

**LES CORPUS ORAUX EN FRANÇAIS LANGUE  
ÉTRANGÈRE : AUTHENTICITÉ ET PÉDAGOGIE**

**Angela Chambers**

University of Limerick (Irlande)

**Résumé**

Le but de cette étude est d'explorer le potentiel des corpus oraux pour une exploitation en français langue étrangère (FLE). S'appuyant d'une part sur les publications dans ce domaine et d'autre part sur des publications pertinentes consacrées à l'apprentissage d'autres langues, dans une démarche à la fois théorique et pratique, l'article tente de répondre à deux questions : 1) Les corpus sont-ils des documents authentiques ? 2) Dans quelle mesure les corpus oraux répondent-ils aux besoins des professeurs et des apprenants de FLE ?

**Abstract**

This study aims to explore the potential of spoken French corpora in the teaching and learning of French as a foreign language, with particular reference to publications in this area which relate not only to French but also to other languages. Taking both a theoretical and a practical approach, the article will address two issues: 1) Can corpora be considered as authentic documents? 2) To what extent do spoken corpora meet the needs of teachers and learners of French as a foreign language?

## Introduction

Les corpus oraux occupent désormais une place importante dans la linguistique et la didactique des langues. En ce qui concerne la langue française, les *Archives de la parole* de la BNF, qui datent de 1911, marquent la première étape d'une longue tradition de recherche sur le français parlé, basée sur des corpus. Au cours du vingtième siècle, les corpus oraux se multiplient, comme on peut le constater en consultant la base de données de Cappeau et Seijido (2005) ou l'état des lieux dressé par Cappeau et Gadet (2007a). Les études se multiplient aussi, principalement phonologiques, syntaxiques et sociolinguistiques, mais aussi dialectologiques et psycholinguistiques.<sup>1</sup> Les corpus représentent « des ressources importantes pour l'apprentissage des langues : phénomènes de collocation et phraséologie, microsyntaxe des entrées lexicales, étude des langues de spécialité, typologie des textes » (Habert *et al.*, 1997 : 15). La pertinence des études linguistiques des corpus oraux pour l'enseignement et l'apprentissage du français langue étrangère (FLE) est donc évidente, puisqu'elles pourraient sensibiliser les enseignants et les auteurs de manuels aux différences entre le français standard, qui se base essentiellement sur l'écrit, et l'oral spontané. La chute du *ne*, par exemple, constatée par beaucoup de chercheurs (Armstrong & Smith, 2002 ; Martineau & Mougeon, 2003), est désormais répandue en France, au Canada, et ailleurs, et date du milieu du 19<sup>e</sup> siècle. Le *ne* est pourtant « massivement maintenu dans les textes qui reflètent le français oral » (Mougeon *et al.*, 2002, cités par Dewaele & Wourm, 2002 : 133). L'enseignement de l'oral spontané représente pour le professeur et l'auteur du manuel un défi à cet égard : « To *ne* or not to *ne*, that is the sociolinguistic question » (Rehner & Mougeon, 1999). La collaboration entre linguistes et auteurs de manuels, ainsi que l'emploi des textes authentiques dans l'enseignement des langues, pourraient permettre aux enseignants de répondre à ce défi (Dewaele & Wourm, 2002 : 133). Mais on pourrait aller beaucoup plus loin et envisager que la collaboration entre linguistes de corpus, auteurs de manuels, enseignants, et formateurs d'enseignants constitue une possibilité d'enrichir l'environnement de l'apprentissage du FLE.

Pour tenter de justifier cette position nous explorerons le potentiel des corpus oraux en français pour l'exploitation au-delà de l'emploi de textes authentiques individuels. Les publications dans ce domaine spécifique étant assez rares, nous prendrons en considération aussi les études pertinentes consacrées à l'apprentissage d'autres langues. Menant à la fois une démarche théorique et pratique, nous tenterons de répondre à deux questions :

1. Les corpus sont-ils des documents authentiques ?
2. Dans quelle mesure les corpus oraux répondent-ils aux besoins des

---

1. Voir, par exemple, l'étude du schwa (Durand & Eychenne, 2004) ou l'analyse quantitative de « parce que » (Debaisieux, 2002).

professeurs et des apprenants de FLE ?

La définition – ou plutôt les définitions – de l'authenticité ayant déjà été l'objet de nombreuses publications, voire d'une querelle bien connue dans le contexte de l'emploi des corpus et des concordances en anglais, nous résumerons brièvement les principaux arguments de Widdowson et d'autres chercheurs, avant de proposer une approche qui accepte l'authenticité de la consultation de corpus sous certaines conditions. Nous verrons ensuite comment trois projets actuels essaient d'intégrer la consultation des données basées sur des corpus oraux dans l'enseignement et l'apprentissage du FLE. Nous nous limiterons à trois domaines susceptibles d'être pertinents pour les enseignants et les apprenants dans l'enseignement secondaire et supérieur, c'est-à-dire la vie quotidienne en France, la langue orale parlée par les jeunes, et le discours de l'enseignant et des élèves dans la salle de classe de FLE, sujet particulièrement intéressant pour les enseignants dont le français n'est pas la langue maternelle. Ceci nous permettra, en guise de conclusion, de mieux connaître les facteurs contribuant à une consultation optimale des corpus oraux par les enseignants et les apprenants de FLE.

La première question qui se pose est évidemment de savoir quels types de corpus correspondent le mieux aux diverses situations des apprenants de FLE. Ces besoins sont tellement diversifiés qu'il est difficile, voire impossible, d'envisager la création d'un seul corpus, si grand soit-il, qui corresponde à tous les besoins que l'on puisse imaginer. Pour certains chercheurs en linguistique (voir, par exemple, Bilger, 2007 : 16), un grand corpus de référence en français formerait un point de repère essentiel, consistant en un ensemble de textes ou d'extraits de textes qui « doivent être hétérogènes et renvoyer aux plus grand nombre d'usages possibles ». Il est certain qu'un tel corpus permettrait à ceux qui étudient un aspect de l'usage dans un domaine spécifique de savoir si cet usage est différent des données dans un corpus de référence. Cependant, Cappeau et Gadet (2007b : 108-109) remettent en question l'idée qu'un grand corpus puisse répondre à tous les besoins des linguistes. Comme nous le verrons, il en est de même en ce qui concerne les enseignants et les apprenants de FLE.

