



Écrire des résumés pour apprendre des cours : un système d'aide à l'apprentissage autorégulé

Virginie Zampa, Philippe Dessus

► To cite this version:

Virginie Zampa, Philippe Dessus. Écrire des résumés pour apprendre des cours : un système d'aide à l'apprentissage autorégulé. 7e Colloque "Technologies de l'Information et de la Communication pour l'Enseignement" (TICE 2010), 2010, Nancy, France. <hal-01495694>

HAL Id: hal-01495694

<https://hal.science/hal-01495694v1>

Submitted on 26 Mar 2017

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.



HAL Authorization

Écrire des résumés pour apprendre des cours : un système d'aide à l'apprentissage autorégulé

Zampa Virginie *, Dessus Philippe **

Virginie.Zampa@u-grenoble3.fr, Philippe.Dessus@umpf-grenoble.fr

* LIDILEM, Université Stendhal, BP 25 - 38040 Grenoble CEDEX 9

** LSE, 1251, av. Centrale, BP 47 Université Pierre-Mendès-France - 38040 Grenoble CEDEX 9

RÉSUMÉ : Cet article présente *Apex 2.1* (pour Aide à la Préparation des EXamens). Il s'agit d'un système prototype donnant des retours automatiques à des étudiants réalisant des résumés de cours, dans un contexte d'enseignement à distance. Nous présentons quelques données théoriques, puis l'architecture d'*Apex 2.1*, pour enfin décrire une étude le validant auprès de trois groupes d'utilisateurs potentiels (étudiants, enseignant ou administrateur), du point de vue de l'utilité, l'acceptabilité et l'utilisabilité de ce système.

Mots clés : Analyse sémantique latente, Apprentissage auto-régulé, Feedback

ABSTRACT: We introduce here *Apex 2.1* (i.e., Assistant for Preparing EXams), a prototype system that gives e-learning students automated feedback on their course summaries. First we present some theoretical underpinnings, then we detail *Apex 2.1* architecture. Eventually, the first results of a validation study involving three groups of stakeholders (students, teacher, administrator) are presented testing the utility, acceptability and usability of the system.

Keywords: Latent Semantic Analysis, Self-regulated Learning, Feedback

1 INTRODUCTION

Cet article présente *Apex 2.1* (Aide à la Préparation des EXamens), un système prototype donnant des retours automatiques à des utilisateurs réalisant des résumés de cours, dans un contexte d'enseignement à distance. *Apex 2.1* permet à des étudiants de choisir un thème sur lequel ils veulent travailler, formulé par un ou plusieurs mot(s)-clé(s). Ensuite, le système leur fournit un ensemble de textes sémantiquement proches de cette requête. À tout moment les étudiants peuvent faire des résumés des textes qu'ils ont lus et compris, et obtenir un feedback sur leurs productions. Cela leur permet ainsi de pouvoir s'entraîner avant leurs examens et de rédiger, vérifier, modifier, et améliorer leur résumé jusqu'à obtenir une évaluation, et pouvoir ainsi avoir une compréhension correcte du thème choisi.

Cet article est l'occasion de décrire les présupposés théoriques sur lesquels ce logiciel est construit, ainsi que son architecture et la méthode d'analyse sémantique (*Latent Semantic Analysis*), sur laquelle il repose. Puis nous décrivons l'étude sur l'utilisabilité d'*Apex 2.1* sur plusieurs catégories d'utilisateurs potentiels et ses premiers résultats.

2 ECRIRE POUR APPRENDRE

Sommairement, la conception d'*Apex 2.1* part du principe que l'on peut « écrire pour apprendre », et qu'on apprend d'autant mieux que l'on réfléchit précisément à ce que l'on a pu apprendre, notamment par l'auto-évaluation. Ces deux assertions, si on les développe, nous permettent de mieux comprendre la possible utili-

sation d'*Apex 2.1* dans un contexte d'enseignement à distance.

Selon l'approche « écrire pour apprendre » (*writing-to-learn*) il est communément admis que l'activité d'écrire entraîne de fait un apprentissage, du seul fait de coucher par écrit une idée, selon un principe d'objectivation, de mise à distance de la pensée [1]. Pourtant, certains chercheurs [2] ont atténué cette idée, et aussi étendu les situations d'apprendre en écrivant, en montrant quatre modèles : *l'écriture spontanée* (*knowledge telling* [3]), *la recherche en avant*, où les rédacteurs externalisent leurs idées par l'écrit, et réalisent de nouvelles révisions de leurs buts à partir des inférences qu'ils effectuent pendant la relecture de l'écrit, *la recherche en arrière*, où les rédacteurs partent des buts de leur rédaction pour les raffiner en sous-but, ce qui leur permet de rédiger les sous-parties correspondantes de leur écrit ; *l'écriture selon un genre*, dans laquelle le rédacteur, en écrivant son texte, se conforme à un genre donné (p. ex., article, synthèse, résumé, notes de cours), ce qui lui permet plus efficacement de relier ce qu'il écrit à ses connaissances antérieures. Ces différents modèles montrent que l'écriture pour apprendre est une activité dans laquelle l'auto-régulation joue un rôle important, et qu'elle doit être souplement assistée, tant les buts du rédacteur sont multiples.

