

Vers un environnement-tuteur d'apprentissage dialogique

Philippe Dessus, Stefan Trausan-Matu, Virginie Zampa, Traian Rebedea, Sonia Mandin, Mihai Dascalu

▶ **To cite this version:**

Philippe Dessus, Stefan Trausan-Matu, Virginie Zampa, Traian Rebedea, Sonia Mandin, et al.. Vers un environnement-tuteur d'apprentissage dialogique. 2e Colloque Echanger pour Apprendre en Ligne (EPAL'09), Jun 2009, Grenoble, France. hal-00404842

HAL Id: hal-00404842
https://hal.science/hal-00404842

Submitted on 17 Jul 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

VERS UN ENVIRONNEMENT-TUTEUR D'APPRENTISSAGE DIALOGIQUE

Philipe Dessus*, Stefan Trausan-Matu, Virginie Zampa***, Traian Rebedea**, Sonia Mandin*, Mihai Dascalu****

*** Laboratoire des sciences de l'éducation (EA 602), Université Pierre-Mendès-France & IUFM, Université Joseph-Fourier, Grenoble**

**** « Politehnica » Université de Bucarest, Roumanie**

***** LIDILEM, Université Stendhal, Grenoble**

Résumé : Cet article présente un environnement-tuteur d'apprentissage pouvant s'insérer dans une plate-forme de FOAD permettant à des étudiants de prendre connaissance et de réviser des contenus de cours. Cet environnement puise son originalité dans le fait qu'il allie une approche cognitive et des références à l'apprentissage dialogique pour permettre : — de fournir aux apprenants les cours les plus pertinents à étudier étant donné leur compréhension du cours ; — de délivrer des *feedback* sur la compréhension d'un cours et le contenu de résumés produits sur ce même cours ; — de renvoyer des *feedback* sur les fils de discussions menées dans des *chats.*

Mots-clés : apprentissage à distance, apprentissage dialogique, analyse de la sémantique latente, modélisation cognitive, *feedback*, activité de résumer

1. Introduction

Les plates-formes actuelles d'enseignement à distance assistent un grand nombre d'activités d'apprentissage et d'enseignement, qui ont d'abord été clairement identifiées et catégorisées (Laurillard, 1993) et font, plus récemment, l'objet de scénarios de plus en plus précis et cohérents (Paquette, 2005). Pour autant, deux points importants nous paraissent à étudier. D'une part, trouver un cadre théorique commun permettant d'unifier les multiples activités de l'élève qui, devant faire preuve d'un haut niveau de littéracie (Bereiter & Scardamalia, 1987), requièrent d'utiliser lecture, écriture et discussion dans un même ensemble d'habiletés en vue de construire des connaissances. D'autre part, déterminer comment ces activités sont compatibles avec une posture réflexive, dont on sait qu'elle peut favoriser une expertise, soit un « processus dialogique dans le réflexif » (Chanier & Cartier, 2006). Des auteurs ont montré que ces activités n'étaient pas toujours bien intégrées dans les plates-formes d'enseignement (Nussbaumer, Steiner & Albert, 2008), mais aussi que l'activité métacognitive elle-même n'y était pas toujours possible (Vovides, Sanchez-Alonso, Mitropoulou & Nickmans, 2007).

La perspective que nous développerons ici est celle de l'apprentissage dialogique (Bakhtine, 1986). Cette perspective permet de rendre compatible de nombreux éléments souvent dissociés dans la littérature. Tout d'abord, elle offre un cadre intégratif des activités de l'élève : tout élève, lorsqu'il dit, lit, écrit ou pense quelque chose, tisse des liens avec le discours des autres (Koschmann, 1999). De plus, l'autorégulation des activités engagées dans l'apprentissage peut se faire à un niveau individuel ou collectif.

Dans une première partie, nous allons présenter le contexte, c'est-à-dire l'idée que l'apprentissage se réaliste tout au long de la vie et nécessite pour cela des outils spécifiques et la théorie dialogique de l'apprentissage. Nous verrons ensuite les outils médiateurs de cet apprentissage ainsi que deux systèmes existants, à partir desquels notre environnement va être implanté. Pour finir, nous détaillerons notre projet d'environnement d'apprentissage dialogique.

2. Contexte et théorie dialogique de l'apprentissage

2.1 L'apprentissage tout au long de la vie dans un dispositif de FOAD

Commençons par expliciter le contexte dans lequel le système que nous projetons d'implanter va pouvoir s'insérer. Les plates-formes actuelles de FOAD (formation ouverte et à distance) ou d'enseignement exclusivement à distance ont des caractéristiques particulières qui autorisent le travail d'étudiants dans un contexte d'apprentissage *tout au long de la vie*. Les étudiants travaillent souvent dans un contexte de travail collaboratif, pour diminuer leur sentiment de solitude (Fischer, 2007 ; Graham & Perin, 2007). De plus, un accès permanent aux cours est préférable. Ensuite, comme beaucoup d'étudiants sont également engagés dans des activités professionnelles, il est nécessaire de mixer des interactions en présence et à distance. Enfin, les étudiants ont différents niveaux de connaissances et compétences, et une aide informatisée doit les prendre en compte. Linn (1996), dans le contexte plus large de l'enseignement à distance, a réalisé une liste complète de compétences que des étudiants engagés dans un contexte académique doivent posséder. Cette liste nous paraît transférable à toute situation d'apprentissage tout au long de la vie :

– Pouvoir déterminer quand, comment et pourquoi les étudiants apprennent tel ou tel nouveau contenu ;

- Pouvoir analyser et déterminer leurs points forts et limites ;
- Pouvoir sélectionner des activités compatibles avec leurs buts, points forts et limites ;
- Pouvoir développer des projets individuels qui correspondent à leurs buts personnels d'apprentissage ;
- Pouvoir prendre en charge leur propre apprentissage ;
- Pouvoir monitorer leur propre progrès ;
- Pouvoir réfléchir sur leur compréhension du matériel enseigné ;
- Pouvoir solliciter l'aide de leurs pairs et de leurs enseignants ;
- Pouvoir mettre en œuvre des activités permettant de pratiquer les habiletés acquises ;
- Pouvoir assimiler les connaissances, pratiques et culturelles, d'un contenu enseigné par des activités d'apprentissage autonome, mais avec un guidage possible (*e.g.*, comprendre les relations entre idées, comparer des alternatives, réfléchir sur son progrès, critiquer des idées) ;
- Pouvoir structurer leur cours en tirant avantage de la nature sociale de l'apprentissage en s'engageant dans un apprentissage collaboratif.

