



HAL
open science

Les dictionnaires en ligne : nouvelles diffusions, nouveaux objectifs

Jean-Luc Manguin

► **To cite this version:**

Jean-Luc Manguin. Les dictionnaires en ligne : nouvelles diffusions, nouveaux objectifs. 2005, Paris, France. edutice-00001422v1

HAL Id: edutice-00001422

<https://edutice.hal.science/edutice-00001422v1>

Submitted on 26 Jan 2006 (v1), last revised 30 Mar 2006 (v2)

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.

Les dictionnaires en ligne : nouvelles diffusions, nouveaux objectifs

Jean-Luc Manguin (CNRS FRE 2805, Université de Caen)

Introduction

Les dictionnaires font partie depuis longtemps des ressources dont disposent le pédagogue, l'élève, et tout producteur d'écrit ; d'une manière générale, on peut même affirmer que ce genre d'ouvrage vient au secours de tout ceux qui rencontrent, à un moment ou à un autre, une difficulté de production ou de compréhension de texte. À l'heure où la maîtrise de la langue (re)devient cruciale, la disponibilité et l'actualité des dictionnaires se révèle capitale. Or, le support papier, de nature matérielle, est difficilement disponible pour plusieurs utilisateurs lorsqu'il se trouve en un seul exemplaire. En outre, malgré l'intérêt qu'il présente dans la fixation de la langue à une période donnée, l'ouvrage imprimé s'avère coûteux à maintenir à jour, aussi bien du côté de l'éditeur que du côté des lecteurs. Le dictionnaire en ligne envisage le problème sous d'autres aspects, en se libérant des contraintes matérielles de la dictionnaire. En effet, dans ce type de ressource, les problèmes de disponibilité, de limitation de place ou d'éditions successives ne se posent plus ; en revanche, c'est en aval, autrement dit après la mise en ligne du dictionnaire, que la résolution des questions dictionnaires devra se faire, dans un échange, le plus souvent indirect, avec les utilisateurs-lecteurs.

A travers l'expérience de notre dictionnaire des synonymes, nous montrerons comment un outil au service de la recherche s'est transformé, par la mise en vitrine du laboratoire sur Internet, en une ressource de référence pour la communauté francophone ; en particulier, nous mettrons en valeur les évolutions imposées par l'adoption empathique du comportement de l'utilisateur vis-à-vis de la ressource en ligne. Ces évolutions ont agi non seulement sur l'ergonomie ou l'accessibilité du dictionnaire, mais aussi sur son contenu, qui en font un outil d'aide à l'écriture.

Nous donnerons également un aperçu de l'outil Alexandria® de la société Memodata, qui est par essence un dictionnaire inséparable de la toile Internet, et nous montrerons comment cette ressource se révèle d'un grand intérêt pour la compréhension des textes.

Enfin, nous évoquerons le projet Alfalex de nos voisins belges de l'Université catholique de Louvain, dans lequel la finalité pédagogique dicte la production et l'exploitation du dictionnaire en ligne et des textes qui ont servi à son élaboration.

1. La dictionnaire Internet

Bernard Quemada (Quemada, 1968) a défini la notion de *dictionnaire* comme une discipline englobant tous les aspects de l'élaboration d'un dictionnaire dans le cadre d'un objectif ; il la distingue ainsi de la *lexicographie*, qui se charge alors de « l'analyse des mots de leurs significations » (Pruvost, 2002) ; néanmoins, comme le rappelle aussi Jean Pruvost dans le même ouvrage, ces deux notions sont intimement liées et complémentaires. Sans détailler les problèmes de dictionnaire qui se posent lors de la conception d'ouvrages imprimés, ou de lexiques électroniques sur cédérom, nous pouvons cependant affirmer que l'apparition des dictionnaires consultables sur Internet a contribué, du fait des spécificités de ce média, au renouvellement des questions et des réponses de la dictionnaire, tout en établissant des propriétés communes aux ressources en ligne, qu'elles soient gratuites ou non.

