Projet ACI École et sciences cognitives 2003-2005

La dynamique des apprentissages : des fonctions cognitives à l'élaboration des connaissances

Le résumé de textes : de l'Analyse Sémantique Latente à l'élaboration d'un tuteur électronique

dirigé par Guy Denhière

Modélisation des processus de hiérarchisation et d'application de macrorègles et conception d'un prototype d'aide au résumé

Contribution du laboratoire des sciences de l'éducation de l'université Grenoble-II, novembre 2005

Maryse Bianco*, Philippe Dessus*, Benoît Lemaire**, Sonia Mandin*, & Patrick Mendelsohn*

*LSE et IUFM, Grenoble **Laboratoire Leibniz-IMAG, Grenoble

PLAN DE LA SECTION

Résumé

Résumer un texte, entre autres tâches, c'est être capable de le comprendre et de sélectionner les phrases les plus importantes du texte (Fayol, 1985). Dans les études suivantes, nous allons nous intéresser à la manière de simuler, par divers modèles, l'activité cognitive de résumer des textes, en la décomposant en différents modules. Le but final étant de proposer un tuteur d'aide à la rédaction de résumés, pour des élèves du secondaire, en utilisant ces modules. La première étude vise à mieux comprendre l'activité experte de résumé de textes, la seconde a pour but de simuler le processus de sélection de phrases importantes d'un texte et la troisième s'intéresse à l'activité de résumé chez des élèves de fin de collège et de lycée. Ces trois études ont abouti à la conception d'un tuteur d'aide au résumé.

1 OBJECTIFS INITIAUX

L'objectif de ce projet, tel qu'il était décrit initialement, était triple : théorique, expérimental et appliqué. Il s'agissait de (1) s'appuyer sur LSA pour bâtir un modèle de l'activité d'identification des informations principales d'un texte et procéder à des simulations sur des textes différents, (2) mettre en place un dispositif expérimental permettant de recueillir des données concernant les règles de résumé enseignées, les résumés produits par des élèves sur des textes de natures différentes, les stratégies utilisées par les élèves à partir des mouvements oculaires, (3) concevoir un tuteur électronique entraînant les élèves à la tâche de résumé.

La contribution décrite dans ce chapitre a pour finalité la conception de l'architecture générale du tuteur. Les choix liés à cette architecture ont fait l'objet de quelques tests, chez des producteurs experts (étudiants) et novices (collégiens et lycéens). Nous commençons par décrire brièvement les deux méthodes d'analyse automatique de textes qui sont au cœur du tuteur, puis différents modèles des processus cognitifs impliqués dans l'activité de résumer des textes, pour finir par décrire l'architecture du tuteur, ainsi que les prolongements possibles de cette recherche.

2 METHODE ET OUTILS EMPLOYES

La contribution grenobloise relève d'une activité de modélisation. Nous avons pour cela eu recours à un outil pour le calcul de similarité sémantique, l'analyse de la sémantique latente (LSA, Landauer & Dumais, 1997), et à un modèle de la compréhension de textes, CI/LSA (Lemaire *et al*., à paraître). Nous les détaillons maintenant.

2.1. Description de LSA

Le principe général de l'analyse de la sémantique latente (*Latent Semantic Analysis*, Landauer & Dumais, 1997) consiste à définir la signification des mots à partir des contextes dans lesquels ils apparaissent au sein de vastes corpus de textes. Il est en effet possible de déterminer le sens d'un mot à partir de son contexte, dès lors que ce mot est rencontré suffisamment souvent. LSA analyse les contextes d'occurrence des mots au sein d'un vaste corpus et réduit le bruit causé par la variabilité de l'emploi de ces mots dans la langue. Chaque mot est représenté par un vecteur dans un espace de plusieurs centaines de dimensions, c'est-à-dire par une suite de centaines de valeurs numériques. Cette représentation vectorielle permet aisément de calculer un vecteur pour une suite de mots, voire un texte entier, en ajoutant simplement les vecteurs des mots qui les composent. Ce formalisme de représentation ne possède pas le côté explicite des représentations symboliques du sens, mais il compense cela par une métrique objective rendant possible des comparaisons de signification entre mots ou groupe de mots, de manière complètement automatique. Par exemple, les deux phrases « J'ai perdu mon chat dans la forêt. » et « Le petit félin a disparu dans les arbres. » sont représentées par deux vecteurs qui sont très proches, indiquant que les significations correspondantes sont voisines. Les mesures d'association sémantique entre mots calculées par LSA corrèlent de manière satisfaisante avec les jugements produits par les humains (Foltz, 1996). D'autre part, de nombreux travaux utilisent LSA dans le cadre de l'analyse de résumés de textes (Wade-Stein & Kintsch, 2004 ; Yeh, Yang, & Meng, 2005).

