



Corpus oraux par et pour l'apprenant

Henry Tyne

► To cite this version:

Henry Tyne. Corpus oraux par et pour l'apprenant. Mélanges CRAPEL, 2009, 31, pp.91-111. hal-00416544

HAL Id: hal-00416544

<https://hal.science/hal-00416544>

Submitted on 21 Sep 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CORPUS ORAUX PAR ET POUR L'APPRENANT

Henry Tyne

Equipe CRAPEL, ATILF–CNRS, Nancy-Université

Résumé

Dans cet article nous présentons les résultats d'une expérience d'enseignement basé sur l'utilisation de corpus oraux. Cet enseignement, quoique dispensé à des étudiants apprenants du français, n'était pas pour autant un cours de langue. Les résultats montrent que, si l'objectif premier reste l'étude de la sociolinguistique, les différentes activités qui contribuent à la mise en place du corpus puis à son analyse peuvent avoir pour effet de promouvoir une approche critique de la langue, de ses variations et de leur fonctionnement social. Ces aspects ont été jugés bénéfiques pour le développement d'une conscience linguistique générale.

Abstract

In this article we present the results from a teaching experience based on the use of spoken corpora. The course, although delivered to student learners of French, was not a language course. The results show that, whilst the primary focus of the course was the study of sociolinguistics, the different activities that contribute to the compilation and the analysis of the corpus can help to foster a critical approach to language and to the social functions of variation. These aspects were reported to have enhanced general language awareness.

Introduction

L'utilisation de corpus, sous une forme ou une autre, dans l'apprentissage des langues n'est pas récente. Avec l'arrivée des nouvelles technologies permettant de travailler sur de grands corpus, est né dans les années 1990 le paradigme « TaLC » (« teaching and language corpora »), qui réunit à la fois un savoir-faire sur corpus et une réflexion sur la place des corpus dans l'apprentissage des langues. En général, on identifie trois types d'approches en TaLC (voir Stewart *et al.*, 2004) : les apprenants comme *sources de données* (corpus et travail sur les données en interlangue) ; les apprenants comme *bénéficiaires de données* (corpus comme support pédagogique) ; les apprenants comme *utilisateurs de données* (travail sur corpus par l'apprenant). Le troisième type d'approche, celui qui voit l'apprenant se servir de corpus, a donné des résultats intéressants dans la réflexion autour de l'autonomisation et de la motivation dans l'apprentissage (par exemple, Seidlhofer, 2002). Mais les corpus utilisés sont très majoritairement des corpus écrits. En effet, comme le remarque Mauranen (2004a), si la communication orale occupe traditionnellement une place importante dans les syllabus, le travail sur des corpus oraux en TaLC est considérablement à la traîne. Les travaux existants sur les corpus oraux dans l'apprentissage, tout comme ceux sur l'écrit, sont souvent orientés vers l'*exploitation* des données attestées (par exemple, Pérez-Paredes & Cantos-Gómez, 2004 ; Jones, 1997) et non sur la *constitution* des corpus par les apprenants. De plus, le champ d'exploitation demeure assez restreint, se limitant le plus souvent à l'étude de la langue (ce qui est vrai des approches sur corpus en général – voir cependant Kettemann et Marko, 2004). Or, comme le souligne Aston (2002), les corpus « faits maison », tout en favorisant l'autonomisation de l'apprenant, offrent de nombreuses possibilités d'exploitation et surtout présentent l'avantage de permettre à l'apprenant de tester des hypothèses spécifiques ou bien d'observer des phénomènes particuliers que l'on ne trouve pas toujours dans les corpus généraux. Ils permettraient également de cibler certains types d'activités ou genres de texte et de prendre conscience de la représentativité des données analysées. Mais les études sur la « fabrication » de corpus par l'apprenant se concentrent typiquement sur les résultats d'analyses de ces corpus, c'est-à-dire des concordances ou bien d'autres observations des données (Aston, 2002). Dans tous les cas, la fabrication du corpus demeure une activité tournée vers la mise en place d'un produit final, le corpus.

Nous nous proposons ici de présenter un enseignement à partir de corpus que nous avons initié à l'Université de Southampton au cours de l'année universitaire 2004-2005 et qui a été dispensé aux étudiants en troisième année de Licence trois années durant. A la différence des études existantes, qui mettent l'accent sur le produit et non le processus, l'approche que nous avons bâtie réunit à la fois le travail en amont de l'analyse (réflexion sur la récolte de données, sur les sujets parlants trop souvent oubliés dans l'exploitation des données) et le travail sur les données

mêmes par l'apprenant (transcription, analyse). Le corpus apparaît ici non pas comme une simple source de données pour l'enseignant ou l'apprenant, pouvant être interrogée par un ensemble d'outils, mais comme une ressource, fabriquée par l'apprenant qui va servir dans l'apprentissage de la langue tout en nourrissant une réflexion sociolinguistique sous-jacente.

1. Sociolinguistique et travail sur corpus oraux par des apprenants

Pour les acquisitionnistes travaillant sur la compétence sociolinguistique, l'appropriation de celle-ci constituerait le dernier « obstacle » dans l'acquisition, tant du fait de l'énormité et de la complexité du phénomène que du fait de la place tardive qui lui est typiquement attribuée dans les méthodes d'apprentissage (voir Mougeon *et al.*, 2002). Et comme l'ont montré les différents chercheurs travaillant dans ce domaine, si les apprenants les plus avancés usent en général des mêmes variables que les locuteurs natifs (par exemple, *ne* de négation variable – voir Mougeon *et al.*, 2002), la variabilité reste typiquement moindre. L'un des défis pour la didactique des langues étrangères ou secondes reste donc la sensibilisation à la variation, chose que soulignent Tarone et Yule (1989 : 11), pour qui l'apprenant doit profiter du contact avec les données pour pouvoir observer le fonctionnement de la langue en contexte. Justement, l'utilisation de corpus peut faire avancer la pédagogie de la variation, comme le souligne Conrad (2004). Et si l'on souhaite que les apprenants abordent cette variation dans des corpus, il paraît alors intéressant, dans un cours de sociolinguistique, de vouloir aller au-delà de la simple description de données fournies par l'enseignant (cf. Zorzi, 2001 ; Conrad, 2004 ; Mauranen, 2004b), de s'intéresser aux étapes qui précèdent l'analyse du corpus, c'est-à-dire à la constitution du corpus. Nous proposons donc que la constitution de corpus, au sein d'une réflexion sociolinguistique, parte de l'enquête sur le terrain : l'apprenant devient ainsi enquêteur et transcripteur, responsable de la fabrication du corpus en amont des analyses (cf. la question de la motivation et du sentiment d'« ownership » des données – Seidlhofer, 2002) et non seulement chercheur¹ travaillant sur des données existantes (par exemple, Gavioli, 2001).