Tout d'abord, le fait d'être des ressources informatiques leur permet d'être munies, comme leurs homologues sur cédérom, de programmes d'aide à la saisie des requêtes de l'utilisateur. Principalement, l'interrogation d'un dictionnaire informatisé ne doit plus se faire en passant forcément par la forme « canonique » du mot, autrement dit son lemme (par ex. le masculin singulier pour les adjectifs), mais peut tolérer l'entrée d'un mot au pluriel ou d'un verbe dans une forme conjuguée ; c'est un programme lemmatiseur qui se charge ensuite de rediriger la requête vers les mots-vedette du dictionnaire. En élargissant ce principe, on peut adjoindre à ce lemmatiseur un correcteur d'erreur qui évite à un utilisateur distrait - ou gêné par une orthographe incertaine - d'être pénalisé en ne recevant aucune réponse à sa demande.

En outre, les dictionnaires en ligne ont, du fait de leur présence sur Internet, des caractéristiques dont certaines leurs sont imposées par les exigences des internautes ; ainsi, leur réponse se doit d'être rapide, puisqu'il est désormais inconcevable, à l'heure des liaisons à haut débit, d'attendre plus d'une ou deux secondes la réponse d'un serveur ; cette exigence est par ailleurs motivée par la comparaison avec les dictionnaires sur cédérom. Du point de vue du diffuseur de la ressource, ce souhait implique des contraintes matérielles importantes. A côté de cet aspect, les ressources en ligne doivent tirer parti de la caractéristique essentielle de leur support, c'est-à-dire l'hypertextualité ; cela se traduit généralement par la présence, dans la page de réponse, de liens qui conduisent à des compléments, et qui peuvent d'une part accélérer la consultation, ou d'autre part orienter l'utilisateur vers des possibilités nouvelles, voire insoupçonnées.

La permanence de la liaison entre le producteur et l'utilisateur constitue un autre aspect majeur des ressources en ligne ; il ne s'agit bien sûr pas de la permanence physique de la liaison informatique, mais du fait qu'à chaque connexion, l'utilisateur est susceptible de rencontrer une nouvelle version du site qu'il consulte, et qu'en plus il renouvelle ou complète les données personnelles qu'il envoie vers le site du producteur. La première propriété règle le problème de l'obsolescence du dictionnaire d'un usager, mais aussi du même coup la question du nombre d'exemplaires nécessaires à une utilisation dans un cadre collectif (qui peut être pédagogique). La seconde propriété vient du fait que tout accès à un site Internet laisse sur celui-ci des traces ; celles qui sont laissées par les usagers d'une ressource sont exploitables par le producteur pour connaître les préférences, mais aussi les lacunes du produit qu'il diffuse ; il peut ainsi en modifier le contenu en conséquence. Ce retour d'information direct renverse complètement le processus dictionnaire habituel qui consiste, par exemple, à établir la liste des mots-vedette d'un dictionnaire en fonction du public visé ; au contraire, dans le cas d'un dictionnaire en ligne, ce sont les mots demandés par les utilisateurs et absents de la ressource qui vont former la liste des futures entrées potentielles.

Nous allons maintenant examiner trois exemples de ressources en ligne, et voir comment ces propriétés sont exploitées dans chacune d'entre elles.

2. Le dictionnaire électronique des synonymes du CRISCO

Le dictionnaire électronique des synonymes du CRISCO résulte de la fusion de sept fichiers contenant les relations présentes dans les dictionnaires qui ont servi de base à cette construction : Lafaye (1958), Guizot (1864), Bailly (1946), Bénac (1954), Grand Larousse de la langue française (1971) et Grand Robert de la langue française (1985), suivant un principe de fusion de dictionnaires des synonymes énoncé dans la thèse d'André Kahlmann (1975). Toutefois, comme il était impossible de conserver les distinctions et divisions d'articles faites dans certains dictionnaires, en raison de l'hétérogénéité des classifications opérées suivant les lexicographes (voir Ploux, 1997), le fichier résultant de la fusion présente donc pour chaque mot-vedette la liste de ses synonymes « en vrac ». Ce principe de base se justifie également par le fait que ce dictionnaire a été créé pour servir de support pratique à certains travaux de recherche, précisément à la construction d'espaces sémantiques dont les fondements théoriques ont

été formulés par Bernard Victorri (Victorri et Fuchs, 1996). Outre cela, les relations synonymiques ont été symétrisées, autrement dit si un dictionnaire attribue à A un synonymes B, nous avons également placé A comme synonyme de B dans notre base ; l'intérêt de cette opération a été discuté dans la thèse de Kahlmann (op. cit.)

