



ASp
la revue du GERAS

15-18 | 1997
Anglais et français de spécialité

Applications des concordanciers à l'enseignement de la grammaire anglaise en DEUG

Joseph Rézeau



Édition électronique

URL : <http://asp.revues.org/2997>
DOI : 10.4000/asp.2997
ISSN : 2108-6354

Éditeur

Groupe d'étude et de recherche en anglais de spécialité

Édition imprimée

Date de publication : 1 décembre 1997
Pagination : 117-132
ISSN : 1246-8185

Référence électronique

Joseph Rézeau, « Applications des concordanciers à l'enseignement de la grammaire anglaise en DEUG », *ASp* [En ligne], 15-18 | 1997, mis en ligne le 21 mai 2012, consulté le 21 février 2017. URL : <http://asp.revues.org/2997> ; DOI : 10.4000/asp.2997

Ce document a été généré automatiquement le 21 février 2017.

Tous droits réservés

Applications des concordanciers à l'enseignement de la grammaire anglaise en DEUG

Joseph Rézeau

Introduction

- 1 Ces dix dernières années ont vu l'émergence concomitante de programmes informatiques permettant la recherche rapide des collocations des mots d'une langue d'une part et de corpus de textes sous forme informatisée d'autre part. Le chercheur en linguistique dispose ainsi des outils et des données nécessaires pour mettre en évidence et analyser les propriétés distributionnelles des mots du point de vue de leur combinatoire sémantique, syntaxique et discursive. Quant à l'enseignant de langues sur le terrain, il peut utiliser ces mêmes outils en mettant à la disposition de ses étudiants un type d'apprentissage s'appuyant sur des données authentiques, ce que Tim Johns (université de Birmingham) appelle *Data Driven Learning*.
- 2 Malgré le nombre croissant de communications et d'articles qui sont consacrés aux concordances (voir références bibliographiques), bien des collègues ignorent encore leur existence, ou sont sceptiques quant à leur utilité. Stevens (1995) disait à ce propos :
Language teachers fall into three groups: those who have never heard of concordances, those who haven't yet taken them seriously, and those who swear by them.
- 3 Dans l'espoir d'amener les lecteurs de cet article au moins à prendre les concordances au sérieux, la première partie de cet article passera en revue les sources de corpus utilisables et les principales caractéristiques des programmes de concordances, tandis qu'une deuxième partie donnera des exemples concrets d'application à l'apprentissage de la grammaire anglaise en DEUG.

Comment constituer un corpus de textes ?

- 4 On trouve de plus en plus de textes sous forme numérique, mais ils sont plus ou moins faciles d'accès et plus ou moins compatibles avec un traitement informatique. Sinclair (1991) consacre un chapitre à la création du corpus. L'enseignant isolé des grands centres de recherche se constituera son propre corpus de manière plus ou moins artisanale. Par exemple, à la fin des années 80, nous utilisons les dépêches que la BBC diffusait sur le Minitel (code 3615 BBC). La capture en était laborieuse, et ne rapportait guère que 1200 mots par jour... Et pourtant, au moment de la Guerre du Golfe, l'accumulation quotidienne de ces dépêches nous a permis d'étudier *in vivo* l'évolution de la détermination du groupe nominal en anglais (Annexe 1).
- 5 Actuellement, Internet offre des possibilités infiniment plus étendues de capture de textes, et dans des genres plus variés. De plus en plus d'œuvres de fiction sont également disponibles sur Internet comme sur cédérom. Sur abonnement (coûteux), il est possible de consulter l'importante *bank of English* de Collins qui compte plus de 200 millions de mots... Mais dans ce cas, l'utilisateur soumet une requête au serveur qui l'exécute et renvoie un résultat et on ne maîtrise pas bien les paramètres de la recherche.
- 6 La solution la plus pratique à défaut d'être la plus complète consiste à utiliser l'un des cédéroms de la presse britannique ou américaine. Nous utilisons habituellement celui du *Times & Sunday Times* qui comporte l'intégralité des articles d'une année, soit 44 millions de mots (en 1996). Est-il bien nécessaire d'avoir accès à un corpus quantitativement aussi important ? Oui, d'après Sinclair (1991 : 18)

