

TAL et EIAH comme aide au travail des enseignants

Georges Antoniadis^{*}, Mathieu Loiseau^{*,**} & Philippe Dessus^{**}

^{*} LIDILEM, université Stendhal, Grenoble

^{**} LSE, université Grenoble-2 et IUFM, université Grenoble-1

•

Atelier TAL et EIAH–Assises GDR I3–Porquerolles, 10 mai 2012

Le TAL commence maintenant à être utilisé dans des EIAH d'aide au travail des enseignants, dans une perspective multi-disciplinaire. L'objet de cette présentation est d'illustrer cette utilisation à partir de nos recherches et de dresser une liste de questions et problèmes qui en résultent.

Traditionnellement, l'étude du travail de l'enseignant est séparée en trois moments principaux, formant un cycle (Riff & Durand, 1993) : la planification (*phase préactive*) : où l'enseignant prépare son cours en fonction de ses objectifs et sélectionne ses ressources ; la *phase interactive*, où l'enseignant, en présence de ses élèves, met en œuvre son enseignement et où les élèves travaillent et apprennent ; la *phase post-active*, où l'enseignant, à nouveau en dehors de la présence de ses élèves, réfléchit à son cours et aux performances de ses élèves, et met en œuvre une nouvelle activité de planification en tenant compte de ses cours précédents.

Pour illustrer l'utilisation de techniques de TAL dans des EIAH promouvant chacune de ces phases, nous avons choisi trois systèmes. La particularité de ces systèmes est triple :

- ils engagent leur utilisateur dans une activité complexe, pour laquelle il n'existe aucun mode opératoire défini ;
- ils utilisent les possibilités du TAL et tiennent compte de la non perfection de ses résultats, en le mettant à contribution pour des « niches » didactiques appropriées ;
- ils utilisent les faiblesses ou les manques connus du TAL pour engager leur utilisateur dans une analyse réflexive de sa production (la question de l'« auto »). Cette analyse, en retour, peut leur permettre une meilleure compréhension de leur propre activité. Cela comble un manque souvent mis en avant dans la littérature (Vovides *et al.*, 2007).

La recherche de documents authentiques par les enseignants de langues pour servir de support à leur enseignement est une partie de l'activité de planification (*phase préactive*). Comme Levy (1997) le signale, le texte authentique joue un rôle essentiel pour l'approche communicative, largement adoptée par les enseignants des langues. Néanmoins, bien que les corpus de textes authentiques soient nombreux, la recherche repose le plus souvent sur de traditionnelles recherches par mots-clefs. En ce sens, il n'existe pas de système lui permettant de formuler sa requête avec des termes de sa problématique (par exemple des textes avec forte occurrence de verbes à un temps donné), en utilisant des concepts pédagogiques qui lui sont familiers.

Le modèle facette/prisme présenté dans Loiseau (2009) essaie de pallier cet inconvénient, en considérant qu'un texte possède plusieurs « facettes » pédagogiquement pertinentes (par exemple nombre de mots, nombre de verbes, etc.) qui peuvent être instanciées en fonction du contexte pédagogique. Dans ce contexte un prisme permet cette instanciation et par transitivité de sélectionner les documents. L'enseignant choisit les prismes pertinents et fournit les informations correspondant au contexte pédagogique et les facettes sont calculées pour chaque texte en prenant en compte ses propriétés intrinsèques entrées soit manuellement, soit automatiquement par calcul (grâce au TAL notamment). La différence fondamentale entre facette et métadonnées, en particulier telles que définies dans les standards de description de ressources pédagogiques, réside dans la manière de l'obtenir et de l'utiliser. Elle se situe dans le fait que les descripteurs contiennent des ensembles de valeurs discrètes,

qui doivent toutes être précisées *a priori*. Les valeurs des facettes peuvent être continues (ou traitées comme telles) et réévaluées en fonction du contexte de la recherche.

Le résultat de la mise en œuvre d'un prisme sur un texte produit une « vue » du texte, destinée à l'enseignant. Ainsi, une vue d'un texte est une instance de ce texte révélant une ou plusieurs facettes, dont les valeurs peuvent être contraintes ou non, dans un contexte pédagogique donné.

