

# **Dictionnaires et Encyclopédies du Français : la Version Électronique**

**Christine Jacquet-Pfau**

Maître de conférences

Collège de France (Paris)

11, place Marcelin-Berthelot

75231 PARIS cedex 05

FRANCE

[ch.jacquet-pfau@college-de-france.fr](mailto:ch.jacquet-pfau@college-de-france.fr)

## **Résumé**

Le développement des technologies dites nouvelles a permis de rendre accessibles toutes sortes de corpus textuels selon des approches nouvelles permises par le développement d'outils d'interrogation qui modifient le rapport du lecteur-utilisateur à l'œuvre, ouvrent des perspectives à exploiter et font naître des exigences particulières. Une petite typologie des dictionnaires sur CD-ROM et DVD-ROM contribuera à montrer le foisonnement de cette nouvelle édition. Les dictionnaires, anciens, récents ou futurs, ont trouvé une nouvelle raison d'être, les outils de recherche qui leur ont été associés permettant des parcours variés, fiables, visant à l'exhaustivité. Ils se trouvent ainsi à nouveau au cœur de la mémoire : mémoire du lecteur, mémoire du document. Mais leur informatisation est encore très perfectible.

## **1. Encyclopédies et dictionnaires : une renaissance**

Le dictionnaire est l'un des corpus les plus consultés, les plus utilisés par un public extrêmement diversifié. Utilisé pour rechercher une information, pour valider ou infirmer une intuition, pour un détail précis ou une connaissance plus large, pour une épreuve imposée, dans le cadre scolaire ou professionnel, ou pour une activité ludique (mots croisés, anagrammes...), le dictionnaire sur support papier est considéré comme la référence par excellence, le document dans lequel on trouvera une réponse à tout... ou presque. Il n'y a pas si longtemps, le passage obligatoire était le volume, les pages tournées, le mot-clé à repérer pour avoir accès à la notice du dictionnaire où découvrir, peut-être, ce que l'on cherchait. Sans mot introducteur, aucune chance de trouver l'éventuelle réponse, à moins d'y passer un temps considérable ou d'avoir beaucoup de chance.

Les technologies informatiques ont révolutionné cette approche du dictionnaire. Ou plutôt, elles ont permis, parallèlement au parcours purement livresque et ordonné du corpus sur papier, un parcours plus structuré, plus logique, plus systématique. D'abord le cédérom puis aujourd'hui le DVD-ROM<sup>1</sup> ont permis de stocker une grande masse d'informations accessibles par l'utilisateur très rapidement et d'y associer des outils de lecture et de recherche qui ouvrent la voie à des parcours diversifiés, plus ou moins élaborés, selon la structuration du dictionnaire et les outils de recherche associés.

Nous n'évoquerons pas dans cet article les 'dictionnaires machine', ainsi dénommés car conçus avant tout pour être interprétés par un langage machine, et donc ininterprétables par une personne qui n'aurait pas connaissance du code qui a été utilisé pour les représenter.

Nous ne nous intéresserons ici qu'aux dictionnaires et encyclopédies lisibles sur écran au sens de compréhensibles par ce même lecteur qui consulte les dictionnaires et encyclopédies sur support papier. De tels dictionnaires présentent trois traits distinctifs principaux :

- ils sont commercialisés ;
- dans l'état actuel de l'art, il s'agit dans la plupart des cas d'une " copie " (rétroconversion) de la version papier du dictionnaire ou de l'encyclopédie;
- la lecture est aidée par l'intégration d'outils de recherche automatique.

Pour un même dictionnaire, les conditions de consultation, et donc de lecture, en sont profondément modifiées.

Nous ferons enfin une précision terminologique préalable, reprenant la distinction établie par Gaston Gross : " il y a en effet une grande différence entre, d'une part, un dictionnaire électronique dont l'objectif est la génération automatique de phrases ou la reconnaissance automatique de phrases, et, d'autre part, la version électronique d'un dictionnaire papier "2. Lorsque nous parlerons de dictionnaires ou d'encyclopédies informatisés, c'est bien des versions électroniques de documents papier dont il sera question.