The only guidance I would give is that a corpus should be as large as possible, and should keep on growing. [...] In order to study the behaviour of words in text, we need to have available quite a large number of occurrences.
- 7 Pour les collègues qui travaillent dans la traduction des médias, l'idéal est d'avoir parallèlement un cédérom comportant le même genre de corpus, par exemple celui du quotidien *Le Monde* ou encore celui de l'*AFP*. Malheureusement, depuis 1995, le texte du cédérom du *Monde* est « codé », et, pour obtenir un corpus d'articles utilisables par un concordancier, il faut au préalable exporter l'ensemble des textes des articles sur disque dur. Le cédérom de l'*AFP*, quant à lui, offre l'avantage de rassembler sur un seul cédérom une sélection de dépêches de l'*AFP* depuis 1988 jusqu'à l'année où vous achetez le produit. Un seul reproche : les textes antérieurs à fin 1991 sont dépourvus des accents du français, ce qui est pour le moins fâcheux. Il faut ajouter que l'*AFP* propose également un cédérom spécial *Sciences*, qui regroupe l'ensemble des dépêches scientifiques émises par le Service scientifique de l'*AFP* depuis 1990 jusqu'à ce jour. Ce produit peut se révéler particulièrement intéressant dans le cadre d'un enseignement spécifique du français à un public scientifique.
- 8 Il est également envisageable de scanner des textes, mais il s'agit d'une opération très longue et qui demande une vérification soignée, car aucun OCR n'est fiable à 100 %. Cette opération peut être envisagée, à défaut d'avoir accès à un corpus déjà informatisé, pour la constitution d'un corpus spécialisé, par exemple d'anglais scientifique, juridique, etc. ou encore pour la constitution de corpus multilingues.
- 9 Enfin, une autre source de corpus pour un travail d'analyse des productions des apprenants consiste tout simplement en ces productions, fournies de préférence sous forme informatique (Osborne 1996).

- 10 À notre connaissance, il n'existe pas actuellement de *recensement* des divers types de documents disponibles sous forme numérique (en dehors du projet Gutenberg). Ce pourrait être une initiative intéressante du GERAS.

Les programmes de concordances

- 11 Avant l'apparition, relativement récente, des « concordanciers » sur le marché, et en dehors des « gros » programmes tournant sur les gros ordinateurs, de petits programmes existaient, programmés en basic, ou en macros de traitements de texte : *Word* (cité par Tribble & Jones 1989) ou *FrameTeach* (Rézeau 1988a et 1988b). Les années 1990 ont vu l'apparition de *Micro concord*, sous DOS, qui n'est plus disponible, et a été remplacé en 1996 par *WordSmith*. Dans l'environnement MacIntosh, citons *MonoConc*. Les fonctionnalités que l'on peut attendre des concordanciers les plus récents sont les suivantes.
- création de concordances de type KWIC (Key Word In Context)
 - tri en ordre alphabétique à droite ou à gauche du mot recherché
 - spécification d'une recherche avec *wildcards*
 - obtention de listes de tous les mots d'un texte, avec tri en ordre alphabétique ou en ordre de fréquence
 - exportation des résultats obtenus pour exploitation ultérieure dans un traitement de texte
- 12 Le *COLLINS English collocations on CDROM* offre des concordances « toutes faites », utiles pour le lexique, mais pas pour la grammaire.

Que peut-on faire avec un concordancier ?