## **1. Les corpus sont-ils des documents authentiques ?**

Les définitions de l'authenticité en didactique des langues mettent l'accent d'une part sur le texte lui-même et d'autre part sur l'objectif de l'apprenant qui lit ou qui écoute le texte. La définition de base, c'est-à-dire un texte qui a été produit pour répondre à un besoin réel de communication, a été remise en question. Il faut souligner cependant qu'elle reste acceptée dans le contexte de l'apprentissage des langues, comme on peut le constater en consultant des descriptions de cours de langue sur le web à partir d'une recherche du terme « texte authentique ». Les critiques de cette définition de base concernent surtout le contraste qu'elle établit

entre la salle de classe et la réalité. « A text is usually regarded as authentic if it is not written for teaching purposes but for a real-life communicative purpose » (Lee, 1995 : 324). Plusieurs chercheurs soulignent cependant que la réalité ne réside pas dans le texte lui-même mais dans le contexte dans lequel il a été produit.

What makes the text real is that it has been produced as appropriate to a particular set of contextual conditions. But because these conditions cannot be replicated, the reality disappears. (Widdowson, 1998 : 712)

Comme nous le verrons en ce qui concerne l'authenticité des textes en général et plus particulièrement l'authenticité des corpus, l'interprétation du mot « contexte » joue un rôle fondamental. Si l'on restreint le contexte aux conditions de la production originale d'un texte, on condamne bien sûr à l'inauthenticité la quasi-totalité de la lecture des textes littéraires, même par les locuteurs natifs. Breen (1985 : 61) accepte l'authenticité de l'interprétation du texte par l'apprenant et l'authenticité de l'activité de l'apprenant, c'est-à-dire qu'il accepte la réalité du contexte de la salle de classe comme l'espace où se joue la vie réelle des enseignants et des apprenants. Cette approche est d'ailleurs acceptée par beaucoup d'autres chercheurs, dont Widdowson (2000 : 7-8).<sup>2</sup>

Le corpus pose problème dans ce débat, tout d'abord parce que l'activité de créer un corpus enlève obligatoirement le texte de son contexte d'origine. Présenter à l'apprenant des données tirées d'un vaste corpus de référence serait donc lui donner accès à ce que Widdowson (2000 : 7) appelle « decontextualized language ».

Le corpus, dès qu'il dépasse une certaine taille, est rarement connu en profondeur. Ce qui est utilisé se limite souvent à des ensembles de lignes qui sont le produit de requêtes formulées. Le plaidoyer en faveur des données réelles et contextualisées débouche de fait sur une vision fragmentée, partielle, de suites de lignes discontinues. (Capeau & Gadet, 2007b : 102)

Lorsque l'on consulte les publications des enseignants/chercheurs qui ont intégré la consultation de corpus dans leur enseignement, il est évident que la plupart préfèrent créer des corpus limités de textes familiers aux apprenants : des manuels qu'ils emploient dans leur apprentissage d'une autre matière (Stevens, 1991), des lectures recommandées (O'Sullivan & Chambers, 2006), ou bien des genres familiers, comme des articles de journaux (Kennedy & Miceli, 2001 : 78) ; c'est-à-dire qu'ils choisissent des textes où les apprenants n'ignorent pas le contexte original. La plupart des études publiées concernent les corpus de textes écrits en anglais, mais il existe aussi quelques études basées sur des textes oraux, en anglais et en français. Parmi les études en anglais, deux sont particulièrement pertinentes par rapport à la question du contexte. Braun (2005), qui a enregistré des anglophones de plusieurs nationalités parlant de leur vie et de leur travail, insiste sur la nécessité de créer des corpus pour répondre aux besoins des professeurs et des apprenants des langues étrangères, de les enrichir d'exercices, et de combiner l'étude de textes complets avec des concordances qui ajoutent d'autres exemples

---

2. Pour une étude plus détaillée des définitions de l'authenticité, voir Mishan (2005).

d'aspects de la langue que les apprenants remarquent dans les textes complets. Sinclair et Coulthard, quant à eux, ont analysé dès 1975 un corpus de discours de la salle de classe en anglais. Les applications d'un tel projet de recherche sont évidentes pour des professeurs qui ne sont pas des locuteurs natifs de la langue. Comme nous le verrons, il existe actuellement des projets en FLE axés sur les genres familiers aux apprenants et aux professeurs et les discours qu'ils veulent maîtriser.

Selon les définitions de l'authenticité qui mettent l'accent sur l'authenticité de l'activité de l'apprenant et l'appréhension du contexte, l'étude d'un texte individuel, écrit ou oral, extrait d'un corpus, représente bien une activité authentique. Par exemple, l'étudiant qui suit une formation pour devenir professeur de français et qui veut que la langue principale de l'enseignement dans ses cours soit le français, mais qui n'est pas locuteur natif, sait très bien quel est le contexte d'un cours de FLE enregistré en France. Certains éléments du contexte francophone lui manquent sans doute, mais les ressemblances sont nombreuses, de telle sorte que ces enregistrements forment une excellente ressource, permettant aux futurs professeurs de les écouter, les analyser, et réfléchir sur leur propre pratique dans le cadre d'un stage ou d'une activité professionnelle à l'avenir. Mais si un enregistrement d'un cours, ou d'un extrait d'un cours, forme un texte authentique, qu'en est-il de la consultation d'une concordance, de petites bribes de phrase discontinues, où le contexte n'est pas clair ? Pour Braun (2005 : 59), ces concordances servent à enrichir l'étude d'un texte complet en fournissant d'autres exemples. Pour certains chercheurs qui créent des corpus limités des textes étudiés dans une autre matière, les concordances, loin de manquer de contexte, représentent même l'accès aux contextes multiples, ce que Cobb (1997 : 303) appelle « multicontextual learning ». Dans ces exemples le problème de manque de connaissance en profondeur des grands corpus, signalé par Cappeau et Gadet (2007b : 102), est largement surmonté. Il s'agit ici de l'importance de la fréquence dans l'apprentissage des langues, un thème longtemps négligé par les chercheurs à une époque où le behaviorisme n'était plus considéré comme une méthode valable, en particulier à cause du manque d'authenticité. Pour Ellis (2002 : 177) la fréquence n'est pas inextricablement liée au behaviorisme, mais a un rôle fondamental à jouer dans l'apprentissage d'une langue : « mindful repetition in an engaging communicative context by motivated learners ». Tenant compte de ces recherches, une approche pédagogique de l'intégration des corpus et des concordances dans l'apprentissage du FLE impliquerait l'étude de textes individuels complets, ou d'extraits de textes dont le contexte est clair, enrichie par des concordances qui fournissent des exemples supplémentaires de certains aspects des textes. Il s'agit ici, non pas de la consultation de vastes corpus de référence, mais de corpus plus limités d'un seul genre ou d'un nombre restreint de genres, (peut-être extraits de corpus de référence) qui permettent aux apprenants de trouver de multiples exemples de certains aspects d'un genre qui leur est déjà familier ou qu'ils veulent maîtriser. Pour l'apprenant qui

n'habite pas un pays francophone, c'est un moyen de recréer certains aspects de la méthode idéale, l'immersion, en trouvant en un instant plusieurs occurrences d'une expression, puisque sa vie courante ne lui donne pas l'accès à de multiples exemples (Cobb, 1997 : 313). Nous verrons quels types de ressources sont disponibles dans trois corpus oraux en français, et quelles données pertinentes le professeur ou l'apprenant peuvent découvrir dans ces ressources.