La mise en ligne de ce dictionnaire correspond à la création du site Internet du laboratoire (<http://www.crisco.unicaen.fr>) en 1998, et au désir de mettre en vitrine les travaux liés à ce dictionnaire ; au départ assez fruste, l'interface de requête s'est rapidement doté de l'hypertextualité en raison du succès que notre ressource a rencontré. Sa gratuité et sa richesse lexicale (49000 entrées, reliées par 400000 relations de synonymie) sont les clés de cet engouement. Le tableau ci-dessous montre la forte croissance du nombre de requêtes qu'il reçoit :

Année	Trafic (en millions de requêtes)	Progression
1999	0,5	
2000	1,9	298 %
2001	4,6	140 %
2002	9,9	114 %
2003	16,2	64 %
2004	24,8	53 %
2005 (prév)	34,2	39 %

Tableau 1 : trafic annuel du dictionnaire des synonymes

L'hypertextualité se manifeste dans notre dictionnaire de manière classique : la page de réponse fournit la liste des synonymes du mot demandé ; chacun de ceux-ci est un lien vers sa propre liste de synonymes. La rapidité de la réponse provient de la structure en arbre binaire du dictionnaire, qui induit une recherche en temps quasi-constant, avec un maximum de 16 comparaisons avant de trouver le mot demandé (ou de déterminer s'il n'est pas dans la base).

L'analyse de la fréquentation de notre ressource nous ayant appris la présence de demandes en provenance de pays non-francophones, nous avons dû régler rapidement la question de la correction des accents, sachant que les claviers informatiques de type français ne sont pas d'un usage universel ; l'adjonction d'un correcteur d'accent a été la première étape vers une interface de saisie qui soit plus au service de l'utilisateur. Par la suite, en effet, l'examen des mots les plus demandés, et plus spécialement de ceux qui ne recevaient pas de réponse de la part du dictionnaire nous a conduit à enrichir le correcteur suivant deux axes. Tout d'abord, nous avons ajouté un lemmatiseur qui ramène certaines requêtes vers leur « forme canonique », et ensuite nous avons enrichi le correcteur avec des procédures qui corrigent les fautes les plus fréquentes (par ex. redoublement de consonne à tort : « gallop » sera corrigé en « galop »). A chaque fois qu'une correction est effectuée, l'utilisateur en est informé dans la page des résultats. L'effet de tous ces ajouts se perçoit à l'examen du nombre de requêtes satisfaites : sur les 80 millions de requêtes que le dictionnaire a reçues, 87 % d'entre elles ont reçu une réponse ; ce chiffre lui-même se décompose en 74 % grâce au dictionnaire seul, et 13 % grâce aux corrections effectuées par l'interface. On peut ajouter que si l'on ne tient pas compte des 5 % de requêtes aberrantes (mots demandés 10 fois ou moins depuis 7 ans), le taux de réponse passe à 92 %. Nous voyons ici en application le processus dictionnaire « inverse » (par rapport à l'édition papier) mentionné au paragraphe

précédent : les traces laissées par les internautes ont permis d'enrichir l'interface du dictionnaire et d'améliorer la satisfaction de ses usagers. Il faut enfin signaler que ce procédé est également à l'œuvre dans l'arrivée de nouveaux mots-vedettes dans le dictionnaire ; nous avons ainsi ajouté des mots comme *festif*, *gratifiant*, *institutionnel*, *s'impliquer*, etc. Les détails de ces enrichissements sont présentés dans Manguin, 2005.