- 13 Sur l'intérêt de l'utilisation des concordances dans l'apprentissage de la langue, nous reprendrons à notre compte les trois arguments avancés par Johns (1988 : 29-31 et 1991 : 1-4)
- authenticité (des données et de l'activité) dans l'apprentissage
 - activité contrôlée par l'apprenant (à des degrés divers...)
 - l'apprentissage conçu comme une recherche : c'est le concept du « data-driven learning » (DDL) :
We simply provide the evidence needed to answer the learner's questions, and rely on the learner's intelligence to find answers.
- 14 Le cadre institutionnel dans lequel nous avons utilisé les concordances est celui de l'enseignement de grammaire anglaise en DEUG 2^e année, où nous poursuivons un double objectif de renforcement des bases et d'initiation à la réflexion linguistique. Le manuel de référence utilisé est la *Grammaire explicative de l'anglais* de Larreya et Rivière. Nous avons observé que les étudiants ont tendance à vouloir appliquer les règles de la grammaire pour leurs besoins de thème grammatical, sans comprendre que ces « règles » proviennent de l'observation de la langue. De ce point de vue, les exemples « fabriqués » que l'on trouve dans le manuel de grammaire leur semblent être faits pour donner raison aux règles, et non l'inverse. Face à cette situation problématique, nous avons formulé l'hypothèse que l'utilisation de données mises en forme sous forme de concordances pourrait impliquer davantage les étudiants dans leur réflexion linguistique, en introduisant une langue authentique et non fabriquée ainsi qu'en sollicitant de leur part des stratégies d'observation et de recherche personnelle.

- 15 Pour la suite de cet article, nous nous contenterons de commenter brièvement les exercices et activités fondés sur des concordances (que l'on trouvera en annexe), ces exercices fournissant d'eux-mêmes l'illustration des objectifs visés dans notre expérimentation.

La détermination du substantif WAR en anglais (annexe 1)

- 16 Comme nous l'avons évoqué plus haut, les dépêches de la BBC capturées sur Minitel au moment de la guerre du Golfe nous ont permis une étude sur le vif des différents degrés de la détermination en anglais.

Les subordonnées nominales en THAT après un nom apparenté à un verbe d'opinion ou de connaissance (annexe 2)

- 17 Tréville et Duquette (1996 : 12) évoquent le concept de « lexique-grammaire », dont l'idée centrale est que ces deux domaines de la langue sont en interaction constante. Cet exercice démontre que, s'il y a des synonymes dans le *lexique* d'une langue, il n'y en a pas dans son *vocabulaire*, c'est-à-dire dans *l'emploi réel* des mots. Pour ce type d'exercice, le recours au *Collins collocations on CDROM* peut s'avérer utile, pour les collocations qu'il donne soit des mots manquants soit des mots de leur contexte immédiat.

Passifs d'état et passifs d'action (Annexe 3)

- 18 Plusieurs indices contextuels « sautent aux yeux ». L'emploi des personnes semble corroborer l'explication b) de Larreya, en effet, *very* est presque toujours employé dans un contexte où le verbe est à la première personne (personne privilégiée pour exprimer « un sentiment, une impression du sujet »), alors que l'on trouve *much* dans des contextes aux deuxième et troisième personnes. En revanche, il est plus difficile de détecter à travers les occurrences fournies si *concerned* et *interested* sont plutôt des adjectifs ou plutôt des verbes. Enfin, un indice évident et qui n'est pas mentionné par Larreya est la fréquence de 100 % d'une forme négative en compagnie de *much* + *interested*. La Figure 1 illustre les résultats d'une recherche portant sur un corpus de 33 millions de mots (*Times & Sunday Times* de 1995) et ayant fourni un total de 125 occurrences de *much/very* dans le contexte immédiat de *concerned/interested*. On y observe le pourcentage faible (pour *interested*) et très faible (pour *concerned*) des occurrences de *much* par rapport à *very*, fait que la grammaire utilisée ne mentionne absolument pas.

Verbes à particule et verbes prépositionnels : l'exemple de *turn on* (annexe 4)

- 19 L'accumulation d'exemples authentiques illustrant une classification permet un retour à l'observation de la langue pour l'analyser... Bien évidemment, la classification elle-même est établie à partir d'observations de faits de langue. Avec des étudiants d'un niveau plus avancé, il conviendrait de partir des occurrences de *turn on* pour aboutir à une classification. On pourrait par exemple proposer la liste des occurrences à des étudiants de licence ou de concours, leur demander d'établir leur propre classification, puis soumettre cette classification et les occurrences à des étudiants de DEUG, avec un retour

éventuel vers les rédacteurs de la classification pour leur faire part des problèmes éventuellement rencontrés par leurs camarades. Notons au passage les citations 3 et 13, qui présentent une ambiguïté intéressante à discuter.