Cette liste met en valeur deux principales capacités nécessaires à tout protagoniste d'une situation d'apprentissage tout au long de la vie. L'apprenant a à adopter une attitude autonome dans son apprentissage, pendant que l'enseignant procure à ce dernier une attitude étayant l'intégration de ses connaissances (*i.e.*, essayer de développer les idées de l'étudiant, aider ce dernier à réfléchir, organiser, connecter ces idées dans une perspective cohérente). Les deux notions d'apprentissage autorégulé (amenant à une compréhension individuelle) et régulé par les pairs, ou collaboratif (amenant à une compréhension collective) sont ainsi centrales. Ces deux points sont les deux principales facettes du processus de construction de connaissances, tel qu'élaboré par (Stahl, 2006).

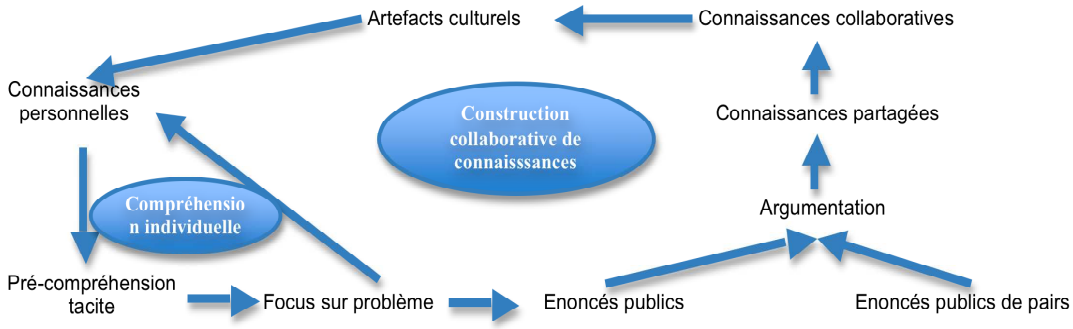


Figure 1 — Processus de construction de connaissances (Stahl, 2006, p. 203)

Dans de tels dispositifs de FOAD, les apprenants construisent des connaissances selon deux processus interreliés (*voir* Figure 1). Un processus individuel, dans lequel ils comprennent progressivement et itérativement un domaine de connaissances par compréhension tacite puis focalisation particulière de leur attention sur un problème à résoudre. Ce problème, une fois formulé en mots, est repris par d'autres apprenants, qui, après discussion et clarification du sens, s'en emparent et en ont une compréhension partagée. Cette double boucle peut fonctionner : — s'il existe des artefacts culturels (*i.e.*, plates-formes, outils) qui la supportent ; — si l'activité d'autorégulation de l'apprentissage évoquée plus haut s'y manifeste. Dans la suite de cet article, nous allons nous focaliser sur deux types d'artefacts et montrer comment réfléchir à leur intégration dans une plate-forme : un artefact promouvant la compréhension individuelle, qui permet à tout étudiant de lire et comprendre des textes issus de cours ; un artefact promouvant la compréhension collective, via un dispositif de *chat*.

2.2 Théorie dialogique de l'apprentissage

Prenons un exemple concret. Un étudiant engagé dans un cours en FOAD doit réaliser un certain nombre d'activités que nous pourrions qualifier de « dialogiques ». Il a en effet à tisser de nombreux liens entre les différents types de discours produits (questions/réponses, texte du cours, exercices, etc.) dans le contexte du cours par les autres protagonistes (enseignants, tuteurs). Ainsi, la compréhension d'un passage d'un cours ne peut se faire de manière isolée, mais des liens doivent être posés entre ce dernier et d'autres éléments antérieurs/postérieurs à ce passage.

Nous empruntons à Bakhtine (1986) sa célèbre théorie socioculturelle montrant que tout texte est dialogique. Le principe est que les activités de lecture, d'écriture, de discussion et d'apprentissage sont autant de « textes », oraux ou écrits, largement interdépendants et concourant à faire exprimer ou étudier par l'élève différentes « voix » (ou types de connaissances) exprimées dans ces textes. Elles consistent à établir des liens entre les propositions lues (*e.g.*, cours) ou écrites (*e.g.*, textes résumés, questions) et les connaissances antérieures des élèves. Ainsi, toute portion de texte lu ou écrit invite à une interaction dialogique entre son auteur et le lecteur, même dans le cas où ils sont une seule et même personne. Elles contribuent toutes à établir chez les élèves des processus générateurs de compréhension (Wittrock, 1990), et par là de construire des connaissances en collaboration (Bereiter, 2002). Comme l'a avancé Volochinov, le collègue de Bakhtine : « Toute vraie compréhension est dialogique en nature » (Volochinov, 1973, p. 102, cité dans Morris, 1994, p. 11).

Une question centrale dans toute théorie sur l'apprentissage est « comment les apprenants acquièrent-ils des connaissances ? ». La réponse à cette question est cruciale pour concevoir des environnements qui puissent accompagner les activités d'apprentissage. Selon les paradigmes de recherche, de nombreuses réponses différentes ont été apportées à cette question : les sciences cognitives sont centrées sur l'idée que toute connaissance peut être décrite, stockée dans une base de connaissances et transférée dans un récipient, qu'il soit un humain ou une machine. La centration, dans ce cas, est sur l'esprit des élèves. Le champ du CSCL (*Computer-Supported Collaborative Learning*), en revanche, est fondé sur l'idée que la connaissance est socialement partagée (Vygotski, 1997) et ensuite internalisée. L'apprentissage y est ainsi un processus social, médiatisé par certains outils symboliques, dans lesquels le langage humain joue un rôle central. Chaque apprenant est un agent actif qui cherche à acquérir des connaissances dans un contexte qui a un sens pour lui.

Sfard (1998) a montré que l'apprentissage pouvait être considéré, de manière également pertinente, comme une acquisition (au sens matériel) de connaissances, ou comme la participation à un discours, dans lequel l'étudiant entre dans une communauté de pratiques dans lequel il devient de plus en plus expert. Cette métaphore a été particulièrement investie dans le champ de la recherche en CSCL (Koschmann, 1999 ; Stahl, 2006 ; Wegerif, 2007). En ce sens, l'apprentissage peut être vu comme directement relié à la construction d'un discours : « L'apprentissage est la co-construction (ou reconstruction) de significations sociales à partir des paramètres d'une activité discursive émergente et socialement négociée » (Hicks, 1996, p. 136)

Il est donc nécessaire de concevoir un « espace » dans lequel ces différentes perspectives peuvent lui être présentées, et qui permet à l'apprenant à la fois de faire le lien avec ses connaissances antérieures (autorégulation interne) et d'interagir avec ses pairs (autorégulation externe). Chacun de ces deux niveaux de compréhension (individuelle, collective) interagit l'un l'autre et chacun est nécessaire à l'autre. Les problèmes que l'étudiant se pose à un niveau individuel (*e.g.*, ai-je bien compris ce que j'ai lu ? quelle est la qualité de ma compréhension ? où faut-il chercher ce que je dois lire ensuite ? ai-je écrit un résumé acceptable du cours ?) peuvent être transférés au niveau collectif et *vice versa*,

l'environnement peut-être analysé à un niveau individuel en favorisant le transfert des problèmes et questions de l'étudiant au niveau collectif.

