

# 3L Université d'été Lyon 2008 : Résumés des cours et des conférences

## 1. Conférences plénières

### *What is language documentation? How does it differ from language description?*

Peter K. Austin,  
Lundi 23 Juin, 10h30

Cette conférence a pour but de discuter de la nature de la linguistique de documentation, et des relations que cette nouvelle sous-discipline de la linguistique entretient avec la linguistique descriptive. Elle abordera aussi des problématiques actuelles de la linguistique de documentation. Pour un résumé plus détaillé de cette conférence, voir la version anglaise.

### *Data management*

Peter K. Austin  
Mardi 24 juin, 9h00

Le coeur d'un projet de documentation d'une langue consiste généralement en un corpus audio et/ou vidéo accompagné de transcriptions, de différentes sortes d'annotations, de traductions dans une langue à large diffusion, et de métadonnées précisant le contexte et l'utilisation du matériel linguistique recueilli. Idéalement le corpus doit être ample, comprendre une variété de genres et de contextes. Il doit pouvoir être complété par la suite. Il doit être facile à consulter, clairement présenté et "préservable" sur le long terme. Il doit également respecter les principes éthiques. Dans cette conférence, nous aborderons les points suivants :

- Que doit contenir un corpus
- Comment les données du corpus doivent-elle être présentées
- Quelles sont les bonnes pratiques de représentation et de traitement des données
- la nature et le rôle des métadonnées
- le traitement du corpus et l'inventaire des différentes tâches à accomplir
- les différents outils pour la documentation des langues

Nous regarderons aussi des échantillons de données afin d'explorer chacun de ces concepts.

### *Archiving*

David Nathan  
Mercredi 25 juin, 9h00

Même si certains considèrent le processus d'archivage comme la phase finale d'un projet de documentation d'une langue, les archives et les principes d'archivage sont néanmoins impliqués dans les diverses étapes du cycle de documentation des données, et ceci dès le début des projets, de sorte que l'archivage ne devrait pas être considéré comme le seul but final de la documentation.

Cette conférence a pour but de présenter différentes sortes d'archives et leur fonctions respectives, de discuter des questions concernant le stockage, la préservation et la sécurisation sur le long terme du matériel archivé, et des problèmes en rapport avec le contrôle de l'accès aux données. Les centres d'archivage modernes travaillent main dans la main avec les "dépositaires" tout au long de leur projets, en leur dispensant des formations et des conseils, en évaluant des échantillons, en les aidant à convertir leur données dans des formats archivables, et en mettant en place des moyens leur permettant de continuer de mettre à jour leurs archives par la suite. Les archives modernes accordent autant d'importance à la préservation qu'à l'accès aux données, afin que leurs collections servent efficacement les besoins des utilisateurs.

Cette conférence discutera des divers types de données qui peuvent être archivées dans des archives digitales modernes et de la sélection et de la préparation du matériel à archiver en abordant notamment les questions de formats et d'encodage (e.g. XML), des moyens d'acheminement des données auprès de l'archive, et enfin les questions concernant les métadonnées.

*Linguistic Anthropology: Linguistic Practices and their Cultural Context*

Lolke van der Veen  
Vendredi 27 juin, 9h00

Lorsqu'un linguiste veut décrire une langue, deux options majeures se présentent à lui. La première consiste à analyser la langue comme un code formel qui comprend à la fois un ensemble de formes lexicales et grammaticales et un ensemble de règles ou de contraintes capables de produire les énoncés bien formés de la langue. Dans ce cas, le linguiste étudie le système de la langue, c.-à-d. les traits structuraux et fonctionnels de cette dernière, et se situe alors sur un plan plutôt abstrait.