Le logiciel EXXELANT (*EXample eXtractor Engine LANguage Teaching*) est un autre type de réponse pour la phase préactive (Antoniadis, Ponton & Zampa, 2010). Il a été développé dans le cadre du projet IDILL (*Integrated DIGital Language Learning*) du réseau européen Kaléidoscope. Il vise l'interrogation et l'exploitation du corpus d'apprenants de FLE (Français Langue Etrangère) FRIDA-bis afin d'en extraire des exemples de phrases comportant certaines caractéristiques, notamment au niveau des erreurs. EXXELANT se destine à trois types d'utilisateurs :

- aux enseignants de langues, pour qu'ils repèrent, par exemple, les difficultés spécifiques des apprenants en fonction du groupe linguistique auquel ils appartiennent, ou les contextes « propices » à tel ou tel type d'erreur ;
- aux chercheurs en TAL, en vue d'élaborer des outils pour la détection et le diagnostic des erreurs ;
- aux apprenants mêmes, pour mettre en évidence leurs erreurs, dans le cadre d'une remédiation par l'enseignant ou pour l'apprentissage par correction des pairs de façon collaborative.

Le corpus est constitué de textes rédigés par des apprenants de FLE ayant des langues maternelles différentes (néerlandais, allemand, japonais, anglais, etc.) et des niveaux de français variables. Il est composé de 764 textes et contient 20 474 erreurs. Au format XML il possède un double système d'annotation : annotation des erreurs à trois niveaux (manuelle), étiquetage (automatique) de l'ensemble des formes (texte, erreurs et correction).

Une interface Web permet son interrogation. L'interrogation peut se faire en précisant des critères comme la langue maternelle des apprenants, la densité d'erreur, la longueur des textes, le type et le contexte d'erreur, etc. L'utilisation d'EXXELANT par l'enseignant donne la possibilité, entre autres, de cerner et quantifier les erreurs typiques des allophones, de visualiser les fréquences d'emploi de mots et de structures, de comparer des productions entre différents groupes d'apprenants, etc. Ces informations peuvent constituer des éléments fort utiles pour élaborer un plan d'enseignement et/ou des activités pédagogiques pertinentes.

À propos de la *phase interactive*, il est à noter que les activités de lecture et d'écriture, concurrentes ou successives, sont les principaux piliers de l'enseignement/apprentissage, dès le secondaire. L'enseignant propose ainsi diverses activités d'apprentissage dans lesquelles l'élève, tour à tour, lit des documents, les comprend, et réalise une activité de production d'écrit (*e.g.*, exercice, prise de notes, synthèse). *Pensum* (système conçu dans le cadre du projet européen LTfLL, *Language Technologies for Lifelong Learning*, voir p. ex. Dupré *et al.*, 2011) est un système qui, tout d'abord, présente à l'élève une série de documents sur un thème donné (choisi par lui ou l'enseignant). L'élève ou l'étudiant lit les documents et, une fois qu'il pense les avoir compris, s'engage librement dans une activité de production d'écrit (en l'occurrence, une synthèse de documents). À tout moment il peut solliciter le système pour avoir une évaluation à trois niveaux, générée par des comparaisons sémantiques par l'analyse sémantique latente (Landauer & Dumais, 1997) :

- *les hors-sujet* : car un cours qui a été mal compris comporte des hors-sujet ;
- le contenu du cours non repris dans la synthèse ;
- *la cohérence inter-phrases de la synthèse*, car des sauts de cohérence sont parfois liés à une argumentation ou compréhension défailtantes.

Une fois ces rétroactions obtenues et analysées, l'utilisateur peut librement apporter des modifications à sa production, quitter le mode écriture pour approfondir la lecture des documents, mais aussi ajuster la sévérité des rétroactions. Enfin, et ce point est lié à la question de la réflexivité, l'utilisateur apprenant peut contester les rétroactions de l'EIAH et peut les justifier (p. ex. en faisant le raisonnement et les actions suivants : la machine a signalé un hors-sujet, mais je sélectionne une section du document de cours qui est reliée à ma production : je n'ai donc pas fait de hors-sujet).

Dans le contexte de recherche actuel, si on trouve bon nombre de travaux concernant *la phase post-active* (en particulier par l'exploitation des traces des apprenants), les indicateurs calculés et/ou fournis à l'enseignant par les différents systèmes manquent souvent de pertinence didactique. L'exposé d'Alexandre Baudrillart, et sa thèse qui vient de débiter, tentent d'apporter une réponse. Force est néanmoins de constater, que l'automatisation de cette phase comporte, sans doute, le plus d'écueils.

