

# Bases lexicales multilingues et objets pédagogiques interactifs : Sensillon pour Papillon

## *Multilingual Lexical Databases and Web-based Interactive Language Learning Objects: Sensillon for Papillon*

Valérie Bellynck\*

\* ARCADE - CLIPS – IMAG, BP 53 - 385 avenue de la Bibliothèque  
38041 Grenoble cédex, FRANCE  
Valerie.Bellynck@imag.fr

### **Résumé :**

Nous présentons quelques outils permettant à des environnements d'apprentissage des langues de tirer parti d'une base de données lexicales multilingue riche. Il s'agit de conduire l'apprenant dans l'acquisition de vocabulaire, en s'appuyant sur les principes lexicologiques et linguistiques structurant cette base lexicale, mais de façon transparente, grâce à des objets pédagogiques interactifs à manipulation directe. Du point de vue de l'EAO, il s'agit d'artéfacts liés aux activités dites « secondaires ». Nous avons ainsi spécifié une fonction d'accès à la base Papillon qui pourrait être utilisée de façon générique par des formateurs-auteurs pour l'apprentissage de n'importe quelle langue supportée par Papillon.

**Mots-clés :** acquisition de vocabulaire, objets pédagogiques interactifs, base lexicale multilingue, environnements d'apprentissage des langues

### **Abstract :**

We present some tools for making language learning environments able to benefit from a rich multilingual lexical database. The aim is to guide learners in assimilating vocabulary, relying on lexicologic and linguistic rules, but transparently, through direct manipulation of interactive pedagogical objects. These tools can be used in secondary activities in CAL. We propose some access functions to the multilingual database Papillon that could be used more generally by teachers to construct tools that would be applicable to any language supported by Papillon.

**Keywords :** vocabulary acquisition, interactive learning objects, multilingual lexical database, language learning environments

## **1. Introduction**

Les distances géographiques se réduisent. Les spécificités culturelles peuvent être préservées à moindres coûts. Avec Internet, de nouveaux besoins d'échange entre humains apparaissent, allant de la communication interpersonnelle à la diffusion de messages à des communautés multilingues, en passant par la « lecture active » de documents en langues étrangères.

Le nombre de langues en question augmente aussi, la part de l'anglais diminuant considérablement. De plus, le nombre de couples de langues à considérer (langue maternelle, langue à apprendre) augmente considérablement : qui aurait pensé au couple <japonais, breton<sup>1</sup>> il y a quelques années ? Ces besoins touchent des langues de plus en plus éloignées géographiquement, et des langues qui n'ont pas jusqu'ici disposé d'assez de ressources pour être dotées de logiciels destinés à leur apprentissage. Ces deux aspects imposent de passer de la situation actuelle, où les logiciels d'apprentissage sont développés de façon spécifique pour quelques langues « riches », à une situation où il serait facile d'ajouter des langues, grâce à des logiciels génériques.

La structure même d'une base de données comme Papillon permet d'envisager tous les accès binaires entre les langues supportées. D'autre part, si plusieurs langues d'une même famille (comme le thaï et le lao) sont contenues dans la base, il est possible de fabriquer des interfaces présentant des n-uplets de termes et de sens, pour apprendre des langues en s'appuyant sur d'autres langues proches déjà apprises<sup>2</sup>.

Du point de vue du contenu pédagogique, les logiciels pour l'apprentissage des langues sont souvent constitués d'une tâche principale en trame de fond orientée vers la communication, accompagnée de tâches secondaires, sous forme d'exercices répétitifs destinés à l'assimilation de concepts linguistiques, syntaxiques ou lexicaux. Nous pensons essentiellement à des exercices qui tireraient parti du multilinguisme offert par la base lexicale multilingue riche qu'est Papillon.

Il s'agit de conduire l'apprenant dans l'acquisition de vocabulaire, en s'appuyant sur les principes lexicologiques et linguistiques structurant cette base lexicale, mais de façon transparente, grâce à des objets pédagogiques interactifs à manipulation directe. Du point de vue de l'ÉAO, il s'agit d'artéfacts liés aux activités dites « secondaires ». Nous avons ainsi spécifié une fonction d'accès à la base Papillon qui pourrait être utilisée de façon générique par des formateurs-auteurs pour l'apprentissage de n'importe quelle langue supportée par Papillon.