La seconde option consiste à étudier les différentes manières dont une langue donnée peut être utilisée au sein dans un cadre culturel spécifique par les locuteurs qui la parlent. Lorsqu'il choisit cette option, le linguiste analyse les pratiques linguistiques effectives d'une communauté linguistique spécifique. Ces pratiques divergent considérablement d'une communauté à l'autre. Les cultures imposent des normes différentes quant à l'utilisation et la forme des langues dans des situations sociales comparables, et par voie de conséquence des pratiques linguistiques qui sont plus ou moins spécifiques à ces langues. Les linguistes anthropologues étudient les productions linguistiques (des textes ou des parties de textes) à la lumière de ces conventions sociales et culturelles spécifiques. Comment ses productions (orales ou écrites) se situent-elles par rapport aux *genres*, aux *types de discours* et aux *pratiques sociales*? Les mêmes constructions/entités linguistiques peuvent prendre des valeurs différentes selon le genre dans lequel elles sont utilisées. Afin de démontrer que les valeurs culturelles et les croyances contribuent de manière substantielle à la construction de la réalité linguistique et que les activités sociales mettent en jeu des relations entre trois sphères différentes (i.e. sémiotique, cognitive et physique), des exemples seront empruntés à un éventail de langues du monde ainsi qu'à une étude de cas consacrée aux salutations des Galwa, communauté ethnolinguistique bantuphone du Gabon.

Fieldwork Methodology And Ethics

Colette Grinevald

Lundi 30 juin 9h00

Cette conférence traitera de l'entrelacement complexe et multidimensionnel des questions de terrain et des questions d'éthique. Elle abordera cette complexité sous plusieurs angles, en considérant, en premier, la nature des questions à traiter (comme celles de qui est responsable de quoi, envers qui, et quand, sur le modèle du *flow chart* de VW-DOBES), puis l'évolution au cours des dernières décennies de la conception des relations de terrains (à la Cameron : de SUR/POUR/AVEC à, aujourd'hui, PAR les locuteurs), tout en soulignant les différences qui peuvent prévaloir d'un continent à l'autre. Le tout sera ensuite présenté dans un cadre temporel de cycles constitués de phases d'AVANT/PENDANT/APRES le terrain, chaque phase étant liée à des questions d'éthique différentes (depuis la question du choix de la langue et du site, la planification de la mission de terrain, l'entrée sur le terrain, la collecte de données, le départ du terrain, et le retour à la vie universitaire et la production de résultats (analyses, archivage, matériaux)). La conférence recommandera comme guide l'IATSI 2000 ('*Guide pour la Recherche Ethique pour des Etudes Indigènes*') d'Australie, avec ses 11 'principes', accompagnés d'explications et d'applications pratiques, organisés en :

- A. consultation négociations, compréhension mutuelle
- B. respect, reconnaissance et engagement
- C. bénéfiques, résultats et accords.

Par-dessus tout, cette conférence soulignera comment, au delà et en de ça des aspects éthiques liés à des questions légales, technologiques et sociopolitiques, être « éthique » est profondément une attitude ou une posture à adopter face au processus qu'est le travail de terrain, à développer pour chaque situation particulière sur le terrain de la rencontre entre un linguiste et une communauté linguistique particulière d'un pays particulier, et à réajuster pour chaque étape d'un projet.

Data collection methods in field-based language documentation

Friederike Lüpke

Mardi 1<sup>e</sup> juillet, 9h00

Cette conférence a pour but de présenter les différentes méthodes de collecte des données sur le terrain pour de la documentation de langues. Elle illustrera comment les différentes façons de collecter des données influe sur les résultats obtenus; elle explorera comment différentes sortes de données peuvent contribuer à l'élaboration d'un corpus représentatif de la structure d'une langue, de l'intuition de ses locuteurs, et de la diversité de leurs répertoires.

Documenting concepts in contact (conceptional transfer vs. conceptional shift)

Matthias Brenzinger

Jeudi 3 juillet, 9h00

La documentation de concepts en situation de contact exige le repérage, l'analyse et la description de concepts authentiques, non-conformistes qui sont inhérentes aux langues parlées par des sociétés non occidentales. Ce travail pourra s'aborder par l'étude de variations intergénérationnelles dans les concepts relatifs à l'espace, aux nombres, aux couleurs, aux taxinomies du vivant, etc. Les observations qui seront présentées dans cette conférence,

s'appuieront sur des recherches de terrain effectuées parmi des communautés de chasseurs-cueilleurs d'Afrique, principalement les Khwe de l'Afrique méridionale.

Toutes les langues parlées il y a environ 12 000 ans se caractérisaient par une vision du monde de type « chasseur-cueilleur ». La naissance d'économies basées sur l'agriculture et l'élevage a fait émerger de nouveaux concepts relatifs à l'environnement, et par voie de conséquence les vocabulaires et les catégories ont évolué pour s'adapter aux nouveaux besoins en matière de communication. En Afrique rurale, le mode de production demeure encore l'une des composantes les plus essentielles de la culture et forme, en même temps, le fondement de la conceptualisation.