2.2. Description de CI/LSA

Nous avons développé un modèle computationnel de la compréhension de textes permettant de suivre, proposition après proposition, la construction des relations associatives entre les éléments du texte, l'incorporation de connaissances extérieures et l'intégration de l'ensemble dans la représentation courante de la signification du texte. L'intérêt de la simulation informatique est d'affiner le modèle par la comparaison avec des données empiriques.

Ce modèle est inspiré du modèle de construction-intégration (Kintsch, 1998), de l'analyse de la sémantique latente (Landauer, 2002) et de l'algorithme de recherche contextualisée des associés sémantique d'une proposition (Kintsch, 2001 ; Lemaire & Bianco, 2003). Son architecture générale est la suivante (*voir figure 1*). Chaque proposition **du texte à analyser** est considérée successivement. La phase de construction permet d'établir une liste d'associés provenant de la mémoire sémantique (représentée par un espace sémantique LSA). Cette phase simule le recours aux connaissances. Par exemple, la proposition « *l'abeille récolte le nectar* » active le terme *miel*. La sélection est contextualisée en ce qu'elle ne garde parmi les voisins d'un terme que ceux qui sont proches d'au moins un des autres mots de la proposition. Par exemple, la proposition « *l'oiseau vole* » va récupérer le mot *ailes* (voisins de *vole* et proche de *oiseau*) mais pas le mot *pilote* (voisin de *vole* mais non de *oiseau*)

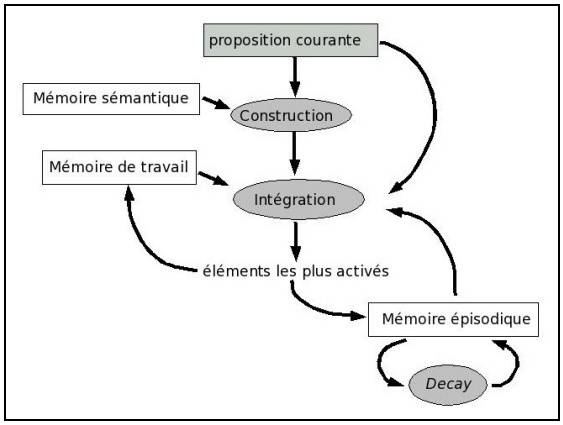

Figure 1 – Architecture du modèle de compréhension de textes.

Ces associés sémantiques, ainsi que la proposition courante, sont ajoutés aux éléments de la mémoire de travail correspondant au traitement des propositions antérieures. Des éléments traités antérieurement, mais non actifs, peuvent également être ajoutés s'ils sont sémantiquement proches des éléments courants. C'est le rôle de la mémoire épisodique. Tous ces éléments forment un vaste réseau duquel on va extraire les éléments les plus pertinents, c'est-à-dire ceux qui sont les plus associés aux autres. C'est la phase d'intégration. Ces éléments vont rejoindre la mémoire de travail et la proposition courante sera considérée comme étant traitée. L'annexe 1 décrit en détail un exemple de traitement.

2.3. Plan de la section

Nous allons maintenant détailler les différentes études que nous avons réalisées en utilisant l'un ou l'autre de ces modèles. Voici leur structure générale :

— *La première étude* vise à comprendre les relations sémantiques pouvant exister entre un résumé produit par des experts (étudiants avancés, dans l'étude 1a, enseignant prescripteur d'une méthode de résumé, dans l'étude 1b) et le texte-source (dorénavant TS). Cette étude permet aussi de tester, sur un grand nombre de résumés (une quarantaine), les capacités de LSA en ce domaine.

— *La deuxième étude* se centre sur une partie du processus de résumé de texte, le processus de hiérarchisation (*i.e.*, de détection des phrases importantes d'un texte). Elle confronte les jugements d'élèves novices (env. 350 élèves de collèges et lycées) à divers modèles computationnels de hiérarchisation. Des corrélations permettent de déterminer, selon le type de texte (narratif *vs*. expositif), le modèle le plus proche des jugements des participants.

— *La troisième étude* analyse une deuxième production des participants de l'étude précédente. Ces derniers ont de plus produit un résumé à partir de l'un des deux textes. Les résumés ont été analysés par une procédure automatique utilisant LSA, afin de déterminer quelles étaient les macrorègles utilisées (Kintsch & van Dijk, 1978).