Ensuite, la question de l'autorégulation de l'apprentissage dans un contexte d'enseignement à distance est cruciale [4]. En effet, dans un tel contexte, l'apprentissage des étudiants ne peut pas aussi fréquemment que dans un contexte en présence faire l'objet de retours, et ces derniers sont laissés beaucoup plus libres dans leur activité. D'où la nécessité de se munir de modèles d'auto-régulation de l'apprentissage,

comme celui décrit en Figure 1. Les étudiants travaillent tout d'abord au niveau d'un objet d'apprentissage donné, puis, une évaluation de cet objet, pouvant être individuelle ou collective, leur permet de se situer à un niveau méta (supervision) à partir duquel ils se munissent d'outils métacognitifs pour mieux contrôler leur activité qui, en une sorte de boucle, est potentiellement redéfinie par les tâches et stratégies mises en œuvre.

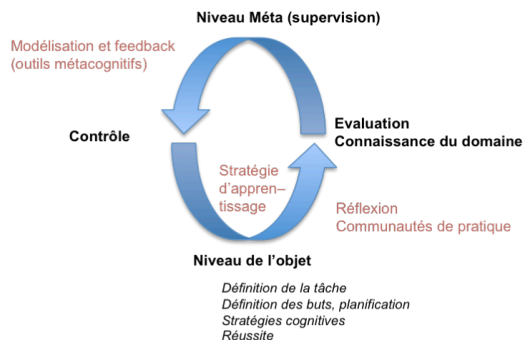


fig 1 : Le modèle de l'auto-régulation de l'apprentissage de Vovides et al. [4]

Nous posons que l'utilisation d'*Apex 2.1* permet aux étudiants de mieux comprendre leurs cours sans avoir recours, au moins dans un premier temps, à des enseignants ou tuteurs. *Apex 2.1* fait partie d'une catégorie de nouveaux services d'enseignement à distance permettant à ses utilisateurs d'être engagés dans des flux d'activité desquels ils peuvent aisément sortir et qui favorisent leur auto-évaluation. Mais l'utilisation d'*Apex 2.1* concerne aussi les enseignants et tuteurs. Elle leur permet d'avoir une vue plus globale de l'activité des étudiants, mais aussi de se centrer sur des interventions plus individualisées. Toutefois, comme l'utilisation de ce type de systèmes est tout à fait nouvelle, il est intéressant de recueillir des données sur leur usage afin d'étudier la manière dont les étudiants se les approprient.

3 DESCRIPTION DU SYSTEME

Apex 2.1 avait déjà fait l'objet d'une première implémentation [5], mais n'avait jamais été testé en situation. Il est développé en PHP et C et fait appel à l'analyse de la sémantique latente (LSA, ou *Latent Semantic Analysis*) [6], que nous présenterons dans une section suivante. Il est accessible sur un serveur dédié à l'URL suivante :

<http://augur.wu-wien.ac.at/apex2/Recovery/progPhp/Apex2.php>.

3.1 Fonctionnement du système

Apex 2.1 fonctionne comme le montre le schéma suivant (cf. Figure 2). Il comprend deux boucles principales (en jaune ou gris) : la première est une boucle de *lecture* (lire une série de textes reliés sémantiquement à une requête), la seconde est une boucle d'*écriture* (écrire un résumé des différents textes qui ont été jugés préalablement compris, et être évalué à leur propos). Pour commencer la session, l'apprenant rentre des mots-clés correspondant au domaine ou thème qu'il

veut revoir, approfondir. *Apex 2.1* lui renvoie un texte à lire. Le choix de ce texte, dans la base de textes du domaine, est fait en utilisant LSA qui sélectionne les textes sémantiquement les plus proches des mots-clés. Les textes sont ensuite fournis à l'apprenant en les lui donnant à lire un par un. Quand le texte est proposé à l'apprenant, le fichier correspondant à un modèle de l'élève simplifié est mis à jour en indiquant que le texte donné a été lu.

