ʕ

ṣ ṣ^ʕ

ẓ ẓ^ʕ

t t^ʕ

r r^ʕ

t +
t^ʕ ..



Louali (1990)

III. La transcription (exemple ...)

1- [nətta lwaħid wənn gizida amaʃbosnəs
enqreta]

2- /ntta lwaħid wnn g-izida amaʃbus^ʃ-nns
inqrit^ʃ-a/

Lui le-seul celui-de dans-ânes- queue-de-lui il-est-dans-l'état-
d'être-coupé

“ Le seul parmi les ânes dont la queue était coupée ”

Pour finir...

“ *produit final* ” → un alignement entre :

-un fichier format.wav (audio)



-un (ou des) fichier(s).text. (transcription/notation)

Merci...

- Nama, Omar et Souliman (Siwi, enquête, mai 2002)
- Abdallah Fili (tachelhit, Massa enquête, juillet 1999)
- Aziz (tachelhit, Asni, enquête, août 2002)

—> Chrisfried Naumann (thèse en préparation sur le Siwi)
Université de Leipzig, Institut Max Plank)