— *Le dernier volet* de ce travail consiste à prendre en compte les résultats des trois études empiriques dans la conception d'un tuteur d'aide à la rédaction de résumés, dont nous présenterons l'architecture.

3 ÉTUDE 1 : L'ACTIVITE DE RESUME EXPERTE

Cette étude préliminaire a consisté en deux tâches, ayant pour but de mieux comprendre en quoi peut consister l'activité experte de résumé de textes. Tout d'abord, il s'agit de comparer avec LSA des résumés produits par des experts au texte-source. Ensuite, nous avons comparé avec LSA une quarantaine de couples texte-source/résumé-type, afin de mettre au jour d'éventuelles régularités.

3.1. Etude 1*a* : L'activité de résumé experte étudiée à partir des résumés d'étudiants

La première analyse porte sur un texte argumentatif tiré de Haxaire (1989). Elle a été réalisée selon le modèle de LSA associé à un corpus (*Le Monde*, 1999). Les résumés sont produits par 8 étudiants (2 en DEA et 6 en thèse en sciences de l'éducation ou en psychologie). Les consignes pour résumer n'imposaient ni contrainte de temps ni contrainte de taille. Les étudiants avaient à leur disposition des brouillons et un ordinateur pour dactylographier leur production. Tous l'ont spontanément rédigée directement sur l'ordinateur. Deux résultats se sont avérés particulièrement intéressants :

— les proximités sémantiques calculées avec LSA entre les résumés des étudiants et le texte-source sont relativement homogènes ($\mu = 0{,}454$; $\sigma = 0{,}072$). Toutefois, deux résumés sont très éloignés du texte et un très proche. Les étudiants se répartissent donc selon 2 types de stratégies : ceux qui résument un minimum d'idées, jugées essentielles, émises dans le texte-source (et s'en éloignent), et ceux qui tentent d'intégrer un maximum d'idées afin de minimiser la perte d'informations (et restent près du texte-source).

— les proximités sémantiques calculées avec LSA entre les résumés des étudiants et chaque paragraphe du texte-source montrent des différences dans la façon de traiter chaque paragraphe. Premièrement, nous observons une forte variabilité dans le traitement des paragraphes du texte. Pour chaque paragraphe du texte-source, les moyennes des indices de proximité sémantique calculés entre les paragraphes et les résumés des huit étudiants varient entre 0,157 ($\sigma = 0{,}069$) et 0,333 ($\sigma = 0{,}125$). Trois paragraphes sur 8 sont particulièrement bien repris et 3 autres paragraphes ne le sont que faiblement. Cependant, ces différences ne semblent pas être dues à la position des paragraphes dans le texte, mais à leur contenu. Deuxièmement, concernant la façon de résumer des étudiants pris isolément, nous constatons deux groupes d'étudiants différents. Les corrélations de Pearson significatives obtenues entre les séries de proximités sémantiques des différents paragraphes pour chaque étudiant et les mêmes valeurs concernant le résumé-type

(Haxaire, 1989) divisent les étudiants en deux groupes. Les différences des résumés portent essentiellement sur trois paragraphes. Deux de ces paragraphes sont fortement éloignés du thème principal et l'autre a une corrélation très forte avec un quatrième paragraphe (r = -0,771). Dans ce dernier cas, plus l'un des paragraphes est traité et moins l'autre l'est. Cela s'explique par le fait qu'ils sont complémentaires, présentant le même objet sous deux angles différents. Alors que certains résumés reprennent la double façon de l'auteur de présenter un même objet, d'autres se contentent de celle allant dans le sens de l'idée principale. Ainsi, il apparaît qu'il existe deux profils d'étudiants concernant la façon de traiter certains paragraphes : ceux qui excluent les paragraphes n'apportant rien à l'argumentation et ceux les résumant tous.

Nous noterons toutefois que, dans les deux profils décrits ci-dessus, les étudiants dont le résumé est particulièrement éloigné ou proche du texte-source, et les étudiants qui traitent particulièrement les paragraphes les plus importants sont les mêmes. L'étudiant proche du texte-source a donc réussi à intégrer un maximum d'informations à son résumé tout en respectant l'importance des paragraphes vis-à-vis du thème dominant.