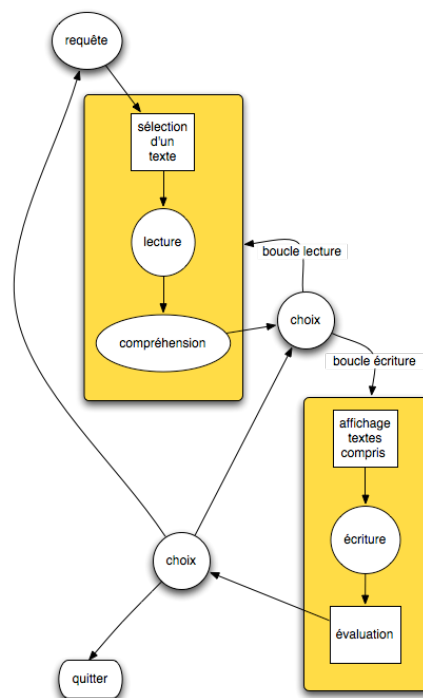


fig 2 : Le fonctionnement d'Apex 2.1. Actions utilisateur/machine

Après avoir lu le texte, l'apprenant indique s'il pense l'avoir compris et s'il est capable ou non de le résumer (cf. Figure 3). Lors de cette étape, le modèle de l'élève est mis à jour en indiquant si le texte est résumable ou non. Tant que l'apprenant indique qu'il n'est pas capable de le résumer et qu'il existe des textes suffisamment proches des mots-clés, *Apex 2.1* continue à lui proposer des textes à lire. Dans le cas où l'apprenant n'aurait compris aucun des textes relevant du domaine souhaité, *Apex 2.1* lui propose de faire une nouvelle recherche par mots-clés.

Une fois que l'apprenant est capable de résumer un texte, il peut soit continuer à lire des textes sur le sujet choisi, soit passer à la phase d'écriture. Dans ce cas, un tableau contenant la liste des textes qu'il a signalés être capable de résumer, lui est fourni (cf. Figure 4). La première colonne lui indique l'ordre d'apparition du texte au cours de ses lectures, la seconde l'identifiant du texte. La troisième lui permet de voir le début du texte et s'il le désire, il peut le revoir en entier en cliquant sur le bouton « + ». Le bouton « Résumer » de la quatrième colonne, permet d'accéder à un champ de rédaction. Si l'élève n'a encore rien écrit sur ce texte, le champ apparaît vide. S'il l'a déjà travaillé, la page

contiendra la dernière version de son texte. Enfin, les deux dernières colonnes lui fournissent, respectivement, la valeur de proximité sémantique entre son résumé et le texte correspondant, renvoyée par LSA, et enfin la comparaison entre cette dernière valeur et l'avis de l'élève concernant sa propre compréhension (p. ex., « *Vous pensez avoir compris le texte et votre résumé le prouve* »).



fig 3 : Ecran de lecture de texte proposé.

Le système actuel fonctionne avec un seul apprenant connecté et comporte un modèle de l'élève simplifié, comme nous l'avons déjà signalé. Il s'agit d'un fichier au format texte comportant autant de lignes qu'il y a de textes dans la base et quatre colonnes : l'identifiant du texte, 1 ou 0 pour indiquer si le texte a déjà été lu ou non, 1 ou 0 pour indiquer si le texte est compris, la valeur de proximité quand le texte a été résumé. Ce fichier est mis à jour lors de chaque action de l'apprenant et réinitialisé lors d'une nouvelle requête.

Après une évaluation l'apprenant peut :

- modifier son résumé et demander une nouvelle évaluation (son résumé est présent dans la page et il peut le modifier à sa convenance) ;
- retourner au tableau des textes compris et débiter ou modifier un autre résumé ;
- lire un nouveau texte ;
- faire une nouvelle recherche par mots-clés ;
- quitter *Apex 2.1*.

Décrivons maintenant LSA, la méthode de comparaison sémantique utilisée par ce système.

3.2 Description du fonctionnement de LSA

L'analyse de la sémantique latente (*Latent Semantic Analysis*) est un outil permettant d'extraire des relations sémantiques à partir de textes [6]. Dans *Apex 2.1*, comme nous l'avons déjà mentionné, nous l'utilisons pour sélectionner et fournir les textes les plus appropriés à l'apprenant ainsi que pour évaluer ses résumés.

Nous avons décidé de recourir à LSA pour sélectionner les textes à lire à partir du ou des mot(s)-clé(s) car cette utilisation a déjà été validée [7, 8]. En effet, LSA apporte un plus puisqu'elle permet de tenir compte de la synonymie, de la polysémie ainsi que des inflexions.

Dans *Apex 2.1*, LSA est aussi utilisée pour évaluer les résumés. Cette utilisation, tout comme la précédente a déjà été validée [9]. Dans *Apex 2.1* nous avons fixé un seuil au-dessus duquel les résumés sont suffisamment proches des textes-sources. Nous l'avons empiriquement fixé à 0,6, et le commentaire « *Vous pensez avoir compris ce texte et votre résumé le prouve* » est retourné si la proximité est supérieure à ce seuil ; si elle est inférieure, « *Vous pensiez avoir compris ce texte et ce n'est pas le cas* » est affiché.