## **2. Petit essai de typologie des dictionnaires et encyclopédies sur cédérom**

L'édition, en 1994, du *Grand Robert électronique*, a constitué le point de départ de l'informatisation des dictionnaires. Mais il faudra attendre environ encore quatre ans pour que s'affirme et se développe l'informatisation des corpus dictionnaires. Aujourd'hui la commercialisation de dictionnaires et encyclopédies informatisés est suffisamment vaste et diversifiée pour nous permettre de proposer une typologie, détaillée dans (Jacquet-Pfau, 2002). Cette classification, qui ne doit pas être linéaire mais permettre de croiser des critères, distingue :

§ les dictionnaires de langue (Les Dictionnaires des XVI<sup>e</sup> et XVII<sup>e</sup> siècles, Les Dictionnaires de l'Académie française, le Grand Dictionnaire universel du XIX<sup>e</sup> siècle de Pierre Larousse, Le Petit Robert, Le Petit Larousse illustré...);

§ les encyclopédies, qui incluent elles-mêmes des dictionnaires de définitions pouvant être appelés à tout moment de la consultation d'un article de l'encyclopédie grâce à un lien prévu entre chaque mot ou certains mots du corpus et le mot correspondant dans le dictionnaire (*L'Encyclopédie de Diderot et d'Alembert ou Dictionnaire raisonné des sciences, des arts et des métiers*; *l'Encyclopædia universalis*, *Encarta*...);

§ les dictionnaires restreints à un type de données linguistiques ou philologiques : dictionnaires de synonymes, dictionnaires de citations, dictionnaires de difficultés, thésaurus... (*Nouveau dictionnaire des difficultés du français moderne* de Hanse et Blampain édité par De Boeck-Duculot (Bruxelles));

§ les dictionnaires d'apprentissage (Dictionnaire du français usuel : 15 000 mots utiles en 442 articles de Picoche et Rolland);

§ les dictionnaires de langue de domaines langagiers géographiquement ou sémantiquement sélectifs (*Dictionnaire suisse romand : Particularités lexicales du français contemporain*, sous la direction de Pierre Knecht, conçu et réalisé par André Thibault; *Tre sor de noms d'oiseaux :  tymologie du lexique europ en par les paradigmes* (A

*Thesaurus of Bird Names Etymology of European Lexis Through Paradigms*) de Michel Desfayes);

§ les dictionnaires et bases de données terminologiques multilingues, constituent la marge extrême du domaine concerné, au seuil du domaine des dictionnaires bilingues ou multilingues (*LeXpro CD-Databank*, édité par la Maison du Dictionnaire ou encore le *CD-ROM de 20 dictionnaires multilingues de 120 000 termes avec définitions*, édité par le Conseil International de la Langue Française);

§ les dictionnaires bilingues;

§ les bases de données textuelles (DISCOTEXT 1 : Textes littéraires français 1827-1923; PORTEXT...).

Cette typologie, qui prend en compte le contenu même du dictionnaire, gagnera à être enrichie par d'autres critères. Nous n'en retiendrons ici que deux, à titre d'exemples :

§ la datation de l'état de la langue

La structure et le contenu métalexico-logique des dictionnaires n'ont pas répondu, jusqu'au développement récent de la lexicographie, à des exigences de cohérence, de hiérarchisation, voire de formalisme qui rendent aisée l'informatisation, c'est-à-dire le repérage systématique des éléments structurants une notice. Un dictionnaire comme le *Littré*, peu structuré, rend ardue la lecture à l'écran et ne permet guère plus qu'une recherche en plein texte.

§ la catégorie du public auquel s'adresse le dictionnaire

Le public très jeune (élèves de maternelle et de primaire) représente une part non négligeable des utilisateurs. Les dictionnaires informatisés qui leur sont destinés sont encore relativement peu nombreux, mais leur qualité suffit à montrer combien il peut être profitable, d'un point de vue didactique, de recourir au support multimédia (*Le Robert Junior*, *l'Encyclopédie multimédia Junior* éditée par Larousse multimédia, *Mon premier dictionnaire super génial*, édité par Nathan). Nous pouvons également prendre en compte le degré de spécialisation du public recherché : la version informatisée des *Dictionnaires de l'Académie* éditée par les éditions Champion électronique s'adresse davantage à un public spécialisé que celle des éditions Redon. Les moyens mis en oeuvre sont proportionnels aux performances du support et des outils de recherche, mais aussi au coût.