## 2. Corpus oraux en français et FLE

### 2.1. Introduction

Trois exemples de projets actuels de création et d'exploitation des corpus oraux illustrent la variété de développements dans ce domaine. ESLO, l'*Enquête sociolinguistique à Orléans*, est un projet particulièrement intéressant pour l'enseignement du FLE, dans un contexte historique ainsi que pour l'avenir. SACODEYL<sup>3</sup> est un projet européen qui a pour objectif de constituer un corpus de la langue parlée par les jeunes dans plusieurs pays européens, créant ainsi une ressource susceptible d'intéresser les apprenants dans l'enseignement secondaire et supérieur. Troisièmement, un corpus du discours de la salle de classe, construit à l'Université de Limerick en Irlande, montre le besoin de construire des corpus oraux limités pour répondre à des besoins spécifiques, dans ce cas la formation des professeurs de français.

### 2.2. ESLO : les Enquêtes sociolinguistiques à Orléans

Bien que les documents sonores collectés dans le passé présentent un intérêt certain pour les linguistes qui veulent se pencher sur l'usage de la langue à une certaine époque, ou sur les changements qui sont survenus depuis cette période, ces données deviennent assez vite dépassées dans le contexte de l'apprentissage des langues. La longévité des ressources linguistiques telles que les dictionnaires, les grammaires et les manuels est limitée. ESLO est un projet particulièrement intéressant dans ce contexte. Le projet se déroule en trois étapes. Entre 1968 et 1971, un groupe d'universitaires anglais a construit un corpus de documents sonores à Orléans :

Au total plus de 300 heures de parole incluant pour moitié des interviews en face à face [environ 200] et pour moitié une gamme d'enregistrements variés (conversations téléphoniques, réunions publiques, transactions commerciales, repas de famille,

---

3. SACODEYL (*System-Aided Compilation and Open Distribution of European Youth Language*) est un projet subventionné par la Commission européenne (Minerva) de 2005 à 2008 pour la promotion de l'enseignement à distance. <http://www.um.es/sacodeyl/> page consultée le 25/02/08.

entretiens médico-pédagogiques, etc.).<sup>4</sup>

Ils avaient pour objectif de créer des ressources pour servir de matériaux pour l'enseignement du FLE s'appuyant sur des textes authentiques. L'objectif de la deuxième étape est la mise en ligne de ces ressources au sein du laboratoire CORAL<sup>5</sup> pour les préserver et les rendre accessibles au public, principalement aux chercheurs en linguistique. La troisième étape consiste en la constitution d'un corpus comparable (ESLO 2) :

L'objectif a été fixé à 400 heures environ de documents sonores répondant à une approche variationniste et conçus à travers des programmes spécifiques. Réunis, ESLO 1 et ESLO 2 formeront une collection de 700 heures d'enregistrement, ce qui est considéré aujourd'hui comme une valeur repère pour les investigations projetées.<sup>6</sup>

Bien que l'objectif principal de ce projet soit d'ordre linguistique, « la réflexion sur l'évolution des modèles et des méthodes en linguistique variationniste » (Baude & Eshkol 2007 : n.p.), elle a sans aucun doute un rôle important à jouer dans le contexte des ressources pédagogiques en FLE. Jetons donc un regard en arrière sur les matériaux qui ont été produits à partir des enregistrements originaux de la fin des années soixante, et aussi un regard vers l'avenir, pour considérer le potentiel des enregistrements contemporains. *Les Orléanais ont la parole* (Biggs & Dalwood), paru en 1976, a marqué un tournant important dans la publication des manuels d'apprentissage non seulement du FLE, mais des langues étrangères en général, surtout en ce qui concerne l'authenticité des textes et le rôle du corpus dans les matériaux pédagogiques d'apprentissage des langues. Jeune enseignante de français dans l'enseignement supérieur britannique en 1976, je me rappelle clairement la parution de ce livre et des bandes qui l'accompagnaient, et les discussions avec mes collègues où nous avons remarqué l'énorme différence entre ces enregistrements et les phrases et les conversations inventées que nous avons l'habitude d'écouter avec nos étudiants. Ignorants des questions concernant les définitions de l'authenticité qui ne seraient publiées que quelques années plus tard, nous avons qualifié ces enregistrements de « réels » ; et ignorants aussi des toutes premières expériences en 1969 avec les concordances tirées de corpus spécialisés dans l'enseignement de l'anglais de spécialité à Birmingham (McEnery & Wilson, 1997 : 12), nous ne nous sommes pas rendus compte qu'il s'agissait de la publication du premier manuel de FLE – et peut-être de langue – à base de corpus. Étant une des premières enseignantes à choisir ce manuel, j'ai été surtout frappée par le défi que représentait l'écoute du français « réel » pour mes étudiants, et pour moi en tant que professeur.

---

4. ESLO, <http://www.univ-orleans.fr/eslo/>, page consultée le 21 février 2008. Voir aussi Bergounioux, (1996) ; Serpollet *et al.* (2007).

5. Centre Orléanais de Recherche en Anthropologie et Linguistique devenu au 1/1/08 le Laboratoire Ligérien de Linguistique par fusion avec les linguistes de l'Université de Tours.