3.2. Etude 1*b* : L'activité de résumé experte étudiée à partir de résumés académiques « types »

L'étude 1*a* ne reposant que sur un seul texte, nous avons choisi d'analyser 43 textes proposés dans Haxaire (1989) avec les résumés types qui leur sont associés. Nous espérions trouver ainsi des régularités qui permettraient de déterminer de bons résumés. Nous avons pu mettre en évidence une dépendance entre les indices issus de la comparaison de chaque phrase des textes-sources à l'intégralité des résumés (quelles sont les phrases des textes-sources dont les idées sont le mieux intégrées aux résumés ?) et les indices issus de la comparaison de chaque phrase des textes-sources à l'intégralité du texte-source lui-même (quelles sont les phrases des textes-sources dont les idées sont les plus proches de l'idée global du texte-source ?). Les résultats obtenus montrent que les résumés reprennent les phrases les plus proches de l'idée générale du texte : 32 documents sur 43 présentent des corrélations de Pearson significatives (*voir figure 2 ci-dessous*). Toutefois, il reste encore des recherches à mener pour parvenir à dégager une base commune à tous ces résumés.

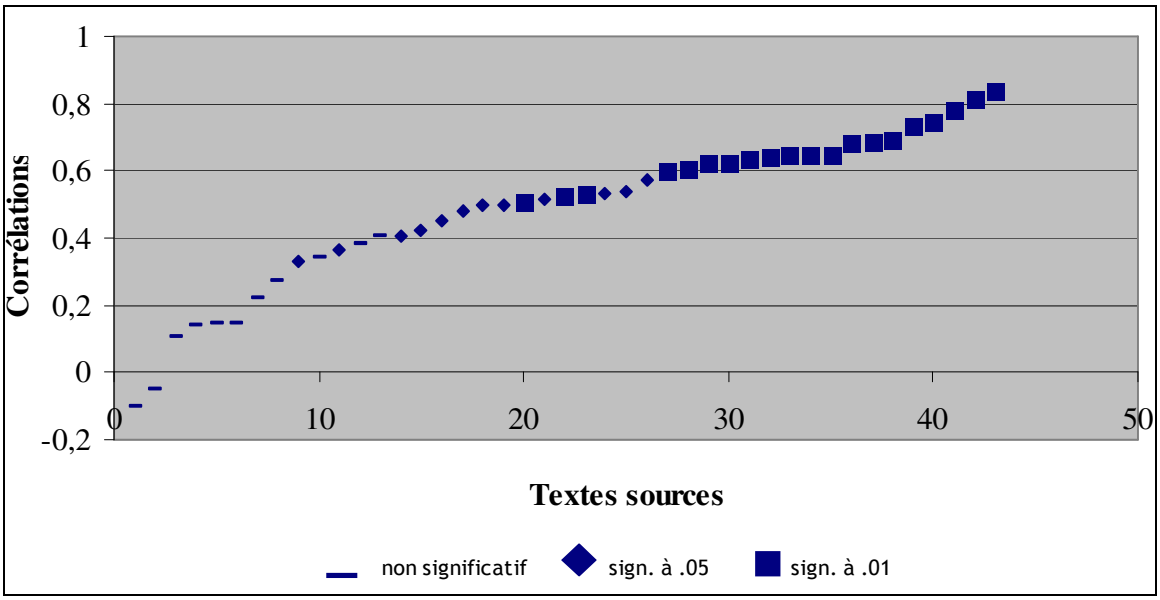

Figure 2 – Corrélations de Pearson entre les comparaisons sémantiques « phrases des TS - résumés » et « phrases des TS - TS entier ».

4 ÉTUDE 2 : MODELISATION DU PROCESSUS DE HIERARCHISATION

D'un point de vue psychologique, résumer un texte comprendrait deux opérations principales : la hiérarchisation des propositions du texte-source puis la sélection de certaines de ces dernières pour produire le résumé, cette sélection pouvant passer par une transformation et/ou une réorganisation des propositions (Fayol, 1985). L'objet de cette deuxième étude (Lemaire, Mandin, Dessus & Denhière, 2005) a été de recueillir des données empiriques sur la manière dont les élèves de niveau secondaire (fin de collège et lycée) résument des textes et de confronter ces données à différents modèles computationnels. Dans cette deuxième étude, nous nous sommes plus particulièrement intéressés à un processus cognitif engagé dans l'activité de résumé : la hiérarchisation des idées d'un texte. La troisième étude, qui suit, s'intéresse au processus de sélection de propositions par l'utilisation de macrorègles, à partir de l'analyse d'une tâche demandée aux mêmes participants.