### **3. Du papier à l'écran**

Le passage du papier à l'écran résulte de l'informatisation d'un document. Ce processus nécessite le recours à un certain nombre d'étapes, qui seront elles-mêmes ultérieurement enrichies d'outils de lecture et de recherche. Rappelons rapidement ces principales phases :

- la saisie des pages du dictionnaire ou de l'encyclopédie papier suivie de relectures pour correction
- le balisage manuel, automatique ou plus généralement mixte
- l'élaboration d'outils de recherche.

La première étape se fait selon des modalités différentes :

- saisie manuelle, caractère par caractère (ce mode de saisie semble encore prévaloir)
- scannérisation et transformation en mode texte par OCR<sup>3</sup> et relectures minutieuses

- scannérisation des images
- scannérisation du texte pour une utilisation en mode image (la page est saisie dans son ensemble, comme une seule image).

L'évolution plus récente des technologies permet maintenant de répondre à une autre question de grande importance pour les dictionnaires et encyclopédies plus anciens : comment rendre compte de l'édition originale le plus fidèlement possible sans perdre le bénéfice des outils de requête. Le développement du mode texte-image, alliant la précision de l'interrogation grâce au mode texte à la fidélité au document papier de référence grâce au mode image, permet de traiter des ouvrages volumineux, qui ne pourraient pas faire l'objet d'une saisie manuelle, pour des raisons évidentes. Ainsi l'utilisateur a-t-il le choix de consulter seulement l'édition imprimée reproduite telle quelle (fac-similé en mode image) ou la version texte, conversion de la première par un logiciel d'OCR. Toutefois, le nombre relativement important de scories contenues dans le *Grand Dictionnaire universel du XIX<sup>e</sup> siècle* de Larousse éditée par les éditions Redon fait entrevoir la qualité peu fiable du traitement OCR, qui, s'il n'est pas accompagné d'un taux de correction manuel élevé, donne ce genre de résultat. La possibilité de disposer des deux versions sur un même écran permet à l'utilisateur de recourir à l'une ou l'autre, de contrôler les données et d'avoir sous les yeux l'original tout en disposant de toutes les possibilités de recherche et d'interactivité du mode texte. La consultation du mode image, telle qu'elle est développée par les éditions Champion électronique (*Dictionnaires des XVI<sup>e</sup> et XVII<sup>e</sup> siècles*, *Dictionnaires de l'Académie française*, *Grand Dictionnaire universel du XIX<sup>e</sup> siècle* de Larousse notamment), est instantanée et relativement souple, surlignant l'occurrence recherchée, autorisant la sélection, la copie, l'impression. Les éditions Champion électronique et Redon ont ainsi réussi à offrir à l'utilisateur deux types de documents bien différents, tant par leurs caractéristiques de forme et de contenu que par leur offre commerciale. Les réalisations des éditions Redon restent accessibles à un public plus large, alors que celles des éditions Honoré Champion, à la fois par leur coût et leur volonté de ne faire aucune concession quant à la lisibilité du document, s'adressent presque exclusivement à des institutions ou à des chercheurs spécialisés. L'intérêt de ce double affichage apparaît immédiatement lorsqu'on évoque les dictionnaires et encyclopédies plus éloignés dans le temps.

L'informatisation des corpus dictionnaires, en-dehors du choix de reproduire la photographie numérique du texte, pose la question du respect de l'édition originale, à savoir l'édition papier à partir de laquelle est faite la rétroconversion. Est-il possible de modifier le document de référence ? Si oui, dans quelles limites ? La réponse est relativement simple : si le texte et la structure, macrostructure et microstructure, doivent être respectés (respect de l'ouvrage original, droits d'auteur...), certains aménagements typographiques peuvent être concédés et la mise en page peut elle aussi être l'objet d'aménagements. La mise en œuvre de telles possibilités reste toutefois soumise à des contraintes éditoriales de coût. Elles ont d'autant plus de chance d'être exploitées qu'elles sont automatisables, c'est-à-dire facilement et exhaustivement réparables.

Plus le dictionnaire est récent, plus on observe la mise en œuvre de tels aménagements. Ils sont essentiellement liés à l'adaptation de la mise en page à la mise en écran (utilisation de la couleur comme repérage métalinguistique, retours à la ligne pour plus de lisibilité...). Ces adaptations sont rendues possibles par la libération de certains

contraintes éditoriales exercées sur la production du livre, telles que le coût de l'impression couleur ou le nombre de pages autorisé par les contraintes financières.