La seconde question importante de la dictionnaire est celle de l'usage qui est fait du dictionnaire, en particulier son adaptation au public visé ; dans le cas de notre dictionnaire, nous avons aussi répondu à cette question par le processus à rebours : ce sont les traces des internautes, et plus spécialement leur répartition dans le temps qui vont nous guider dans notre analyse. Tout d'abord, il faut remarquer que la fréquentation de notre ressource baisse de manière très importante durant le week-end (cf. tableau 2) :

Jour	Lundi	Mardi	Mercredi	Jeudi	Vendredi	Samedi	Dimanche
Part de trafic %	16,7	17,9	18,7	18,3	15,6	5,8	7,1

Tableau 2 : répartition hebdomadaire du trafic du D.E.S.

Ceci nous laisse supposer que le dictionnaire de notre laboratoire est un outil à vocation professionnelle ; la confirmation de cette hypothèse vient de la courbe de répartition du trafic en fonction de l'heure qu'il est chez celui qui consulte notre site (une part importante de nos consultations provient en effet du Québec) :

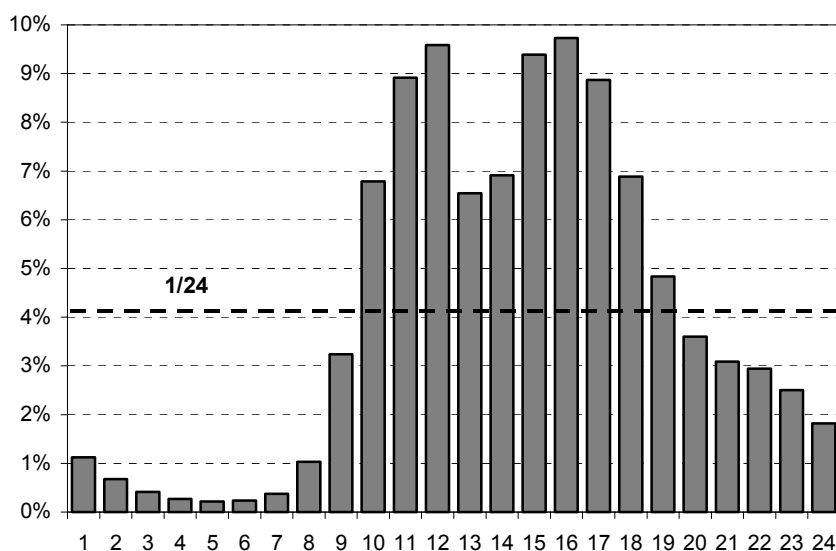


Figure 1 : histogramme journalier du trafic suivant l'heure client

On voit sur ce graphique que la majeure partie des requêtes est émise aux heures d'ouverture des bureaux ; d'ailleurs, le gouvernement québécois et l'administration helvétique sont parmi les « clients » les plus importants de notre ressource. Enfin, des témoignages écrits (par ex. du bureau de traduction des Nations-Unies) nous permettent d'affirmer que notre dictionnaire électronique des synonymes est bel et bien un outil d'aide à la production de textes ; autrement dit, il est avant tout utile à ceux qui maîtrisent déjà notre langue. Cependant, des témoignages d'enseignants ou de service

de formation nous conduisent à penser qu'il est également utile aux pédagogues lors de la préparation d'exercices de production.

En dernier lieu, il convient de noter que ce dictionnaire garde toujours un lien avec les principes théoriques qui l'ont fondé, et surtout une relation avec les travaux de recherche en linguistique, tant dans notre laboratoire que dans d'autres unités de recherche. Ces travaux qui partent d'une modélisation en graphe (cf. Ploux, 1997, Ploux & Victorri 1998, Manguin 2005) se retrouvent dans la visualisation des espaces sémantiques qui est accessible sur notre site à partir des requêtes « ordinaires ». Il y a là un autre aspect éducatif, certes qui concerne un public plus étroit, qui est celui de l'enseignement supérieur en Sciences du langage, mais qui a son importance à l'heure où les technologies de l'information et de la communication font partie des thèmes de recherche soutenus par les institutions.