4.1. Données expérimentales

Nous avons fait passer un questionnaire à des élèves de collège et lycée, comprenant deux tâches, d'une part, souligner les 3 à 5 phrases les plus importantes d'un autre texte et, d'autre part, résumer un texte, ces deux tâches étant contrebalancées dans chaque classe et pour deux élèves voisins, afin d'éviter d'éventuelles copies (*voir l'annexe 3 qui reprend les résultats bruts de cette tâche*). Afin de ne pas transformer l'activité de résumé en activité de comptage de mots, la consigne de résumé était peu contraignante. Elle demandait précisément de « *Rédiger un résumé qui raconte la même histoire que le texte ci-dessous* », sans contrainte formelle de longueur (*voir l'annexe 2*). En revanche, des contraintes de situation ont été introduites : rédiger le résumé en 30 min, sans l'aide de dictionnaires ni brouillon, sur une page A4. Il était également possible de souligner ou d'annoter la page du texte à résumer. De plus, les élèves devaient compléter un bref questionnaire leur demandant quelques informations à propos de leur niveau scolaire, et sur les enseignements au résumé de textes éventuellement reçus. Nous avons fait passer cette tâche à divers niveaux de secondaire (*voir tableau ci-dessous*), de manière à observer d'éventuels effets du niveau scolaire sur la performance.

Tableau 1 – Principales caractéristiques des classes ayant réalisé la tâche de l'étude 2.

Niveau de classe	Nombre de classes par niveau	Nombre d'élèves
4e	2	50
3e	3	79
CAP	1	13
2e BEP	2	44
2e	4	108
1re	1	41
Total	14	345

Nous avons sélectionné deux textes (*voir protocole en annexe 2*) selon les critères suivants :

— textes de niveau 4e (Pfeffer, 1989 ; Vidal, 1984), de manière à ce qu'ils puissent être compris de tous les participants ;

— proposer aux élèves un texte narratif (*Miguel*) et un texte expositif (*Les pharmacies des éléphants*). Le texte narratif, en proposant une structure linéaire d'événements, est censé être plus aisément résumable que le texte expositif, censé décrire des concepts plus abstraits (Brewer, 1980).

— peu d'exemples sont présents. En effet, un test préliminaire sur quelques élèves nous a montré que les exemples peuvent être totalement laissés de côté par les élèves, les amenant à réaliser un résumé de quelques phrases.

Voici les principaux traitements réalisés à partir de ces données. Les phrases principales repérées (*i.e.*, les plus importantes) par les élèves seront comparées avec diverses données obtenues avec nos deux modèles computationnels.

— Tout d'abord, chaque phrase soulignée pourra être comparée avec l'ensemble du texte-source, la comparaison s'effectuant avec LSA entre vecteurs représentant les phrases. Notre hypothèse étant que plus une phrase est importante, plus elle est sémantiquement reliée avec le texte-source *en entier*.

— Ensuite, nous vérifierons si les ruptures de cohérence entre deux phrases consécutives d'un texte permettent de prédire leur importance. En effet, il est possible de considérer que, si deux phrases consécutives d'un texte sont cohérentes, c'est qu'elles expriment des événements causalement reliés, et également qu'elles peuvent être jugées comme plus importantes que les autres (Trabasso & Sperry, 1985).

— Enfin, le texte-source pourra être traité par le modèle couplant CI (Kintsch, 1988) et LSA. Ici, nous vérifierons que les phrases restant activées une fois le texte « lu » sont bien celles soulignées par les élèves.

Il est à noter que ces trois modèles utilisent de manière de plus en plus contraignante l'ordre des phrases du texte-source. Le premier modèle donne les mêmes résultats quel que soit l'ordre des phrases du texte, le second considère les phrases contiguës du texte-source (désormais TS) deux à deux ; enfin le troisième s'intéresse aux phrases du texte dans l'ordre exact de son écriture.

Chacun de ces traitements a été réalisé dans deux espaces sémantiques différents. Un espace censé représenter des textes auxquels sont exposés des enfants (issu d'un corpus de 3,3 millions de mots comprenant des productions d'enfants, des contes, des textes encyclopédiques, des manuels scolaires et deux dictionnaires), un deuxième espace représente des textes auxquels sont exposés des adultes (issu d'un corpus de 13 millions de mots, comprenant le corpus précédent ainsi que deux corpus de 5 millions de mots chacun provenant de textes littéraires et d'articles du quotidien *Le Monde*). Ces deux corpus permettent de simuler des connaissances différentes, et nous avons vérifié, dans les tests avec les participants humains, quel est le corpus le plus en adéquation avec leurs performances (Denhière & Lemaire, 2004a).