#### **4. Le dictionnaire à l'épreuve de l'informatisation**

Toute informatisation d'un corpus textuel soumet le texte traité à l'épreuve de plusieurs tests qui, par la possibilité de faire apparaître dans un espace temporel et spatial très réduit des informations qui, à la lecture du support papier apparaissent très éparées, font brusquement surgir lacunes et incohérences. Les critères de ces tests se trouvent définis d'une part par les exigences a priori de l'utilisateur, d'autre part par l'exigence de qualité et d'exhaustivité que va susciter naître inévitablement le mode d'interrogation associé à la lecture. Nous avons choisi quelques mots-clés pour donner accès aux principales spécificités des versions informatisées des dictionnaires et encyclopédies.

##### **4.1. Accessibilité**

Parmi les exigences les plus fréquentes formulées a priori par l'utilisateur figurent l'accessibilité et la maniabilité. L'édition électronique est d'autant plus appréciée que le support papier est difficilement accessible (disponible seulement dans quelques bibliothèques spécialisées) et peu maniable (les éditions successives du *Dictionnaire de l'Académie française*, les dix-sept volumes du *Grand Dictionnaire universel* de Pierre Larousse ou la vingtaine de volumes de l'*Encyclopædia universalis*, réactualisée régulièrement par l'ajout de volumes supplémentaires). La rapidité de l'accès aux informations recherchées, la simplicité de l'utilisation sont également des critères qui rendent plus performante l'édition électronique. Mais cette exigence de rapidité peut avoir son revers : la recherche d'occurrences d'une fréquence trop élevée conduit parfois à l'échec de l'accès aux informations, le système ne pouvant gérer les informations que jusqu'à un certain seuil.

La lisibilité du texte à l'écran et de l'interface est l'un des avantages de ce type d'outil. La vision synoptique qu'offre le support papier est compensée par une fonction "zoom" qui permet de modifier la taille des caractères et, dans certains dictionnaires comme *Le Petit Robert*, par une fonction qui permet de sélectionner certains types d'informations à afficher à l'écran (plan de l'article, étymologie, citations...), regroupant sur un seul écran des informations affichées successivement sur deux, voir trois écrans.

##### **4.2. Cohérence**

A côté de ces exigences de confort visuel, se développent des exigences de qualité que l'on peut percevoir comme conséquences de l'informatisation des dictionnaires et des encyclopédies. Le mode hypertexte permet par son système de renvoi d'un mot à l'autre, d'une article à l'autre, de saisir très rapidement des informations redondantes, circulaires, voire incohérentes. Ce qui n'apparaît qu'au hasard d'une recherche sur le support papier est détecté, par le rétrécissement de l'espace et du temps, très rapidement sur un support informatisé. Ainsi une définition circulaire, une absence d'information, deux informations contradictoires apparaîtront-elles immédiatement.

L'utilisation du métalangage est soumise également à une plus grande rigueur, essentiellement sur deux plans. Le métalangage est une source importante pour élaborer une

requête. Il peut être en effet utilisé par le concepteur de l'outil dans le cadre du balisage du dictionnaire ou de l'encyclopédie, pour repérer des informations grammaticales ou étymologiques, des domaines, la nature de l'information (exemples, citations, homonymes...), mais il peut – et c'est actuellement plus souvent le cas – être en soi un outil de recherche pour l'utilisateur : ainsi pourront être cherchées dans le dictionnaire ou l'encyclopédie toutes les variantes orthographiques, toutes les féminisations des noms de métier ou encore tous les mots signalés comme appartenant au domaine de la botanique, si ces informations sont signalées dans le texte par une marque textuelle et si cette marque correspond à une notation constante, sans aucune variation, dans tout le texte.

### 4.3. Mémoire

Le support informatique, cédérom et mieux DVD-ROM pour éviter la manipulation de plusieurs supports de lecture, doit offrir un mode de consultation aisé. Lieu où la mémoire stockée est parfaitement adaptée à un corpus du type encyclopédie de la taille du *Grand Dictionnaire universel* de Pierre Larousse (dix-sept volumes de 24036 pages au total), il permet également, dans la plupart des cas, de stocker une mémoire interactive, celle qui gardera les parcours du lecteur et les annotations dont il souhaitera commenter les articles consultés. Outil de consultation et de recherche, le dictionnaire devient par excellence l'un de ces " lieux de mémoire " dont parle Pierre Nora<sup>4</sup>.