La transcription comme activité en didactique des langues, sous des formes différentes, n'est pas récente, qu'il s'agisse de dictées, d'exercices de « texte à trous », de travail sur la compréhension d'un document (par exemple, Boulon, 1998) ou d'activités basées sur la confrontation et la réécriture (par exemple, Lynch, 2001, 2007). Dès les années 1970, Biggs et Dalwood (1976) proposaient des exercices basés sur la transcription d'enregistrements effectués dans la ville d'Orléans comme

1. Nous prêtons à ce terme un sens plus large que celui, très connu, formulé par Johns (par exemple, 1997) pour qui l'apprenant-chercheur est un « détective » qui apprend à reconnaître et interpréter les indices fournis par le concordancier.

première prise de contact avec la langue parlée². Or, dans le cadre d'un cours de sociolinguistique, nous avons imaginé une approche davantage portée sur l'accès aux données authentiques récoltées sur le terrain, en privilégiant la réflexion oral/écrit et une prise en compte du rôle de l'écrit dans une langue de littératie ou dans la description traditionnelle de la langue pour apprenants. Une place centrale a ainsi été accordée à une tâche typiquement peu théorisée dans la constitution de corpus³. L'apprenant se voit alors réfléchir à la meilleure façon de représenter les données orales, en se donnant des conventions de transcription, aux questions éthiques de la représentation de l'autre, etc.

2. Recherche-action

Le cours de sociolinguistique que nous avons mis en place visait à explorer la notion de variation à partir d'un approfondissement du *pourquoi* de la variation et d'une meilleure compréhension du fonctionnement social de la langue en étayant les observations par l'étude en parallèle de données récoltées par les étudiants eux-mêmes. Ainsi, on peut schématiser l'articulation des différents aspects de ce cours de la manière suivante :

- théorie, lectures, découverte (comprendre qu'il existe de la variation) ;
- description, compte-rendu de travaux de sociolinguistique (observer ce qu'ont trouvé d'autres chercheurs) ;
- recherche personnelle en parallèle (observer la variation dans deux corpus différents ; retour sur les résultats de travaux existants ; expliquer les convergences/divergences) donnant suite à un dossier comportant corpus (bande son et transcription) et analyses.

Cette expérience, qui a débuté en 2004-2005, s'est renouvelée à deux reprises (les années universitaires 2005-2006 et 2006-2007), permettant ainsi d'affiner le fonctionnement du cours, notamment grâce aux questionnaires d'évaluation communément utilisés dans les établissements en Angleterre. Les résultats présentés dans les sections suivantes comprennent les réponses aux

-
2. Dans l'introduction de Biggs et Dalwood (1976 : 5) on peut lire : « It is first of all the transcription exercise which directs your attention to the full detail of the words used, and often to the manner in which they are spoken... If you hear the speaker's voice sounding in your head long after the study session, you are putting the recording already to one of its best uses. »
 3. Outre le fait que les analyses de données attestées par les sciences sociales se sont traditionnellement portées sur le contenu plus que sur la forme (autre que phonique), la transcription *per se* a souffert d'un simple manque d'intérêt pour ce qui est essentiellement vu comme une simple « tâche » (quoique longue et fastidieuse) par laquelle le chercheur (ou quelque autre subalterne) est obligé de passer afin d'avoir des résultats. Comme le rappellent Beaud et Weber (2003 : 247), « la transcription est le plus souvent considérée comme une phase pratique du travail, ingrate et ennuyeuse, d'ailleurs fréquemment sous-traitée à des 'petites mains' auxquelles on donne simplement quelques consignes de transcription. L'analyse du matériel, la partie noble du travail, ne commencerait qu'une fois l'entretien transformé en 'texte' ».

questionnaires généraux d'évaluation des trois promotions successives ainsi que les réponses à un questionnaire plus détaillé rempli par les étudiants de la dernière promotion à suivre ce cours⁴.

Les étudiants devaient réunir deux petits corpus oraux différents de français parlé par des locuteurs natifs⁵ (5 à 10 minutes de parole continue pour chaque corpus) afin de procéder à une analyse comparative. La différenciation des corpus s'effectuait en fonction d'hypothèses formulées à partir d'aspects théoriques vus en cours et à partir des résultats d'autres études. Ainsi, par exemple, un étudiant pouvait choisir d'enregistrer un même locuteur dans deux situations différentes (considérations diaphasiques : par exemple, un étudiant lors d'un exposé oral et lors d'une conversation ordinaire avec ses colocataires), ou bien de comparer différents locuteurs (considérations diastratiques ou même diatopiques). Avant de procéder à l'analyse des corpus, les étudiants devaient ainsi mettre en place une enquête, récolter des données, les transcrire. Les travaux dirigés ont été consacrés à la mise en place du travail personnel des étudiants avec des exercices notamment ciblant le passage de considérations théoriques à une application pratique. L'accès aux données occupait une place centrale dans le travail : aborder la langue, certes, mais aussi comprendre les facteurs extralinguistiques pouvant entrer en jeu (paradoxe de l'observateur, questions de spontanéité, questions d'éthique, de description du terrain, etc.).

Les données, une fois récoltées, ont été transcrites à l'aide du logiciel *Praat*⁶. Pour certains étudiants il s'agissait d'une réutilisation du logiciel qu'ils avaient découvert auparavant en cours de phonétique. Néanmoins, deux sessions d'une heure dans une salle informatique ont permis de s'assurer qu'ils sachent tous utiliser le logiciel. Et si chaque étudiant travaillait à l'élaboration de son corpus personnel, l'ensemble des données récoltées allaient pouvoir être interrogées par la classe au terme de l'enquête et le dernier cours a été consacré à des observations générales sur le corpus collectif.

-
4. Sachant que l'année 2006-2007 allait être notre dernière année à l'Université de Southampton, nous avons souhaité sonder les étudiants de manière plus approfondie sur certains points précis en rapport avec ce cours en particulier. Il est regrettable que le groupe cette année-là n'ait pas été aussi important que les années précédentes. Cette baisse d'effectifs s'expliquerait en particulier par le petit nombre d'étudiants dans la promotion combiné à un changement de prérequis dans le parcours linguistique et langues. Nous tenons d'ailleurs à remercier tous les étudiants d'avoir pris le temps de répondre aux questionnaires, tout comme nous remercions les collègues à l'Université de Southampton qui nous ont toujours soutenu lors de la mise en place de ce cours.
 5. Sans entrer dans le débat sur la définition de la parole native, nous dirons simplement que ce travail visait à explorer la parole authentique (c'est-à-dire non produite pour des besoins pédagogiques) de locuteurs francophones natifs. En général, les étudiants ont exploité leurs réseaux de connaissances parmi les nombreux étudiants et enseignants francophones présents sur le campus.
 6. *Praat* est un logiciel de transcription et d'analyse développé par Boersma et Weenink à l'Université d'Amsterdam au Pays-Bas (<http://www.fon.hum.uva.nl/praat/>, page consultée le 04/06/08).