Cannot bear/stand + N/TO+INF/V-ING (annexe 5)

- 20 Nous observons que *cannot bear* a un taux d'occurrence relativement faible dans nos deux corpus et que cette collocation est prioritairement accompagné d'un Groupe Nominal. La plus grosse surprise vient du très faible taux d'occurrence de la forme en *-ING* (et plus particulièrement de la structure *N + V-ING*) dans tous les cas de figure. Cette constatation est contraire à la présentation des faits dans Larreya (p. 278) comme chez Berland-Delépine (§462). Nous en tirons la leçon que les grammaires déforment la réalité, d'abord en accordant autant de place (voire plus) aux exceptions qu'aux régularités, ce qui est une évidence, mais dans le cas présent, par omission. On ne mentionne *can't bear* que dans le contexte du verbe qui les suit, à la forme infinitive ou en *-ING*, alors que ces verbes sont principalement suivis d'un substantif. L'utilisation d'une concordance permet de rétablir une image statistiquement plus fidèle à la réalité...

I spent all my life/I have spent all my life? (annexe 6)

- 21 Le point de départ de cette recherche de concordance a été la phrase de thème grammatical proposée par un collègue : « J'ai passé toute ma vie à étudier la philosophie, mais je ne suis guère plus avancé ». Le corrigé proposait : « I spent all my life studying philosophy », ce qui nous semblait être soit une erreur soit un américanisme. Une recherche dans notre corpus habituel a donné les résultats 1 à 6 dans lesquels les emplois respectifs du *Present Perfect* et du *Simple Past* correspondent à ce que nous attendions. En effet, il ressort clairement de ces citations que le *Simple Past* correspond à une partie antérieure de la vie de celui qui parle, partie achevée et nettement coupée de la partie actuelle, tandis que l'emploi du *Present Perfect* correspond aux cas où la partie de la vie évoquée par le locuteur se prolonge jusqu'au moment de l'énonciation. Dans le cas de notre phrase de départ, il nous semblait que, puisque la vie du locuteur n'est pas terminée et qu'il l'a consacrée toute entière à l'étude de la philosophie, l'emploi du *Present Perfect* s'imposait. Mais la citation 7 semble contredire ce point de vue, puisque le locuteur parle également d'une vie entière, et pourtant utilise le *Simple Past*. L'explication de cet emploi est pourtant tout à fait claire dans le contexte très particulier de ce locuteur, transsexuel, pour qui toute cette vie antérieure, passée à courir après des ombres, est bel et bien terminée. Quelle meilleure preuve de l'intérêt des concordances que cette citation, totalement authentique, et qu'on ne risquerait guère de trouver comme exemple dans un manuel de grammaire, où elle paraîtrait fabriquée ! Nous rejoignons ici la remarque de Johns (1991 : 3) :

[the learners] will often notice things that are unknown... to the standard works of reference on the language. It is this element of challenge and of discovery that gives DDL its special flavour and stimulus.

Although (annexe 7)

- 22 Un développement récent dans le domaine des concordances est celui des concordances parallèles, auxquelles s'intéresse entre autres le projet *Lingua*, *Multiconcord* développé à

Birmingham, en lien avec plusieurs partenaires européens (dont Nancy 2 et Rennes 2 pour la France). Ce type de programme fonctionne dans un premier temps comme un concordancier monolingue, en recherchant dans le texte source les occurrences d'un mot donné. Puis, le programme essaie de retrouver en parallèle les phrases du texte cible (traduction) équivalant aux phrases où se trouvent les occurrences trouvées... Se posent alors de difficiles problèmes d'alignement des textes du corpus. Parmi les objectifs de l'utilisation pédagogique de ce type de concordancier nous citerons l'analyse contrastive appliquée et la lutte contre la tendance à la traduction mot à mot. Au total, le concordancier parallèle constitue un très intéressant complément aux dictionnaires et grammaires traditionnels. Ajoutons-y la possibilité de se constituer des corpora de langue spécialisée à condition, bien entendu, de disposer de textes bilingues (ou plurilingues) du domaine, ce qui est pose d'autres problèmes, de copyright en particulier.