4.3.1. *Mémoire du lecteur.* La consultation d'un document informatisé, rapide, facilité par de nombreux outils de lecture et de recherche, ne saurait échapper aux préoccupations liées à la mémoire. Bien au contraire, la mémoire est au cœur de la thématique. Mémoire de l'utilisateur, mémoire du document, mémoire des sources écrites, mémoire des documents. Plus que jamais sollicité par une lecture où le renvoi à une autre lecture est facile et rapide, le lecteur va être confronté à la difficulté de garder le fil du déambulement dans le labyrinthe d'une lecture hypertextuelle. La plupart des dictionnaires/encyclopédies informatisés ont incorporé à leurs outils des fonctions de retour aux étapes successives du parcours de lecture depuis l'ouverture d'une session de travail et des outils d'annotation, permettant de conserver en mémoire, de consultation en consultation, des commentaires personnels, voire de les classer et d'en constituer des documents à part entière dans les encyclopédies techniquement les plus évoluées.

4.3.2. *Mémoire du document.* Mémoire du lecteur, mais aussi mémoire du document. Si la conservation du livre est bien maîtrisée, il n'en va pas de même aujourd'hui du support électronique, Cédérom ou DVD-ROM. Supports encore fragiles, ils exigent de conserver plusieurs sauvegarde, mais surtout l'évolution des techniques informatiques contraint à de réguliers changements de supports d'accès aux données. Ainsi *DISCOTEXT 1*, corpus de trois cents œuvres littéraires (1827-1923) – soit un total de trente-cinq millions d'occurrences –, premier et seul cédérom conçu à partir de la base Frantext du CNRS, n'est accessible que dans un environnement DOS, faute d'avoir été réactualisé par l'éditeur<sup>5</sup>. Le même problème se posera de façon cruciale pour les différentes versions, toutes uniques, témoins de l'état de la langue à un moment donné, de tous ces dictionnaires. Si le transfert sur les nouveaux supports de lecture n'est pas régulièrement assuré, nous n'aurons plus, dans un délai très bref, le moyen d'accéder à cette bibliothèque électronique. La question posée

par un titre le *Monde interactif* du 2 juin 1999 : " Comment conserver la mémoire numérique? " semble toujours d'actualité.

Il faut ici attirer l'attention sur la mémoire qui s'instaure entre les deux types de support, papier et informatisé : le rythme des rééditions des versions informatiques ne correspondent pas toujours à celui des versions papier (ainsi *Le Petit Robert* papier connaît une réédition annuelle alors que le *cédérom*, édité en 1996, n'a fait l'objet d'une première mise à jour qu'en 2001, *Le Nouveau Petit Robert électronique* ; *l'Encyclopædia universalis* propose une mise à jour annuelle).

4.3.3. *Mémoire bibliographique*. Enfin, comment conserver sous forme écrite la mémoire de l'histoire de chaque dictionnaire informatisé ? Autrement dit, comment construire la mémoire bibliographique ? La tendance est de considérer le numérique dans sa phase la plus récente, la plus nouvelle. Or il est important de veiller à conserver et dater les différentes éditions d'un dictionnaire sur *cédérom*, comme celles d'un dictionnaire papier. Les éditions informatisées participent à la mémoire de la langue et de la technique, aussi. Il est important de définir des normes de répertoriage et d'archivage de manière à en assurer la pérennisation.

#### 4.4. Recherche / Exhaustivité

Nous terminerons ce synopsis des valorisations liées au passage du papier à l'écran en abordant le domaine spécifique de la recherche d'informations. Il s'agit là de l'aspect innovant des corpus dictionnaires informatisés, puisque tout corpus de ce genre est associé à des outils de requête. Toutefois ces outils sont de performance très différente d'un dictionnaire à l'autre, et cela, essentiellement pour deux raisons.