3. L'outil Alexandria®

Alexandria® est un outil gratuit développé par la société Memodata® (<http://www.memodata.com/>), spécialisée dans la production de ressources lexicales mono- et multilingues ; il s'agit d'un ensemble de petits programmes qui vont interroger de manière « intelligente » un dictionnaire multilingue (actuellement 15 langues sont dans la base) lui-même hébergé sur les serveurs de la société conceptrice du produit. Son principe très simple a déjà donné lieu à plusieurs publications (voir par ex. Dutoit, 2005) : le but est de faciliter la compréhension des textes présents sur le Web en permettant à l'utilisateur d'interroger un dictionnaire par un seul double-clic de souris sur le mot qui lui pose un problème de compréhension. Derrière cette idée concise se cachent plusieurs propriétés que nous allons détailler.

Tout d'abord, c'est un système extrêmement simple à mettre en œuvre : il suffit en effet d'inclure 4 lignes (par un simple copier-coller) dans le code source d'une page HTML pour accéder aux fonctionnalités de l'outil sur cette page. Ces lignes contiennent en fait les ordres de téléchargement des programmes d'interrogation sur le site de Memodata® ; ainsi, lorsqu'un internaute consultera cette page « outillée », les programmes de consultation vont se charger sur son poste de manière invisible, et il lui suffira ensuite de double-cliquer sur un mot pour faire apparaître dans une petite fenêtre les informations lexicales (définition, traduction, et éventuellement synonymes) concernant ce mot. Cette nouvelle fenêtre est elle-même hypertextuelle : tous les mots qui y figurent sont des liens vers leurs données lexicales. L'interrogation est aussi pourvue d'un lemmatiseur, ce qui est normal puisqu'elle se fait à partir des mots d'un texte réel.

L'intérêt du multilinguisme apparaît, dans la version la plus simple d'Alexandria®, sous deux formes : tout d'abord, la fenêtre de cet outil permet de sélectionner la langue cible, et de donner au locuteur un résultat dans la langue de son choix parmi les 16 proposées¹. Un locuteur serbe, par exemple, mais qui posséderait un navigateur Internet dans une version américaine (toujours par exemple), pourrait disposer d'informations dans sa propre langue. Ensuite, le multilinguisme apparaît de manière automatique et relativement « intelligente » : le système détecte quelle est la langue du navigateur, et va automatiquement diffuser les informations dans cette langue. Ainsi, un néerlandais possédant sur son ordinateur un navigateur dans sa langue, s'il consulte une page d'un site français (la rubrique « Langue française » de TV5, par exemple), obtiendra la traduction néerlandaise d'un mot français qui lui semble difficile en cliquant simplement dessus.

¹ allemand, anglais, bulgare, coréen, espagnol, estonien, français, grec, italien, néerlandais, portugais, roumain, serbe, tchèque et turc.

Mais le multilinguisme et les fonctionnalités d'Alexandria® ne s'arrêtent pas là : moyennant un partenariat, Memodata® autorise le site qui a installé son outil à profiter des fonctionnalités avancées, précisément la saisie d'un mot dans la fenêtre de l'application, et le choix de la langue source (qui est normalement celui de la page d'appel). L'outil que nous décrivons devient ainsi un dictionnaire multilingue dans lequel l'utilisateur peut employer toutes les combinaisons de traduction.

L'intérêt d'un tel outil d'aide à la compréhension peut encore s'élargir si l'on prend en compte le fait qu'il est possible, même dans la version gratuite de base, de redéfinir un nombre illimité de mots « réservés » dont les définitions seront prioritaires sur celles proposées par Memodata®. Malgré la richesse des dictionnaires de la société conceptrice, cette possibilité revêt un intérêt non négligeable dans le cas de pages très pointues, ou celui de mots (noms propres, par exemple) pour lesquels le rédacteur a des compétences particulières ; on peut également envisager l'ajout d'informations grammaticales ou d'exemples d'illustration. L'extrême simplicité du principe de cet outil le rend ainsi indépendant de son contenu de départ (un dictionnaire multilingue), mais l'accroche profondément à son support (le document écrit hypertexte) en y apportant un enrichissement qui en multiplie la dimension.