L'intégration de techniques de TAL dans les EIAH pose nombre de questions qui attendent réponse et qui seront discutées.

- Quelles sont les « niches » didactiques les plus appropriées pour l'utilisation du TAL ?
- Comment traduire les données et résultats du TAL en termes et concepts didactiquement pertinents ?
- Faut-il construire de systèmes TAL spécifiques pour les EIAH ? L'exposé et le système de Mourad Mars, apporte une première réponse à cette question.
- Quel type d'interaction promouvoir et sur quel mode ? Le mode sur lequel les rétroactions sont délivrées (e.g., « Tu t'es trompé » ou « je pense que ») ne sont pas sans conséquences sur la compréhension que l'utilisateur se fait du système et de ses possibilités.
- Quelle transparence de fonctionnement offrir à l'utilisateur ? Le fonctionnement des techniques de TAL est souvent complexe et fournir des niveaux d'explication trop ou trop peu détaillés peut rendre l'utilisation de ces EIAH opaque. De plus, le manque de transparence dans le fonctionnement du système induit des attentes de l'utilisateur infondées (p. ex., certains utilisateurs attendaient de *Pensum* qu'il corrige leur style).
- Quelle est l'issue des boucles d'activités (p. ex., lecture, écriture) dans lesquelles l'utilisateur est engagé ? Il est manifeste que c'est uniquement à ce dernier d'estimer la qualité du résultat des activités réalisées, et non au système. C'est également à lui seul de déterminer quand il peut changer d'activité.

Si le TAL peut indéniablement apporter des solutions pour les systèmes d'EIAH (Antoniadis, 2008), son utilisation soulève des interrogations et des défis aussi bien pour l'association didactique/EIAH/TAL que pour lui-même. Sa place et sa nature précises restent à déterminer. Néanmoins, pour nos systèmes comme pour d'autres, certaines techniques de TAL intégrées dans des EIAH peuvent être utilisées pour guider leur utilisateur dans des activités réflexives, en boucles, pouvant favoriser leur apprentissage.

Références

- Antoniadis, G. (2008), Du TAL et son apport aux systèmes d'apprentissage des langues : Contributions, mémoire de l'habilitation à diriger des recherches, juin 2008, université Stendhal de Grenoble, France
- Antoniadis G., Ponton C., Zampa V. (2010). Exxelant et Mirto : Deux exemples d'environnement d'ALAO intégrant des outils TAL. In Ismaïl Biskri, Adel Jebali (eds.) *Multilinguisme et traitement des langues naturelles* (pp. 151-166). Montréal : PUQ.
- Dessus, P., Lemaire, B., Loiseau, M., Mandin, S., Villiot-Leclercq, E., & Zampa, V. (2011). Automated free-text assessment: Some lessons learned. *International Journal of Continuing Engineering Education and Life-Long Learning*, 21(2/3), 140–154.
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: the Latent Semantic Analysis theory of acquisition, induction and representation of knowledge. *Psychological Review*, 104(2), 211–240.

- Levy, M. (1997). *Computer-Assisted Language Learning, context and conceptualization*. Oxford : Oxford University Press.
- Loiseau, M. (2009), Elaboration d'un modèle pour une base de textes indexée pédagogiquement pour l'enseignement des langues, thèse de Doctorat, université Stendhal de Grenoble, décembre 2009
- Loiseau, M., Dupré, D., & Dessus, P. (2011). *Pensum*, un système d'aide à la compréhension de cours à distance. In M. Bétrancourt, C. Depover, V. Luengo, B. D. Lièvre & G. Temperman (Eds.), *Conférence EIAH 2011* (pp. 287–299). Mons: Presses de l'université de Mons.
- Riff, J., & Durand, M. (1993). Planification et décision chez les enseignants. *Revue Française de Pédagogie*, 103, 81–107.
- Vovides, Y., Sanchez-Alonso, S., Mitropoulou, V., & Nickmans, G. (2007). The use of e-learning course management systems to support learning strategies and to improve self-regulated learning. *Educational Research Review*, 2(1), 64–74.