Nous commençons par décrire les spécificités des objets pédagogiques interactifs pour l'apprentissage des langues. Nous expliquons ensuite quel type d'activité pédagogique peut être basé sur une exploitation de la base lexicale multilingue Papillon. Enfin, nous proposons deux interfaces : l'une pour l'apprenant, l'autre pour le formateur-auteur.

## **2. Outils pédagogiques interactifs pour l'apprentissage des langues**

La place de l'ordinateur dans l'apprentissage a beaucoup évolué depuis ces dernières années. On n'a plus seulement affaire à des logiciels ne permettant soit que d'accéder à une information « en ligne », sorte de transposition du livre à l'hypermédia, soit que de rabâcher des activités répétitives, mais à des applications de plus en plus intelligentes (micro-mondes, simulations, tutoriels intelligents) basées sur une activité de construction ou de résolution de problème. Cette démarche est issue des théories d'apprentissage socio-constructivistes : « J'apprends en faisant, en communiquant et en échangeant ».

On appelle objet pédagogique interactif (OPI) un objet réel ou simulé, accessible depuis un poste informatique, et permettant un apprentissage actif. Les logiciels pédagogiques peuvent être classés dans une typologie qui est définie par les fonctions pédagogiques qu'ils remplissent et les types de

---

<sup>1</sup> L'auteur est bretonne et aimerait bien communiquer en japonais...

<sup>2</sup> Voir par exemple les projets européens Galatea et Galanet d'apprentissage parallèle de plusieurs langues romanes à partir d'autres langues romanes.

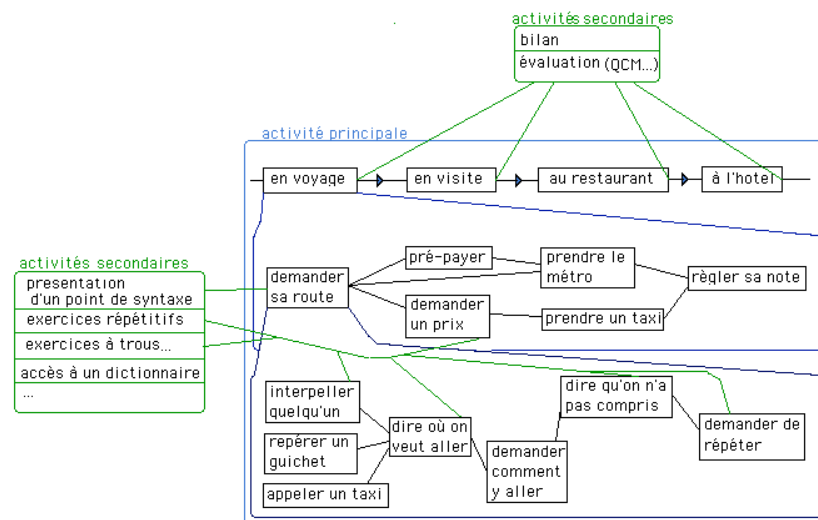
tâches demandés aux apprenants : *Présenter de l'information* (« support de cours » ou « tutoriel »), *Dispenser des exercices* (« machines à enseigner »), *Véritablement enseigner* (« tuteur intelligent »), *Captiver l'attention et la motivation de l'élève* (« jeux éducatifs »), *Fournir un espace d'exploration* (« hypermédia » ou « multimédia »), *Fournir un environnement pour la découverte de lois naturelles* (« simulation »), *Fournir un environnement pour la découverte de domaines abstraits* (« micro-monde »), *Fournir un espace d'échange entre élèves* (« logiciel d'apprentissage collaboratif ») [6].