Les concepts du monde se transforment sans cesse, tant sur le plan des individus que sur le plan des communautés. Cette transformation s'accompagne d'une évolution incessante des manières d'envisager et d'expliquer les environnements sociaux et physiques changeants. Ces processus évolutionnaires suivent des schémas de transfert conceptuel unidirectionnels qui vont dans le sens « d'une plus grande abstraction sur le plan des signifiés » (Heine 1997, p. 7).

De tels processus internes de transfert conceptuel s'opposent à d'autres processus – des phénomènes de contact – qui peuvent être caractérisés comme des remplacements conceptuels. Des concepts qui reflètent des points de vue traditionnels autonomes de cultures marginalisées sont rapidement remplacés par des pensées dominantes sur le plan mondial : les membres de telles communautés peuvent adopter les stratégies conceptuelles de la culture dominante ou même, de manière plus radicale, faire leurs des concepts-cibles entiers. Les concepts « qui dominent sur le plan mondial » se répandent au travers de systèmes éducatifs uniformisés et contrôlés par les gouvernements (avec des concepts provenant le plus souvent de cadres scientifiques), ou au travers d'œuvres missionnaires religieuses (avec des concepts provenant de différents systèmes de croyance). Les locuteurs copient des idées abstraites, comme les concepts relatifs aux numéraux, aux directions cardinales, à l'opposition gauche-droite, aux termes de couleurs ou aux noms de personne, etc. A partir de ce qui par ailleurs auraient été des cibles de transferts conceptuels, ils recherchent des termes dans leur lexique pour véhiculer les nouveaux concepts, ce qui signifie un retour aux sources. Des concepts peuvent déjà exister, pour lesquels aucune terminologie (institutionnalisée) n'a encore été inventée : des cibles sans sources.

L'effort qui vise à en apprendre plus sur les structures conceptuelles non-conformistes dans les langues des sociétés non occidentales, nous amène inévitablement à nous rendre compte que ces pensées spécifiques, celles qui divergent des concepts répandus par culture mondiale, constituent les atouts les plus vulnérables de ces langues, et leur documentation devrait recevoir une priorité élevée.

## 2. Ateliers de terrain

### *Phonologie*

Enseignant : Lolke van der Veen (DDL, Université de Lyon)

Horaires : 23 et 24 juin, 14h00 à 15h30 (français)  
25 et 26 juin, 14h00 à 15h30 (anglais)

*Première partie* : introduction. Un aperçu des questions les plus fondamentales comme le choix des informateurs, le choix du type de corpus à constituer (mots, énoncés, textes oraux), la manière de mener une session de collecte de données, la conduite à mener pour rester en bons termes avec les informateurs, le besoin d'une connaissance minimale des aspects morphosyntaxiques de la langue, etc.

*Seconde partie* : travaux pratiques avec un informateur bantophone du Gabon.

- ✓ La collecte de vocabulaire de base à l'aide d'une liste de mots.
- ✓ La perception et la transcription des sons inconnus (segments et traits suprasegmentaux).
- ✓ La transcription étroite.
- ✓ La gestion des corrections.
- ✓ La gestion de la variation (intra- et inter-locuteurs).
- ✓ La morphosyntaxe de la langue.
- ✓ Le recours à des jeux langagiers.
- ✓ La mise à jour du cahier de collecte et du journal de bord.
- ✓ L'enregistrement de séries limitées d'items bien choisis contenant des sons peu habituels ou problématiques en vue d'analyses instrumentales (par ex. PRAAT).

### *Tonologie*

Enseignant : Sophie Manus (DDL, Université de Lyon)

Horaires : 23 et 24 juin, 14h00 à 15h30 (anglais)  
25 et 26 juin, 14h00 à 15h30 (français)

Cet atelier est une présentation générale des questions de tonologie qui se posent au linguiste descriptiviste, depuis le premier terrain jusqu'à l'analyse finale.