6. ESLO, <http://www.univ-orleans.fr/eslo/>, page consultée le 21/02/08.

Feuilletant *Les Orléanais ont la parole* en 2008, l'on est toujours frappé par la modernité de ce manuel dans son approche pédagogique, surtout parce qu'il permet aux apprenants d'écouter des textes authentiques, l'oral spontané dans la mesure du possible (les personnes enregistrées sont bien sûr conscientes du fait que leurs paroles sont enregistrées). Mais il est certain que ce manuel est aujourd'hui bien dépassé. Les personnes interviewées sont maintenant bien plus âgées que les parents de la plupart des étudiants dans l'enseignement supérieur. Les vêtements, les coiffures et les photos d'appareils électroménagers sont d'un intérêt historique. En outre, dans le contexte linguistique, les chercheurs de l'équipe ESLO découvrent dans des études comparatives avec ESLO 2 des changements qui sont survenus en français oral depuis la fin des années 1960. Par exemple :

Words used to represent society have also evolved since ESLO1. Words such as 'classe moyenne' (middle-class), 'favorisée' (advantaged), 'agriculteurs' (farmers), 'ruraux' (rural) have replaced 'classe bourgeoise' (bourgeois), 'de ceci' 'de cela' (these/those people), 'classe prolétaire' (working class), 'ouvrier' (worker), 'paysan' (peasant). The decline in Marxism is also reflected in changes in the lexicon of class. (Chesneau, 2007 : 5)

Dans le contexte pédagogique, il sera intéressant à l'avenir de comparer *Les Orléanais ont la parole* et les matériaux produits à partir des enregistrements de ESLO 2. Les développements dans la disponibilité des corpus et des ressources pédagogiques en ligne, l'annotation, et surtout l'existence d'une communauté – ou des communautés – de discours qui partage les ressources pédagogiques devraient permettre aux enseignants d'avoir accès à des corpus, voire à des textes enrichis avec des concordances déjà préparées.

### **2.3. SACODEYL**

Ce projet actuel permet de voir les développements en linguistique de corpus et apprentissage des langues qui sont survenus depuis ESLO 1. Prenant comme thème la vie des jeunes, l'objectif de la création des corpus SACODEYL est de créer des matériaux authentiques pour l'apprentissage des langues, suivant l'approche que nous avons esquissée plus haut, c'est-à-dire l'étude de textes authentiques enrichie d'exercices à base de données tirées d'un corpus (Braun, 2005). Les jeunes, âgés de 13 à 18 ans, sont enregistrés sur vidéo dans leur langue maternelle (allemand, anglais, espagnol, français, italien, lituanien et roumain). Les questions posées dans les entretiens concernent la famille, la vie quotidienne, les loisirs, les vacances, l'école, le travail, les projets pour l'avenir, et des questions à l'ordre du jour telles que l'interdiction de fumer dans les lieux publics, ou ce que l'Europe signifie pour eux. Le projet vise à rendre accessible aux professeurs de langue des corpus annotés pour les thèmes, la grammaire, le lexique, les marqueurs de discours, et le niveau selon les critères du *Cadre européen commun de référence pour l'apprentissage et l'enseignement des langues* (Conseil de l'Europe, 2001). En

outre, le projet a pour objectif de fournir aux professeurs des matériaux concernant le contexte culturel et linguistique, ainsi que des exercices et des activités pertinentes. « The overall objective is to help teachers and learners to proceed from decontextualised textual data to context-embedded discourse construction and thus to facilitate and promote authenticated language learning » (Kohn, 2008 : 173).

Un corpus de 13 entretiens en français<sup>7</sup> (la version finale en contiendra 22) nous permet déjà de voir l'utilité de ce genre de ressource. Par exemple, une jeune de 17 ans parle de ses loisirs<sup>8</sup>. Ce texte forme une ressource très riche pour un professeur et des apprenants de FLE. Une étude détaillée dépasserait de loin le cadre de cette étude ; nous nous limiterons donc à considérer deux aspects du texte, le choix des verbes pour désigner la pratique des loisirs et l'expression « mais bon ».

Les apprenants de FLE apprennent à employer deux prépositions différentes avec *jouer* en parlant des sports et des instruments de musique, et à employer *faire* pour parler d'activités comme la natation, la pêche, etc. A écouter les jeunes dans ce corpus, on se rend compte aussi qu'une variété d'expressions est employée, surtout avec le verbe *faire*, pour parler des sports, des instruments, et des loisirs. Ainsi dans le texte que nous avons cité on trouve les expressions suivantes<sup>9</sup> :

Oui, alors je fais du sport. Au lycée, je fais du badminton.  
Et sinon, je pratique la guitare depuis dix ans.  
on faisait de l'instrument et j'ai choisi la guitare  
Donc pendant pendant sept ans, j'ai fait de la guitare classique.  
Après quand j'ai arrêté, j'ai un peu fait de la guitare électrique  
il va falloir que je joue un morceau de guitare et j'ai choisi un morceau de guitare  
classique  
donc je travaille surtout la guitare classique, en ce moment.  
je me suis donc lancée dans la guitare électrique  
et je me suis dit qu'afin d'évoluer, c'était bien de jouer en groupe, de composer,  
Et c'est comme ça que j'ai pu trouver mes deux groupes avec lesquels j'ai joué  
Et j'aime bien faire de l'ordinateur.

La question se pose de savoir si cette personne est la seule dans le corpus à favoriser l'emploi de *faire* en parlant des loisirs. Avec quelques recherches de mots-clés concernant les sports et les loisirs dans les 13 entretiens, on trouve le verbe *faire* ainsi que d'autres expressions pour désigner les sports et les loisirs.

Je suis pas licencié de club mais je fais du surf.  
Oui, je fais du football dans un club.  
Ben, eh ben, je fais les quatre appareils : le sol, le saut, les barres asymétriques et la  
poutre  
Ben tout ce qui est du sport, généralement, j'aime bien. Je fais de la piscine en  
dehors, enfin avec le lycée,  
Le mercredi, je fais du rock et de la salsa.  
et la salsa, j'en fais avec une amie parce que c'est vrai c'est lié quand même au  
monde espagnol,

- 
7. Je tiens à remercier les membres du projet SACODEYL, surtout Sabine Braun, Petra Hoffstaedter et Kurt Kohn de m'avoir permis de travailler sur ces documents.
  8. Voir annexe 1. Une longue discussion sur l'épreuve de musique au baccalauréat et sur les badges que porte la jeune a été omise dans ce texte.
  9. Dans les concordances citées ici nous avons supprimé les annotations (à l'exception de <break/>) pour faciliter la lecture par les apprenants.

Ben déjà au niveau du sport, je fais beaucoup d'escrime. Quatre heures par semaine et aussi du tennis  
je fais de la piscine. Sinon, si vous voulez que j'énumère tous les sports que j'ai faits, on n'a pas fini.

Je vais me mettre à la guitare  
Ben, j'ai une préférence pour la poutre et le sol.  
Je pratique du football, ça fait huit ans que j'en pratique. Et c'est mon sport préféré.  
Ça va faire un an maintenant que je pratique la natation.  
Alors, ça fait six ans que je joue au football.