Aussi, il est important de développer la perspective dialogique, même au niveau individuel. Par exemple, l'étudiant ayant pour but de comprendre un cours doit être habitué à considérer de possibles questions pouvant être posées par ses pairs et, bien sûr, leurs possibles réponses ou les dialogues explicatifs. Nous pensons donc que l'utilisation de sessions de *chat* pour la compréhension et négociation *collaboratives* peut également avoir comme effet latéral l'induction de cette perspective dialogique *individuelle* (*voir* Figure 1).

3. Outils médiateurs d'apprentissage

Dans cette section nous décrivons les outils permettant de faire le lien entre l'individuel et le collectif. Tout d'abord, les outils médiateurs que sont l'écriture et l'informatique, puis les environnements que nous avons précédemment implantés qui, se centrant sur l'activité d'écriture (résumé de cours et *chat*) pour l'apprentissage.

3.1 Deux outils médiateurs privilégiés : l'écriture et l'informatique

L'une des caractéristiques importantes de l'apprentissage, qu'il soit individuel ou collectif, est qu'il peut être facilité, médiatisé, par des facteurs internes ou externes. Les dispositifs de FOAD et d'enseignement à distance figurent bien sûr parmi ces derniers. Nous nous centrons ici sur deux outils médiateurs principaux : l'écriture et les outils informatisés, et les relions à leur effet sur la construction de connaissances.

Dans le contexte de l'apprentissage tout au long de la vie décrit plus haut, les liens entre l'activité d'écriture et d'apprentissage sont cruciaux, et doivent à notre avis être recherchés, développés, ou encore évalués. Même si tout programme d'enseignement universitaire utilise intensivement l'écriture, le lien entre cette activité et la compréhension ou l'apprentissage n'est pas toujours explicite. La mise en évidence d'effets bénéfiques de l'écriture sur l'apprentissage n'est pas non plus évident, tant cette activité est intégrée à la démarche même d'apprentissage (prise de notes, évaluations, etc.). Toutefois, la question de savoir comment des textes écrits (*i.e.*, essais) sont adéquats comme bases d'évaluation des connaissances acquises par les étudiants a été posée (Valenti, Neri & Cucchiarelli, 2003 ; Phillips, 2007).

Certains auteurs (Emig, 1977 ; Bangert-Drowns, Hurley & Wilkinson, 2004) ont mis en avant les correspondances entre écriture et apprentissage. Ces deux activités — émergent de processus cognitifs symboliques ; — sont permises et favorisées par les *feedback* autoévaluatifs ; — utilisent ou procurent des connexions entre les concepts en jeu ; — sont efficaces si elles sont délibérées et autorégulées. Ces similitudes font de l'écriture une activité centrale dans l'apprentissage, non seulement pour évaluer ce dernier, mais aussi pour engager les étudiants à réfléchir sur leurs propres stratégies d'apprentissage. Écrire à propos d'un contenu académique donné : — permet sa révision ; — autorise une compréhension plus approfondie ; — entraîne un monitorage réflexif. Ainsi, l'écriture est une activité suffisamment générique pour être une source et un indicateur du processus d'apprentissage ou de construction de connaissances. Notre but est de concevoir un environnement d'apprentissage informatisé qui guide entièrement cette activité et son évaluation.

Selon la perspective bakhtinienne décrite plus haut, le *feedback* peut être considéré comme une autre sorte de dialogue, ou un catalyseur du dialogue, introduisant la présence d'une autre personne, d'une autre voix. Il n'est pas rare, dans un dispositif de FOAD, que les étudiants se plaignent du manque d'interaction et de *feedback* de la part de leurs tuteurs ou enseignants, mais aussi que ces derniers considèrent ces demandes comme intrusives (Villiot-Leclercq &

Dessus, 2009). Les *feedback* sous forme de dialogues sont donc partie intégrante de tout processus d'interaction humaine.

Les outils informatisés sont bien sûr l'un des moyens de favoriser ce *feedback* et cet apprentissage autorégulé (Vovides, Sanchez-Alonso, Mitropoulou & Nickmans, 2007; Winters, Greene & Costich, 2008). Stahl (2006) a montré que ces artefacts faisaient partie de l'ensemble des outils médiateurs entre la compréhension individuelle et collective. Au même titre que la discussion elle-même, stockée dans un système externe, ces outils permettent à la fois l'articulation individuelle de connaissances et l'échange, la négociation collective d'arguments menant à une construction collaborative de connaissances (*voir* Figure 1). Décrivons maintenant plus précisément de tels outils médiateurs.

3.2 Les systèmes existants : *Apex* et *Polyphony*

Nous avons montré que, dans un contexte d'apprentissage tout au long de la vie, la compétence d'apprentissage autorégulé est centrale. Notre question de recherche est la suivante : quel flux d'activités proposer à l'étudiant pour favoriser cette compétence ? Comme la production d'autorégulations nécessite une charge cognitive et une implication importantes de la part des enseignants et des tuteurs, peut-on automatiser en partie leur évaluation et leur gestion ? Nous avons conçu, indépendamment l'un de l'autre, deux systèmes informatiques incorporant des techniques de TAL, l'un favorisant une autorégulation individuelle, l'autre sur une autorégulation collective.