Les formateurs accèdent à un rôle de création, adaptation et exploitation de situations actives d'apprentissage. Ces activités sont conçues pour aider à assimiler des concepts du domaine d'enseignement, dans le contexte d'une progression pédagogique (apprentissage planifié et contrôlé). La conception de ces activités est basée sur la notion de scénario. Le travail à fournir pour réaliser des scénarios doit être réutilisable, et éventuellement mettre à contribution plusieurs OPI différents. Des travaux de normalisation sont en cours dans le domaine des environnements informatisés des Apprentissages humains (EIAH) visant l'interopérabilité des OPI.

La spécialité de l'équipe ARCADE de Grenoble est de fournir aux formateurs des modèles, méthodes et outils informatiques pour la production, l'adaptation et l'exploitation d'OPI. On voit apparaître de nouveaux produits issus de la recherche en EIAH et basés sur les technologies XML qui permettent de définir le contrôle sur l'utilisation d'un OPI comme le parcours d'un document hypermédia. En particulier, pour le traitement des documents de type encyclopédique, et l'industrialisation de la conception de supports pédagogiques, des outils sont proches de sortir sous forme de produits commerciaux [1].

Les environnements pour l'apprentissage des langues sont basés sur une *activité principale*, porteuse de la progression définie par le formateur-auteur comme d'une trame de scénario, et d'*activités secondaires* imposant ou proposant des bilans sur les éléments linguistiques rencontrés, des exercices d'assimilation, des aides structurelles ou lexicologiques...

L'acquisition de quatre compétences complémentaires est visée : la *compréhension* de la langue *orale*, de la langue *écrite* et l'*expression* dans la langue orale et écrite, avec une approche qui peut être communicative, conceptualisatrice, ou structurale par exemple. Les environnements d'aide à l'apprentissage de langues sont plus souvent orientés vers la compréhension, s'appuyant sur les « appels à l'aide » pour motiver les autres besoins, qu'ils proposent dans le cadre d'activités secondaires. Par l'exemple, l'aide peut fournir des représentations partielles du dialogue entendu par l'apprenant (résumé, liste de mots, accès à un dictionnaire...) ou des explications sémantico-structurelles, des exercices d'aide à l'assimilation. D'autres types d'activités secondaires, prévues à la fin d'étapes dans le scénario, peuvent être le support de bilans ou d'évaluations.



Il existe aujourd'hui des environnements qui permettent de concevoir des supports d'exercices ou de cours qui seront portés par les activités secondaires. Des services Web comme HotPotatoes<sup>3</sup> ou Quia<sup>4</sup> permettent de fabriquer des exercices d'assimilation, comme un texte à trous, un jeu de mémoire, un questionnaire à choix multiples, des questions/réponses, un jeu d'appariement, des mots croisés...

Dans ces environnements, les ressources sont issues d'une seule langue. En fait, les activités secondaires pourraient reposer sur des ressources multilingues, au niveau grammatical et au niveau lexical. Dans le cadre de cet article, nous nous focalisons sur une proposition d'artéfacts interactifs qui s'appuieraient sur la base lexicale multilingue Papillon.

### 3. Objectifs pour des activités secondaires basées sur Papillon

Nous voulons proposer des outils pédagogiques pour les environnements d'apprentissage des langues dont l'activité principale est basée sur une *communication à travers le Web*, par exemple une « tchatche » (« chat ») avec un but (création de document ou jeu de rôle). À plus long terme, nous envisageons une utilisation in situ, l'activité principale devenant alors la situation vécue, et les artéfacts proposés en activités secondaires sur des supports interactifs mobiles.

Nous avons parlé plus haut de *généricité* des logiciels pour pouvoir traiter toutes les langues d'une base comme Papillon sans faire de développement spécifique lourd pour chaque nouvelle langue. Cela sera facilité par le caractère générique des schémas XML des dictionnaires monolingues de Papillon.