- 23 À titre d'exemple, nous proposons en annexe un travail sur la traduction (ou plus exactement les multiples traductions possibles) de *although* en français. La variété des traductions trouvées dans un corpus pourtant relativement restreint (un ou deux romans), par rapport aux résultats trouvés dans les dictionnaires suffit à convaincre de la richesse de l'apport de ce type d'outil de travail. Un autre avantage du concordancier parallèle ou multilingue par rapport au concordancier unilingue, s'il est utilisé pour l'étude de la grammaire de la langue-cible est que, puisque l'étudiant dispose du texte traduit dans sa langue maternelle, il n'y a plus de problème de compréhension. Les lecteurs intéressés trouveront (sur nos pages Internet) d'autres exercices de ce type, et en particulier des exercices tentant de résoudre le délicat problème d'emploi contrasté de *while/whereas*, d'une part et de *though/although* d'autre part...

Conclusion

- 24 Cette conclusion commencera par une série de mises en garde. En effet, l'utilisation régulière de concordanciers et de corpus de texte suppose un certain nombre de prérequis. L'utilisateur doit avoir un accès facile à des données en volume suffisant, faute de quoi les résultats risquent de ne pas être significatifs, voire d'être faussés. Il doit en outre disposer d'un équipement informatique relativement puissant, et en particulier d'un lecteur de cédérom rapide et de beaucoup de mémoire vive. Il ne faut pas cacher que la recherche d'occurrences de collocations, leur tri, leur sélection, leur importation dans un traitement de texte pour la fabrication d'exercices, demande beaucoup de temps... Par ailleurs, il ne faut pas perdre de vue l'influence du type et du genre des textes utilisés sur les résultats. Nous utilisons abondamment la presse sur cédérom pour des raisons de commodité, mais il est bien évident qu'un corpus plus équilibré et comprenant en particulier un pourcentage appréciable de langue parlée serait souhaitable. Au moment où nous écrivons ces lignes, on nous promet une édition grand public et sur cédérom du fameux BNC (British National Corpus) qui comporte un assortiment de différents genres de langue.
- 25 Quel que soit le domaine ou la spécialité de son enseignement, ou le niveau des étudiants auxquels il s'adresse, tout enseignant d'anglais peut trouver une utilité à l'utilisation des concordanciers. Si l'équipement de son institution le permet, il est évidemment souhaitable, après une initiation au fonctionnement, de mettre les concordanciers directement à la disposition des étudiants. À défaut, l'enseignant peut fabriquer des

exercices papier tels que ceux que nous présentons dans cet article, ou encore les mettre sur Internet.

- 26 Nous espérons avoir convaincu les lecteurs que les concordances sont à prendre au sérieux, et nous les invitons à aller faire, à leur tour, un voyage dans le pays des concordanciers pour y observer, selon l'expression de Johns « the company that words keep ».