Pour un francophone ces exemples n'ont rien de particulièrement intéressant. Mais pour un professeur de FLE, francophone ou locuteur non natif, la disponibilité de textes où des jeunes parlent de leurs sports préférés permet de créer un environnement beaucoup plus authentique et motivant pour les apprenants que des phrases inventées dans un manuel. Ces textes et les exemples supplémentaires que l'on peut facilement rajouter, permettent de combler ce « manque de 'naturel' des interactions verbales du non-francophone en français » (Debrock *et al.*, 1999 : 46). En outre, comme c'est souvent le cas lorsqu'on consulte un corpus, on trouve des exemples que l'on n'aurait jamais inventés, comme cette jeune qui fait beaucoup de sports et qui observe : « Sinon, si vous voulez que j'énumère tous les sports que j'ai faits, on n'a pas fini. »

Une recherche du mot *bon* dans le corpus s'avère aussi très intéressante pour un professeur ou un apprenant. Contrairement à ce que nous enseignons souvent aux débutants, l'emploi le plus courant de ce mot n'est pas comme adjectif, mais comme marqueur linguistique. Sur les 33 occurrences, deux sont prononcés par l'intervieweur, jouant le rôle de marqueur linguistique :

Bon on va évoquer un troisième sujet, ce sont les études.  
Tu pourrais nous décrire une journée quotidienne, banale si je puis dire mais bon, qui se répète, ton emploi de temps, ce que tu peux faire

Sur les 31 occurrences prononcées par les jeunes, 10 sont des adjectifs et 21 des marqueurs linguistiques, dont 10 occurrences de « mais bon ».

on a quelques petites restrictions par rapport aux années d'avant mais bon.  
peut-être un peu trop un peu trop hollywoodien peut-être mais bon.  
Non, en fait, pas trop mais bon la culture espagnole, d'Amérique latine et tout ça, ça m'a toujours intéressée  
déjà y avait pas de chanteur mais bon on a composé quelques trucs et tout  
cette ville te plaît ? REPONSE : Oui, c'est vrai mais bon mais à la longue, ça devient peut-être un peu ennuyant d'habiter  
Voilà. En quelque sorte, c'est un peu comme un documentaire mais bon d'un autre côté avec un personnage principal quand même ça  
J'espère, j'espère mais bon c'est en Europe et c'est /ce n'est quand même pas  
y a vraiment une charge de connaissances mais bon en évitant de tout apprendre la veille, on peut quand même bien s'e  
je suis pas spécialement brillante, mais bon quand même, ça va, je suis quand même arrivée à m'en sortir. Après, en math  
quelque chose d'assez d'assez stressant, et d'assez mais bon si on s'organise bien on peut quand même arriver à gérer et je pense que ça

Lorsque l'on cherche des exemples dans un corpus pour créer des matériaux

pédagogiques pour des apprenants, il est essentiel de distinguer cette activité de la recherche linguistique sur la fréquence basée sur de vastes corpus de référence. Comme l'observe Aston (2001 : 75), l'emploi d'un corpus pour trouver des exemples est une activité valable, pourvu que l'on évite toute tendance à trop généraliser.

#### **2.4. Le discours de la salle de classe**

Le troisième projet que nous allons considérer concerne les langues de spécialité. Il s'agit d'un corpus de cours de français, enregistrés pour former une ressource pour de futurs professeurs de français. Puisqu'il est de coutume dans plusieurs pays, dont l'Irlande, de recommander que la langue principale de communication dans le cours de langue soit la langue cible, il est important de donner aux futurs professeurs une formation pertinente. A l'Université de Limerick dans un programme d'études qui forme des professeurs de deux langues (allemand, anglais langue étrangère, espagnol, français ou japonais), un corpus de cours de français, où la langue de l'enseignement est le français, est employé comme ressource pour permettre aux futurs professeurs de français d'observer la pratique de professeurs de français anglophones et francophones. La version finale du corpus<sup>10</sup> contiendra 15 cours :

- cinq cours de français en Irlande où le professeur est anglophone et les élèves sont anglophones ;
- cinq cours de français en Irlande où le professeur est francophone et les élèves sont anglophones ;
- cinq cours de français en France où le professeur et les élèves sont francophones.<sup>11</sup>

L'analyse de ces cours servira de support authentique à la préparation des futurs professeurs pour une vie professionnelle où la langue principale de l'enseignement est le français.

L'enregistrement et la transcription des premiers cours nous ont permis de les analyser et de les rendre accessibles aux étudiants. Par exemple, l'emploi des marqueurs de discours dans deux cours de français dans ce corpus et dans deux cours équivalents d'espagnol a été analysé par Amador Moreno *et al.* (2006). Les cours que nous avons analysés sont les suivants :

- un cours de français en Irlande où le professeur est francophone et les élèves sont anglophones ;

---

10. Stéphanie O'Riordan crée ce corpus pour servir de ressource aux étudiants qui suivent ce programme d'études. Elle analyse le corpus et les applications pédagogiques dans sa thèse de doctorat.

11. Cette dernière catégorie permet aux futurs professeurs et à leurs élèves d'entendre le français parlé par des élèves dont le français est la langue maternelle.

- un cours de français en France où le professeur et les élèves sont francophones ;
- un cours d'espagnol en Irlande où le professeur est hispanophone et les élèves sont anglophones ;
- un cours d'espagnol en Espagne où le professeur et les élèves sont hispanophones.

Comme nous le verrons, les résultats de cette étude sont à la fois très intéressants et très encourageants pour les futurs professeurs. Nous avons observé beaucoup de ressemblances entre les résultats dans les deux langues, mais nous nous limiterons ici à quelques observations sur l'analyse des cours de français et les réactions des étudiants à la consultation de ces données.

Le choix des marqueurs de discours est facile à justifier. Bien que la définition des marqueurs de discours soit assez complexe (Brinton, 1996 ; Schiffrin, 2001), les chercheurs considèrent en général qu'ils sont un moyen de structurer le discours (Hansen, 1997 : 185). Or, la communauté de discours qu'est la salle de classe est caractérisée par le rôle dominant du professeur, c'est-à-dire que c'est lui qui structure le discours. Les marqueurs de discours en français ont déjà fait l'objet de plusieurs études (Cadiot *et al.*, 1985 ; Heisler, 1996 ; Hansen, 1997 ; Moine, 2005), mais la salle de classe n'en est pas le contexte principal. Le tableau 1 résume l'emploi des marqueurs de discours par les deux professeurs de français (Amador Moreno *et al.*, 2006 : 90).

Marqueurs de discours	Professeur		Élèves	
	Cours 1 (Irl)	Cours 2 (Fr)	Cours 1 (Irl)	Cours 2 (Fr)
OK	48	0	0	0
alors	23	10	0	0
donc	14	11	0	0
bon	16	2	0	2*
allez	9	0	0	0
d'accord	0	2	0	0
bien	1	0	0	0
c'est-à-dire	1	0	0	0
enfin ('fin)	1	0	0	0

Tableau 1. Marqueurs de discours employés par deux professeurs de français.  
\* Prononcés dans de courtes interactions entre élèves.