4. L'environnement Alfalex

Le projet Alfalex de l'Université catholique de Louvain (<http://www.kuleuven.be/alfalex/>) s'inscrit dans la politique didactique de cette université qui « vise à stimuler l'apprentissage semi-autonome et le développement d'outils informatiques qui favorisent ce type d'apprentissage » (Verlinde, Selva & Binon, 2005). En fait, les recherches menées dans cette université sur la lexicographie pédagogique ont trouvé là un champ d'application lié à l'exploitation de ressources informatisées, ainsi qu'un enjeu qui n'est rien moins que la mise à niveau en français des étudiants néerlandophones qui doivent par nécessité suivre un cours en langue française. Cette mise à niveau est rendue nécessaire par les disparités observées parmi les étudiants dans la maîtrise de la langue française.

Ce projet s'articule autour de deux ressources étroitement liées : un corpus journalistique contemporain de plus de 75 millions de mots, provenant de plusieurs années des journaux *Le Monde* et *Le Soir*, et un dictionnaire général du français à but pédagogique, le DAFLES (Dictionnaire d'Apprentissage du Français Langue Etrangère ou Seconde) dans lequel sont décrits plus de 6600 mots et 14000 sens. La liaison entre ces deux ressources est non seulement faite lors de leur participation à l'environnement Alfalex, mais encore parce que les sens qui sont décrits dans le dictionnaire sont ceux qui ont été observés dans le corpus.

On peut dire ici que le processus dictionnaire « normal » a été respecté lors de la construction du Dafles ; en effet, ce dictionnaire a été conçu comme une base de données relationnelle, dont les informations ont été retenues en fonction de leur pertinence dans le projet pédagogique. Ainsi, le Dafles contient des groupes d'informations dont la structure permet de générer dans Alfalex des séries d'exercices, et qui concernent entre autres :

- la morphologie des mots : terminaisons des noms et adjectifs (pluriel, féminin), conjugaison des verbes, dérivation.
- la combinatoire des mots : les prépositions, les collocations.
- les champs sémantiques et les associations : synonymes et antonymes, schémas actanciels.
- les équivalents néerlandais.

Bien que le dictionnaire Dafles soit consultable en ligne comme n'importe quel équivalent, son originalité et son intérêt se trouvent dans son mode de diffusion via

l'environnement Alfalex. La construction des exercices se base en effet sur la liste des entrées du Dafles, et sur les données liées à chaque entrée, tandis que les phrases d'exemples sont directement puisées dans le corpus journalistique. Ces exercices sont de plus munis d'un coefficient de difficulté, dont nous signalerons l'application dans les paragraphes qui suivent.

Les exercices en ligne se présentent généralement sous la forme de phrases « à trous » que l'apprenant doit compléter ; en voici quelques exemples d'exercices disponibles (qui ne sont jamais les mêmes) :

Morphologie ; féminin des adjectifs :

L'opinion réagit, dit -elle, on l'a senti avec les grèves de décembre 1995 et avec le mouvement des chômeurs. "

{**public**}

(L'adjectif à employer est indiqué pour chaque phrase)

Dérivation :

créer

Une opportunité s'ouvrirait donc pour qu'une orientation nouvelle, de centaines de milliers d'emplois, puisse être mise en oeuvre.

Soumis à un marché trop saisonnier, les fabricants de jouets cherchent à convaincre durablement avec des jeux ou des matériaux authentiques.

Derrière ces préoccupations, il n'est cependant pas question de toucher à la loi de 1946, qui a permis la d'EDF.

Pour les jeunes femmes qui voient plutôt bergère que princesse, qui veulent une robe qui leur ressemble et qui soit plus de fête que strictement de mariée, Zelia des modèles à la carte. (*présent*)

Son nom arrive très largement en tête, sans doute à cause de sa double activité : il est à la fois un original et un pédagogue qui a mis l'accent sur l'acteur et sa formation.