D'autre part, nous désirons que nos artéfacts puissent être intégrés (sans programmation de bas niveau) par des formateurs-auteurs dans les différents scénarios d'apprentissage qu'ils construisent dans tel ou tel générateur d'EIAH. Pour cela, nous visons l'*interopérabilité*, c'est-à-dire la possibilité pour deux applications d'échanger des données correctement sans avoir besoin de concevoir du code

<sup>3</sup> <http://web.uvic.ca/hrd/hotpot/mactutor/intro.htm>

<sup>4</sup> <http://www.quia.com/findout.html>

pour s'adapter à des formats spécifiques de chaque application, mais en s'appuyant sur un format normalisé, caractérisant la manière dont sont organisées les données.

Les objectifs précédents correspondent aux besoins d'EIAH des langues vis-à-vis de Papillon. À l'inverse, il paraît souhaitable de solliciter la contribution des formateurs-auteurs. En effet, Papillon est une base lexicale multilingue contributive : en échange de sa consultation, les utilisateurs sont sollicités à participer à l'amélioration de la base elle-même par l'ajout d'annotations qui ajustent et affinent, corrigent et complètent les informations associées aux sens de mots (« lexies ») et à leurs liaisons (« axies »), ces informations et liaisons provenant elles-mêmes d'entrées récupérées automatiquement à partir des dictionnaires bilingues et monolingues du domaine public qui sont en ligne sur la Toile.

Il faut enfin que nos OPI soient agencés de telle sorte que les formateurs-auteurs soient encouragés à contribuer, non pas seulement de façon isolée, mais encore en liaison forte avec leur communauté, ce qui est possible en principe grâce à l'organisation du serveur Papillon, mais reste bien sûr à implémenter de façon concrète, et si possible attrayante.

## **4. Des propositions**

Nous présentons deux propositions d'interface et d'outils :

- une interface pour rechercher un mot isolé ou une unité de sens, ne nécessitant pas de connaissance linguistique a priori, ni de connaissance sur les concepts utilisés dans Papillon pour gérer les données,
- une interface pour le formateur-auteur afin qu'il fabrique les interfaces avec lesquelles il veut faire travailler ses élèves.

Il conviendrait de compléter ces propositions par quelques idées d'outils pour exploiter les résultats fournis par Papillon du point de vue de l'apprentissage (apprentissage lexical, transfert interlingue), ainsi que pour favoriser la communication entre les utilisateurs de Papillon avec un objectif d'apprentissage (formateurs-auteurs ou apprenants).

### **4.1. L'interface de base de Papillon**

La partie de l'interface du site de Papillon qui permet de définir une requête se présente comme le montre la copie d'écran ci-dessous :

The screenshot shows the Papillon web interface in a browser window. The address bar displays `http://bushido.imag.fr/papillon/Consult.po`. The page features a butterfly logo and the word "Papillon". Navigation links include "Information", "Consultation", "Edition", and "Contacts". On the left, there are links for "Register", "Login", "Clips", "NII", "Old Site", and "English français". The main search area contains three sections: "Source:" with a dropdown set to "French", "Targets:" with a dropdown set to "English", and "Resources:" with a dropdown set to "FeM". Below these, there is a "Lemmatize the entry:" checkbox, a "Headword:" field containing "amour", a "Pos:" field, and an "Any part:" field. A "Go" button is at the bottom of the search area.

Pour les paramètres entrés dans les champs dans l'illustration précédente, la requête calculée est :

`http://bushido.imag.fr/papillon/Consult.po?SOURCE=fra&TARGETS=eng&RESOURCES=FeM&vocontains=amour&poscontains=&anycontains=`

Et le résultat de cette requête est présenté de la manière suivante :

[FeM](#) [Papillon](#) [XML](#) [DEFAULT](#) [DEC](#) [JMDict](#) [ThaiDict](#) [VietDict](#)  
**amour** //amou-r//  
 n.m. love pour l'amour de for the sake of faire l'amour to make love  
 n.f. pl. de belles amours poét. wonderful love

Les résultats fournis par Papillon sont proposés sous plusieurs formes, dont une forme XML. Certaines présentations sont proches de ce qu'on trouve habituellement dans les dictionnaires. La forme XML peut être exploitée par les environnements d'apprentissage pour permettre au formateur-auteur de concevoir une présentation adaptée au public des apprenants visés.