En tout, sur les trois ans, 65 étudiants ont suivi ce cours, dont 40 ont rendu le questionnaire d'évaluation (tableau 1).

année et nombre de réponses obtenues	structure et contenu	enseignement	progrès de l'étudiant	soutien	ressources pour apprendre	appréciation générale
2004-05 (13/29)	3,7	3,3	3,2	3,9	3,8	3,5
2005-06 (23/26)	4,2	4,5	3,8	4,6	4,1	4,2
2006-07 (4/10)	4,3	4,5	3,8	4,5	4,3	4,0

Tableau 1. Résultats (scores moyens sur 5 : 5 = excellent ; 1 = très mauvais) des questionnaires d'évaluation 2004-2007.

3. Résultats

3.1. Suivi général : questionnaires d'évaluation

Le tableau 1 présente les scores moyens obtenus pour les questionnaires généraux d'évaluation distribués aux étudiants afin d'obtenir des informations permettant d'améliorer le cours d'année en année. Une bonne partie de ces informations ne porte pas spécifiquement sur le travail sur corpus, étant surtout destinées au service de surveillance de la qualité de l'enseignement. Sans trop nous attarder sur ces données⁷ nous remarquerons seulement qu'il y a une nette augmentation du score moyen après la première année, ce qui reflète sans doute les réglages effectués suite à une première mise en place du cours. Par ailleurs, nous avons été frappé par le nombre de remarques positives accompagnant ces questionnaires qui mettaient en relation les techniques développées dans ce cours et celles plus conventionnelles utilisées dans d'autres cours. C'est surtout suite à ces remarques que nous avons souhaité mettre en place un questionnaire plus détaillé, portant plus précisément sur certains aspects du travail sur corpus et sur la transcription.

3.2. Questionnaire détaillé

Les questions posées sous format électronique à remplir en ligne (tableau 2) ont été organisées autour de trois axes : les avantages perçus de la transcription et du travail sur corpus pour l'apprentissage du français langue étrangère (FLE) ; le

7. Nous signalons aussi que les questions posées aux étudiants ainsi que les modalités de réponses (à la maison, en cours, en ligne) n'ont pas été constantes d'une année à l'autre. Le fait que le questionnaire ne soit pas obligatoire rend également difficile toute généralisation à partir des réponses obtenues.

rapport à la tâche particulière que représente la transcription ; les autres possibilités d'utilisation de transcription ou de travail sur corpus.

Effets bénéfiques pour l'apprentissage du FLE
1. Pensez-vous que votre français s'est amélioré grâce à la transcription ?
2. Pensez-vous que votre français s'est amélioré grâce à l'étude de corpus ?
3. Avez-vous eu l'impression de faire des découvertes concernant votre connaissance du français grâce à la transcription ou au travail sur corpus ?
4. Avez-vous eu le sentiment que la transcription ou le travail sur corpus vous permettait de confirmer des choses que vous saviez (ou que vous pensiez savoir) déjà en/sur le français ?
Comment la tâche de transcription est vécue par l'apprenant
5. Avez-vous trouvé le travail de transcription difficile ?
6. Avez-vous aimé faire de la transcription ?
Peut-on faire plus ? A quel niveau ?
7. Pensez-vous qu'il serait utile/approprié d'introduire le travail sur corpus dans d'autres aspects de l'apprentissage de la langue ?
8. Pensez-vous que la transcription serait un exercice adapté à des apprenants moins avancés ?
9. Pensez-vous que le travail sur corpus serait un exercice adapté à des apprenants moins avancés ?

Tableau 2. Questions posées aux étudiants de la promotion 2006-2007.

Les résultats sont présentés dans la figure 1. Ces résultats montrent que l'expérience a globalement été vécue de manière positive par les 10 étudiants qui ont suivi le cours en 2006-2007, aussi bien la transcription que l'étude de corpus.

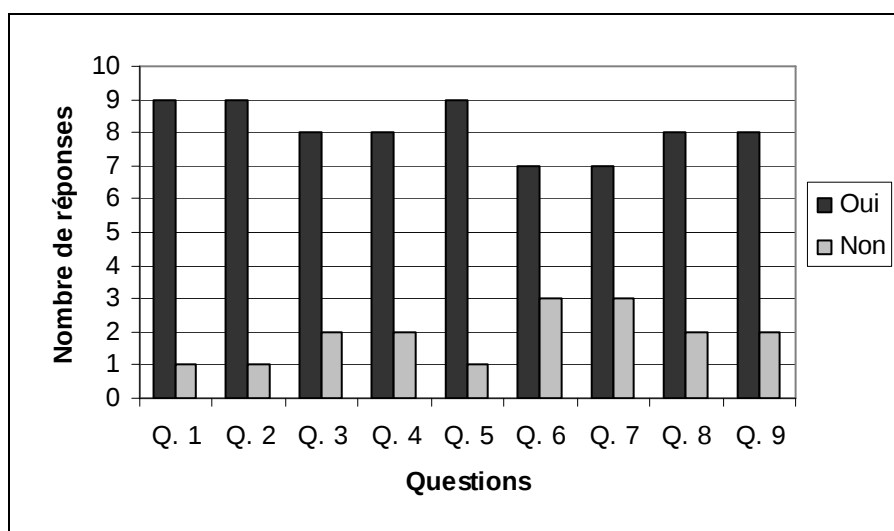


Figure 1. Réponses aux questions portant sur le l'enquête et l'analyse.

Outre les réponses en *oui/non*, consigne a été donnée de fournir des explications ou des justifications détaillées des choix de réponse. Nous

présenterons ces résultats dans les sections à suivre.

3.3. Réponses

3.3.1. Effets bénéfiques pour l'apprentissage du FLE

- *Question 1. Pensez-vous que votre français s'est amélioré grâce à la transcription ?*

En tout, neuf étudiants sur dix ont trouvé que leur français s'était amélioré grâce à la transcription (figure 1). Il y a sans doute des valeurs subjectives propres à chaque étudiant concernant l'impression d'amélioration en français. Or, ce qui nous intéressait à ce stade c'était simplement ce que les étudiants ressentaient à la fin du cours : ont-ils ou pas l'impression d'avoir progressé ? Nous n'avons mis en place aucun dispositif pour tenter de jauger cette amélioration.