(Tous les mots sont des dérivés du verbe placé en titre)

Prépositions :

L'Australie et la Nouvelle-Zélande pensent peu souffrir la crise asiatique

Synonymes :

exigence = r (Les professeurs ont remis une liste d'exigences à la direction de l'école.)

On remarque que la première lettre du synonyme demandé (ici *revendication*) est indiquée.

Collocations :

Pour autant, le maire ne compte pas suite au préalable formulé jeudi par Jacques Toubon et par Claude Goasguen, chef de file des élus UDF ralliés à l'ancien ministre de la justice.
{verbe support de **suite**, infinitif}

[\[modifier\]](#) [\[première lettre\]](#)

Dans le cas des collocations, l'apprenant à la possibilité de demander la première lettre du verbe support, mais cela lui réduit sa note. D'autre part, le lien « modifier » permet de changer la phrase, tout en gardant la même collocation.

Chaque série d'exercices s'accompagne d'un bouton d'envoi qui génère la page des résultats ; sur celle-ci, l'étudiant va savoir si ces réponses sont correctes ou non, pourra connaître les réponses correctes, et recommencer des exercices sur la ou les phrases qui lui ont posé problème. Etant donnée la taille du corpus associé à cet environnement d'apprentissage, les exercices sont très rarement constitués des mêmes phrases, ce qui met l'apprenant dans des situations variées ; à titre d'exemple, les exercices sur le genre des mots sont basés sur un lexique de 3700 unités.

L'intérêt de l'hébergement centralisé de ce système se situe ailleurs que dans la production pure et simple d'exercices aléatoires qui pourrait être réalisée « en local » sur le poste de l'apprenant par un environnement installé à partir d'un cédérom. Tout d'abord, les résultats de l'étudiant sont gardés en mémoire, et il est toujours possible de les consulter ; certes, cela aussi pourrait se faire localement, mais dans le cas d'Alfalex les résultats deviennent également accessibles au professeur qui suit l'étudiant.

De plus, l'intérêt de centraliser les réponses est de permettre par la suite des traitements statistiques sur les erreurs commises ; les erreurs les plus fréquentes ou systématiques sont ainsi repérées et peuvent donner lieu à la conception de petits aide-mémoires très ciblés insérés dans les cours de langues. En outre, ce repérage des fautes a une action rétro-active sur le classement des exercices selon leur difficulté ; en effet, chaque étudiant qui profite de cet environnement commence par passer un test qui situe son niveau et le palier de difficulté de ses premiers exercices ; lors de sa progression, il se verra proposer des phrases « classées » comme plus difficiles. Il va de soi que lors de l'élaboration de ce projet, les phrases ont dû être classées selon des avis d'enseignants ; le retour d'information donne un critère de classement beaucoup plus objectif.

Enfin, le retour d'information a permis aux concepteurs de faire une constatation intéressante en faveur du système et de la personnalisation de l'enseignement qu'il permet ; en effet, si les résultats globaux des exercices apparaissent corrélés aux résultats du test de niveau, il n'en va pas de même des résultats de chaque batterie d'exercices où cette corrélation disparaît complètement. Autrement dit, un étudiant qui obtient une bonne note au test initial peut parfaitement être d'un niveau faible sur des problèmes de conjugaison ; inversement, un mauvais résultat au test de départ peut s'accompagner par la suite de très bons résultats sur un aspect particulier de la langue.

Conclusion

A travers les trois exemples que nous avons détaillés, nous avons pu mesurer les avantages des dictionnaires en ligne évoqués dans le premier paragraphe de cet article. Dans tous ces cas, la gratuité de l'outil, l'accessibilité universelle, la simplicité d'utilisation fortement liée à l'hypertextualité, l'interactivité et l'évolution transparente – mais

conditionnée par les retours d'informations- assurent le succès des projets et les hissent au niveau des références de leur domaine, comme c'est le cas pour le dictionnaire des synonymes du Crisco. Toutefois, en raison de la dématérialisation des ressources présentes sur Internet, la crainte légitime de l'utilisateur concernant leur pérennité doit trouver une réponse, et celle-ci doit forcément provenir des institutions ; il faut en effet en ce domaine que les intérêts commerciaux soient clairement distingués de ceux de la communauté : la disparition du site du dictionnaire Hachette® de la francophonie, qui concurrençait probablement l'édition papier du même dictionnaire, est là pour nous le rappeler.