Ces données confirment bien le rôle dominant du professeur dans la salle de classe. Elles révèlent aussi que ces deux locuteurs natifs de français (comme leurs homologues espagnols) emploient un nombre relativement restreint de marqueurs de discours. De plus, le nombre des fonctions des marqueurs de discours dans les quatre cours est aussi limité, cinq en tout : introduire un nouveau sujet, activité ou question ; attirer l'attention des élèves ; répéter ou préciser ce qui a été fait ; motiver ou encourager les élèves ; et corriger ou reformuler ce que le professeur ou l'élève vient de dire (Amador Moreno *et al.*, 2006 : 92).<sup>12</sup> Puisqu'il ne s'agit que de quatre cours de langue, cette étude ne permet pas des conclusions générales concernant l'emploi des marqueurs de discours par l'ensemble des professeurs de langue. Mais elle fournit aux futurs professeurs de langue des exemples pertinents, susceptibles de provoquer une réflexion sur l'activité d'enseigner dans une langue qui n'est ni leur langue maternelle ni celle de leurs élèves. Dans les réponses au questionnaire et à l'entretien semi-dirigé, cette réflexion était évidente. Les étudiants ont discuté les différences entre les enregistrements de ces deux cours et les manuels pour les professeurs de français qui recommandent des phrases utiles, souvent des impératifs : ouvrez vos cahiers, etc. L'impératif ne figure que rarement dans ces deux cours. Les étudiants, qui comprenaient bien la complexité de leur situation de locuteurs non natifs, ont conclu qu'il faudrait écouter et analyser aussi des enregistrements de cours par des professeurs dont le français n'est pas la langue maternelle avant de tirer des conclusions concernant leur propre pratique.

## **Conclusion**

Dans une certaine mesure, l'intégration des corpus oraux dans l'apprentissage du FLE est bien avancée. Les chercheurs en linguistique d'ESLO 2 et d'autres projets créent des corpus qui seront disponibles au public et publient des analyses des données qui peuvent influencer le contenu des cours et des manuels de FLE. Cependant, en ce qui concerne l'enseignement du FLE et de toutes les langues étrangères, il reste des obstacles majeurs à l'intégration de cette nouvelle ressource. Dans l'enseignement secondaire, la quasi-totalité des professeurs de langue ne consultent pas les corpus, et dans l'enseignement supérieur il semble que les professeurs qui présentent des textes extraits de corpus et des concordances à leurs apprenants sont eux-mêmes des chercheurs dans ce domaine. Le jour où la consultation de corpus par l'ensemble des professeurs de langue sera aussi fréquente que la consultation du dictionnaire ou de Google semble être toujours bien loin. Boulton et Wilhelm (2006) soulignent en outre qu'il devrait être possible de permettre aux apprenants, surtout au niveau avancé, de chercher dans des corpus les réponses aux problèmes qu'ils rencontrent, c'est-à-dire de jouer un rôle actif dans

---

12. Quelques exemples de ces fonctions figurent dans l'annexe 2.

le processus d'apprentissage. Parmi les obstacles qui empêchent ou qui retardent cette intégration, les plus importants sont la disponibilité des corpus oraux et le temps qu'il faut consacrer à la maîtrise du concordancier. Comme nous l'avons vu, des projets comme ceux que nous venons de considérer fourniront aux professeurs et aux apprenants des corpus oraux pertinents. En outre, utilisant des ressources en ligne comme *LexTutor*,<sup>13</sup> un enseignant ou un apprenant peut créer des concordances sans une formation préalable. La dissémination de l'information sur l'existence de tels outils et ressources permettra aux enseignants et à leurs apprenants d'en profiter, soutenus par des stages et des cours dans les programmes de formation des enseignants de FLE. Des projets de recherche évaluant l'efficacité de ces interventions et les réactions des enseignants et des apprenants devraient bien sûr accompagner de telles initiatives, pour enfin permettre l'intégration des corpus dans l'apprentissage du FLE.

## ANNEXES

### ANNEXE 1. SACODEYL<sup>14</sup>

QUESTION : Parlons un peu maintenant de tes loisirs.<sup>15</sup> Et ce que tu fais pendant tes loisirs ? Est-ce que tu as des activités particulières ?

REPONSE : Oui, alors je fais du sport. Au lycée, je fais du badminton. Et sinon, je pratique la guitare depuis dix ans.

QUESTION : Alors parle-nous un peu de cette expérience, longue expérience.

REPONSE : Avec la guitare, ben en fait, c'est quand je suis arrivée à Montpellier, mes parents, ils m'ont inscrite dans une école où on pouvait <break/> où les après-midi, on faisait de l'instrument et j'ai choisi la guitare et j'ai continué jusqu'au lycée. Là, j'ai arrêté donc je suis plus au conservatoire mais je continue moi toute seule <break/> personnellement.

QUESTION : Et c'est quel type de guitare que tu joues ?

REPONSE : Au conservatoire, j'ai appris la guitare classique. Donc pendant pendant sept ans, j'ai fait de la guitare classique. Après quand j'ai arrêté, j'ai un peu fait de la guitare électrique et j'ai cherché à faire des groupes, donc j'en ai fait quelques-uns mais là, j'en ai pas en ce moment, enfin des groupes de <break/> de groupes de rock. Et là, comme je vais passer le bac musique [...]

QUESTION : Alors je voudrais revenir sur une expérience que tu as évoquée, c'est que tu as joué dans des groupes de rock pendant trois années si je ne m'abuse, plus ou moins.

REPONSE : Voilà.

QUESTION : Et pourrais-tu nous parler de ces expériences ? Qu'est-ce qui t'a poussée ? Est-ce que c'est toi qui a créé les groupes ? Quelle musique vous jouiez ? Est-ce que vous avez joué devant <break/> dans des fêtes particulières ?

REPONSE : En fait, c'est <break/> après avoir été au conservatoire et la guitare classique, je me suis donc lancée dans la guitare électrique et je me suis dit qu'afin d'évoluer, c'était bien de jouer en groupe, de composer, je trouvais ça intéressant et enrichissant. Donc j'ai cherché des groupes, j'ai passé des annonces au lycée. Donc j'ai collé des papiers en disant que je cherchais un

---

13. [www.lextutor.ca/](http://www.lextutor.ca/), page consultée le 06/06/08.

14. Puisque le corpus n'est pas encore dans sa version finale, les noms des deux interlocuteurs ont été remplacés par « Question » et « Réponse ».