Parmi les différentes raisons données, on voit qu'il y a, au-delà du simple fait de mettre l'apprenant en position de fournisseur de ses propres données (une chose gratifiante en soi comme en ont témoigné plusieurs étudiants), un travail de prise de conscience de divers facteurs relevant du fonctionnement de la langue : il y a un contact particulier avec les données, permettant d'être aux premières loges quant à l'observation des phénomènes (exemples 1 et 2), mais il y a également un début de travail d'analyse des données afin de comprendre ce qui est dit pour pouvoir le transcrire (exemples 2 et 3 – nous reviendrons plus loin sur la question de la compréhension orale, section 3.3.4) pouvant donner lieu à l'utilisation des phénomènes observés (voir le cas de l'étudiant E6 qui dit avoir employé des particules énonciatives, exemple 2) :

1. E4⁸ **You get to hear French, how it is really spoken instead of how it is taught in the classroom**, which are two very different things.
2. E6 **Transcribing French has made me far more aware of things such as discourse markers**, which might not be of semantic importance, however, being a non-native speaker of French they are features that can help make my French sound more authentic. **Transcribing has enabled me to tune into other features** such as variable liaison, the significance of which was not apparent to me beforehand. **Transcription was a thoroughly worthwhile task**, particularly as I had to transcribe data from two very different contexts... As I said above, **I have found myself using discourse markers more**, which (hopefully!) will make my French sound a bit more authentic.
3. E10 **I enjoyed finding interesting variations in the language**, such as schwa tagging and liaisons in the formal setting and unusual pronunciations of certain words.

8. Les différents étudiants apparaissent ici de manière anonyme (E1, E2, E2, etc.). Les extraits des questionnaires dans le texte sont en anglais (langue originelle du questionnaire) avec des passages en caractères gras afin de faire apparaître les informations importantes. Des éléments de contextualisation sont fournis entre crochets carrés.

Il y a également la question du type d'accès à la langue que peut procurer ce genre d'exercice. Dans le cas de l'exemple 4, on voit comment l'étudiant prend conscience à travers la transcription de son statut de locuteur non natif et des conséquences pour le comportement des interlocuteurs :

4. E2 **...I was listening to native speakers speak in a way I don't normally have access to, as my presence as a non-native speaker normally means they often modify their language** but as I wasn't there but just left the recorder they spoke as they would between themselves.

- *Question 2. Pensez-vous que votre français s'est amélioré grâce à l'étude de corpus ?*

Comme pour le travail de transcription, neuf étudiants sur dix (voir la figure 1 ci-dessus) ont trouvé le travail sur corpus bénéfique pour leur apprentissage du français. En ce qui concerne l'analyse des données, il s'agissait avant tout de vérifier des hypothèses fonctionnelles suivant les éléments abordés en cours et les différents travaux cités. Par exemple, en tenant compte des observations d'Armstrong (2001) sur quelques phénomènes de variation en français parlé, il s'agissait pour chaque apprenant de vérifier dans son corpus si l'on trouvait les mêmes types de phénomènes, les mêmes distributions. Les étudiants pouvaient ainsi formuler des hypothèses en fonction des études existantes, se fixant sur un nombre limité d'éléments variables (voir la liste des éléments abordés en annexe). Ces aspects de vérification et d'explication des résultats ont été appréciés, d'autant plus qu'ils ont permis aux étudiants de mieux réfléchir sur la présentation traditionnelle de la langue et sur leurs propres productions orales (voir en particulier le cas de l'étudiant E2 dans l'exemple 5 qui arrive à mieux cerner l'hétérogène – cf. l'exemple 10 plus loin) :

5. E2 **Our investigation focused on areas of the language that I have noticed in the past but I have never known their specific significances before, such as variable 'ne de négation' where I now see that I was over-generalising my use, and variable liaison which I used randomly before.**
6. E5 **[You] discover how different words/registers etc. are used in real language**, something which you cannot always get from textbook study.
7. E6 **The transcription made me aware of certain features and the analysis enabled me to understand their significance more fully.**

- *Question 3. Avez-vous eu l'impression de faire des découvertes concernant votre connaissance du français grâce à la transcription ou au travail sur corpus ?*

Grâce à la transcription ou à l'analyse de corpus, huit étudiants sur dix (figure 1) ont déclaré avoir fait des découvertes. Lorsqu'on regarde les commentaires qui accompagnent les réponses, on s'aperçoit que ces découvertes sont surtout positives, étant souvent en rapport avec le mode de production oral des données (voir la remarque sur les « erreurs » dans l'exemple 8 ; voir également l'exemple 9) :

8. E2 **[I discovered] how sometimes even native speakers make errors**, [and I discovered] use of certain aspects of the language for very specific reasons.
9. E4 **People do not naturally speak slowly, in formal French with perfect pronunciation and no accent – I had a real shock when I first realised that the French I spoke was not really French at all.**
10. E6 I was completely unaware of the way variable liaison is used and as such it has made me more conscious of how and when I use it. **The difference between the two contexts regarding speech makes complete sense, however it wasn't something I had really ever thought about before.**
- 11.E10 **The informal speech had various utterances that were said without any verbal constructions**, but still get the meaning across to the other speaker.

Cependant, les découvertes n'ont pas toujours été vécues de façon positive et certains étudiants ont dit avoir été surpris par les découvertes qu'ils ont faites :

12. E1 I realised that **most of what I thought was being said was completely wrong.**
13. E5 **Didn't realise that I would find it so hard to distinguish the different sounds** in the transcription.

- *Question 4. Avez-vous eu le sentiment que la transcription ou le travail sur corpus vous permettait de confirmer des choses que vous saviez (ou que vous pensiez savoir) déjà en/sur le français ?*

Si huit étudiants sur dix ont déclaré avoir fait des découvertes grâce à la transcription ou à l'analyse de corpus, le même nombre a aussi déclaré avoir eu confirmation de choses qu'ils savaient (ou dont ils se doutaient) déjà. Ce qui nous paraissait important ici c'était de voir à quel point des observations ou des hypothèses, parfois latentes, pouvaient trouver confirmation dans ce genre d'exercice. Et dans les réponses on remarque qu'il s'agit surtout de prises de conscience de certaines choses remarquées auparavant mais dont le fonctionnement précis dans le langage était ignoré ou tout simplement méconnu (voir en particulier les exemples 16 à 18 ; voir aussi l'exemple 5 et l'exemple 57 plus loin) :

14. E3 **[I saw confirmation of] the variation which can occur between two different language settings.**
15. E4 **Very slight differences in sound which we think are unimportant, or can't hear, can make you look very silly when you get them wrong...** (I cannot differentiate between elbow, neck and bum in French, much to my boyfriend's amusement!).
16. E5 **Found that I knew a fair amount about differences in formal and informal language** (however, I think it did also increase my knowledge of this).
17. E9 Yes it taught me much more informal phrases and **reminded me of things that I'd learnt on my year abroad but had forgotten when I got back.**
- 18.E10 **The corpus work confirmed my knowledge of 'ne de négation' in French and how it is naturally used in a more formal setting.** Also the variation of lexis depending on the situation (formal or informal)... Confirmed in the informal setting the use of 'tu' and more uses of generalisations such as 'quoi' and 'et tout ça'.

3.3.2. Comment la tâche de transcription est-elle vécue par l'apprenant ?

- Question 5. Avez-vous trouvé le travail de transcription difficile ?
- Question 6. Avez-vous aimé faire de la transcription ?