Apex (Dessus & Lemaire, 2002 ; 2004) est un système qui engage un élève dans un contexte de révision de cours et de réalisation de résumés. Il s'appuie sur un jugement de compréhension par l'élève de textes lus. *Apex* fonctionne de la façon suivante. L'apprenant donne des mots-clés sur les thèmes ou notions qu'il souhaite travailler. LSA (*Latent Semantic Analysis*), dont nous expliquons le principe plus loin, sélectionne dans une base de textes, les textes sémantiquement les plus proches de ses mots-clés. Ensuite, l'interaction de l'élève avec le système comprend deux phases, effectuées en boucle : une phase de lecture de textes, et une phase de production d'un résumé des textes lus. Après chaque lecture de texte, l'élève doit indiquer s'il lui semble possible de résumer le texte. Lorsqu'il le souhaite, l'élève passe à la phase de production : il doit résumer, un par un, l'ensemble des textes qu'il a estimé avoir suffisamment compris pour le faire. C'est là qu'intervient pleinement LSA : la comparaison sémantique entre les résumés produits et les textes d'origine permet de montrer à l'élève l'écart entre ce qu'il a cru avoir compris et ce qu'il a vraisemblablement compris. Pour réduire cet écart, l'élève peut de nouveau retravailler son résumé et pour cela peut aussi relire le texte source. À tout moment, il peut de nouveau repasser à la phase lecture où de nouveaux textes, là encore sélectionnés par LSA en fonction de ce qu'il a déjà lu et de ce qu'il est supposé avoir compris, lui sont proposés. La Figure 2 montre un des textes lus par l'étudiant (écran de gauche) et le résumé d'un élève (écran de droite). La Figure 3 montre les commentaires d'*Apex* à ce sujet.

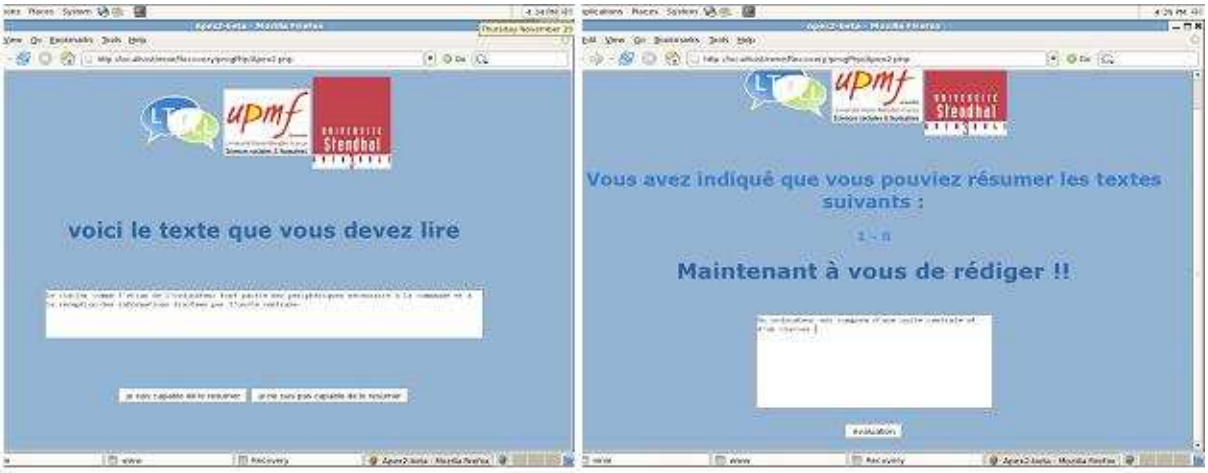


Figure 2 – Copie d'écran d'*Apex*. Image de droite : boucle de lecture. Image de droite : boucle d'écriture.

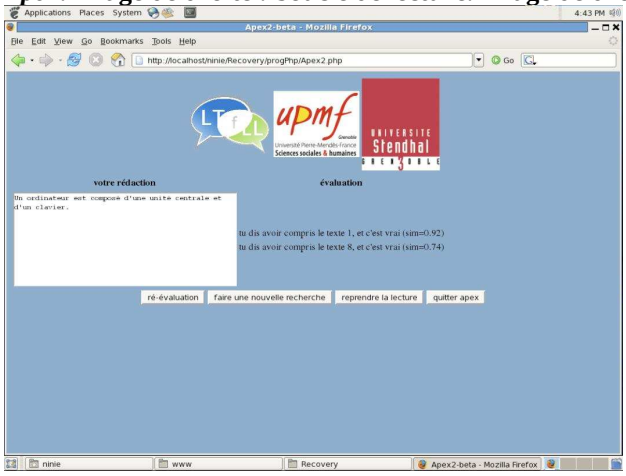


Figure 3 – Copie d'écran d'*Apex*, évaluation du résumé.

LSA (*Latent Semantic Analysis*, Landauer & Dumais, 1997) est un procédé de traitement automatique des langues naturelles. Le principe est le suivant. Dans un premier temps, LSA construit une matrice d'occurrences à partir des textes qui lui sont fournis (une ligne par terme, une colonne par paragraphe). Puis la matrice est réduite par une décomposition en valeurs singulières. Cela permet de représenter les textes et les graphies dans un espace vectoriel à environ 300 dimensions. Chaque graphie, chaque paragraphe et chaque texte sont représentés par un vecteur dans cet espace multidimensionnel. La proximité entre deux textes, deux paragraphes, deux termes est fournie par la valeur du cosinus de l'angle que forment leurs vecteurs dans l'espace.

Le système conçu par Trausan-Matu et son équipe (Trausan-Matu & Stahl, 2007) a un but de visualisation des contributions des apprenants dans un environnement de chat existant (*Virtual Maths Teams*). Cet environnement fait partie d'un projet plus large, *Math Forum*, dans lequel des élèves (de collège et lycée), discutent et tentent de résoudre divers problèmes de mathématiques et peuvent recevoir des *feedback* de tuteurs ou d'enseignants. Différents services en ligne sont offerts (Stahl, sous presse) : — le « problème de la semaine », qui pose, semaine après semaine, un nouveau problème que les élèves peuvent résoudre et dont les meilleures solutions seront publiées, — « *Demande à Dr Maths* » qui est un service de demande d'aide en mathématiques, — l'accès à divers outils mathématiques, — un forum de discussion à l'intention des enseignants, — l'organisation d'ateliers de formation des enseignants, — et l'outil de chat VMT, structuré en différentes « pièces » thématiques, permettant d'aider à résoudre les « problèmes de la semaine ». Trausan-Matu et Rebedea

(sous presse) ont ajouté à ce système de chat classique un outil permettant de structurer et de visualiser les différents fils de discussion. À tout moment, les élèves peuvent relier deux énoncés par un lien explicite (*voir* Figure 4), montrant le lien entre les idées de ces énoncés. De plus, un second type de lien, implicite (Figure 5), est automatiquement calculé par le système (Figure 6), selon une technique décrite dans Trausan-Matu, Rebedea, Dragan et Alexandru (2007). Ces deux types de liens permettent de visualiser les fils de discussion et de quelle manière la discussion évolue au cours des énoncés. De plus, une évaluation automatique de la « force » de la contribution de chaque élève est également réalisée, à partir de l'influence supposée de chaque énoncé (l'influence d'un énoncé est d'autant plus grande que des étudiants s'accordent sur ce dernier, et s'il fait référence à un énoncé lui-même important).