4.2. Hiérarchisation par comparaison avec le texte

Commençons par vérifier si l'indice de proximité d'une phrase avec le TS est un bon prédicteur de son importance. Le tableau 2 ci-dessous reprend les corrélations données concernant nos deux textes. La valeur de LSA que nous avons comparée au pourcentage du soulignement de chaque phrase est la similarité de cette phrase à l'ensemble du texte. Le fait d'ajouter des connaissances du monde à l'espace de LSA augmente systématiquement la similarité élèves/LSA. Cela peut être dû au fait que le corpus enfants contient des textes pour des enfants de 7 à 11 ans, donc plus jeunes que les participants de notre étude. L'autre fait marquant est que les corrélations sont meilleures pour le texte expositif que pour le texte narratif.

Tableau 2 – Valeurs de corrélation entre l'importance des phrases évaluée par les élèves et leur indice de proximité avec le texte-source.

Textes	Sélection par LSA/élèves	4e sp.	4e	3e	CAP	2e BEP	2e	1re	Totalhtech	Total
Miguel	Corpus enfants	0,41*	0,24	0,04	0,16	0,06	0,21	0,22	0,19	0,20
	Corpus adulte	0,45*	0,37	0,18	0,25	0,21	0,34	0,28	0,31	0,33
Eléphants	Corpus enfants	-0,26	0,40	0,62**	0,42	0,26	0,48*	0,45	0,53*	0,48*
	Corpus adulte	-0,33	0,52*	0,70**	0,48*	0,38	0,59*	0,58*	0,64**	0,60**

* p < 0,05 ; ** p < 0,01
Légende : Total htech : sans les classes « techniques »

4.3. Hiérarchisation par détection de rupture de cohérence interphrases

Le niveau de compréhension d'un texte lu est lié à sa cohérence, qui exprime à son tour le niveau de relation causale entre deux phrases (Langston & Trabasso, 1999). Foltz *et al.* (1998) ont utilisé les capacités de LSA à évaluer le degré de relation sémantique entre phrases adjacentes pour mesurer la cohérence textuelle. Ils sont partis de deux études antérieures évaluant les performances de compréhension de sujets selon le niveau de cohérence, manipulé, des textes qu'ils lisaient. LSA compare les phrases contiguës de chaque texte deux à deux, la moyenne de ces comparaisons donnant un score global de cohérence. Ces scores moyens de cohérence corrèlent très fortement ($r = 0,99$) avec la compréhension des lecteurs.

Le deuxième test que nous avons réalisé à partir des phrases soulignées par les élèves est de déterminer si des indices de cohérence pouvaient permettre de prédire l'importance d'une phrase. L'idée est que, tout comme les premières et dernières phrases d'un texte semblent souvent appartenir à son résumé, il pouvait en être de même pour les premières et dernières phrases d'un bloc cohérent. Nous avons donc recherché les ensembles de phrases cohérents, c'est-à-dire des successions de similarités interphrases élevées, au-dessus d'un seuil arbitraire. Par exemple, si les phrases 12, 13, 14 et 15 ont des similarités interphrases respectives de .1 (entre la 12 et la 13), .5 (entre la 13 et la 14) et .1 (entre la 14 et la 15), on dira que les phrases 13 et 14 forment un bloc cohérent. Les phrases appartenant à un bloc cohérent ont été codées avec la valeur 1 et les autres avec la valeur 0. Aucun bloc n'étant supérieur à trois phrases, nous n'avons pas eu à déterminer comment coder l'intérieur du bloc. Encore une fois, les performances des sujets à propos du texte Miguel ne sont pas correctement prédites par le modèle, quel que soit le corpus. En revanche, pour le texte Eléphants, le lien entre le choix des phrases selon leur importance et leur cohérence est assez élevé.

Tableau 3 – Valeurs de corrélation entre la cohérence des phrases des textes évaluée par LSA et leur importance évaluée par les élèves.

R	Corpus	4e sp.	4e	3e	CAP	2e BEP	2e	1re	Totalhtech	Total
Miguel	enfant	0,22	0,02	-0,01	0,05	0,16	0,13	0,05	0,07	0,09
	adulte	-0,05	0,05	0,03	-0,25	0,22	0,18	0,03	0,10	0,10
Eléphants	enfant	-0,08	0,17	0,34	-0,33	0,03	0,13	0,29	0,23	0,18
	adulte	0,13	0,44	0,49*	-0,21	0,24	0,32	0,48*	0,43	0,39

* p < 0,05
Légende : Total htech : sans les classes « techniques »

4.4. Hiérarchisation par simulation du processus de Construction/Intégration

Nous avons ensuite utilisé le modèle de compréhension de textes décrit plus haut au § 2.2. Deux types de simulation sont possibles avec ce modèle. La première concerne l'utilisation du modèle pour tenter de reproduire la sélection des phrases importantes (première tâche de

l'expérimentation décrite précédemment). La seconde vise à rendre compte de la compréhension de texte qui a précédé la production du résumé par les élèves (seconde tâche). Nous n'avons pour le moment effectué des simulations que pour la première tâche.