BIBLIOGRAPHIE

- Flowerdew, J. 1996. « Concordancing in language learning ». In Pennington, Martha (dir.), *The Power of CALL*. Houston, TX : Athelstan.
- Johns, T. 1988. « Implications et applications des logiciels de concordance dans la salle de classe ». *Les Langues modernes* 5.
- Johns, T. 1991. « Should you be persuaded: two samples of data-driven learning ». *ELR Journal* 4.
- Kettemann, B. 1995. « On the use of concordancing in ELT ». *Arbeiten aus Anglistik und Amerikanistik* 20/1, 29-41. <http://www.uni-graz.at/bernhard.kettemann>.
- Larreya, P. et C. Rivière. 1991. *Grammaire explicative de l'anglais*. Paris : Longman France.
- Lehmann, D., C. de Margerie et A. Pelfrène. 1988. « LECTICIEL : l'autonomie ? ». *Le français dans le monde*, numéro spécial Nouvelles technologies et apprentissage des langues.
- Osborne, J. 1996. « Sensibilisation lexico-grammaticale à l'aide d'un outil de concordance ». *ASp* 11-14.
- Rézeau, J. 1988a. « Que faire avec un outil professionnel en EAO des langues ? ». *Les langues modernes* 5.
- Rézeau, J. 1988b. « De l'utilisation d'un progiciel professionnel en EAO des langues ». *Le français dans le monde*, numéro spécial Nouvelles technologies et apprentissage des langues.
- Sinclair, J. 1991. *Corpus Concordance Collocation*. Oxford : Oxford University Press.
- Stevens, V. 1995. « Concordancing with language learners: Why? When? What? ». *CAELL Journal* 6/2, 2-10. <http://www.vancestevens.com/papers/archive/concordance1996.htm>
- Tréville, M-C. et L. Duquette. 1996. *Enseigner le vocabulaire en classe de langue*. Paris : Hachette.
- Tribble, C. & G. Jones. 1989 [1997 : new edition by Athelstan]. *Concordances in the Classroom*. Londres : Longman.
- Tribble, C. 1996. « Feature review: WordSmith Tools ». *CALL Review, The Journal of the IATEFL Computer SIG*, sept.

Références Internet et cédéroms

- Tim John's Data-Driven Learning page : <http://sun1.bham.ac.uk/johnstf/timconc.htm>
- CALL bibliography : <http://sun1.bham.ac.uk/johnstf/biblio.htm>
- Joseph Rézeau's Home page : <http://joseph.rezeau.pagesperso-orange.fr/>

Pour télécharger une version de démonstration de WordSmith sur Internet, aller à l'adresse :
<http://www.oup.co.uk/cite/oup/elt/software/>

Le Monde sur cédérom : Office Central de Documentation, 33 rue Linné F-75005 Paris.

Les cédéroms de l'AFP : Agence France-Presse, 13 Place de la Bourse B.P. 20 F-75061 Paris Cedex 02.

The Times & Sunday Times on CDROM : News Multimedia PO box 5060, Leighton Buzzard, LU7 7YS
England tel +44 (0) 1525 852813. www.newsmultimedia.co.uk

Collins Cobuild English Collocations on CD-ROM : HarperCollins publishers www.harpercollins.com

Broadcast News : sélection de scripts d'émissions de télévision et de radio de l'année écoulée (plus quelques photos) des chaînes & stations de radio américaines: ABC News CNN National Public Radio PBS Networks ; Research Publications US - Primary Source Media - c/o Josiane Stern 5 avenue Pasteur F-92400 Courbevoie tél 01 47 88 19 99 fax 01 43 33 30 18.

RÉSUMÉS

Les programmes de recherche des collocations des mots d'une langue sont récemment devenus conviviaux et abordables ; les corpus de textes sous forme numérique existent. L'utilisation didactique de ces outils reste néanmoins marginale ou épisodique. Cet article donne des exemples concrets et commentés d'application des concordanciers à l'enseignement de la grammaire anglaise à l'université et une ébauche de typologie d'exercices.

Concordancing programs have recently become more user-friendly and affordable on small computer systems; text corpora have likewise become more readily available in computer-readable format. However, such tools are still either unknown to or ignored by language teachers. The purpose of this article is to give a number of concrete examples of applications of concordancers to the teaching – and learning – of English grammar at university level, and to provide the basis for a typology of concordance-based exercises.

INDEX

Keywords : authenticity, autonomy, collocation, concordance, corpus, grammar, vocabulary

Mots-clés : authenticité, autonomie, collocation, concordance, corpus, grammaire, vocabulaire

AUTEUR

JOSEPH RÉZEAU