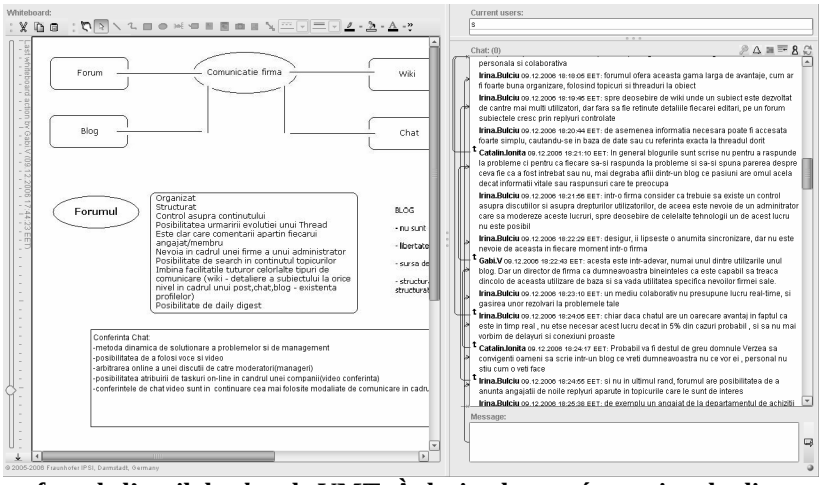


Figure 4. Interface de l'outil de *chat* de VMT. À droite, la représentation des liens explicites.

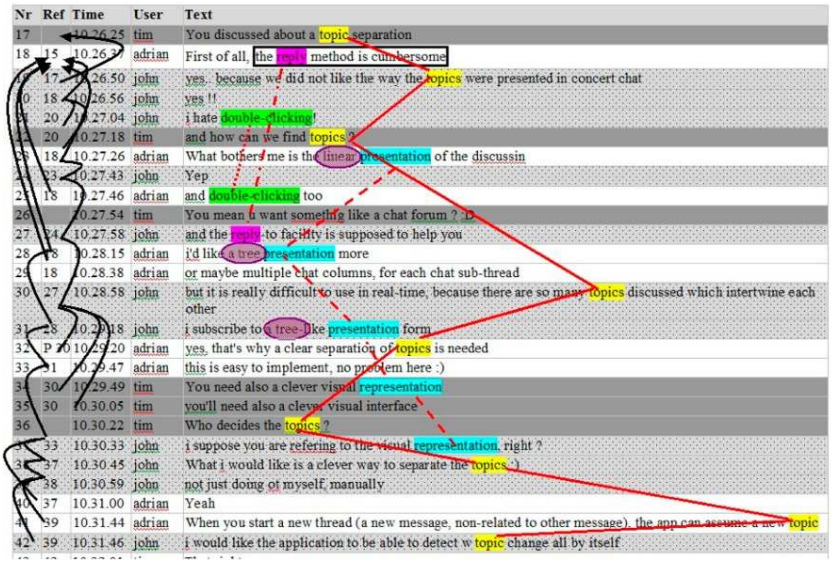


Figure 5 — Représentation visuelle d'une discussion et visualisation des liens explicites (lignes courbes) et implicites, détectés automatiquement (les liens entre mots répétés sont représentés avec des lignes droites).

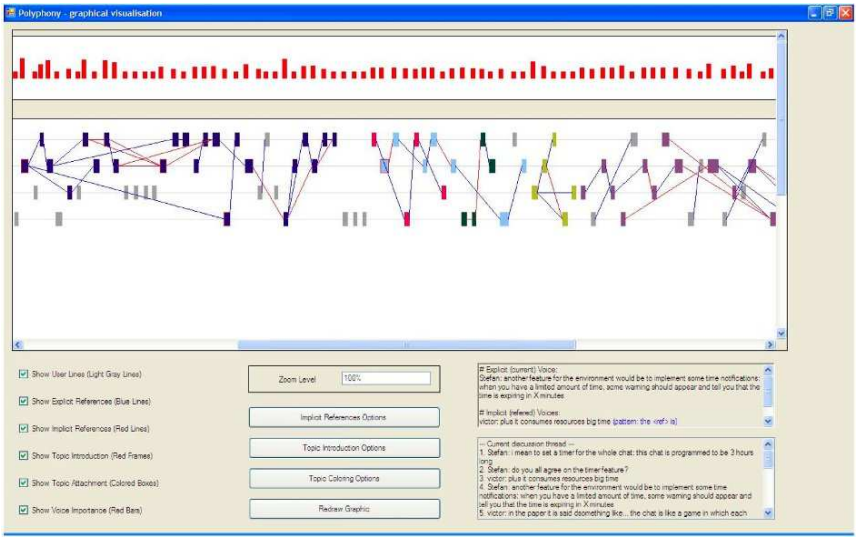


Figure 6 — Copie d'écran de l'interface de *Polyphony*. Représentation visuelle des énoncés de quatre étudiants et des interrelations polyphoniques (explicites, en bleu ; implicites, en rouge). La largeur du rectangle code la taille de l'énoncé en caractères.

4. Présentation de notre projet d'environnement d'apprentissage dialogique

Il s'agit ici d'un projet d'environnement-tuteur pouvant s'insérer dans une plate-forme d'enseignement à distance. Notre but est d'utiliser conjointement les deux outils précédemment présentés afin de permettre un apprentissage dialogique. En effet, un apprenant peut lire ou rédiger tout seul, mais à tout moment il peut aussi discuter avec ses pairs de différentes notions abordées afin de confronter son point de vue à celui des autres. L'apprenant va ainsi pouvoir développer ses connaissances en créant des liens entre des notions. Il fera émerger ces liens en utilisant les deux outils. Ce service utilisera intensivement des techniques d'analyse automatique des langues, notamment l'analyse du discours et l'analyse sémantique latente (LSA, Landauer & Dumais, 1997). Comme vu plus haut, LSA détecte les régularités de l'utilisation des mots dans des paragraphes pour en déterminer une partie du sens. Ainsi, tout texte de la plate-forme peut être analysé par ce service et ce dernier peut procurer un certain nombre d'indications et de *feedback* à l'étudiant. LSA peut être utilisé pour faire ressortir les liens entre les cours, les discussions, les notes prises, etc. Un des buts importants de notre recherche est de montrer comment on peut utiliser LSA pour détecter automatiquement des « liens dialogiques » au sein de discours en contexte d'enseignement-apprentissage. Nous en ébauchons plus bas quelques pistes, en partant des liens sémantiques entre mots, entre les répliques d'un *chat* mais aussi, plus généralement, en essayant de rendre compte de l'intertextualité (Bazerman, 2004).