[FeM](#) [Papillon](#) [XML](#) [DEFAULT](#) [DEC](#) [JMDict](#) [ThaiDict](#) [VietDict](#)  

```

<fem-entry>
  <ENTRY>amour</ENTRY>
  <FRENCH_PRON>amou-r</FRENCH_PRON>
  <FRENCH_CAT>n.m.</FRENCH_CAT>
  <ENGLISH_EQU>love</ENGLISH_EQU>
  <FRENCH_PHRASE>pour l'amour de</FRENCH_PHRASE>
  <ENGLISH_PHRASE>for the sake of</ENGLISH_PHRASE>
  <FRENCH_PHRASE>faire l'amour</FRENCH_PHRASE>
  <ENGLISH_PHRASE>to make love</ENGLISH_PHRASE>
  <FRENCH_CAT>n.f. pl.</FRENCH_CAT>
  <FRENCH_PHRASE>de belles amours</FRENCH_PHRASE>
  <FRENCH_LABEL>poét.</FRENCH_LABEL>
  <ENGLISH_PHRASE>wonderful love</ENGLISH_PHRASE>
</fem-entry>

```

Dans l'interface de base de Papillon, tous les choix sont offerts à travers une seule interface (langues sources, langues cibles, ressources lexicales utilisées) et cela suppose d'avoir une idée précise du mot cherché, dans la langue source. L'interprétation des résultats demande aussi d'avoir une idée précise des fonctions syntaxiques et lexicales du mot cherché.

#### **4.2. Un exemple : notre interface « Sensillon »**

Le premier outil nécessaire est celui qui permet de rechercher un sens, indépendamment de la langue dans laquelle on veut l'exprimer. Dans le cadre de l'apprentissage des langues, quelles sont les exigences particulières d'un tel outil ?

Deux types d'utilisateurs interviennent : les formateurs-auteurs et les apprenants. Les premiers doivent pouvoir concevoir l'outil utilisé par les seconds, mais n'ont pas de compétence a priori en informatique, et pas de motivation non plus à en acquérir par un investissement lourd. Nous proposons au formateur-auteur de « brider » l'usage des seconds, à partir d'outils prédéfinis, plutôt que de ne lui offrir qu'une seule forme « programmable ».

La particularité des apprenants est qu'ils ne connaissent pas la terminologie utilisée en linguistique et en lexicologie pour désigner les concepts qui permettent de caractériser la base lexicale Papillon, et dont certains la régissent. L'interface doit donc montrer l'usage des concepts manipulés dans le contexte de ce qu'un apprenant a déjà saisi, et lui proposer les choix restants en fonction de ceux qui ont déjà été effectués.

Enfin, pour suggérer le sens d'un mot qu'on ne trouve pas [8], différentes méthodes sont possibles. Nous en avons déduit 4 propositions :

- une entrée par la forme d'utilisation désirée,
- une par une paraphrase,
- une par un exemple éventuellement à trou,
- et une par un ensemble de mots proches mais non satisfaisants.

L'illustration suivante présente une interface pour l'apprenant, comme aurait pu la définir un formateur-auteur.

Cette interface est constituée de

- une zone de titre,
- une zone d'entrée multiple du sens à chercher,
- une zone d'envoi de la requête.

La zone de titre permet de choisir les langues source et cible. Ces langues sont à prendre parmi les langues gérées par Papillon. Une même langue peut être choisie à la fois en langue source et en langue cible, de façon à permettre de s'appuyer sur Papillon pour trouver un ou des mots dans sa propre langue, à partir de son sens. La langue dite « Papillon » permet d'accéder aux axes. La zone de titre permet aussi de choisir la langue de l'interface elle-même, qui peut n'être ni la langue source, ni la langue cible.

La zone d'entrée multiple contient 4 modalités d'entrée pour le sens de mot recherché, et toutes les modalités ne sont pas possibles conjointement ou selon les langues choisies en source et cible. On pourrait en imaginer d'autres, et surtout, le formateur-auteur doit pouvoir en définir de nouvelles ou limiter celles qui sont proposées ici.

