

# INTRODUCTION

Ce numéro 3 du Bulletin PFC a été confié à Julien Eychenne et Géraldine-M. Mallet. Ce numéro du bulletin présente une importance particulière : il constitue la mise au point écrite la plus précise depuis le Bulletin n°1 et le montage du site PFC à Paris X :

<http://infolang.u-paris10.fr/pfc>.

## 1. Bilan et perspectives

Le lecteur trouvera tout d'abord dans ce numéro une mise en perspective du projet établie par J. Eychenne et G.-M. Mallet : où en sont les enquêtes ? quels points prévus restent à faire ? quelle sera la structure du site internet ? Quel type d'accès aux données est prévu ? Sur ce dernier point, R. Walter, responsable du site PFC fournit des précisions indispensables pour bien comprendre les développements futurs.

## 2. Les formats

En second lieu, le Bulletin précise les conventions actuelles concernant le format des rendus (J. Durand, B. Laks, et C. Lyche révisé par J. Eychenne, P. Hambye et G.-M. Mallet). Il explique aussi comment les fichiers textes correspondant aux listes de mots doivent être constitués. Il est essentiel de se souvenir que, dans la première étape du projet, il n'était pas exigé d'avoir un alignement texte-son pour les listes de mots. Le développement d'outils d'analyse automatique rend ce travail aujourd'hui absolument nécessaire. En conséquence, le format des rendus a dû être amendé.

Ce nouveau format des rendus doit être examiné de très près par tous les participants, et appliqué, faute de quoi leurs enquêtes finales ne seront pas aux normes du projet et ne pourront pas être validées pour la base finale.

Il faut noter que certaines tâches recommandées dans le protocole d'enquête de 1999 ne sont plus exigées depuis au moins 2002. Par exemple, les Directions d'analyse (Durand, Laks, Lyche, Bulletin PFC n°1) présentaient des fiches phonologiques pour défricher les inventaires phonémiques des témoins. Ces fiches ne sont pas requises et ne l'étaient déjà pas dans le Format des rendus du Bulletin de 2002.

Tout ce qui est explicitement exigé dans la remise d'une enquête est spécifié dans le Format des rendus 2004 que l'on trouvera ci-après. On constatera qu'en dehors des enregistrements et des transcriptions orthographiques (liste de mots, texte, 5 mn guidé, 5 mn libre), le seul travail d'analyse à effectuer est le codage du schwa et celui de la liaison dans les conditions précises explicitées dans le Format.

## 3. Les codages

En ce qui concerne ces codages, la liaison n'a pas posé de problèmes insurmontables aux transcrip-teurs et ne nécessite donc aucune remarque particulière. En revanche, le codage du

schwa est un exercice plus difficile, qui exige une bonne connaissance des conventions. P. Hambye a donc rédigé un ensemble de remarques qui guident l'utilisateur chaque fois que les conventions de codage sont imprécises voire ambiguës. J. Durand et J. Eychenne proposent aussi quelques recommandations à la lumière des transcriptions dans les enquêtes qu'ils ont coordonnées.

#### **4. Enregistrements**

A propos des recommandations, il faut rappeler que la qualité des enregistrements reste une exigence importante. Certaines enquêtes sont difficiles à interpréter par des outils automatiques. R. Espesser a donc pris le soin de rappeler quelques règles essentielles pour obtenir des enregistrements de qualité, acceptables pour les enquêtes à venir.

#### **5. Outils**

Un ensemble d'outils a été mis au point au sein du projet PFC. Ces outils n'ont pas tous le même statut. Certains visent à aider les participants au projet dans la constitution des corpus sous Praat : Automatisation de la segmentation des listes de mots (C. Auran), Transpraat (A. Meqqori), SedGrid (J. Eychenne), Validateur (J. Eychenne), Précodeur (C. Auran et J. Eychenne). D'autres permettent de comparer ou d'analyser les données : comparateur (A. Meqqori), classeurs schwa et liaison (A. Meqqori), classeur-codeur (J. Eychenne), outil d'extraction semi-automatique des chartes formantiques (R. Espesser et N. Nguyen). En dehors du dernier outil en cours de développement au LPL par R. Espesser et N. Nguyen, tous ces outils sont disponibles sur le site central PFC et sont dans leur version finale. Hormis la correction de « bugs » possibles, ces outils ne seront pas développés plus avant. On notera que ces utilitaires ne fonctionnent pas en réseau et que des outils similaires mais plus puissants sont prévus dans la base internet en construction à Paris X. On trouvera dans ce Bulletin PFC les manuels ou spécifications correspondant à chacun des utilitaires opérationnels ou prévus.

#### **6. Prosodie**

Les participants au projet PFC savent que, depuis 2002, la prosodie est au coeur de nos préoccupations sous la houlette de Geneviève Caelen du LPL qui a bien voulu coordonner les débats. Ce bulletin s'en fait l'écho. Aux yeux des Directeurs du projet, la recherche sur la prosodie dans PFC peut se situer à trois niveaux dont les deux premiers sont prioritaires. Il est tout d'abord souhaitable de disposer d'un cadre comparatif général pour mieux analyser les enquêtes. A cet égard, dans les publications PFC à venir, nous souhaitons (pour un sous-ensemble de points d'enquête à définir) offrir des analyses qui portent sur :

- (a) les inventaires phonologiques et les principales réalisations phonétiques (avec l'aide d'extracteurs automatiques de chartes formantiques),
- (b) la liaison,
- (c) le schwa,
- (d) les principaux patrons prosodiques de la variété étudiée.

Une stratégie commune doit permettre d'obtenir des résultats comparables dans ce domaine. En deuxième lieu, il est souhaitable de disposer de codages prosodiques pour une partie des passages transcrits et codés pour le schwa et la liaison. A ce jour, peu d'études ont mené de front une analyse phonologique classique et une analyse prosodique. Sans ce travail, les assertions habituellement faites sur le rôle déterminant de la prosodie pour le fonctionnement de la phonologie d'une langue sont sans fondement. Puisque le schwa et la liaison sont tous deux des phénomènes qui dépassent les frontières de mots, un complément

de codage prosodique semble indispensable (dans des conditions à déterminer car tout codage est onéreux). Enfin, la base PFC peut et doit servir de terrain d'exploration pour des études théoriques de la prosodie, mais ce troisième niveau reste à l'appréciation des spécialistes avec qui nous collaborons.

Quoiqu'il en soit, pour une enquête donnée, on ne passera au codage prosodique que lorsque les codages du schwa et de la liaison seront terminés. Le lecteur trouvera donc dans ce volume un ensemble de réflexions et de propositions concernant la prosodie par G. Caelen, C. Auran, A.-C. Simon, A. Lacheret-Dujour, Ch. Lyche et M. Morel. Avant la fin de 2004, le projet devra disposer de conventions uniformes dans le domaine de la prosodie.

Enfin, pour clore ce numéro, H. Andreassen a rédigé un article de recherche sur son enquête PFC dans le canton de Vaud. Cet article démontre que le travail minutieux de récolte, de transcription et de codage n'exclut pas la dimension théorique. Tout au contraire, des données fiables permettent à la théorie de s'épanouir.

Les Directeurs du projet PFC sont fiers de présenter aux participants ce troisième numéro du Bulletin et remercient chaleureusement Julien Eychenne et Géraldine-M. Mallet d'avoir accepté la responsabilité d'en coordonner la fabrication et d'en assurer l'édition.

Jacques Durand  
Bernard Laks  
Chantal Lyche