15. Les auteurs du corpus m'informent qu'ils ont opté pour la ponctuation traditionnelle pour rendre les transcriptions facilement lisibles par des professeurs et des apprenants qui ne sont pas des spécialistes en linguistique de corpus.

groupe. Et c'est comme ça que j'ai pu trouver mes deux groupes avec lesquels j'ai joué donc ça a jamais été terrible.

QUESTION : Alors pourquoi ?

REPONSE : Ben le premier, on était deux guitaristes et un batteur donc <break/> déjà y avait pas de chanteur mais bon <break/> on a composé quelques trucs et tout mais ça a pas duré pas duré longtemps et c'était pas très sérieux en fait donc on pouvait pas <break/> ça pouvait vraiment aboutir à quelque chose de construit. Et puis pareil, la deuxième fois, c'était pire. C'était juste un autre guitariste qui m'avait contacté donc on n'était que deux guitares et on a essayé de trouver un bassiste et un batteur mais on n'a pas réussi à trouver donc après on s'est disputé et c'est tombé à l'eau.

QUESTION : C'est tombé à l'eau. Et comme autre passe-temps, qu'est-ce que tu aimes bien faire ?

REPONSE : J'aime bien écouter de la musique. Et j'aime bien faire de l'ordinateur.

## ANNEXE 2. Discours de la salle de classe<sup>16</sup>

### 1. Introduire un nouveau sujet, activité ou question

PROF : Inquiet ?

ÉLEVE : Avoir soucis.

PROF : Oui être inquiet ça veut dire avoir des soucis. Très bien oui. Être inquiet, avoir des soucis.

OK ? **Bon.** D'autres questions ? ... OK. Alors dernière question c'était...

### 2. Attirer l'attention des élèves

PROF : Ça peut être aussi quelqu'un de

ÉLEVE : [tousse]

PROF : de froid qui est indifférente.

ÉLEVE : Ouais.

PROF : On peut dire [des élèves bavardent] **D'accord.** Donc [des élèves bavardent] vous avez la démarche...

### 3. Répéter ou préciser ce qui a été fait

ÉLEVE : Avec une barbe.

PROF : Avec une barbe. Qui avec une barbe ?

ÉLEVE : Un homme.

PROF : Un homme. Ouais un homme barbu c'est un homme avec une barbe. Oui ? **Alors** [dessinant au tableau] voilà un homme. Le voilà barbu... Oui ? Un homme barbu c'est un homme qui a une barbe.

### 4. Motiver ou encourager les élèves

PROF : Mais on a fait la deuxième partie de la question quatre ? On a répondu à « Pourquoi Antoine voulait-il voir Daniel » ? [Les élèves acquiescent] Oui ? OK. **Alors on y va.** Question numéro cinq.

### 5. Corriger ou reformuler ce que le professeur ou l'élève viennent de dire

PROF : Quelqu'un qui a des pouvoirs, c'est quelqu'un qui est puissant. Ouais. **C'est-à-dire** c'est quelqu'un qui peut faire ce qu'il veut. OK. Donc il peut faire la plupart des choses qu'il veut faire. OK ? [...] Pourquoi vous riez c'est vrai. C'est normal. **'fin** [enfin] c'est pas normal mais c'est c'est la bonne traduction ouais [Petits rires dans la classe]

(Amador Moreno *et al.*, 2006 : 91-93)

16. L'auteur du corpus a opté pour la ponctuation traditionnelle pour rendre les transcriptions facilement lisibles par de futurs professeurs qui ne sont pas des spécialistes en linguistique de corpus.

## BIBLIOGRAPHIE

AMADOR MORENO, Carolina, Angela CHAMBERS & Stéphanie O'RIORDAN, 2006. « Integrating a corpus of classroom discourse in language teacher education: the case of discourse markers. » *ReCALL*, 18/1, p. 83-104.

ARMSTRONG, Nigel & Alan SMITH, 2002. « The influence of linguistic and social factors on the recent decline of French *ne*. » *Journal of French Language Studies*, 12/1, p. 23-41.

ASTON, Guy, 2001. « Text categories and corpus users: a response to David Lee. » *Language Learning & Technology*, 3/1, p. 73-76.

BAUDE, Olivier & Iris ESHKOL, 2007. « Constitution et exploitation d'un grand corpus de 'données situées'. Problèmes et solutions pour les *Enquêtes socio-linguistiques à Orléans (1968-2008)*. » Communication présentée au Colloque de Fribourg. <http://www.univ-orleans.fr/eslo/spip.php?article65>, page consultée le 25/07/08.

BERGOUNIOUX, Gabriel, 1996. « Etude sociolinguistique sur Orléans. » *Revue française de linguistique appliquée*, 1/2, p. 87-88.

BIGGS, Patricia & Mary DALWOOD, 1976. *Les Orléanais ont la parole*. Londres : Longman.

BILGER, Mireille, 2007. « Réflexions sur un obscur objet de désir : le corpus. » *Cahiers de l'Association for French Language Studies*, 13/1, p. 2-17. <http://www.afls.net/Cahiers/13.1/13.1%20Bilger.pdf>, page consultée le 25/07/08.

BOULTON, Alex & Stephan WILHELM, 2006. « Habeant Corpus – they *should* have the body. Tools learners have the right to use. » *ASp*, 49-50, p. 155-170.

BRAUN, Sabine, 2005 « From pedagogically relevant corpora to authentic language learning contents. » *ReCALL*, 17/1, p. 47-64.

BREEN, Michael, 1985. « Authenticity in the language classroom. » *Applied Linguistics*, 6/1, p. 60-70.

BRINTON, Laurel J., 1996. *Pragmatic Markers in English: Grammaticalization and Discourse Functions*. Berlin/New York : Mouton de Gruyter.

CADIOT, Anne, Oswald DUCROT, Bernard FRADEN & Than Binh NGUYEN, 1985. « *Enfin* marqueur métalinguistique. » *Journal of Pragmatics*, 2/3, p. 199-239.

CAPPEAU, Paul & Françoise GADET, 2007a. « L'exploitation sociolinguistique des grands corpus : maître-mot et pierre philosophale. » *Revue française de linguistique appliquée*, 12/1, p. 99-110.

CAPPEAU, Paul & Françoise GADET, 2007b. « Où en sont les corpus sur les français parlés ? » *Revue française de linguistique appliquée*, 12/1, p. 129-133.

CAPPEAU, Paul & Magali SEIJIDO, 2005. *Les corpus oraux en français* (inventaire 2005 v.1.0).

[http://www.culture.gouv.fr/culture/dglf/recherche/corpus\\_parole/Inventaire.pdf](http://www.culture.gouv.fr/culture/dglf/recherche/corpus_parole/Inventaire.pdf), page consultée le 25/02/08.

