

# **Méthodologie liée à l'Utilisation de Grandes Ressources Linguistiques dans le Cadre de l'Apprentissage : le cas du TLFi en Français au Cycle 3**

**Chrysta Péliissier, Christiane Jadelot, Jean-Marie Pierrel**

ATILF (Analyse et Traitement Informatique de la Langue Française)

UMR 7718 – CNRS & Université Nancy 2

chrysta.pelissier@atilf.fr, christiane.jadelot@atilf.fr, jean-marie.pierrel@atilf.fr

## **Résumé**

Dans cette communication, nous présentons les trois problématiques qui semblent se poser dans le cadre de l'exploitation des grandes ressources linguistiques pour l'apprentissage : l'identification des connaissances liées à l'apprentissage, la caractérisation des ressources linguistiques et la détermination des activités caractérisées pouvant être proposées. Pour illustrer chacune de ces problématiques, nous prenons le cas du Trésor de la Langue Française (TLFi), dictionnaire de langue, développé au sein de l'ATILF, CNRS & Université de Nancy 2, et de l'apprentissage de la lecture au cycle 3 (CE et CM) du primaire.

Plus particulièrement, nous présentons d'abord les savoirs qu'un enfant au cycle 3 doit acquérir dans le cadre de son apprentissage de la lecture. Ensuite, nous décrivons les ressources de manière à faire apparaître les possibilités d'utilisation pédagogique, puis, nous proposons différentes activités pour l'enseignant et pour l'apprenant qui mettent en œuvre les connaissances liées à l'apprentissage et les ressources linguistiques telles qu'elles sont aujourd'hui disponibles. Certaines de ces activités ont fait l'objet d'une expérimentation dont nous présenterons les résultats.

## **1. Les Dictionnaires dans l'Apprentissage**

Dans le cadre de l'apprentissage, différents types d'environnements informatiques peuvent être utilisés (Demaizière & al., 1992 ; Demaizière & Dubuisson, 1989 ; Bruillard, 1997). Nous avons choisi de nous intéresser plus particulièrement à l'utilisation d'un type particulier d'outil informatique, les grandes ressources linguistiques. Ces grandes ressources peuvent prendre la forme par exemple de dictionnaires monolingues, bilingues ou plurilingues réalisés par des linguistes dans des laboratoires de recherche ou des éditeurs, et mis à disposition d'auteurs, écrivains, traducteurs, linguistes et autres d'experts de la langue.

Il existe de nombreux dictionnaires pour enfants. Mais ces dictionnaires ne sont pas toujours adaptés aux besoins des élèves (Péliissier, 2002). Nous faisons l'hypothèse que ces ressources linguistiques, telles qu'elles sont aujourd'hui présentées peuvent être utilisées en situation de formation. L'exploitation de telles ressources dans le cadre de l'apprentissage, soulèvent différentes problématiques :

1. Identification des connaissances liées à l'apprentissage d'un domaine. Il s'agit de définir les savoirs et savoir faire qu'un apprenant doit acquérir dans le cadre de son apprentissage d'un domaine tel que le français, les langues étrangères, etc.
2. Caractérisation des ressources. Le but de caractériser les ressources linguistiques de manière à laisser apparaître des possibilités d'exploitation pédagogique.

3. Définition des activités mettant en jeu ces ressources. L'objectif est de proposer des activités pouvant se caractériser par les connaissances liées à l'apprentissage qu'elles mettent en jeu.

Nous illustrerons ces problématiques en prenant le cas du dictionnaire du Trésor de la Langue Française informatisé (TLFi), et de son exploitation dans l'apprentissage du Français au cycle 3.

## **2. Identification des Connaissances du Domaine**

L'identification des connaissances liées à l'apprentissage d'un domaine passe par l'analyse des savoirs et savoir faire que l'apprenant doit acquérir.

Au primaire, la maîtrise du langage et de la langue française est une priorité. En effet, le maniement du français oral et du français écrit doit permettre à l'élève, au collège, d'être autonome. D'après les programmes du Ministère de l'éducation nationale (Ministère, 2002), les compétences liées à la maîtrise de la langue au cycle 3 sont réparties sur plusieurs champs disciplinaires. Nous nous attachons plus spécifiquement à deux d'entre eux, celui de la littérature (dire, écrire, lire) et l'observation réfléchie de la langue française (grammaire, conjugaison, orthographe, vocabulaire). Les compétences associées à ces deux champs disciplinaires sont de trois types : compétences en communication, compétences liées à la lecture et compétences liées à production de texte.

Notre objectif étant uniquement d'illustrer comment des ressources linguistiques peuvent être exploitées, nous choisissons de nous attacher, pour ce travail, aux compétences liées à l'apprentissage de la lecture.