Le but de LSA est de représenter dans un espace multidimensionnel (300 dimensions) les mots de la langue. Pour construire cet espace, LSA prend un ensemble de textes en entrée et construit une matrice d'occurrences qui est réduite par le biais d'une analyse statistique. Cela permet ainsi de faire ressortir les relations sémantiques entre mots, paragraphes ou textes. Grâce à cette méthode, deux mots peuvent être considérés sémantiquement proches, même s'ils n'apparaissent jamais conjointement dans un texte. Il suffit qu'ils soient utilisés dans des contextes similaires. Le contexte d'un mot est ici défini comme l'ensemble des mots qui apparaissent conjointement avec lui. La méthode fonctionne si un nombre suffisant de textes est utilisé. Il ne s'agit pas simplement d'un comptage d'occurrences, il faut aussi disposer d'une procédure pour établir les liaisons entre mots. Cette procédure est la réduction de la matrice d'occurrences à 300 dimensions [10].

Dans l'espace multidimensionnel ainsi obtenu, chaque mot et chaque paragraphe est ainsi représenté par un vecteur à 300 dimensions. La proximité de sens entre deux mots ou deux paragraphes ou un mot et un paragraphe correspond à la proximité entre les vecteurs.

Vous avez indiqué que vous pouviez résumer les textes suivants :					
Rang d'apparition	Num texte	Texte		Evaluation	Commentaires
3	20	L'outil NEDERLEX génère un texte qui ... +	Résumer	0,91	Vous pensez avoir compris ce texte et votre résumé le prouve
1	29	Norme de saisie Il n'existe pas à l'heure ... +	Résumer		
2	30	Lemmatisation La plupart des mots sont susceptibles ... +	Résumer		

fig 4 : Tableau fourni à l'apprenant au début de la phase de rédaction

4 VALIDATION DU SYSTEME

La validation d'*Apex 2.1* a été réalisée en utilisant un corpus de soixante-six textes portant sur les domaines du traitement automatique des langues (TAL) et/ou des environnements informatiques pour l'apprentissage des langues étrangères (EIAL). Il s'agissait d'une collection de textes homogènes en nombre de mots (moyenne d'environ 650 mots), comportant des cours et des articles de recherche. Cette homogénéité en nombre de mots permet de ne pas introduire un biais lié à l'effet de la longueur des textes sur la qualité des résumés. Ce corpus a constitué la base de documents à lire par les étudiants. Nous avons ajouté à cette base un corpus plus général et plus important afin que des connaissances générales sur la langue soient également présentes. Ce corpus « général » comporte 101 123 formes différentes. Il est construit à partir de textes de forme, style et vocabulaire variés : des articles du journal *Le Monde*, des textes écrits par des enfants, des textes pour enfants, et a fait l'objet de validations [11]. Afin de tester *Apex 2.1* dans des conditions les plus proches possible de la réalité, nous avons constitué trois groupes de participants.

- un groupe « utilisateurs », composé de 11 étudiants en master ou en doctorat spécialisés dans le domaine du TAL et/ou des EIAL à l'université Grenoble-3. Ce groupe a utilisé *Apex 2.1* dans les mêmes conditions que des étudiants d'un cours à distance.
- un groupe « démo », composé de 28 étudiants de licence en sciences de l'éducation de l'université Grenoble-2. Ce groupe a suivi une démonstration d'une séance-type d'utilisation d'*Apex 2.1* et a répondu ensuite à un questionnaire.
- un groupe « enseignant/administrateur », composé de deux personnes : le responsable du département TICE de l'IUFM de Grenoble-1, et le responsable de l'espace numérique de travail du même établissement. Ce groupe a effectué la même tâche que le groupe précédent.

La raison de la présence de ces deux derniers groupes de participants (qui ne testent donc pas réellement *Apex 2.1*) est de vérifier dans quelle mesure un tel outil est

jugé ou non acceptable par des groupes possibles d'utilisateurs ou de prescripteurs.