La première, comme nous l'avons déjà évoqué ci-dessus, est liée au degré de structuration plus ou moins hiérarchisée du texte d'origine, à la richesse et la cohérence du métalangage, bref à tous les repères qui peuvent être utilisés pour un balisage automatique ou tout au moins semi-automatique du texte. Nous donnerons ici deux exemples représentatifs, celui du *Littré*, où l'article ne présente pour ainsi dire aucune structure, et celui du *Petit Robert*, dans lequel, au contraire, les repères indiquant les différents champs (sous-entrées, niveaux sémantiques, exemples, citations...) sont nombreux et cohérents, de même que le métalangage, facilitant ainsi le travail de balisage.

La seconde raison est dépendante des choix faits par l'éditeur, choix linguistiques et lexicographiques, que les contraintes technologiques obligent encore à choisir entre la qualité, le coût, la rapidité, entre autres. S'il est indéniable que l'utilisateur exige l'exhaustivité, il exige également la rapidité. Et veut un outil accessible financièrement.

C'est en jouant de tous ces critères souvent difficiles à concilier, parfois contradictoires, que l'éditeur, en fonction du public qu'il pense conquérir, programme le produit.<sup>6</sup> L'exhaustivité suppose que la recherche des informations repose sur des critères rigoureux et précis, ce qui, même dans les meilleurs dictionnaires, échappe parfois, volontairement (pour des raisons de moyens qui seraient peut-être trop lourds à implémenter) ou non, même aux meilleurs dictionnaires. Les outils de requête visibles par l'utilisateur (éléments *jokers*, opérateurs booléens, sélections des champs où rechercher l'information, transcription phonétique...) ne suffisent pas : il est nécessaire que le traitement du contenu soit également performant. Pour illustrer ce propos qu'il nous suffise de décrire un seul

exemple. Il concerne le traitement des caractères non alphabétiques, que nous n'avons pas toujours trouvés reconnus en tant que tels, par exemple dans *Le Petit Robert*, qui présente pourtant par ailleurs des outils performants. Les caractères diacrités sont neutralisés (si l'on souhaite extraire tous les adjectifs terminés par *-é*, le résultat du tri donne la liste de tous les adjectifs terminés par un *-e* ou un *-é*, soit plus de 26 000 entrées ! : beaucoup de bruit pour une information relativement facile à trier !). Autre exemple : les traits d'union ne sont pas considérés comme des caractères et font échouer la recherche : si l'on recherche *cd-rom* (les minuscules et les capitales sont neutralisées), qui existe bien sous cette forme dans la nomenclature, la réponse sera " Il n'y a pas de réponse pour cette recherche "; en revanche, si l'on fait la même requête en texte intégral, les réponses seront exactes... Ce qui signifie que les critères de tri utilisés ne sont pas les mêmes dans ces deux cas. Ce qui implique également – mais n'est-ce pas là le propre de toute recherche ? – que l'on doit toujours être prudent sur l'exhaustivité des données recueillies.

#### 4.5. Universalité

Cette même exigence d'exhaustivité, renforcée par le besoin d'actualiser toujours plus vite le savoir, a trouvé dans l'informatisation des encyclopédies un moyen idéal pour tenter d'approcher ce rêve d'une connaissance toujours plus grande. Il a suffi aux concepteurs d'encyclopédies sur support informatique de proposer des liens avec des bases de données Internet (*l'Encyclopædia Universalis*, en est, dès 1998, le premier exemple). Ainsi des données " en ligne " peuvent-elles compléter des données " hors ligne " <sup>7</sup>, qui constituent des compléments d'information et des mises à jour destinés à combler le temps passé d'une édition informatisée à l'autre, déjà très court (un an pour *l'Encyclopædia Universalis*).

#### Notes

1. Convendrait-il d'harmoniser la graphie de *DVD-ROM* (Digital Versatile Disc - Read Only Memory) avec la graphie recommandée par l'Académie pour *CD-ROM* (Compact Disc - Read Only Memory), *céderom*, et de préférer alors écrire *dévédérom* ? *Le Petit Robert*, dans son édition de 2000, propose *céderom* comme variante de *CD-ROM*, mais donne pour seule graphie de *DVD-ROM* le sigle. La même source fait mention de la recommandation officielle de *disque optique compact* abrégée en *doc*, mais aucune pour *DVD-ROM*, alors que le terme est attesté depuis 1996.
2. Gaston Gross, " La grammaire, les dictionnaires et l'informatique ", *Les dictionnaires de langue française et l'informatique. Actes du Colloque La Journée des dictionnaires 1995, Cergy-Pontoise, Centre de Recherche Texte/Histoire*, juin 1997, pp. 55-64.
3. Optic Character Recognition. Technique utilisée pour convertir des fichiers numérisés (obtenus par scanner) en fichiers ASCII.
4. Pierre Nora, "Le Dictionnaire de pédagogie de Ferdinand Buisson, cathédrale de l'école primaire" , *Les lieux de mémoire. I. La République*, 1984, pp. 353-378.
5. Éditeur : Hachette, développeur : Van Dijk, 1992.
6. Voir Christine Jacquet-Pfau (1998).
7. Alain Vuillemin, "L'édition électronique", <http://www.uottawa.ca/academic/arts/astrolabe/articles/art0026.htm>.