Ne requérant pas de connaissance préalable en tonologie, cet enseignement s'attachera à présenter brièvement ce que l'on entend par "ton" (langues tonales du monde; langue tonale vs langue accentuelle; différents types de tons; tons lexicaux et tons grammaticaux; divers modes de transcription des tons chez les amérindianistes, les africanistes et les sinisants; éléments phonétiques et acoustiques).

On évoquera ensuite un certain nombre d'aspects spécifiquement liés à l'étude de terrain d'une langue à ton, tels les protocoles de recherche et le choix des informateurs ainsi que l'enregistrement, la transcription et l'analyse phonologique (selon différents modèles) de données de première main.

On s'appuiera sur de nombreux exemples transcrits et des échantillons sonores pris dans des familles de langues très diverses issues de tous les continents.

## ***Morphologie***

Enseignant : Denis Creissels (DDL, Université de Lyon)

Horaires : 30 juin et 1<sup>e</sup> juillet, 14h00 à 15h30 (français)  
2 et 3 juillet, 14h00 à 15h30 (anglais)

La pratique de l'élicitation de données pour une analyse morphosyntaxique se fera grâce à un travail fait avec un locuteur natif du Q'anjob'al, langue maya des montagnes du Guatemala. L'atelier se concentrera / focalisera sur la structure du syntagme nominal, en particulier sur les constructions possessives. Les cours alterneront entre démonstration d'élicitation de données et analyse de corpora distribués aux étudiants pour faire avancer la discussion. Des lectures choisies sur la morphosyntaxe des syntagmes nominaux seront indiquées en temps voulu.

## ***Lexiques spécialisés, références culturelles***

Enseignant : Lolke van der Veen (DDL, Université de Lyon)

Horaires : 30 juin et 1<sup>e</sup> juillet, 14h00 à 15h30 (anglais)  
2 et 3 juillet, 14h00 à 15h30 (français)

*Première partie* : introduction. Un bref aperçu de questions cruciales concernant notamment le choix des informateurs (âge, sexe, statut au sein du groupe, etc.), la préparation d'une collecte de vocabulaire spécialisé (formation, préparation de matériels à utiliser, approche), l'utilisation d'un ensemble de stimuli bien choisis, le choix du meilleur site et du moment le plus approprié pour les activités de collecte de données de ce type.

*Seconde partie* : travaux pratiques avec un informateur bantophone du Gabon.

✓ La collecte de vocabulaire spécialisé (maladie, troubles pathologiques).

La collecte de données culturelles en rapport direct avec le type vocabulaire spécialisé recherché (les principes de dénomination, le rôle du tabou, la perception et la catégorisation culturelles, les vocabulaires parallèles (initié vs non initié, opaque vs transparent, degrés de spécialisation, etc.).

## **3. Ateliers techniques**

### ***Logiciel de traitement des données Field Linguist's Toolbox***

Enseignant : Antoine Guillaume (DDL, Université de Lyon)

Horaires : 23 et 24 juin, 16h00 à 17h30 (français)  
25 et 26 juin, 16h00 à 17h30 (anglais)

Toolbox/Shoebox est un logiciel de gestion et d'analyse de bases de données textuelles et lexicales, à l'usage des linguistes de terrain, disponible gratuitement à <http://www.sil.org/computing/toolbox>. Il est particulièrement utile pour (1) le découpage morphologique et le glosage automatique de textes et (2) la gestion de fiches lexicales. Le but de cet atelier est une introduction et un entraînement **pour des débutants** à ces deux

fonctions principales du logiciel, à partir de données d'une langue Amazonienne de Bolivie, le reyesano, et de données apportées par les participants, si elles sont disponibles. Nous apprendrons comment créer un projet, entrer des mots, des phrases et des textes entiers, utiliser les fonctions de découpage morphologique et de glosage automatique, réaliser des fiches lexicales pour les morphèmes du corpus, et générer un dictionnaire. Nous passerons aussi en revue les diverses options de navigation et d'interrogation de la base de donnée, telles que l'affichage sous forme de liste, le tri, la recherche, le filtrage, etc.