4.1 Description des tâches des participants

Les participants du groupe « utilisateurs » avaient une moyenne de 24,8 ans (écart type de 4,26) et étaient cinq étudiants de Master 1, deux étudiants de Master 2 et quatre doctorants. Avant de commencer l'expérimentation, ils devaient prendre connaissance du protocole de travail et le signer, puis ils devaient remplir un questionnaire portant sur leur(s) méthode(s) de révision de cours. Ensuite, ils ont utilisé le logiciel en entrant les mots-clés de leur choix. Ils devaient réaliser le résumé de cinq textes de leur choix (qu'ils avaient lus et jugés compris), sans limite de durée. Leur parcours dans le logiciel a été enregistré (p. ex. : mots-clés saisis, nombre de recherches, textes lus, durée de lecture de chaque texte, textes compris ou non, durée de chaque boucle, etc.). À la fin de cette tâche, ils devaient remplir un second questionnaire portant sur leur opinion à propos du logiciel. L'ensemble des documents fournis à ce groupe est disponible à l'URL suivante :

<http://augur.wu-wien.ac.at/apex2/Expe/groupe-utilisateur.pdf>

Le groupe « démo ». Les participants de ce groupe (âge moyen : 23,4 ans ; écart type : 4,5) ont eu à remplir le même questionnaire que le groupe utilisateur concernant leur(s) méthode(s) de révision. Puis ils ont suivi une démonstration d'*Apex 2.1* et ont répondu à un second questionnaire sur leur opinion à propos d'*Apex 2.1*, disponible à l'URL suivante :

<http://augur.wu-wien.ac.at/apex2/Expe/groupe-demo.pdf>

Le groupe « enseignant/administrateur ». Il s'agit d'un deux personnes respectivement âgés de 42 et 60 ans. Ils ont suivi une démonstration de *Apex 2.1* et ont répondu à un questionnaire disponible à l'URL suivante :

<http://augur.wu-wien.ac.at/apex2/Expe/groupe-admin.pdf>

4.2 Hypothèses

Nous avons élaboré les hypothèses suivantes, en nous fondant sur les critères d'utilité, d'utilisabilité et d'acceptabilité des EIAH [12, 13] :

a) L'hypothèse d'utilité, qui suppose que l'utilisation du système induit un gain pour l'utilisateur (en temps, intérêt, etc.). Elle se décompose en deux sous-hypothèses :

- *hypothèse de requête-lecture* : *Apex 2.1* peut être utilisé comme moteur de recherche pour trouver des textes qui lui conviennent sur les thèmes qu'il choisit. Les textes ayant été fournis par les enseignants, ils sont donc adaptés et issus de sources fiables. Le système permet donc aux étudiants de lire des textes qui leur conviennent et qui vont être compris par eux ;
- *hypothèse autorégulation/évaluation* : Les étudiants peuvent librement rentrer dans chacune des deux boucles (requête-lecture et rédaction-évaluation), ce qui leur permet de réguler au mieux leur apprentissage ; de plus *Apex 2.1* leur procure des retours sur leurs résumés valides.
- b) L'*hypothèse d'utilisabilité*, qui correspond à la maniabilité d'*Apex 2.1*, se joue au niveau de l'interface et de la navigation. Il s'agit d'évaluer, d'un point de vue ergonomique, la facilité à prendre en main le système et à réaliser avec lui les buts que l'on s'est fixés.
- c) L'*hypothèse d'acceptabilité* correspond à « la valeur de la représentation mentale (attitudes, opinions, etc. plus ou moins positives) à propos d'un système, de son utilité et de son utilisabilité » [12, p. 396]. Cela correspond à un avis général sur les changements (en termes d'usage) induits par le système ; pour nous, cela correspond à savoir si les étudiants trouvent *Apex 2.1* facile à utiliser et s'ils l'utiliseraient pour travailler s'ils le pouvaient.

4.3 Résultats

Nous présentons ici les résultats pour les trois groupes de participants. Les hypothèses ci-dessus ne seront scrutées que pour le groupe « utilisateurs ».

4.3.1 Le groupe « utilisateurs »

Pour ce groupe, comme nous l'avons signalé, nous avons deux types de données : les traces (comportement des utilisateurs en termes de requêtes, lecture et écriture) et les réponses aux questionnaires. L'analyse des traces a mis en évidence les points suivants :

a) *Hypothèse d'utilité/requête-lecture*, sept utilisateurs sur 11 n'ont effectué qu'une seule requête (cf. Tableau 1), et ont donc travaillé sur le même thème pour leurs textes, deux ont saisi deux requêtes, un en a fait trois et le dernier en a fait quatre (moyenne du nombre de requêtes : 1,64 ; écart type : 1,03). Concernant la lecture, les 11 sujets ont lu entre 5 et 13 textes (moyenne à 8 ; écart type : 3,46). Nous constatons cependant que la plupart des textes jugés « non compris » et donc non résumés n'ont été que survolés. En effet, sur ces 33 textes, 5 ont été lus en plus de 2 minutes, et la moyenne de la durée de lecture pour les 28 restants est de 30 secondes. Il semblerait ainsi que, lorsque les sujets indiquent s'ils ont compris ou non le texte et sont capables de le résumer, il ne s'agisse pas en réalité d'une évaluation de compréhension mais plus d'une appréciation d'intérêt. Si le texte les intéresse, ils le lisent et le résument, sinon ils le survolent et passent au texte suivant. *Apex 2.1* leur servirait ainsi surtout à sélection

ner les textes correspondant à leurs attentes. Ce qui valide notre première hypothèse.