## Références

- Jacquet-Pfau, Ch.** "Bibliographie des dictionnaires du français sur cédérom", Musée virtuel des dictionnaires,  
[http://www.u-cergy.fr/dictionnaires/biblio/mvd\\_bibliographie.html](http://www.u-cergy.fr/dictionnaires/biblio/mvd_bibliographie.html)
- Jacquet-Pfau, Ch.** "A propos des dictionnaires sur cédérom", Musée virtuel des dictionnaires, [http://www.u-cergy.fr/dictionnaires/biblio/mvd\\_bibliographie.html](http://www.u-cergy.fr/dictionnaires/biblio/mvd_bibliographie.html)
- Jacquet-Pfau, Ch.** 2002. "Les dictionnaires du français sur cédérom", *International Journal of Lexicography*, Oxford University Press, mars 2002, pp. 98-104.
- Jacquet-Pfau, Ch.** 2000. "Les dictionnaires de langue française sur cédérom : Analyse comparative", in *Cahiers du CIEP : Dictionnaires électroniques et télématiques : nouveaux rapports à la langue, nouveaux accès aux savoirs, Actes du séminaire du Centre International d'Etudes Pédagogiques (CIEP), Sèvres, 18 février 1999*, Paris : Didier, pp. 43-62.
- Jacquet-Pfau, Ch.** 1998 (nov.). "Une autre lecture du dictionnaire de langue : le CD-ROM", *La Tribune Internationale des Langues Vivantes*, n° 24, pp. 63-71.
- Pruvost J.** (2000). *Dictionnaires et nouvelles technologies*, Paris : Presses Universitaires de France, Coll. "Ecritures électroniques".
- Pruvost J.** (2003). *Dizionario e nuova tecnologia*. Traduzione, adattamento e cura di Maria Rosaria Ansalone, Naples : Liguori Editore, La Collana del Re Sole 2.
- Rey Ch.** (2003). *Dictionnaires électroniques, dictionnaires informatisés ou dictionnaires-machines?* [bibliographie analytique], Mémoire de DEA sous la direction de André Valli, Université de Provence Aix-Marseille I (Lettres et Sciences Humaines) [disponible à l'adresse suivante : [www.up.univ-mrs.fr/lettres.html](http://www.up.univ-mrs.fr/lettres.html)]

## Bibliographie sélective des dictionnaires du français sur cédérom

Dictionnaire de l'Académie française - édition de 1694, Champion électronique, 1998.- CD-ROM PC/MAC.

Le Dictionnaire de l'Académie française : La série complète des huit éditions depuis 1694 présentées par Isabelle Turcan, Redon, 2000.- CD-ROM PC.

Dictionnaires des XVI<sup>e</sup> et XVII<sup>e</sup> siècles, Champion électronique, 1998.- CD-ROM PC, version 1.0.

- § Robert Estienne, Dictionnaire françois-latin, Paris, 1549
- § Jean Nicot, Thresor de la langue françoise, Paris, 1606
- § Randle Cotgrave, A dictionarie of the French and English tongues, London, 1611
- § Gilles Ménage, Les origines de la langue françoise, Paris, 1650
- § Pierre Richelet, Dictionnaire françois, Genève, 1680
- § Antoine Furetière, Essais d'un dictionnaire universel, Amsterdam, 1687
- § Antoine Furetière, Dictionnaire universel, La Haye et Rotterdam, 1690
- § Gilles Ménage, Dictionnaire étymologique de la langue française, Paris, 1694
- § Dictionnaire de l'Académie française, Paris, 1694
- § Thomas Corneille, Le Dictionnaire des Arts et des Sciences, Paris, 1694.