4.1 Fonctionnement général de l'environnement dialogique

Un étudiant connecté à un tel système peut librement travailler dans les boucles d'apprentissage décrits plus bas et exercer également ses capacités d'auto-régulation (Figure 7). De plus, il peut à tout moment changer de boucle. S'il peut bien évidemment poser des questions (à ses pairs) dans la troisième boucle, il est à noter qu'il lui est également possible de s'en poser dans les deux premières.

– *Boucle de lecture* : Après avoir saisi une requête, l'étudiant lit une série de textes de cours proposés par l'environnement-tuteur. À tout moment, l'étudiant peut auto-évaluer son niveau de compréhension du texte, mais aussi raffiner sa requête pour obtenir de nouveaux textes.

– *Boucle d'écriture* : L'étudiant écrit un résumé ou une synthèse du ou des texte(s) qu'il a compris. À tout moment, il peut demander l'évaluation automatique de certaines caractéristiques du cours ou de son texte (*e.g.*, la détection des idées principales, le diagnostic des possibles stratégies utilisées pour résumer).
– *Boucle de discussion en ligne* : L'étudiant interagit avec ses pairs ou l'enseignant, afin de discuter des points précédemment lus ou rédigés. Pour cela, il utilise un système de messagerie instantanée de type *chat*, qui permet de lier chaque réplique aux précédentes (Holmer, Kienle & Wessner, 2006) et, en conséquence, permet la mise en évidence des multiples fils de discussion en parallèle, où plusieurs voix (selon Bakhtine) interagissent de manière polyphonique. À la fin des discussions, les participants peuvent visualiser les fils par l'affichage d'un graphe et recevoir des mesures qui quantifient les contributions de chacun.

Ce système est en cours d'implémentation et certains de ses composants sont en cours de test en tant que prototypes, dans les domaines de l'informatique et de la médecine.

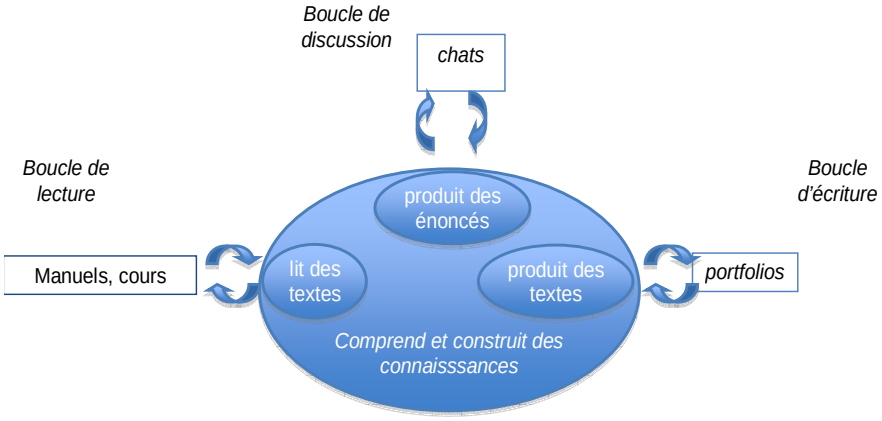


Figure 7 — Représentation des trois boucles d'activités de notre projet d'environnement d'apprentissage dialogique.

4.2 Architecture de notre environnement

Nous présentons maintenant l'architecture de notre environnement en décrivant les différentes étapes et en détaillant la façon dont les données collectées au cours des trois boucles sont analysées (Figure 8).

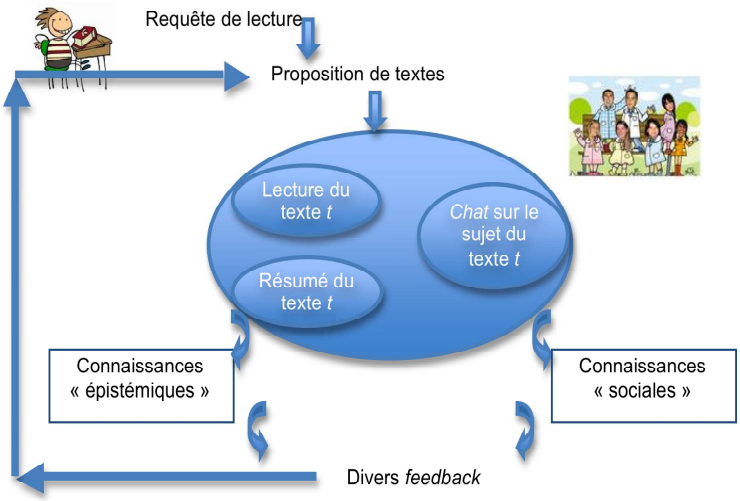


Figure 8 — Architecture de notre projet d'environnement d'apprentissage dialogique

Notre environnement est fondé sur l'architecture d'un tuteur intelligent (Wenger, 1987). Il comprend donc un modèle du domaine, un modèle de l'apprenant et un modèle pédagogique. Ces trois modules interviennent dans chacune des trois boucles. De plus, différents types de *feedback* pourront être présentés aux apprenants (Dessus, Mandin & Zampa, 2008).

Le modèle du domaine est représenté par les cours mis par les enseignants et des textes additionnels afin que les apprenants puissent approfondir les connaissances qu'ils désirent. Des textes plus généraux sont aussi ajoutés à ce corpus afin d'obtenir des connaissances plus précises sur la langue. Tous ces textes sont traités par LSA afin d'obtenir un espace sémantique multidimensionnel permettant ainsi de calculer des proximités entre les différents éléments textuels. Afin de s'adapter à l'apprenant, le système à besoin d'une représentation de ce dernier, *un modèle de l'apprenant* qui évoluera en fonction de ce qu'il lit et écrit mais aussi en fonction des *feedback* qui lui sont renvoyés. Dans un premier temps, le système établit le modèle de référence de l'étudiant (sorte d'initialisation) à partir d'un large corpus de textes généraux (*e.g.*, à partir de journaux ou d'entrées encyclopédiques). Cela permet de se doter d'un modèle générique de l'étudiant. Au fur et à mesure des utilisations, le modèle est mis à jour en fonction de ce que l'apprenant lit, de ce qu'il rédige (et des *feedback* associés), mais aussi en fonction de ses participations aux différents *chats*. Pour cela, l'espace sémantique est recalculé avec LSA en fonctions des traitements effectués sur les textes. Si un contenu donné d'un cours est lu *et* résumé *ou* discuté entre pairs, les textes correspondants au cours et à la discussion sont ajoutés dans le modèle avec une pondération de 2 (*i.e.*, ils sont ajoutés deux fois dans le corpus avant recalcul de l'espace sémantique). En revanche, si un cours est seulement lu, il sera ajouté avec une pondération à 1 (ce qui peut avoir un impact moins grand sur les relations sémantiques entre mots du cours que pour une pondération de 2). Nous pensons améliorer la compréhension de l'apprenant en lui donnant des occasions d'exprimer ce qu'il a compris plutôt qu'en lui demandant de relire les mêmes textes : ce qui est lu est deux fois (valeur arbitraire) mieux compris si cela est résumé ou discuté entre pairs. Les informations du modèle de l'apprenant (*i.e.*, les informations épistémiques, stratégiques et sociales) pourraient être présentées à ce dernier sous la forme de cartes conceptuelles de connaissances, incluant les thèmes discutés dans les *chats.*