b) *hypothèse d'utilité/autorégulation-évaluation*. Les données suivantes ne concernent que les textes lus et résumés. Concernant la variation inter-participants en moyenne, les participants ont consacré entre 3 min 24 s et 14 min 19 s pour lire un texte (moyenne du temps de lecture par texte : 8 min et 2 s ; écart type de 2 min 50 s). La variation inter-participants est donc également très élevée. Concernant la variation intra-participant, les écarts types sur les durées de lecture vont de 41 s à 5 min et 52 s avec une moyenne de 2 min 40 s. Les participants ont tous suivi la consigne et écrit au moins 5 résumés en y passant entre 1 min 49 s et 27 min 40 s pour rédiger (moyenne du temps de rédaction par texte : 7 min et 5 s, avec un écart type de 4 min et 16 s).

tab. 1 : Données générales sur l'activité de lecture et d'écriture dans Apex 2.1

Etu- diant	Nb de re- quêtes	Nb de textes lus et jugés compris	Nb de textes lus et jugés non compris	Nb total de textes lus
1	1	5	0	5
2	1	5	0	5
3	1	5	0	5
4	1	5	0	5
5	1	5	0	5
6	1	5	4	9
7	3	5	7	12
8	1	5	5	10
9	4	5	1	6
10	2	5	8	13
11	2	5	8	13

La moyenne des évaluations des résumés est à 0,73 ; seulement 6 des 55 résumés ont une note inférieure au seuil (0,6) et 3 des 6 correspondent à un même participant. Uniquement un résumé a été réécrit. Les étudiants n'ont donc pas modifié leur résumé après avoir obtenu leur évaluation. Ils n'ont pas tenu compte du feedback et n'ont pas cherché à les améliorer. Cela est peut être dû au fait que pour la majorité le feedback était correct, cela leur suffisait. Un feedback plus fin, et non juste en terme de vrai/faux aurait certainement plus d'impact.

Passons maintenant à l'analyse des deux questionnaires. Le *premier questionnaire* portait essentiellement sur l'utilisation des ordinateurs pour réviser les cours. La plupart des participants disent compléter leurs cours avec des documents cherchés sur Internet (7 sur 11 le font « une fois par semaine » ou « presque tous les jours »). De plus pour réviser ils utilisent plutôt papier-crayon et ordinateur simultanément (8 sur 11 répartis entre « presque chaque jour » et « plusieurs fois par jour »). Ils ont donc l'habitude de réviser avec leur ordinateur. Ce premier questionnaire nous a permis de vérifier qu'ils ont l'habitude de travailler sur ordinateur et donc qu'un problème d'utilisabilité serait

directement imputable à *Apex 2.1* et non à un manque d'expertise en navigation, prise de note ou recherche avec ordinateur. Le second questionnaire portait sur l'utilisation d'*Apex 2* et comportait deux parties.

c) *hypothèse d'utilisabilité*. De la première intitulée « difficultés d'utilisation » et issue du test NASA-TLX [14], nous pouvons dire que pour la majorité des participants, *Apex 2.1* :

- n'entraîne pas de pression physique,
- entraîne une pression liée à l'expérimentation,
- est facile à utiliser.

De la seconde partie intitulée « les fonctionnalités d'*Apex 2.1* », nous pouvons dire que, pour la majorité des participants :

- les textes fournis correspondaient à la requête,
- les textes fournis étaient adaptés, précis, etc.,

Cela conforte notre seconde hypothèse. Enfin, la majorité des participants utiliserait *Apex 2.1* de temps en temps si son accès était possible, ce qui confirme la validation de notre hypothèse d'acceptabilité.

4.3.2 Le groupe « démo »

Le groupe « démo » a répondu à un questionnaire comportant deux parties. Concernant l'utilisation de l'ordinateur pour réviser les cours, les participants de ce groupe vont moins fréquemment compléter leur cours avec des documents issus d'Internet (13 sur 28 « une fois par semaine environ » et 9 sur 28 « une fois par mois environ ») et utilisent plus le papier-crayon seul pour réviser (15 sur 28 « presque chaque jour »). Leurs réponses pour l'usage de l'ordinateur et du papier-crayon simultanément sont assez réparties : 1 « jamais », 9 « une fois par mois », 10 « une fois par semaine », 6 « presque chaque jour », 2 « plusieurs fois par jour ».