### ***Audio recording, best practices***

Enseignant : David Nathan (SOAS, Université de Londres)

Horaires : 23 et 24 juin, 16h00 à 17h30 (anglais)  
25 et 26 juin, 16h00 à 17h30 (anglais avec aide pour francophones)

L'enregistrement audio est au cœur de la documentation des langues. Dans le contexte de la documentation de langues en danger, il est crucial d'effectuer les meilleurs enregistrements possibles, dans la mesure où l'on n'est pas certain de pouvoir enregistrer à nouveau. Effectuer un bon enregistrement dépend de multiples facteurs : s'être défini des buts précis, disposer de bons équipements et savoir s'en servir, avoir un minimum de connaissances dans le domaine de l'acoustique. Les deux facteurs qui influencent le plus la qualité d'un enregistrement sont le choix et l'usage que l'on fait des microphones et le contrôle de la prise de son afin de s'assurer que l'on obtient ce que l'on a anticipé.

Cet atelier discutera des questions d'environnement acoustique, des effets de psycho-acoustiques, de l'utilisation de microphones et d'enregistreurs audio, et des propriétés du signal audio.

### ***Praat***

Enseignant : Emmanuel Ferragne (DDL, Université de Lyon)

Horaires : 30 juin et 1<sup>e</sup> juillet, 16h00 à 17h30 (français)  
2 et 3 juillet, 16h00 à 17h30 (anglais)

Ces dix dernières années, le logiciel Praat est devenu le programme d'analyse des sons de la parole le plus largement utilisé par la communauté scientifique. Il offre une large palette de fonctions incluant l'annotation et la manipulation du signal de parole, et, bien sûr, la possibilité d'effectuer des mesures acoustiques. En plus de cela, ce logiciel est gratuit, mis à jour régulièrement, et peut être adapté aux besoins de l'utilisateur grâce à son langage de programmation. Quiconque travaillant sur des données orales peut donc sérieusement songer à l'utiliser.

Les cours sur Praat se décomposent en deux modules :

1. Fonctions élémentaires;
2. Fonctions avancées.

Nous nous efforcerons d'adapter les cours aux besoins des participants ; par conséquent, le plan de cours présenté ici n'expose que les grandes lignes. L'exploration des fonctions élémentaires comprendra, entre autres :

- L'ouverture, la fermeture d'un fichier ; la visualisation du signal et comment se déplacer dans un fichier, etc. ;
- L'édition et l'annotation du signal ;
- Les différentes représentations graphiques (spectrogramme, spectres, courbe intonative, etc.)
- Le calcul de mesures acoustiques fiables.

Les paramètres acoustiques que nous étudierons seront déterminés ultérieurement. Le cours sur les fonctions avancées sera l'occasion d'aborder le langage de programmation du logiciel (aucune connaissance en matière de programmation n'est requise). Il s'adresse aux personnes dont les travaux impliquent des tâches répétitives (ouverture et écoute de plusieurs fichiers les uns après les autres, mesure de la durée d'un son sur des centaines de fichiers, etc.). A la fin de ce cours, les participants seront en mesure de créer leur propre programme, ce qui leur permettra de confier à l'ordinateur les tâches les plus rébarbatives.

### ***Elicitation à partir de stimuli***

Enseignant : Friederike Lüpke (SOAS, Université de Londres)

Horaires : 30 juin et 1<sup>e</sup> juillet, 16h00 à 17h30 (anglais)  
2 et 3 juillet, 16h00 à 17h30 (français)

Ce cours a pour but d'introduire une méthode particulière d'élicitation de données linguistiques qui consiste à faire parler les locuteurs en utilisation des stimuli non-verbaux (images, clips vidéos, jeux de cartes). Le but de ce type d'élicitation est d'obtenir des données plus naturelles que celles obtenues par les méthodes d'élicitation traditionnelles tout en permettant au chercheur de contrôler les résultats obtenus.

### ***La production de vidéos sur le terrain pour l'analyse linguistique***

Enseignant : Lorenza Mondada (Laboratoire ICAR, CNRS & Université de Lyon)

Horaires : 30 juin, 17h30 à 18h30 (anglais)  
2 juillet, 17h30 à 18h30 (français)

L'enregistrement vidéo de données linguistiques est de plus en plus courant, grâce à la disponibilité d'équipements techniques de plus en plus miniaturisés et efficaces mais aussi grâce à des intérêts théoriques croissants pour les dimensions multimodales de la parole en action.