Enfin, *le modèle pédagogique* comporte deux sous-parties : la sélection des cours et le *feedback*. Cette sélection peut se faire de deux manières :

- *par requête formulée par l'apprenant*. Ce dernier peut décider de travailler sur un thème donné. Pour cela il lui suffit de formuler une requête en langage naturel et LSA sélectionne les textes du cours les plus proches de cette requête.
- *en fonction du travail de l'apprenant*. Les textes peuvent être sélectionnés par le système en fonction de ce que l'apprenant a lu et compris/non compris. Par exemple, si une notion est entièrement comprise, LSA lui fournit des textes qui ne sont ni trop proches ni trop éloignées de ce qu'il connaît afin d'optimiser son apprentissage (Zone de développement proche (Vygotski, 1997). En revanche, pour les notions qui ne sont pas comprises, LSA choisit les textes les plus proches qui l'orienteront peut-être vers un contenu sensiblement différent. Enfin les textes peuvent être sélectionnés par LSA en fonction des discussions dans le *chat* en complément des thèmes abordés.

Des *feedback* sont générés à partir des différentes actions de l'apprenant dans l'environnement. Pour chacun d'eux LSA est utilisé pour indiquer les liens sémantiques entre les différents éléments textuels (cours, résumés, *chats*). Nous distinguons des *feedback* :

- *Entre thèmes de discussion* (dans le *chat*). Le système permet une visualisation des différentes interventions des apprenants inter- et intrafil(s) de discussion portant sur les mêmes thématiques.

- *Entre thèmes de discussion et cours*. De la même façon, le système permet de faire ressortir les liens sémantiques entre les fils de discussion et les cours et de proposer ces derniers à la lecture.
- *Entre textes lus par chaque apprenant*. Ce qui permet à l'apprenant de visualiser les liens entre les différentes parties des cours lus.
- *Entre résumés produits et cours lus*. Il s'agit ici de l'évaluation du contenu sémantique des résumés produits.
  Entre apprenants, selon ce qu'ils ont lu. Ce *feedback* permet le regroupement des apprenants ayant travaillés sur les mêmes cours afin de les inciter à discuter sur le *chat* des différentes notions abordées. Le système indique à chaque apprenant ceux qui ont compris le cours et ceux qui ne l'ont pas compris.

5. Conclusion

Nous avons décrit dans cet article un projet d'environnement-tuteur d'apprentissage dialogique. La conception de cet environnement repose sur une vue intégrée des différentes activités possibles dans les plates-formes de FOAD : lire des contenus de cours, en réaliser des résumés et les discuter. Les relations d'intertextualité entre ces différents types de discours sont prises en compte avec la théorie de Bakthine et ses collègues. Notre hypothèse, qu'il faudra valider, est que le fait de proposer sur une même plate-forme des outils de lecture, écriture et discussion offrant des *feedback* automatiques favorise la construction de connaissances autorégulée.

Note des auteurs

Cet article a été écrit dans le cadre du projet européen LTfLL (*Language Technologies for LifeLong Learning*, 7e PCRD). Une partie de ce dernier est tirée de Trausan-Matu *et al*. (2008).

Bibliographie

BAKHTINE M., *Speech genres and other late essays*, University of Texas, Austin, 1986.

BANGERT-DROWNS R. L., HURLEY M. M., WILKINSON B., 2004, « The effects of school-based writing-to-learn interventions on academic achievement: A meta-analysis », *Review of Educational Research*, vol. 74, n° 1, pp. 29-58.

BAZERMAN C., 2004, « Intertextualities. Volosinov, Bakhtin, literary theory, and literacy studies », dans A. F. Ball, S. Warshauer Freedman (dir.), *Bakhtinian perspectives on language, literacy, and learning*, Cambridge University Press, Cambridge, pp. 53-65.

BEREITER C., *Education and mind in the knowledge age*, Erlbaum, Mahwah, 2002.

BEREITER C., SCARDAMALIA M., *The psychology of written composition*, Erlbaum, Illsdale, 1987.

CHANIER T., CARTIER J., 2006, « Communauté d'apprentissage et communauté de pratique en ligne : le processus réflexif dans la formation des formateurs », *Revue internationale des technologies en pédagogie universitaire*, vol. 3, n° 3, pp. 64-82.

DESSUS P., LEMAIRE B., 2002, « Using production to assess Learning : An ILE that fosters Self-Regulated Learning », dans S. A. Cerri, G. Gouardères, F. Paraguaçu (dir.), *Intelligent Tutoring Systems (ITS 2002)*, Springer, Berlin, pp. 772-781.

DESSUS P., LEMAIRE B., 2004, « Assistance informatique à la correction de copies », dans E. Gentaz, P. Dessus (dir.), *Comprendre les apprentissages : sciences cognitives et éducation*, Dunod, Paris, pp. 205-220.

DESSUS P., MANDIN S., ZAMPA V., 2008, « What is teaching? Cognitive-based tutoring principles for the design of a learning environment », dans S. Tazi, K. Zreik (dir.), *Common innovation in e-learning, Machine learning and humanoid (ICHSL'6)*, IEEE/Europia, Paris, pp. 49-55.

EMIG J., 1977, « Writing as a mode of Learning », *College Composition & Communication*, vol. 28, n° 2, pp. 122-128.

FISCHER G., 2007, « Designing socio-technical environments in support of meta-design and social creativity », dans, *Proceedings of the Conference on Computer Supported Collaborative Learning (CSCL '2007)*, Rutgers University, p. 1-10.

GRAHAM S., PERIN D., 2007, « A meta-analysis of writing instruction for adolescent students », *Journal of Educational Psychology*, vol. 99, n° 3, pp. 445-476.