Dans la première modalité, l'apprenant écrit le mot qu'il cherche dans la langue source ainsi que sa catégorie. L'interface présente des choix les régimes [8], par exemple « amour de X pour Y » où X et Y sont des individus.

La deuxième modalité consiste à préciser le sens cherché au moyen d'une paraphrase. Le service doit calculer le sens, et au besoin interpellé l'apprenant pour qu'il précise sa pensée, l'exprime



correctement et sans ambiguïté. Cette modalité doit s'aider d'autres ressources logicielles issues du traitement automatique de la langue. Par exemple, la saisie de la paraphrase peut s'appuyer sur une édition synchrone d'un énoncé et de son graphe sémantique.

Dans la troisième modalité, l'apprenant doit saisir un exemple. Les différentes modalités devant être synchronisées autant que possible, l'édition de l'exemple peut être accompagnée d'un remplissage automatique de l'entrée de la première modalité. Le mot cherché dans l'exemple peut-être remplacé par un « ? ». Cette modalité cherche à rendre sensible la notion de vecteur conceptuel [4].

La quatrième modalité proposée permet de chercher un mot par ressemblance/dissemblance sémantique. L'apprenant entre des mots dont le sens est proche, et des mots dont le sens est éloigné, puis déplace les mots saisis pour que leurs positions relatives reflètent leurs proximités avec le mot cherché.

La zone d'envoi de la requête peut permettre de choisir entre plusieurs formes de présentation des résultats, dans un menu surgissant qui apparaît lorsque la souris reste appuyée un certain temps.

Pour différentes raisons, la requête peut ne pas aboutir : elle n'est pas assez précise, la langue cible n'est pas assez représentée dans Papillon... Dans le cas où il y aurait trop de choix, le service peut demander à l'utilisateur de préciser sa requête en complétant une nouvelle modalité, ou en modifiant une des modalités d'entrée.

#### **4.3. Une interface « Auteur »**

Nous avons vu que nous proposons au formateur-auteur de “brider” les possibilités d'un service ouvert plutôt que de l'obliger à s'atteler à une activité de nature plus programmatoire. L'interface que nous présentons dans ce paragraphe est un outil pour créer un « sensillon », c'est-à-dire un service Internet du type de celui du paragraphe précédent.

Un formateur peut être amené à en concevoir plusieurs pour un même environnement d'apprentissage, selon qu'il veut obliger l'apprenant à mobiliser telle ou telle ressource cognitive à certaines étapes du scénario. Par exemple, il peut vouloir

- favoriser l'exploitation par l'apprenant de ses connaissances dans d'autres langues pour accélérer de nouvelles acquisitions lexicologiques, ou
- limiter les recherches de vocabulaire au domaine de l'étape courante du scénario pour constituer un lexique ontologique<sup>5</sup> (les nombres, les plats d'un menu, les objets de la table), ou
- l'aider à se familiariser avec les concepts lexico-linguistiques et lui apprendre la terminologie utilisée dans ce domaine, ou encore
- l'obliger à décrire un concept en le situant intuitivement et géographiquement par rapport à d'autres,
- lui faire comprendre que les mots ne se correspondent pas bijectivement d'une langue à une autre, mais par l'intermédiaire de « non-formes », idéalement des concepts, et moins ambitieusement des « acceptions interlingues », approchées dans Papillon par les « axes », qui ne sont que des liens interlingues entre « lexies » (sens de mots).

Dans son activité de conception d'un sensillon, la tâche du formateur consiste à minimiser les choix à effectuer par l'apprenant, à lui offrir les moyens de décrire les mots qu'il cherche, et à lui présenter les résultats sous une forme abordable.

---

<sup>5</sup> le lexique cognitif ainsi constitué est structuré et permet un accès comme avec une ontologie.

L'interface que nous proposons se présente comme une boîte à outils à l'aide de laquelle le formateur-auteur définit la présentation d'un sensillon en plaçant dans une vue « résumée » l'organisation des outils qu'il souhaite fournir aux apprenants, ainsi que leurs limitations.