Il faut noter que, tel quel, ce modèle n'est probablement pas le plus approprié puisque c'est un modèle de compréhension alors que la tâche exige une phase de compréhension suivie d'une tâche de sélection. En particulier, le modèle donne une valeur d'activation plus importante aux dernières phrases, parce qu'il modélise l'oubli dans le *buffer* épisodique, alors que les sujets ont vraisemblablement corrigé cette différence lors de la sélection des phrases importantes. Une des finalités du projets est de coupler la simulation de la compréhension avec la simulation de la sélection, ce qui reviendra à réaliser les mesures de similarité sémantiques décrites précédemment, non pas entre chaque phrase et le texte entier, mais entre chaque phrase et le résultat de la compréhension du texte entier.

Les simulations ont été effectuées avec l'espace sémantique issu du corpus Enfants et celles issues du corpus Adultes. Elles ont consisté à comparer les pourcentages moyens de sélection des phrases par les enfants avec les valeurs d'activation des propositions du texte à l'issue du traitement du texte entier. Les principaux paramètres sont : nombre de voisins = 3, .3 < poids < .7, récupération en mémoire épisodique si activation > .5. Les résultats détaillés sont présentés dans le tableau 4 suivant. En résumé, il n'y a pas de corrélations pour l'espace enfants. En revanche, avec le corpus adultes, les corrélations valent $r(24)=0{,}42$ pour le texte narratif (*Miguel*) et $r(18)=0{,}40$ pour le texte expositif (éléphants).

Ces différents modèles suivent une progression dans la manière dont les phrases sont évaluées. Dans le premier modèle, les phrases sont comparées au texte-source en entier. Le modèle 2 procède à une analyse plus fine, car les phrases sont comparées à d'autres phrases du texte-source, sans prendre en compte leur position dans le texte. Le troisième modèle consiste à un nouveau raffinement, en ce qu'il est dépendent de l'ordre des phrases, en extrayant automatiquement la macrostructure du texte. Ce dernier modèle est assez différent des autres. Il est fondé sur une représentation de différentes structures cognitives et est capable d'activer des concepts qui ne sont pas effectivement présents dans le texte. Ces caractéristiques sont spécialement nécessaires pour traiter des textes narratifs, car les connexions interphrases de ces derniers sont souvent moins évidentes que pour des textes expositifs. Toutes les phrases de notre texte expositif (*les éléphants*) étaient à propos du même sujet, alors que le domaine du texte narratif était plus large. Cela pourrait expliquer pourquoi le dernier modèle donne de meilleurs résultats avec le texte narratif.

Tableau 4 – Valeurs de corrélation entre l'importance des phrases évaluée par les élèves et les valeurs d'activation des propositions dans le modèle.

Textes	Sélection par LSA/élèves	4e sp.	4e	3e	CAP	2e BEP	2e	1re	Totalhtech	Total
Miguel	*r* activations corpus enfant	-0,11	-0,04	0,03	-0,15	0,02	-0,02	-0,02	-0,01	0,02
	r activations corpus adulte	0,31	0,42*	0,42*	0,27	0,16	0,39	0,41*	0,43*	0,42*
Eléphants	*r* activations corpus enfant	-0,23	0,28	0,18	0,11	0,28	0,26	0,25	0,25	0,27
	r activations corpus adulte	0,13	0,3	0,31	0,35	0,44	0,39	0,37	0,37	0,4

* $p < 0,05$; ** $p < 0,01$
Légende : Total htech : sans les classes « techniques »

5 ETUDE 3 : MODELISATION DE L'UTILISATION DES MACROREGLES

Comme pour la première étude, nous avons comparé les textes-sources aux résumés produits par des élèves de collèges et de lycées. Cette étude consiste en l'analyse du deuxième volet de l'expérimentation décrite dans la sous-section précédente. L'analyse sémantique utilise LSA avec le même corpus du *Monde* 1999. Là aussi les proximités sémantiques sont relativement élevées. Tous niveaux scolaires confondus, leur moyenne est de 0,52 ($\sigma = 0,14$) pour le texte narratif et de 0,58 ($\sigma = 0,14$) pour le texte expositif. Toutefois, dans le cas du texte narratif, nous observons que les moyennes des classes de 4e et de 2e sont significativement différentes ($t(92) = 3,043$; $p = 0,003$) et que celles de 4e et de 1re tendent aussi à l'être ($t(48) = 1,762$; $p = 0,084$). Avec l'âge, il apparaît donc que les résumés se distancient du texte-source. Nous pensons que cela est dû à moins de copies au profit de davantage de restructuration du texte-source.