HICKS D., 1996, « Contextual inquiries: a discourse-oriented study of classroom learning », dans D. Hicks (dir.), *Discourse, learning, and schooling*, Cambridge University Press, Cambridge, pp. 104-144.

HOLMER T., KIENLE A., WESSNER M., 2006, « Explicit referencing in learning chats: Needs and acceptance », dans W. Nejdl, K. Tochtermann (dir.), *Proc. EC-TEL 2006*, Springer, Heidelberg, pp. 170–184.

KOSCHMANN T., 1999, « Toward a dialogic theory of learning: Bakhtin's contribution to understanding learning in settings of collaboration », dans C. Hoadley, J. Roschelle (dir.), *Computer Support for Collaborative Learning (CSCL'99)*, Palo Alto, Mahwah.

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*, vol. 104, n° 2, pp. 211-240.

LAURILLARD D., *Rethinking university teaching*, Routledge, Londres, 1993.

LINN M. C., 1996, « Cognition and distance Learning », *Journal of the American Society for Information Science*, vol. 47, n° 11, pp. 826-842.

MORRIS P. (dir.), 1994, *The Bakhtin reader: Selected writings of Bakhtin, Medvedev, Voloshinov*, Arnold, Londres.

NUSSBAUMER A., STEINER C., ALBERT D., 2008, « Visualisation tools for supporting Self-Regulated Learning through exploiting competence structures », *Proc. Int. Conference on Knowledge Management*, Graz (Austria).

PAQUETTE G., 2005, « Apprentissage sur Internet : des plateformes aux portails d'objets à base de connaissance », dans S. Pierre (dir.), *Innovations et tendances en technologies de formation et d'apprentissage*, Presses de l'école polytechnique de Montréal, Montréal, pp. 1-30.

PHILLIPS S. M., 2007, *Automated essay scoring: a literature review*, TASA Institute, Society for the advancement of excellence in education, Kelowna.

SFARD A., 1998, « On two metaphors for learning and the dangers of choosing just one », *Educational Researcher*, vol. 27, n° 2, pp. 4-13.

STAHL G., *Group cognition. Computer support for building collaborative knowledge*, MIT Press, Cambridge, 2006.

STAHL G., sous presse, « The VMT vision », dans G. Stahl (dir.), *Studying virtual math teams*, Springer, New York.

TRAUSAN-MATU S., DESSUS P., LEMAIRE B., MANDIN S., *et al.*, 2008, Deliverable D5.1 - Support and Feedback Design, OUNL, Research report of the LTfLL Project, Heerlen.

TRAUSAN-MATU S., G. STAHL, 2007, « Polyphonic inter-animation of voices in chats », *Proc. CSCL'07 Workshop on Chat Analysis in Virtual Math Teams*, New Brunwick (NJ).

TRAUSAN-MATU S., REBEDEA T., DRAGAN A., ALEXANDRU C., 2007, « Visualisation of learners' contributions in chat conversations », dans J. Fong, F. L. Wang (dir.), *Blended learning*, Pearson/Prentice Hall, Singapour, pp. 217-226.

TRAUSAN-MATU S., REBEDEA T., sous presse, « Polyphonic inter-animation of voices in VMT », dans G. Stahl (dir.), *Studying virtual math teams*, Springer, New York.

VALENTI S., NERI F., CUCCHIARELLI A., 2003, « An overview of current research on automated essay grading », *Journal of Information Technology Education*, vol. 2, n°, pp. 319-330.

VILLIOT-LECLERCQ E., DESSUS P., 2009, « Les contraintes de l'activité de tutorat à distance », dans V. Guéraud, P. Leroux (dir.), *Atelier « Instrumentation des activités du tuteur » de la conférence EIAH'09*, Le Mans.

VOLOCHINOV V. N., *Marxism and the philosophy of language*, Cambridge University Press, Cambridge, 1973.

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*, vol. 2, n° 1, pp. 64-74.

VYGOTSKI L., *Pensée et langage*, La Dispute, Paris, 1997.

WEGERIF R., *Dialogic education and technology: Expanding the space of learning*, Springer, New York, 2007.

WENGER E., *Artificial intelligence and tutoring systems*, Morgan Kaufmann, Los Altos, 1987.

WINTERS F. I., GREENE J. A., COSTICH C. M., 2008, « Self-regulation of learning within computer-based learning environments: A critical analysis », *Educational Psychology Review*, vol. 20, n°, pp. 429-444.

WITTROCK M. C., 1990, « Generative processes of compréhension », *Educational Psychologist*, vol. 24, n° 4, pp. 345-376.

Notices biographiques

Philippe Dessus est professeur en sciences de l'éducation à l'IUFM de Grenoble (Université Joseph-Fourier) et chercheur au Laboratoire des sciences de l'éducation de la même ville. Il s'intéresse à la modélisation cognitive des activités engagées dans l'enseignement et l'apprentissage, notamment à l'aide de l'analyse sémantique latente. Mél. : Philippe.Dessus@upmf-grenoble.fr

Stefan Trausan-Matu est professeur d'informatique à l'université « polytechnique » de Bucarest (Roumanie) et chercheur à l'académie roumaine d'intelligence artificielle. Ses thèmes de recherche sont la construction collaborative de connaissances, la fouille de données et l'analyse de discours. Mél. : Stefan.Trausan@cs.pub.ro

Virginie Zampa est maître de conférences à l'Université Stendhal de Grenoble au laboratoire LIDILEM (LInguistique et DIdactique des Langues Etrangères et Maternelles). Elle s'intéresse aux environnements informatiques pour l'apprentissage, notamment en utilisant l'analyse sémantique latente ainsi qu'à la création d'ontologie a-domaine par l'utilisation des jeux. Mél. : Virginie.Zampa@u-grenoble3.fr

Traian Rebedea est doctorant en informatique à l'université « polytechnique » de Bucarest (Roumanie). Il s'intéresse à l'analyse dialogique de *chats* dans un contexte d'apprentissage collaboratif. Mél. : trebedea@gmail.com

Sonia Mandin est doctorante au Laboratoire des sciences de l'éducation de Grenoble. Elle s'intéresse à la modélisation computationnelle de l'activité de résumer et a conçu un EIAH d'aide à l'apprentissage de cette activité, pour un public de lycéens. Mél. : Sonia.Mandin@msh-alpes.fr

Mihai Dascalu est étudiant en licence d'informatique à l'université « polytechnique » de Bucarest (Roumanie). Il s'intéresse à l'analyse dialogique de *chats* dans un contexte d'apprentissage collaboratif. Mél. : mikedascalu@yahoo.com