Par exemple, c'est le formateur-auteur qui décide des langues source et cible, des ressources utilisées, et de la langue associée au dialogue de l'interface. Pour cela, il s'appuie sur une ébauche du sensillon (représentation condensée de l'interface et de ses différentes zones) et choisit les langues qu'il laisse à l'apprenant dans les zones de titre et d'envoi. La zone d'envoi de l'ébauche permet également de proposer plusieurs présentations des résultats, définies a priori par le formateur, de la même façon.



L'interface du « créateur » (de sensillons) permet également au formateur de

- sélectionner des modalités parmi les quatre modalités de base présentées dans la section précédente,
- les repositionner dans l'ordre qu'il veut,
- les paramétrer éventuellement de plusieurs façons pour une même modalité, et de proposer se différentes variantes,
- concevoir de nouveaux types d'entrées non initialement prévus.

Toutes les possibilités de sélection et de position des modalités sont offertes à travers les boutons présentés dans l'illustration sous le texte « Demander un sens ». Appuyer sur un bouton déclenche l'ouverture d'une zone dans l'ébauche de sensillon, repositionnable a posteriori.

En ce qui concerne le paramétrage des différentes modalités que nous avons proposées, et plus particulièrement de la modalité « Paraphrase », le formateur pourrait associer à l'édition du texte de la paraphrase un éditeur visuel de la forme sémantique, synchronisée avec une vue textuelle de la phrase en cours.

Pour la première (mot+régime), le formateur pourrait choisir de guider l'apprenant pour l'aider à préciser le régime du mot cherché, sans faire apparaître le mot « régime », mais en proposant un exemple pour chaque catégorie.

À l'opposé, l'outil de création pourrait remplir automatiquement un exemple à partir de la première modalité (mot+régime).

Et, dans l'idéal, on pourrait imaginer une interface capable de repérer le type d'entrée inconsciemment désiré par l'utilisateur : par exemple, on exprimerait son idée et on soulignerait le mot ou le groupe de mots dont on n'est pas satisfait, ou on désignerait le « trou » laissé à la place du mot. Le service calculerait les paramètres nécessaires pour construire la requête à soumettre à la base lexicale, en ne dérangeant l'utilisateur que pour lui demander confirmation sur les ambiguïtés. Ce serait une option « a posteriori ».

L'option « a priori » consiste à entrer la recherche de l'expression d'un sens, de façon volontaire. Dans ce cas, chacune des « entrées » doit être possible, conjointement ou seule, le nombre de réponses dépendant de la précision de l'entrée.

#### **4.4. Présentation des résultats**

Dans cette section, nous présentons quelques possibilités d'exploitation pédagogique des interfaces proposées dans l'article, directement ou en relation avec d'autres outils.

La première utilisation pédagogique d'un sensillon est d'être un outil pour la lecture active, finalisée, qui conduit à apprendre et comprendre en même temps qu'on lit. On peut imaginer que, pour les langues à écriture éloignée (caractères, règles de coupures des mots), le système propose des gloses et donne des indications supplémentaires.

Du point de vue des logiciels pour l'apprentissage des langues plus « classiques », une ressource lexicale utilisée seule permet de contribuer à l'acquisition de vocables et d'expressions. On peut adapter les activités générées par les moteurs à exercices « classiques » aux résultats fournis par la base lexicale Papillon.

Il serait aussi utile de mémoriser les résultats présentés à l'apprenant au cours d'une session (ce qui correspondrait à une leçon, et du point de vue des progressions pédagogiques, à un thème). Des exercices de type question/réponse ou des exercices d'association peuvent être générés automatiquement en fonction des textes entrés par l'apprenant dans chacune des modalités.

On peut aussi imaginer des activités qui profitent de l'insertion de l'application dans Internet pour rechercher des exemples d'utilisation du mot trouvé, et demandent à l'apprenant de l'écrire éventuellement sous la forme fléchie adaptée, dans un champ à remplir, ou de choisir entre plusieurs formes fléchies présentées dans un menu déroulant.