Si les résultats montrent que la transcription et le travail sur corpus ont été reconnus comme étant des pratiques bénéfiques pour l'apprentissage du FLE (voir les réponses aux questions 1 et 2, figure 1), la transcription même (le moment où les apprenants « se salissent les mains ») a également été appréciée par la majorité des étudiants (voir la question 6, figure 1). Ou du moins il y a une reconnaissance de l'acte de transcrire, ou de la gratification du fait de produire des données (exemple 21), mais sans réel engouement (la majorité des étudiants ont tout de même trouvé la tâche ardue – voir la question 5, figure 1), comme le résumant ces remarques :

- 19. E5 **Quite pleasing that you improve** as you go along.
- 20. E9 It wasn't exactly enjoyable but **it was definitely very satisfying and it was a real learning experience.**
- 21. E3 It is an extremely tedious task, although **there is a huge amount of satisfaction when you have finished.**
- 22. E2 Yes [it was enjoyable] because **I enjoy a challenge**, and also **it was good to see the transcript develop as I made the changes...** It was frustrating and sometimes I could hardly bear to listen to the same stuff again.
- 23.E10 **Quite time-consuming and frustrating at times**, as well as trying to understand exactly what the speakers are talking about, **but also fun** at the same time and comparing transcriptions with other members of the group.

3.3.3. Peut-on faire plus ?

- Question 7. Pensez-vous que qu'il serait utile/approprié d'introduire le travail sur corpus dans d'autres aspects de l'apprentissage de la langue ?

La question 7 visait à déterminer si les étudiants avaient conscience des possibilités d'exploitation de corpus au-delà de ce cours dans différents domaines de l'apprentissage. Et si oui, lesquels ? Et si non, pourquoi pas ?

En tout, sept étudiants sur dix (voir la figure 1) ont estimé qu'il conviendrait de se servir d'approches sur corpus dans d'autres domaines de l'apprentissage. Il est intéressant de noter que c'est le travail sur la grammaire qui revient le plus souvent (exemples 24 à 26), et ce pour des raisons essentiellement liées à une appréciation du rôle social de la langue et une prise de conscience des phénomènes non standard (exemple 27) :

- 24. E2 **To see how aspects of grammar are used in flowing conversation.**
- 25. E5 Yes, **you can see how textbook rules are applied in spoken language**, could be used testing how well you know rules, for example, spotting patterns of language, spotting where language differs from the standard, etc.
- 26.E10 Useful to incorporate corpus investigations into grammar because **it can give realistic examples of language and highlight variables.**
- 27. E6 **...it is easy just to accept the prescriptive approach and not even entertain the idea that there is a descriptive approach...** It is important,

particularly at degree level, that what is taught in the classroom reflects what happens in everyday life in the relevant country(ies).

Cependant, parmi les réponses négatives, on trouve une certaine réticence liée à la difficulté associée au travail sur corpus, surtout quand il s'agit d'aborder la grammaire :

- 28. E1 No, because I think **it would just cause a lot of confusion and bring up a lot of questions with no real answers.**
- 29. E3 I think **it would be confusing as the use of grammar can often be adapted to the situation and this would make learning the basics extremely difficult.**
- 30. E7 I think **there are more effective methods of language learning that are less time-consuming.**

On détecte dans ces réponses une certaine nostalgie des règles, des données « propres », ou un sentiment d'insécurité face à la variation (cf. les problèmes pour les enseignants dans Conrad, 2004). On trouve également un phénomène d'amalgame où oral devient synonyme de langage familier, comme en témoigne l'exemple 31, où l'étudiant, tout en louant par ailleurs le travail sur corpus, et en particulier la transcription (cf. exemple 36), met en garde contre les « mauvaises habitudes »⁹ que peuvent alors prendre les étudiants une fois exposés aux données orales :

- 31. E4 **It creates bad habits** and now I find myself speaking so much 'français familier' **I completely forget the correct French.** I shorten everything (like 't'as' – 'tu as') and I don't know when things are meant to have apostrophes any more and when they shouldn't.

Les étudiants qui ont suivi ce cours, quoique tous en L3, n'avaient pas le même niveau en français. En effet, pour ceux combinant l'étude de plusieurs langues, il était possible que le français soit leur deuxième langue (par exemple, niveau 5 sur une échelle de 1 à 7)¹⁰. C'était d'ailleurs un souci constant pour l'enseignant (même si les résultats d'année en année n'ont pas donné confirmation de cette inquiétude) : les étudiants les moins avancés allaient-ils être à même de profiter de ce cours aux côtés d'autres étudiants dont le niveau en langue était plus élevé ? Des pratiques particulières telles que la double écoute, la vérification par les pairs et la réflexion sur les stratégies de transcription ont été mises en place afin de contourner cet obstacle, de faire en sorte que le niveau en langue ne soit pas vécu comme un handicap au moment de mettre sur pied une enquête ou de procéder à la

9. Hales (1997 : 221-222) cite le cas d'un enseignant en formation qui, confronté à des transcriptions d'échanges entre professeurs et élèves, s'indigne devant les productions orales (hésitations, phénomènes non standard, etc.) des personnes que l'on laisse enseigner : « and these people are out there teaching English ? »

10. Dans les cours de langue dispensés à l'Université de Southampton, le niveau 1 (plus ou moins équivalent du niveau A1 du CECR – *Cadre européen commun de référence* ; Conseil de l'Europe, 2001) correspond aux apprenants débutants et le niveau 7 (plus ou moins équivalent du niveau C2 du CECR) correspond aux apprenants très avancés. En dehors des différences « évidentes » (séparation des étudiants par niveau), il y a également des différences au sein d'un seul et même

transcription. Cependant, nous voulions savoir si, au-delà de ces différences inhérentes à chaque groupe au sein d'une promotion donnée en L3, les étudiants pensaient qu'il serait possible d'introduire la transcription ou l'analyse de corpus auprès d'apprenants de niveaux moins avancés.