Les recherches théoriques sur l'apprentissage de la lecture, ont permis dans le cadre du projet AMICAL (Architecture Multi-agents Interactive Compagnon pour l'Apprentissage de la Lecture) de Clermont-Ferrand (Chambreuil et al., 2000) d'organiser les connaissances mises en jeu dans l'apprentissage de la lecture suivant quatre constituants de la langue écrite, la lettre, le mot, la phrase et le texte. Pour le mot, les connaissances qui lui sont associées 3 peuvent être regroupées en quatre ensembles : les connaissances liées à la représentation écrite, les connaissances liées à la correspondance entre le système de l'écrit et le système de l'oral, les connaissances liées aux aspects syntaxiques, les connaissances liées aux aspects sémantiques.

L'enfant doit d'abord se construire sa propre représentation du système de l'écrit. Il doit élaborer une représentation cognitive du fonctionnement de l'écrit et acquérir des règles du code alphabétique et orthographique dans la langue donnée. Ensuite, l'élève doit acquérir des règles de correspondances entre le système de l'écrit et le système de l'oral. Ces règles lui permettent d'associer à chaque chaîne écrite, une chaîne orale. L'élève doit enfin acquérir des stratégies de compréhension lui permettant de construire la signification à partir de l'écrit. Ces stratégies peuvent faire intervenir des connaissances liées à des aspects syntaxiques et des aspects sémantiques de la langue (Rolland, 1998).

A titre d'exemple, parmi les connaissances liées aux aspects syntaxiques, nous avons la connaissance « Mot est associé à une catégorie grammaticale ». Dans cette connaissance le mot est présenté comme une unité appartenant à un groupe d'unités de même type grammatical (par exemple, 'loup' appartient à la catégorie 'nom'). Cette connaissance

permet à l'enfant de rentrer dans la composition de la phrase dans laquelle le nom et le verbe joue un rôle important.

### **3. Caractérisation des Ressources**

L'objectif est de caractériser les ressources informatiques de manière à voir apparaître des possibilités d'exploitation pédagogiques. Cette caractérisation peut se faire d'une part sur les données que le système présente et d'autre part sur les fonctionnalités de consultation mises à disposition de l'utilisateur.

Le TLF (Trésor de la Langue Française) est un dictionnaire des XIX<sup>ème</sup> et XX<sup>ème</sup> siècles en 16 volumes. Il a été rédigé au sein de l'INaLF entre 1971 et 1994 et regroupe 100 000 mots avec leur histoire, 270 000 définitions, 430 000 exemples. Le TLFi est la version informatisée du TLF. Il est en accès libre en ligne (<http://atilf.fr/tlfi>) et a été développée à l'ATILF (successeur de l'INaLF depuis le 01/01/2001).

Les données linguistiques sont les différentes informations que le système informatique met à la disposition de l'utilisateur. Dans le TLFi, nous trouvons des données synchroniques et des données diachroniques. Les données synchroniques correspondent aux informations associées au mot à un moment donné alors que les données diachroniques font référence aux informations associées au mot dans son évolution. Parmi les données liées à la synchronie nous trouvons l'entrée et l'article. L'entrée se compose du mot vedette et de son code grammatical. Le mot vedette est le mot du dictionnaire qui a été sélectionné par l'utilisateur et le code grammatical correspond à la catégorie grammaticale de ce mot. Parmi les données diachroniques nous trouvons la phonétique du mot vedette, son histoire, son étymologie, sa fréquence dans les œuvres littéraires et une bibliographie.

D'un point de vue pédagogique, par exemple, avoir le code grammatical du mot permet à l'élève de définir la catégorie grammaticale à laquelle ce mot appartient. La connaissance pouvant être rapprochée de cette donnée est donc « Mot appartient à une catégorie grammaticale ».

Parmi les fonctionnalités laissées à la disposition de l'utilisateur dans le TLFi pour consulter les données linguistiques, nous distinguons celles qui permettent de : rechercher un mot dans le dictionnaire, mettre en valeur une information particulière dans un ensemble d'informations, se déplacer d'un mot à un autre, établir une liste de mots.

Chacune de ces fonctionnalités, comme les données précédemment abordées, peut être analysée d'un point de vue de l'apprentissage de la lecture. Par exemple, la recherche d'un mot dans un dictionnaire peut s'effectuer par l'écriture en lettres de ce mot. L'utilisateur peut sélectionner un mot du dictionnaire en tapant à l'aide du clavier de l'ordinateur la séquence de lettres du mot qu'il cherche. Cette séquence peut comporter certaines erreurs comme l'omission d'une lettre, l'absence d'accents et le remplacement d'un graphème par un autre qui correspond au même phonème. L'apprenant, au cycle 3, peut oublier une lettre lorsqu'il tape un mot au clavier. Il peut aussi rencontrer des difficultés à déterminer l'accent (grave, aigu ou circonflexe) à mettre sur les lettres de certains mots. Proposer à l'élève de taper le mot qu'il veut en oubliant une lettre ou en omettant un accent par exemple est une fonction intéressante. La connaissance associée à cette fonctionnalité est « Mot est une structure de lettres ». Par ailleurs, la possibilité d'obtenir un mot du dictionnaire même si l'élève a remplacé un graphème par un autre graphème correspondant au même phonème est

également une fonction intéressante. Elle permet à l'enfant de taper le mot qu'il veut à partir du mot oral qu'il connaît. La connaissance correspondante à cette fonction est « Mot oral correspond à un mot écrit ».