Concernant les fonctionnalités d'*Apex 2.1*, la majorité des participants pense qu'*Apex 2.1* leur permettrait d'acquérir des connaissances aussi facilement voire plus facilement qu'avec leur méthode habituelle de révisions (19 réponses, réparties entre 11 « plus facilement » et 8 « aussi facilement »). De plus, pour eux, les connaissances ainsi acquises sont aussi précises (pour 15 participants sur 28) et aussi étendues (14 sur 28) qu'avec leur méthode habituelle. Enfin, si cela était possible vingt d'entre eux utiliseraient *Apex 2.1* de temps en temps pour réviser leurs cours.

4.3.3 Le groupe « enseignant/administrateur »

Les deux participants de ce groupe ont tous deux jugé que l'utilisation d'*Apex 2.1* serait un moyen d'acquérir de nouvelles connaissances plus facile que des moyens traditionnels (papier-crayon), et que ces dernières seraient également plus précises. Ils en préconiseraient tous deux l'accès fréquent à leurs étudiants. Ils ont estimé qu'*Apex 2.1* permettrait aux étudiants de mieux se focaliser sur l'information nécessaire à l'apprentissage d'un contenu de cours. L'administrateur a de plus signalé que ce logiciel permet l'accès par les étudiants à un grand ensemble d'informations, tout en restant simple d'emploi.

4.3.4 Synthèse des résultats

En faisant passer l'expérimentation à ces trois groupes, nous avons fait une double évaluation : une « évaluation par inspection » avec le groupe « enseignant/administrateur » et une « évaluation empirique » avec les deux autres, et cela pour les critères d'utilité, d'utilisabilité et d'acceptabilité.

Concernant l'évaluation par inspection, les deux enseignants-administrateurs jugent qu'*Apex 2.1* répond à ces trois critères. Concernant l'évaluation empirique, le groupe « démo » valide les trois critères. En revanche, avec le groupe « utilisateurs », l'hypothèse d'utilité concernant l'évaluation des textes n'est pas entièrement validée. En effet, quand les résumés étaient jugés corrects, les participants n'ont pas essayé de les améliorer et quand ils n'étaient pas jugés corrects, ils ne les ont pas retravaillés. Cependant, la deuxième partie de cette hypothèse est validée et les deux autres également.

5 CONCLUSION

Dans cet article, nous avons décrit le fonctionnement et une première validation d'*Apex 2.1*. Ce logiciel, dédié à la préparation aux examens permet différentes utilisations par l'apprenant.

Tout d'abord, *Apex 2.1* peut être utilisé comme base de données de textes avec moteur de recherche intégré. Les avantages de cette utilisation par rapport à un moteur de recherche sur Internet résident dans le fait que la base de textes, ayant été construite par l'enseignant, ne contient que des textes adaptés (ce qui permet de ne pas se perdre dans des recherches) et qui plus est, des textes fiables (ce qui n'est pas nécessairement le cas sur Internet). Le moteur de recherche associé, utilisant LSA, permet de faire en partie disparaître les problèmes liés à la polysémie, la synonymie ainsi que les inflexions [6].

Au cours de notre validation, nous avons pu constater que les participants sélectionnent les textes qu'ils désirent lire et résumer. En effet, la majorité des textes qu'ils signalent n'être pas capables de résumer n'a seulement pas été lue.

Apex 2.1 est développé afin d'aider les apprenants à acquérir des connaissances par le biais de la lecture et de l'écriture (approche « écrire-pour-apprendre »). L'avantage est d'apporter une approche souple dans laquelle l'utilisateur évolue selon ses envies. De plus, il apporte un feedback sur des contenus « réels » et non sur des réponses à des QCM ou autres exercices fermés.

Avec cette première étude sur la première version du système, nous constatons que la majorité des participants indique qu'*Apex 2.1* est facile à utiliser et qu'ils l'utiliseraient de temps en temps si son accès était possible. Cela nous conforte dans l'idée de continuer notre développement.

Cette version comporte certaines limites. L'une principale est le manque de raffinement du feedback après l'évaluation. En effet, indiquer à l'apprenant qu'il dit

avoir compris le texte le laisse supposer que cela est vrai et ne le pousse pas à retravailler son résumé pour chercher à encore l'améliorer. À l'opposé, dire que le texte n'est pas compris mais sans apporter aucune suggestion n'est pas très constructif.

C'est pourquoi, dans la version suivante en cours d'implémentation, nous avons pour but de remédier à ces problèmes en apportant différentes sortes de feedback portant sur des points tels que :

- la cohérence au sein du résumé (afin de détecter les ruptures de cohérence) ;
- le lien entre le résumé et le texte source (afin de signaler à l'utilisateur les phrases potentiellement hors-sujet, ou encore celles pouvant être reprises dans le résumé).