- *Question 8. Pensez-vous que la transcription serait un exercice adapté à des apprenants moins avancés ?*

Si les étudiants ont globalement trouvé la transcription dure, ils y ont trouvé de l'intérêt pour l'apprentissage, en particulier pour ce qui est du rapport particulier avec les données qu'elle permet. Huit étudiants sur dix (voir la figure 1 ci-dessus) ont estimé qu'il conviendrait d'introduire la transcription de données orales auprès d'apprenants moins avancés, et ce pour plusieurs raisons différentes : originalité de la tâche (exemples 32 et 33), contact avec les données authentiques (exemple 34), amélioration de la compréhension orale (exemple 35), ou bien le simple plaisir de transcrire (exemple 36) :

- 32. E2 Yes because **it's quite a novel thing to do** and I think that **modern language teaching could benefit from using a more varied range of material to keep students interested.**
- 33. E5 Might be useful as **a less tedious form of learning grammar rules and for picking out new vocab.**
- 34. E3 It would benefit them as **they would see how the language is actually used** and their listening ability would improve greatly.
- 35. E6 I think **transcription is a very useful tool for honing listening skills and the earlier this can be done, the better.**
- 36. E4 **...the earlier the better for people who enjoy it. I can't really think of a reason why not to.**

Plusieurs étudiants ont cependant insisté sur l'importance du choix des données à transcrire :

- 37. E8 It would be useful provided it was introduced at **quite a slow pace, starting perhaps with formal speech**, then introducing informal gradually and at quite slow speed.
- 38. E9 Yes but the piece used for transcription **would have to be chosen very carefully.**

Il semble y avoir un consensus sur le niveau (assez avancé – voir les remarques plus haut) auquel il conviendrait d'introduire la transcription, c'est-à-dire quelque part entre le niveau baccalauréat et le niveau Licence 1 à l'université (à peu près B1-B2 selon le CECR – Conseil de l'Europe, 2001) :

- 39. E2 I think I would have appreciated it **from A-level on.**
- 40. E6 Perhaps the **first year of degree study** would be a good time to start.
- 41. E8 I feel that **too much informal speech for translation [i.e. transcription?] before Stage 4 [i.e. plus ou moins le niveau 'A-level'] might be counter-productive**, but after that can only serve to help with listening and speaking and exams.

42. E9 It would perhaps be **best for uni level** as the finished result is more interesting than a normal listening task.

Nous pouvons noter que les deux réponses négatives (exemples 43 et 44) s'accompagnent de commentaires sur la difficulté de la transcription (surtout telle qu'elle a été vécue par ces deux étudiants). Nous reviendrons sur ce point lors de la discussion ci-dessous.

43. E1 I think **it would just dishearten them**, listening comprehension is hard enough as it is!
44. E7 I think **it would be too challenging**.

- *Question 9. Pensez-vous que le travail sur corpus serait un exercice adapté à des apprenants moins avancés ?*

Huit étudiants sur dix (voir la figure 1) ont estimé qu'il conviendrait d'introduire l'analyse de données orales auprès d'apprenants moins avancés. On remarque surtout dans ces réponses l'importance accordée à l'aspect découverte : ce sont des étudiants de niveau avancé qui ont l'impression d'avoir découvert certaines choses sur le tard qu'ils auraient souhaité aborder, savoir ou comprendre plus tôt (voir également la prise de conscience de l'utilisation aléatoire d'un phénomène de variation dans l'exemple 5 ci-dessus).

45. E4 Yes because **when I went to France for the first time I couldn't understand a word of what was being said**.
46. E5 At the lower levels of study words are sometimes learnt in isolation, as lists, perhaps **it would be beneficial to use a corpus to gain a better understanding of how such words are used in different contexts**.
47. E7 Yes, where the transcription is provided **looking for specific features such as variable liaison**.
48. E8 In particular **for oral and aural classes it would be very beneficial. It would give a 'feel' for the language**, especially in informal speech. For me, analysing 'ne de négation' was the easiest to evaluate and **would help a learner at quite a basic level from both the practical aspect of understanding the differences between formal and informal French, and also self-confidence in usage**.

Comme pour la question 8 traitant de la transcription, l'importance du choix de corpus (cf. l'exemple 38) a été soulignée :

- 49.E10 ...more interesting and fun to study spoken French, that which is used day to day in France. **Important to find simple corpora**.

3.3.4. Autres aspects commentés par les apprenants

Nous avons relevé d'autres commentaires sur le travail de transcription ou d'analyse. Ceux-ci concernent surtout la compréhension orale et l'obligation d'écoute minutieuse que requiert la transcription. Ceci aurait également pour effet de faire travailler l'orthographe (exemple 52) :

50. E1 **General listening ability and understanding improved**.

51. E3 **My listening ability has greatly improved** because of the need to deduce the exact words/sentences in French from the data.
52. E4 **You hear each individual sound** that someone says as opposed to all the sounds blurring into one, which makes it seem like they are speaking slower – which obviously means you understand a lot more and don't miss as much. It means that **when you hear a word you've never heard, you know how to spell it. Also, little differences in sounds** (as in 'dessous' and 'dessus') which are hard for English people to pick up on, **become more noticeable**.
53. E9 It focuses your attention onto what language people use and how they use it. Also, it means that **you're doing a substantial amount of listening on a regular basis which can only be a good thing...** Just listening to the foreign language a lot can only be beneficial.
- 54.E10 Initially listening both to the formal and informal datas, they were quite hard to transcribe exactly and every time **listening to it over and over again began to improve my understanding of the speech and the exact words being said**.

Nous avons également trouvé des remarques sur l'utilité d'une telle expérience pour l'apprentissage d'autres langues (le cas de l'italien dans l'exemple 55) ou même pour mieux appréhender sa propre langue (exemples 56 et 57) ou les langues en général (exemples 56 et 57) :

55. E8 I'm not sure what benefits I have gained from doing the transcriptions in French, but I feel that **I would find something like this very useful in Italian, as I am learning the language in the conventional, theoretical fashion**. This, I am sure, would help immensely with understanding spoken and television language, which I do struggle with in Italian.
56. E9 **Showing differences in people, sex and generation and exposing how situation affects language use regardless of which language it is...** Also **it's interesting to notice language slips people make and to compare them to those you make in your own language**.
57. E6 In my opinion **transcription allows you to really engage with the language, be it your native language or otherwise, in a way that simply hearing it and speaking it does not...** **It makes you more aware of features you didn't necessarily realise were present** (or that, as a non-native speaker of French you used subconsciously without knowing why).

4. Discussion

Si le fait de placer l'apprenant en position de chercheur, fournisseur de ses propres données, semble tout à fait indiqué dans une conception autonomisante de l'apprentissage, notre expérience a montré que cette position n'a pas toujours été vécue comme un aspect positif par tout le monde. Certes, la majorité de nos étudiants ont parlé d'activités utiles, bénéfiques, enrichissantes, etc., mais il semble néanmoins y avoir des attitudes opposées que l'on ne peut ignorer. Nous pensons surtout aux étudiants E1 et E7. Et quand on regarde leurs réponses, très souvent négatives, on s'aperçoit que c'est avant tout l'aspect enquête-transcription qui a posé problème :

58. E1 I found it **extremely frustrating and confusing!** ...I think **it is too much to think about when you aren't fluent in a language** (and possibly when you are!).