Après avoir donné quelques repères sur l'histoire des pratiques vidéo en sciences humaines et sociales ainsi qu'en linguistique, le cours présentera quelques (bons / mauvais) exemples de vidéos tournées sur le terrain, en fondant sur eux la discussion de problèmes techniques liés au tournage vidéo, ainsi que leurs conséquences théoriques et analytiques. De cette manière, il vise à montrer qu'il n'y a pas de questions purement techniques, mais qu'elles sont toujours porteuses d'enjeux épistémologiques. Ainsi, le cours réfléchira sur la spécificité des images, des cadrages, du positionnement des caméras et des micros en fonction des intérêts analytiques des chercheurs et plus largement en fonction de leurs modèles théoriques.



### 3. Cours aréaux

#### *Thèmes sur les langues africaines (cushitiques)*

Enseignant : Maarten Mous (Université de Leiden)

Horaires : Lundi à jeudi 23 au 26 juin, et Lundi à jeudi 30 juin au 3 juillet, 14h00 à 15h30 (8 séances, en anglais)

I discuss a number of topics that are important for many African languages and beyond. For each of the topics I present an overview of the situation in Cushitic including the analytical issues involved. This I contrast to the facts in other African languages. The aim is two-fold: to learn about the topics involved and to get an overview of Cushitic and to some extent of the linguistic diversity of Africa.

#### *Overview of topics:*

In the first lecture I present an introduction to African linguistic landscape: classification, typology, sociolinguistic profiles, research traditions and developments.

*Tone or pitch accent:* There is a wide variety of systems within Cushitic from clearly tonal to accent realised by pitch and this enables us to go deeper into the essential difference between tone and accent and contrast this to reduced tone systems in other African languages.

*Gender and number:* Gender and number are interrelated in Cushitic with plural as a value for gender. This brings us to category definition and nominal classification in Niger-Congo.

*Verbal conjugation:* Comparing verb conjugations crucially involves a double approach from to function and function to form. On the basis of Cushitic we discuss the issues in dealing with verb conjugation typologically and descriptively.

*Verbal derivation and argument structure:* Cushitic has passives of intransitive, impersonal constructions, indirect causatives, middle with a wide semantic range. This serves as the basis for discussing valence changing verbal derivations in Africa and brings us to a discussion of verbal diathesis and argument structure including languages using labile verbs and serialisation.

*Case and adpositions:* Cushitic languages have a marked nominative case system which has properties of both nominative-accusative and ergative-absolutive. The comparative discussion focuses on the coding of other case relations such as instrumental which is in Cushitic either by clitics or by multipurpose adpositions (sometimes stacked).

*Topic and focus:* Topic and focus drive Cushitic syntax and this has attracted considerable attention which I position into the vast literature on other African languages.

### ***Traits typologiques des langues mayas***

Enseignants : Prof. Colette Grinevald, et Marc Peake (DDL, Université de Lyon)

Horaires : Lundi à jeudi 23 au 26 juin, et Lundi à jeudi 30 juin au 3 juillet, 16h00 à 17h30 (8 séances, en anglais)

This course will cover the major typological features of the Mayan family of languages. This language family, spoken in Guatemala and parts of Southern Mexico, has been extensively studied since the 70ies, and is rather well known today for its many contributions to general and typologically oriented linguistics. After an introduction to the language family and a brief history of the development of Mayan linguistics, we will discuss the major phonological, morphological, and syntactic features of Mayan languages. The course will first quickly present phonological inventories; major word classes; noun morphology (possession, classifier systems); relational nouns vs. prepositions, verb-initial word order with rigid vs. flexible systems. It will then give an overview of the very rich verbal morphosyntax (TAM, and the multiple markings of transitivity), and treat in turn well known themes of this family, such as ergativity and various alignments, multiple voice systems (several types of passives and antipassives, applicative and agentive/inverse constructions). While still rare for most families of languages of the Americas, the study of the basic syntax of complex sentence formation is rather extensive today for Mayan languages; topics such as relativisation, subordination, coordination, and considerations of ergative syntax will be considered in turn. Time permitting, the last session(s) will consider nominal classification systems (particularly the rare noun classifier system of Q'anjob'alan languages, and the expression of space (through categories of positionals and directionals, in the Tzeltalan and Q'anjob'alan branches).