Cela permettrait ainsi à l'apprenant de savoir sur quels points retravailler son résumé ou sa synthèse. Ces feedback permettraient d'améliorer à la fois la technique d'écriture de résumé (p. ex., en se centrant sur la cohérence entre paragraphes, etc.) mais aussi sur le fond (grâce à la mise en évidence de liens sémantiques entre textes-sources et synthèses). De plus, afin de mieux s'appropriier les textes lus, l'apprenant aurait la possibilité de « stabilobosser » les phrases importantes, de prendre des notes dans un bloc-notes dédié. Enfin, il pourra garder un historique avec les différentes versions de ses résumés et les notes/commentaires associés à chacun.

Cet article montre qu'il est possible de proposer à des étudiants à distance un outil évaluant automatiquement certaines caractéristiques sémantiques de leur production écrite, et donc leur compréhension. Cet outil, malgré quelques imperfections, a été bien reçu par trois catégories de ses possibles utilisateurs.

Note des auteurs

Ce travail a été partiellement réalisé dans le cadre du projet européen LTfLL (*Language Technologies for Lifelong Learning*), 7^e PCRD, ICT-STREP. Nous remercions vivement tous les participants de l'étude de validation, ainsi que la MSH-Alpes pour son soutien matériel.

Bibliographie et webographie

- [1] Emig, J., "Writing as a mode of learning", *College Composition & Communication*, vol. 28, (1977), pp. 122-128.
- [2] Klein, P. D., "Reopening inquiry into cognitive processes in writing-to-learn", *Educational Psychology Review*, vol. 11, (1999), pp. 203-270.
- [3] Bereiter, C., Scardamalia, M., *The psychology of written composition*. Ithaca: Erlbaum, 1987.
- [4] Vovides, Y., Sanchez-Alonso, S., Mitropoulou, V., Nickmans, G., "The use of e-learning course management systems to support learning strategies and to improve self-regulated learning", *Educational Research Review*, vol. 2, (2007), pp. 64-74.

- [5] Dessus, P., Lemaire, B., "Using production to assess learning: An ILE that fosters Self-Regulated Learning," *Intelligent Tutoring Systems (ITS 2002)*, in S. A. Cerri, G. Gouardères, and F. Paraguaçu (Eds.), Berlin, Springer, 2002, pp. 772-781.
- [6] Landauer, T. K., Dumais, S. T., "A solution to Plato's problem: the Latent Semantic Analysis theory of acquisition, induction and representation of knowledge", *Psychological Review*, vol. 104, (1997), pp. 211-240.
- [7] Dumais, S. T., "Latent Semantic Indexing (LSI) and TREC-2," *The Second Text RE-trieval Conference (TREC2)*, in D. Harman (Ed.) Gaithersburg, National Institute of Standards and Technology, 1994, pp. 105-116.
- [8] Dumais, S. T., "Using Latent Semantic Indexing for information retrieval, information filtering and other things", *Cognitive Technology Conference*, 1997.
- [9] Foltz, P. W., "Latent semantic analysis for text-based research", *Behavior Research Methods, Instruments and Computers*, vol. 28, (1996), pp. 197-202.
- [10] Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., Harshman, R., "Indexing by Latent Semantic Analysis", *Journal of the American Society for Information Science*, vol. 41, (1990), pp. 391-407.
- [11] Denhière, G., Lemaire, B., Bellissens, C., Jhean-Larose, S., "A semantic space for modeling children's semantic memory," *Handbook of Latent Semantic Analysis*, in T. K. Landauer, D. McNamara, S. Dennis, and W. Kintsch (Eds.), Mahwah, Erlbaum, 2007, pp. 143-165.
- [12] Tricot, A., Plégat-Soutjis, F., J-F., C., Amiel, A., Lutz, G., Morcillo, G., "Utilité, utilisabilité, acceptabilité : interpréter les relations entre trois dimensions de l'évaluation des EIAH," *Actes de la conférence EIAH 2003*, in C. Desmoulins, P. Marquet, and D. Bouhineau (Eds.), Paris, INRP/ATIEF, 2003, pp. 391-402.
- [13] Bétrancourt, M., "L'ergonomie des TICE : quelles recherches pour quels usages sur le terrain ?," *Regards croisés sur la recherche en technologie de l'éducation*, in B. Charlier and D. Peraya (Eds.), Bruxelles, De Boeck, 2007, pp. 77-89.
- [14] Hart, S. G., Staveland, L. E., "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research," *Human Mental Workload*, in P. A. Hancock and N. Meshkati (Eds.), Amsterdam, North-Holland, 1988, pp. 139-183.