59. E7 I think **there are more effective methods of language learning** that are less time-consuming... **Took so long** and got to the point where the changes weren't particularly evident and so less satisfying. Plus **listening to the same thing over and over again is irritating.**

Il paraît important de nous attarder sur au moins trois points (au-delà du problème de l'inquiétude que peut générer l'évaluation finale du dossier et que nous n'aborderons pas ici). Premièrement, tout apprenant n'est pas nécessairement bon chercheur et n'apprécie pas ce type de contact avec le terrain. Deuxièmement, il se peut que l'apprentissage par découverte soit plus adapté à certains « types » d'apprenants, question que se posent également Boulton et Tyne (2008). L'étude menée par Lewis (2006) montre que les apprenants qui préfèrent un apprentissage « conventionnel » trouvent le travail sur corpus trop déroutant. Comme le signale par ailleurs Harmer (2001), la culture particulière d'apprentissage, formée par les habitudes des enseignants et des apprenants, peut être plus ou moins favorable à l'application d'une approche inductive. Finalement, dans le cadre limité d'un cours de sociolinguistique, le recours à des pratiques telles que l'enquête pilote n'est pas toujours facile à mettre en place (contraintes liées au temps consacré au dossier – voir la note 11) et le risque que les étudiants se retrouvent avec des données difficiles à transcrire (débit, bruits ambiants, chevauchements, registre inconnu, etc.) est réel. Se pose aussi la question de la transcription pure et simple d'éléments en langue cible que les étudiants ne connaissent pas ou n'ont pas encore vus. Plutôt que de recourir à une présélection de données « appropriées » (ce qui aurait été contraire aux principes du cours, bâti sur l'enquête menée par l'étudiant), nous avons préféré insister sur les stratégies de transcription (comment s'y prendre), sur la réalité subjective de l'écoute, les hésitations légitimes du transcripateur, l'importance de la double écoute. Il semble néanmoins que ces écueils n'aient pas toujours suffi à pallier les problèmes de transcription et on constate que mêmes les étudiants les plus enthousiastes éprouvent parfois des difficultés pour aborder la langue. Difficultés toutefois surmontables :

60. E2 **I found myself getting frustrated** because I would sometimes listen over and over to the same parts and still not catch what was said. This was particularly the case when two or more people spoke at once. Also, **knowing that ultimately it is impossible to accurately reproduce what was said was hard for me** because I'm a perfectionist, and I found some of the inevitable compromises you have to make unsatisfactory. I'd say **it did get easier as I went along because I got more used to the way the speakers spoke.**

Ces problèmes semblent en partie liés à l'individu et son contact avec la langue cible. Ainsi, un seul étudiant (E8), qui a passé beaucoup de temps au contact de la communauté cible, dit ne pas avoir trouvé la transcription difficile, même celle de données provenant de situations mettant en jeu plusieurs locuteurs dans un cadre informel.

Un souci récurrent que l'on relève dans la littérature sur l'utilisation de corpus est celui de la représentativité des données. La taille du corpus ainsi que les

différents types de productions qui le constituent sont des facteurs à prendre en considération. Dans le cas des corpus oraux ce problème paraît d'autant plus évident que le rapport entre le temps de travail consacré à la constitution du corpus et le nombre de mots obtenus n'est pas du tout comparable à celui pour la constitution d'un corpus écrit. Il paraît en fait très difficile de demander à des apprenants (même à des apprenants très avancés) de s'attaquer à des transcriptions allant au-delà de quelques minutes seulement¹¹. Dans notre cas, nous avons demandé aux étudiants de réunir entre 1000 et 2000 mots (production d'un seul locuteur dans chaque corpus) dans deux situations distinctes, c'est-à-dire entre cinq et dix minutes de temps d'enregistrement pour chaque situation afin de procéder à une analyse comparative. Et si certaines des hypothèses n'étaient pas toujours vérifiables (par exemple, difficulté pour travailler sur *ne* de négation variable dans un corpus de 1000 mots s'il y a peu de constructions négatives), dans l'ensemble, les résultats ont été au rendez-vous (par exemple, débit, liaison variable¹², emploi de lexique non standard ou spécialisé, phénomènes d'hésitation et de co-construction, etc.). Mais comme le résume cette remarque de l'étudiant E6 (exemple 61), c'est surtout le fait qu'il *puisse* y avoir des phénomènes de variation qui est important, et que ceux-ci *puissent* s'expliquer à partir de facteurs extralinguistiques. Ainsi, ce n'est pas tant l'observation des analyses effectuées sur le corpus qui constitue le moteur dans l'apprentissage (des analyses ponctuelles peuvent même ne rien signaler de particulier, ce qui ne veut pas dire pour autant que le corpus est « mauvais ») mais c'est plutôt le fait de s'attendre à ce qu'il y ait un résultat particulier puis d'expliquer pourquoi on obtient ce qu'on obtient (voir aussi la conclusion de Bondi, 2001). Une telle approche permet à l'apprenant de développer un regard critique sur la langue et de chercher à comprendre pourquoi les choses sont comme elles sont :

61. E6 **It can help you become a more 'critical' language learner, by not just accepting that that is what happens in spoken language, but by asking yourself why this happens.**

Conclusion

Dans un premier temps, nous avons remarqué que les observations générales des étudiants qui ont suivi ce cours étaient particulièrement positives et motivantes.

-
11. En ce qui concerne la taille du corpus, il y a également des contraintes d'ordre institutionnel. En effet, le temps de travail estimé de cette tâche devait également être validé par les instances compétentes de l'université, et être obligatoirement comparable à celui estimé pour les dossiers soumis à évaluation dans d'autres cours au même niveau d'étude.
12. On peut noter que l'analyse de phénomènes récurrents tels que la liaison ou le schwa dans le projet PFC sur la *Phonologie du français contemporain* (<http://infolang.u-paris10.fr/pfc/> – page consultée le 4 juin 2008) se base sur des extraits de corpus de quelques minutes seulement. Par ailleurs, d'autres chercheurs (par exemple, Ghadessy *et al.*, 2001) ont insisté sur l'utilité de petits corpus pour l'apprentissage des langues étrangères.

Dans un deuxième temps, nous avons souhaité avoir davantage d'informations concernant certains aspects précis du cours au travers d'un questionnaire sur les attitudes des apprenants. Comme l'avoue Maia (1997), à propos de son expérience d'enseignement sur des corpus fabriqués par des apprenants, il est difficile de tout planifier, de tout prévoir : si les buts escomptés sont clairs, les résultats concrets en termes d'apprentissage sont difficilement vérifiables, dépendant en large partie de la manière dont l'étudiant entre dans la tâche, et n'étant vraiment visibles qu'à travers des comptes-rendus d'apprentissage et des sondages d'attitudes et d'opinions.

Il reste encore beaucoup de chemin à parcourir dans l'évaluation des possibilités de l'apprentissage sur corpus de cette nature, où les apprenants sont amenés à fabriquer eux-mêmes leurs corpus oraux, et en ce qui concerne les éventuels problèmes de démotivation ou de confusion face aux données orales. Néanmoins, les résultats présentés ici semblent indiquer qu'une telle démarche peut permettre à des apprenants relativement avancés de prendre conscience de certains aspects du fonctionnement de la langue cible en contexte et de développer des pratiques utiles pour appréhender les langues en général.