#### **4. Définition des Activités Pédagogiques mettant en jeu ces Ressources**

A partir d'une part des connaissances liées à l'apprentissage de la lecture et d'autre part à la description des ressources linguistiques présentes dans le TLFi, nous proposons de décrire et de caractériser des activités mettant en œuvre ces connaissances et ce dictionnaire.

Parmi les activités pouvant être envisagées avec le TLFi, nous distinguons des activités pour les enseignants et les apprenants. Pour les apprenants, au cycle 3, nous avons par exemple l'activité qui consiste pour l'apprenant à déterminer la catégorie grammaticale de certains mots présents dans un texte vu en classe. En ce qui concerne l'enseignant, nous pouvons proposer une activité dans laquelle l'enseignant demanderait au système informatique de lui fournir une liste de mots contenant les lettres 'mp', 'mm' ou 'mb'. Cette liste pourrait ensuite lui permettre de constituer un exercice qu'il donnerait à ses élèves pour leur faire acquérir la règle « Devant les lettres b, m, p, j'emploie m (et non n) : temps, ensemble, emmener etc. Exceptions : bonbon, bonbonne, bonbonnière, néanmoins, embonpoint ».

Chaque activité proposée peut se caractériser par l'acteur (l'apprenant ou l'enseignant), la connaissance lecture centrale mise en jeu, le type de cette connaissance, une fiche qui explique comment utiliser la ressource pour réaliser une activité particulière. Une telle caractérisation fait apparaître les relations qui peuvent être établies entre les connaissances que l'apprenant doit acquérir et les ressources linguistiques mises à sa disposition.

#### **5. Résultats d'Expérimentation**

Nous avons réalisé une expérimentation dans une classe de CE2 (école des Trois Maisons, à Nancy). Nous avons établi une séquence de six activités utilisant le TLFi. Dans l'une de ces activités, on présente à l'élève un texte, dans lequel différents mots sont soulignés. L'apprenant doit chercher chacun des mots soulignés dans le TLFi et écrire la catégorie grammaticale associée à chacun d'eux.

Dans le TLFi, la catégorie « nom » est indiqué par « Subst. » (pour substantif). Ce terme a été expliqué par l'enseignante le matin, en classe, comme étant l'autre mot utilisé pour désigner les « noms ». Les élèves, lors de l'expérimentation, ont recopié sur leur feuille d'exercice le mot « subst. » et lorsqu'on leur a demandé ce que cela signifiait, ils ont répondu : « c'est comme le nom, c'est pareil ». Il semblerait donc que le vocabulaire présenté dans le TLFi, après avoir été expliqué, ne pose pas un problème majeur aux élèves.

#### **Conclusion**

Nous avons présenté brièvement les trois grandes problématiques qui semblent se poser dans le cas de l'utilisation d'une grande ressource linguistique dans l'apprentissage et l'enseignement d'un domaine : l'identification des connaissances liées à l'apprentissage, la caractérisation des ressources linguistiques et la détermination des activités caractérisées pouvant être proposées. Une expérimentation en classe a permis de valider une partie de ces

activités proposées. L'utilisation du TLFi par des élèves de cycle 3 semble donc être intégrée dans le cadre de l'apprentissage du français.

## Références

- BO**, 2002 : Programmes de l'école primaire – cycle des approfondissements, Bulletin officiel n°1, 14 février 2002, hors série, pp 66-69 et pp 72-76
- Bruillard, E.** 1997 : *Les machines à enseigner*, Hermes
- Chambreuil M. & al.**, 2000 : *Le projet AMICAL : Architecture Multi-agents Interactive Compagnon pour l'Apprentissage de la Lecture*, Chambreuil Michel et l'équipe de recherche AMICAL, rapport interne, Laboratoire de Recherche sur le Langage, Université Blaise Pascal – Clermont-Ferrand 2
- Demaizière & al.**, 1992 : *De l'EAO aux NTF - Utiliser l'ordinateur pour la formation*, Françoise Demaizière, Colette Dubuisson et Odile Blanvillain, Ophrys
- Demaizière, F. and Dubuisson, C.**, 1989 : « Panorama critique », *Langue Française*, n°83, septembre 1989, pp 3-13
- Dendien, J. and Pierrel, J-M.**, 2003 : « Trésor de la Langue Française informatisé », *Traitement Automatique de la Langue*, vol. n°44-n°2/2003, Hermes Sciences Edition, 28 p.
- Ministère**, 2002 : « Qu'apprend-on à l'école élémentaire ? – Les nouveaux programmes », Ministère de l'éducation nationale, CNDP Editions
- Pélissier, C.** 2002. « Analyse de Mon Premier Dictionnaire », revue *ALSIC (Apprentissage des Langues et Systèmes d'Information et de Communication)*, rubrique Analyse de logiciel, décembre 2002, volume 5, numéro 2, pp 269-286
- Rolland, M-C.** 1998 : « Cycle III – Une première culture en sciences humaines », collection dirigée par Gilbert PY, Formation des personnels de l'éducation nationale, Paris : ellipses