Les Dictionnaires de l'Académie française (XVII<sup>e</sup> et XVIII<sup>e</sup> siècles), Champion électronique, 2000.- DVD-ROM PC.

- § Dictionnaire de l'Académie françoise (de A à Neuf), Paris, 1687
- § Grand Dictionnaire de l'Académie françoise (de A à Aversion), Francfort, 1687
- § Grand Dictionnaire de l'Académie françoise (de A à Confiture), Francfort, 1687
- § Dictionnaire de l'Académie françoise dédié au Roy, Paris, 1694
- § Nouveau Dictionnaire de l'Académie françoise dédié au Roy, Paris, 1718
- § Dictionnaire de l'Académie françoise, Paris, 1740
- § Dictionnaire de l'Académie françoise, Paris, 1762

§ Dictionnaire de l'Académie française, Paris, 1798.

Pierre Larousse, Grand Dictionnaire Universel du XIX<sup>e</sup> siècle, Champion électronique, 2001 [réimpression de l'édition de Paris, 1866-1890, en mode texte-image].- CD-ROM PC/MAC, version monoposte ou version multiposte.

Le Littré, Dictionnaire de la langue française classique, Marsanne : Redon, 1999.- CD-ROM PC, version 1.3.

Grand Atelier historique de la langue française, Marsanne : Redon, 2003.- CD-ROM PC.

§ La Curne de Saint-Palaye, Dictionnaire historique, édition Favre, 1876

§ Furetière, Dictionnaire, Rotterdam, édition Leers, 1690

§ Dictionnaire de l'Académie française, édition de 1762, 4<sup>ème</sup> édition

§ Dictionnaire Le Littré, édition de 1872 et du supplément de 1876

§ Antoine Oudin, Curiosités françaises, 1640

§ Voltaire, Dictionnaire philosophique, édition de Kehl, 1765, et compléments

§ F. Guizot. Dictionnaire universel des synonymes de la langue française, 1822

§ Jean Nicot, Le Thresor de la langve francoyse [...], 1606

§ Pierre Richelet, Dictionnaire français contenant les mots et les choses, 1680

§ Gilles Ménage, Dictionnaire étymologique, 1694

§ Thomas Corneille, Dictionnaire des arts et des sciences, 1694

§ Dictionnaire universel françois et latin de Trévoux (1743-1752)

§ Abbé J.F. Féraud, Dictionnaire critique

§ Abbé J.F. Féraud, Dictionnaire grammatical portatif de la langue française.

Le Grand Robert électronique, Le Robert-Havas interactive, 1994.- CD-ROM PC/MAC.

Le Petit Larousse 2001, Larousse Multimedia-Liris interactive, 2000.- 3 CD-ROM PC; dernière édition : Le Petit Larousse 2004, CD-ROM PC, Larousse/ VUEF, 2003.

Le Petit Robert. Dictionnaire analogique et alphabétique de la langue française : Nouvelle éd. du Petit Robert de Paul Robert, texte remanié et amplifié sous la direction de Josette Rey-Debove et Alain Rey, Le Robert - Liris interactive, 1996.- CD-ROM PC/MAC, version 1.2.

Le Petit Robert. Dictionnaire de la langue française. Nouvelle édition, 2001, Le Robert/VUEF, CD-ROM PC/MAC, version 2.0.

Le Robert Junior, Dictionnaires Le Robert - Havas interactive, 1998.- CD-ROM PC.

Encyclopédie multimédia Junior, Larousse multimédia, 1999.- CD-ROM PC/MAC ou DVD-ROM.

L=Encyclopédie de Diderot et d'Alembert ou Dictionnaire raisonné des sciences, des arts et des métiers, Redon, 1999.- CD-ROM PC, version 1.01.

DISCOTEXT 1 : Textes littéraires français 1827-1923, CNRS - Hachette supérieur - Van Dijk, 1992.- CD-ROM PC (DOS).

PORTEXT, Réalisation technique Etienne Brunet, sous la direction de Carlos Maciel, INALF-CNRS, JNICT-CNPq \* Bases, corpus et langage +, Faculté des Lettres, Nice, 1999.- CD-ROM PC pour la version française, MAC pour la version portugaise.