ANNEXE

Liste des phénomènes de variation pouvant être analysés

Phonétique et phonologie

- accent, débit, rythme
- e muet (insertion/effacement ; « schwa-tagging »)
- simplification des groupes et chute de // final
- liaison (obligatoire, variable, erratique, sans enchaînement)

Lexique

- type de lexique (registres de langue)
- créativité lexicale (truncation, emprunts, composition, verlan, etc.)

Morphosyntaxe

- conjugaisons, temps
- groupe nominal (pronominal, lexical, relatives)
- ordre des mots, détachements
- négation
- formes interrogatives
- subordonnées
- extensions de *que*
- parataxe vs. hypotaxe

Discours

- appui sur des particules énonciatives (*tu vois, enfin*, etc.)
- utilisation des axes syntagmatique et paradigmatic et phénomènes de construction (répétition d'un schéma, retour en arrière, extensions à droite, etc.)

BIBLIOGRAPHIE

- ARMSTRONG, Nigel, 2001. *Social and Stylistic Variation in Spoken French: A Comparative Approach*. Amsterdam : John Benjamins.
- ASTON, Guy, 2001. « Learning with corpora: an overview. » In G. Aston (éd.), *Learning with Corpora*. Bologna : Clueb / Houston : Athelstan, p. 1-45.
- ASTON, Guy, 2002. « The learner as corpus designer. » In B. Kettemann & G. Marko (éds.), *Teaching and Learning by Doing Corpus Analysis*. Amsterdam : Rodopi, p. 9-25.
- BEAUD, Stéphane & Florence WEBER, 2003. *Guide de l'enquête de terrain*, nouvelle édition. Paris : La Découverte.
- BIGGS, Patricia & Mary DALWOOD, 1976. *Les Orléanais ont la parole*. Londres : Longman.
- BONDI, Marina, 2001. « Small corpora and language variation: reflexivity across genres. » In M. Ghadessy, A. Henry & R. Roseberry (éds.), *Small Corpus Studies and ELT*. Amsterdam : John Benjamins, p. 135-174.
- BOULON, Joline, 1998. « La création de pages Internet dans le cadre de parcours individualisé : la place de la méthodologie dans l'apprentissage des langues. » *ALSIC*, 1/2, p. 147-154.
- BOULTON, Alex & Henry TYNE, 2008. « Learning with corpora: changing learning practices. » Communication au 5^e colloque international *IVACS (Inter-Varietal Applied Corpus Studies)*. Université de Limerick, Irlande, 13-14 juin.
- CONRAD, Susan, 2004. « Corpus linguistics, language variation and language teaching. » In J. Sinclair (éd.), *How to Use Corpora in Language Teaching*. Amsterdam : John Benjamins, p. 67-85.
- CONSEIL DE L'EUROPE, 2001. *Cadre européen commun de référence pour l'apprentissage et l'enseignement des langues : apprendre, enseigner, évaluer*. Paris : Didier.
- GAVIOLI, Laura, 2001. « The learning as researcher: introducing corpus concordancing in the classroom. » In G. Aston (éd.), *Learning with Corpora*. Bologna : Clueb / Houston : Athelstan, p. 108-137.
- GHADESSY, Mohsen, Alex HENRY & Robert ROSEBERRY (éds.), 2001. *Small Corpus Studies and ELT*. Amsterdam : John Benjamins.
- HALES, Tom, 1997. « Exploring data-driven language awareness. » *ELT Journal*, 51/3, p. 217-223.
- HARMER, Jeremy, 2001. *The Practice of English Language Teaching*. Harlow : Longman.

JOHNS, Tim, 1997. « Contexts: the background, development and trialling of a concordance-based CALL program. » In A. Wichmann, S. Fligelstone, T. McEnery & G. Knowles (éds.), *Teaching and Language Corpora*. Londres : Longman, p. 100-115.

JONES, Randall, 1997. « Creating and using a corpus of spoken German. » In A. Wichmann, S. Fligelstone, T. McEnery & G. Knowles (éds.), *Teaching and Language Corpora*. Londres : Longman, p. 146-156.

KETTEMANN, Bernhard & Georg MARKO, 2004. « Can the L in TaLC stand for literature? » In G. Aston, S. Bernardini & D. Stewart (éds.), *Corpora and Language Learners*. Amsterdam : John Benjamins, p. 169-194.

LEWIS, Jonathan, 2006. *Connecting Corpora to Learner Style*. Mémoire de Master, Université de Porto, Portugal. <http://nautilus.fis.uc.pt/cec/teses/>, page consultée le 18/06/08.

LYNCH, Tony, 2001. « Seeing what they meant: transcribing as a route to noticing. » *ELT Journal*, 55/2, p. 124-132.

LYNCH, Tony, 2007. « Learning from the transcripts of an oral communication task. » *ELT Journal*, 61/4, p. 311-320.

MAIA, Belinda, 1997. « Making corpora: a learning process. » In G. Aston, S. Bernardini & D. Stewart (éds.), *Proceedings of Corpus Use and Learning to Translate*. <http://www.sslmit.unibo.it/cultpaps/paps.htm>, page consultée le 04/06/08 via <http://web.archive.org/>.

MAURANEN, Anna, 2004a. « Speech corpora in the classroom. » In G. Aston, S. Bernardini & D. Stewart (éds.), *Corpora and Language Learners*. Amsterdam : John Benjamins, p. 195-211.

MAURANEN, Anna, 2004b. « Spoken corpus for an ordinary learner. » In J. Sinclair (éd.), *How to Use Corpora in Language Teaching*. Amsterdam : John Benjamins, p. 89-105.

MOUGEON, Raymond, Terry NADASDI & Katherine REHNER, 2002. « Etat de la recherche sur l'appropriation de la variation par les apprenants avancés du FL2 ou FLE. » *AILE*, 17, p. 7-50.

PEREZ-PAREDES, Pascual & Pascual CANTOS-GOMEZ, 2004. « Some lessons students learn: self-discovery and corpora. » In G. Aston, S. Bernardini & D. Stewart (éds.), *Corpora and Language Learners*. Amsterdam : John Benjamins, p. 247-258.

SEIDLHOFER, Barbara, 2002. « Pedagogy and local learner corpora: working with learner-driven data. » In S. Granger, J. Hung & S. Petch-Tyson (éds.), *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Amsterdam : John Benjamins, p. 213-234.

STEWART, Dominic, Sylvia BERNADINI & Guy ASTON, 2004. « Introduction: ten years of TaLC. » In G. Aston, S. Bernardini & D. Stewart (éds.), *Corpora and Language Learners*. Amsterdam : John Benjamins, p. 1-18.

TARONE, Elaine & George YULE, 1989. *Focus on the Language Learner*. Oxford : Oxford University Press.

ZORZI, Daniela, 2001. « The pedagogic use of spoken corpora: learning discourse markers in Italian. » In G. Aston, (éd.), *Learning with Corpora*. Bologna : Clueb / Houston : Athelstan, p. 85-107.