Enfin, Internet favorise un type de travail particulier grâce à la mise en réseau des ressources lexicales : il s'agit du travail collaboratif. Ce genre de ressource autorise les utilisateurs à apporter leur propre contribution à l'ouvrage ; l'exemple le plus connu est l'encyclopédie en ligne Wikipedia®, où chaque internaute peut déposer ou modifier un article. Dans un domaine aussi mouvant que celui de la langue, il est clair qu'un dictionnaire dont les principes ont été fixés au départ, ne peut pas fonctionner selon un mode collaboratif totalement « ouvert ». Néanmoins, si l'on applique un contrôle à l'admission des collaborateurs, il est clair que la confection d'un dictionnaire multilingue peut avoir lieu via ce principe. Il semble cependant que les projets de ce type ne soient pas légions ; on peut citer le projet Papillon (<http://www.papillon-dictionary.org/>) qui semble avoir beaucoup de difficultés à être mis en œuvre, et qui malgré son intérêt et les soutiens dont il bénéficie, reste beaucoup moins riche tant en vocabulaire qu'en nombre de langues que l'outil Alexandria® dont nous avons parlé.

Pour conclure, nous pouvons dire avec certitude que la mise en ligne des dictionnaires de langue a clairement fait apparaître la nécessité de concevoir ces ressources en fonction du nouveau média et de ses caractéristiques. Le passage des ouvrages papier aux cédéroms avait déjà mis en évidence l'absurdité de la simple « informatisation » des pages sans tirer parti des atouts de l'informatique (notamment en matière de saisie). L'étape qui conduit aux versions en ligne aboutit au même constat : un dictionnaire en ligne ne peut pas être pensé comme son équivalent sur cédérom ; il lui faut profiter de la liaison avec ses utilisateurs, mais aussi comprendre qu'Internet est un milieu propice à la propagation rapide, que ce soit par création de liens (le dictionnaire des synonymes du Crisco a ainsi été révélé), ou bien grâce à l'insertion possible de son interface d'accès par les créateurs de pages eux-mêmes (c'est la méthode choisie pour diffuser Alexandria®).

Bibliographie

- DUTOIT Dominique (2005), « Intratexte : quand les textes aident leur propre lecture », *Technologies internationales*, 116, 37-40.
- KAHLMANN André (1975), *Traitement automatique d'un dictionnaire de synonymes*, Stockholm, Université de Stockholm.
- MANGUIN, Jean-Luc (2005) : « La dictionnaire Internet : l'exemple du dictionnaire des synonymes du CRISCO », *CORELA - Cognition, Représentation, Langage*, Numéro spécial.
- POUX Sabine (1997). « Modélisation et traitement informatique de la synonymie ». *Linguisticae Investigationes*, XXI (1), Amsterdam, John Benjamins.
- POUX Sabine & VICTORRI Bernard (1998), « Construction d'espaces sémantiques à l'aide de dictionnaires de synonymes », *Traitement Automatique des Langues*, Vol 39/1, pp. 161-182.
- PRUVOST, Jean (2002) : *Les dictionnaires de langue française*. Paris, PUF, Collection « Que sais-je ? ».
- QUEMADA, Bernard (1968), *Les Dictionnaires du français moderne (1539-1863). Étude sur leur histoire, leurs types et leurs méthodes*, Paris, Didier.

VERLINDE S., SELVA T. & BINON J. (à paraître 2005), « Dictionnaires électroniques et environnement d'apprentissage du lexique ». *Revue française de linguistique appliquée*, Vol X, 2.

VICTORRI B. & FUCHS C. (1996), *La Polysémie : construction dynamique du sens*, Paris, Hermès.