Nous avons déjà parlé précédemment du couplage possible des ressources linguistiques multilingues lexicales et sémantiques, avec la proposition d'un éditeur synchronisé pour l'édition des paraphrases. La manipulation directe d'une représentation visuelle du sens d'un énoncé, synchronisée avec son « résultat » textuel, peut conduire à la conception de nouveaux outils intéressants pour les personnes qui ont une mémoire visuelle développée. De telles interfaces pourraient s'appuyer sur le concept de micro-mondes de création d'énoncés.

### Note

Les interfaces que nous avons proposées dans cet article sont accessibles sur Internet, à l'URL : <http://membres.lycos.fr/papisens>. Mais ce ne sont que des faces avant, pour l'instant sans relation effective et active avec la base Papillon. Nous les avons publiées dans l'idée d'ouvrir à la discussion publique.

## 5. Conclusion et perspectives

Nous avons donc présenté quelques artéfacts permettant à des environnements d'apprentissage des langues de tirer parti d'une base de données lexicales multilingue riche comme la base Papillon. Ils permettront de conduire l'apprenant dans l'acquisition de vocabulaire, en s'appuyant sur les principes lexicologiques et linguistiques structurant cette base lexicale, tout en le faisant de façon transparente, grâce à des objets pédagogiques interactifs à manipulation directe.

Nous avons ainsi spécifié une fonction d'accès à la base Papillon qui pourrait être utilisée de façon générique par des formateurs-auteurs pour l'apprentissage de n'importe quelle langue supportée par Papillon.

L'ouverture au milieu de l'enseignement de l'utilisation de la base lexicale multilingue contributive Papillon permettrait aux formateurs de confronter leur expertise sur l'usage des mots dans les langues dont ils sont experts, et à la base de s'enrichir d'annotations lexicales importantes.

La création d'environnements d'apprentissage des langues basés sur des ressources dans lesquelles la langue cible est un paramètre permettra aux didacticiens de l'enseignement des langues de décorrélérer leurs travaux du point de vue technique, pour se centrer plus facilement et plus directement sur les spécificités des diverses langues et méthodes d'apprentissage, en permettant des comparaisons transversales immédiates sur les langues.

Les études proposées dans cet article permettent d'envisager de nouvelles problématiques didactiques, comme l'exploitation didactique des bases de données, dont les buts principaux sont de comprendre et de retrouver.

## 6. Bibliographie

1. Crozat S. (2001) *Eléments pour la conception industrialisée des supports pédagogiques numériques*. Thèse soutenue à l'UTC le 22 mars 2001.
2. David J.-P., Guilloux C., Flament A. (2002) *GenDoc : un outil auteur générique d'objets pédagogiques*. ITS'2002, à paraître.
3. Duval E. (2001) *Standardized Metadata for Education: a Status Report*. Ed-Media'2001, Tampere, Finland.

4. Lafourcade M., Prince V. (2001) *Synonymies et vecteurs conceptuels*. In *proc. of Traitement Automatique du Langage Naturel (TALN'2001)*. Tours, France, juillet 2001, pp. 233-242.
5. Poussard C. (2000) *Compréhension de l'anglais oral et technologies éducatives*. Thèse d'études anglophones, Paris 7.
6. de Vries E. (2001) *Les logiciels d'apprentissage : panoplie ou éventail ?* Revue Française de Pédagogie, 137, 105-116. (deVries\_RFP.pdf)
7. Mel'tchuk I. (1984, 1988, 1992) *DEC: Dictionnaire Explicatif et Combinatoire du français contemporain, recherches lexico-sémantiques I, II et III*. Presses de l'Université de Montréal, Montréal (Québec), Canada, 172 p., 332 p. et 323 p.
8. Polguère A. (1998) *La théorie Sens-Texte*. Dialangue, Vol. 8-9, Université du Québec à Chicoutimi, 9-30.
9. Zock M., Fournier J.-P. (2001) *How can computers help the writer/speaker experiencing the Tip-of-the-Tongue-Problem ?* Proc. Of RANLP, Tzigov Chark, pp. 300-302.