CHESNEAU, Annie, 2007. « Going back over the survey: 'Language: the looking glass of society'. » In M. Davies, P. Rayson, S. Hunston & P. Danielsson (éds.), *Proceedings of the Corpus Linguistics Conference CL2007*. University of Birmingham, Grande Bretagne, 27-30 juillet 2007.

<http://ucrel.lancs.ac.uk/publications/CL2007/>, page consultée le 25/07/08.

COBB, Tom, 1997. « Is there any measurable learning from hands on concordancing? » *System*, 25/3, p. 301-315.

CONSEIL DE L'EUROPE, 2001. *Cadre européen commun de référence pour l'apprentissage et l'enseignement des langues : apprendre, enseigner, évaluer*. Paris : Didier.

DEBAISIEUX, Jeanne-Marie, 2002. « Le fonctionnement de 'parce que' en français contemporain : étude quantitative. » In C.D. Pusch & W. Raible (éds.), *Romanistische Korpuslinguistik – Romance Corpus Linguistics*. Tübingen : Gunter Narr Verlag, p. 349-362.

DEBROCK, Mark, Danièle FLAMENT-BOISTRANCOURT & Raymond GEVAERT, 1999. « Le manque de 'naturel' des interactions verbales du non-francophone en français : analyses de quelques aspects à partir du corpus LANCOM. » *Faits de Langue*, 13. *Oral-Ecrit : Formes et théories*, p. 46-56.

DEWAELE, Jean-Marc & Nathalie WOURM, 2002. « L'acquisition de la compétence sociopragmatique en langue étrangère. » *Revue française de linguistique appliquée*, 7/2, p. 129-143.

DURAND, Jacques & Julien EYCHENNE, 2004. « Le schwa en français : pourquoi des corpus ? » *Corpus*, 3. *Usage des corpus en phonologie*.

<http://corpus.revues.org/document246.html>, page consultée le 25/02/08.

ELLIS, Nick, 2002. « Frequency effects in language processing: a review with implications for theories of implicit and explicit language acquisition. » *Studies in Second Language Acquisition*, 24, p. 143-188.

*ESLO : enquête sociolinguistique à Orléans*. <http://www.orleans.fr/eslo/>, page consultée le 21/02/08.

HABERT, Benoît, Adeline NAZARENKO & André SALEM, 1997. *Les linguistiques de corpus*. Paris : Armand Colin.

HANSEN, Maj-Britt Mosegaard, 1997. « 'Alors' and 'donc' in spoken French: a reanalysis. » *Journal of Pragmatics*, 28/2, p. 153-88.

HEISLER, Troy, 1996. « OK – a dynamic marker in Montreal French. » In J. Arnold,

R. Blake, B. Davidson, S. Schwenter & J. Solomon (éds.), *Sociolinguistic Variation: Data, Theory and Analysis*. Stanford University : CSLI Publications, p. 293-312.

KENNEDY, Claire & Tiziana MICELI, 2001. « An evaluation of intermediate students' approaches to corpus investigation. » *Language Learning & Technology*, 5/3, p. 77-90.

KOHN, Kurt, 2008. « Telos Language Partner: 'do it yourself' authoring for content-based language learning. » In A. Gimeno (éd.), *Computer Assisted Language Learning: Authoring Tools for Web-Based CALL*. Valencia : Universidad Politécnica de Valencia, p. 157-174.

LEE, Winnie Yuk-chun, 1995. « Authenticity revisited: text authenticity and learner authenticity. » *ELT Journal*, 49/4, p. 323-328.

MARTINEAU, France & Raymond MOUGEON, 2003. « A sociolinguistic study of the origins of 'ne' deletion in European and Quebec French. » *Language*, 79/1, p. 118-152.

McENERY, Tony & Andrew WILSON, 1997. « Teaching and language corpora. » *ReCALL*, 9/1, p. 5-14.

MISHAN, Freda, 2005. *Designing Authenticity into Language Learning Materials*. Bristol : Intellect.

MOINE, André, 2005. *The Role of Discourse Markers in the Structuring of Discourse: A Study of the Use of the Word 'Alors' in the French Language*. Queenston : Edwin Mellen Press.

MOUGEON, Raymond, Terry NADASDI & Katherine REHNER, 2002. « État de la recherche sur l'appropriation de la variation par les apprenants avancés du FL2 ou FLE. » *AILE*, 17. *L'Acquisition de la variation par les apprenants du français langue seconde*, éds. J.M. Dewaele & R. Mougeon, p.17-50.

O'SULLIVAN, Ide & Angela CHAMBERS, 2006. « Learners' writing skills in French: corpus consultation and learner evaluation. » *Journal of Second Language Writing*, 15/1, p. 49-68.

REHNER, Katherine & Raymond MOUGEON, 1999. « Variation in the spoken French of immersion students: to 'ne' or not to 'ne', that is the sociolinguistic question. » *La Revue canadienne des langues vivantes*, 56, p. 124-154.

SACODEYL: *System-Aided Compilation and Open Distribution of European Youth Language*. <http://www.um.es/sacodeyl/>, page consultée le 25/02/08.

SCHIFFRIN, Deborah, 2001. « Discourse markers: language, meaning and context. » In D. Schiffrin, H.E. Hamilton & D. Tannen (éds.), *Handbook of Discourse Analysis*. Malden, MA : Blackwell.

SERPOLLET, Noëlle, Gabriel BERGOUNIOUX, Annie CHESNEAU & Richard

WALTER, 2007. « A large reference corpus for spoken French: ESLO 1 and 2 and its variations. » In M. Davies, P. Rayson, S. Hunston & P. Danielsson (éds.), *Proceedings of the Corpus Linguistics Conference CL2007*. University of Birmingham, UK, 27-30 juillet. <http://ucrel.lancs.ac.uk/publications/CL2007/>, page consultée le 25/02/08.

SINCLAIR, John & Malcolm COULTHARD, 1975. *Towards an Analysis of Discourse: The English Used by Teachers and Pupils*. Londres : Oxford University Press.

STEVENS, Vance, 1991. « Concordance-based vocabulary exercises: a viable alternative to gap-fillers. » In T. Johns & P. King (éds.), *Classroom Concordancing: English Language Research Journal*, 4. Birmingham : Centre for English Language Studies, University of Birmingham, p. 47-63.

WIDDOWSON, Henry, 1998. « Context, community and authentic language. » *TESOL Quarterly*, 32/4, p. 705-716.

WIDDOWSON, Henry, 2000. « On the limitations of linguistics applied. » *Applied Linguistics*, 21/1, p. 3-25.