Dans le cas du texte expositif, ce sont des différences entre les 4es et chaque autre niveau qui sont remarquées ($F(1, 64) = 19,227$; $p < 0,001$ entre les 4es et les 3es ; $F(1, 110) = 12,735$; $p = 0,001$ entre les 4es et les 2es ; $F(1, 45) = 5,322$; $p = 0,026$ entre les 4es et les 1es ; $F(1, 31) = 7,292$; $p = 0,011$ entre les 4es et les CAP). Ces différences peuvent être dues au fait que le texte expositif est un texte dont le schéma est moins bien acquis. Ces résultats nous incitent à analyser plus finement les résumés. Pour cela, nous avons comparé chaque phrase des textes-sources à chaque phrase des résumés afin de détecter (Kintsch & van Dijk, 1978) :

— *les suppressions*, c'est-à-dire les phrases du TS dont les idées ne sont pas reprises dans le résumé (suppressions) : ce sont les phrases du texte-source dont l'indice de comparaison aux phrases du résumé est toujours faible.

— *les répétitions*, c'est-à-dire les phrases du TS dont les idées sont reprises une ou plusieurs fois dans le résumé. Ce sont les phrases du texte-source dont l'indice de comparaison aux phrases du résumé est élevé plus d'une fois.

— *les élaborations*, c'est-à-dire les phrases du résumé qui ne reprennent aucune idée du texte-source : ce sont les phrases du résumé dont l'indice de comparaison aux phrases du texte-source est toujours faible.

— *les généralisations*, c'est-à-dire les phrases du résumé qui reprennent les idées de plusieurs phrases du TS. Ce sont les phrases du résumé dont l'indice de comparaison aux

phrases du texte-source est élevé plusieurs fois. Le résultat pourrait être affiné en tenant compte des proximités sémantiques entre les phrases du texte-source susceptibles d'être reprises. Il s'agirait d'une généralisation que si les phrases n'étaient pas similaires.

Bien entendu, pour déterminer si les indices sont faibles ou élevés, il faut tenir compte d'un seuil. Les résultats qui suivent tiennent compte d'un seuil variable en fonction de chaque résumé. Ce seuil est équivalent à la moyenne des indices obtenus lors des comparaisons avec LSA plus l'écart type. D'autres seuils ont été testés (0,2 ; moyenne ; moyenne + 0,5 écart type), mais c'est avec celui-ci que nous obtenons les résultats les plus intéressants. Dans le cas du texte expositif, nous obtenons des différences à propos des suppressions. Les troisièmes effectuent davantage de suppressions que les secondes et premières ($t(57) = 2{,}325$; $p = 0{,}024$ entre les 2es et les 3es ; $t(122) = 3{,}147$; $p = 0{,}002$ entre les 1res et les 3es). Nous pouvons expliquer cela par un maintien de la paraphrase chez les 4es et une utilisation plus variée des différentes stratégies possibles chez les 2es et les 1res. Les 3es sont dans une phase où ils ne se contentent plus de paraphraser. Ils contractent désormais le texte aussi par suppression d'informations.

Un autre point est à remarquer dans l'analyse du texte expositif. Concernant les généralisations, les suppressions et les répétitions, les CAP obtiennent des moyennes significativement différentes de toutes les autres classes à l'exception de la 3e générale dont ils se rapprochent le plus. Ils réalisent peu de généralisations et de répétitions pour adopter une stratégie basée quasi exclusivement sur la suppression d'informations. Ainsi, alors que pour le texte narratif, tous les élèves utilisaient majoritairement la suppression, pour le texte expositif, les CAP sont les seuls à l'utiliser. Il semblerait donc qu'ils aient des difficultés à adapter leurs stratégies au type du texte lu. Enfin, il nous faudra parvenir à distinguer les élaborations pertinentes d'élaborations non pertinentes, que LSA ne peut actuellement parvenir à distinguer. Un recours à des juges humains experts pourra permettre cette distinction (*voir § 7.2*).

Tableau 5 – Nombre de phrases de chaque texte ayant subi une généralisation, élaboration, suppression ou répétition lors du résumé, selon notre modélisation.

	Texte narratif					**Texte expositif**			
	Génér.	Elabo.	Suppr.	Répét.		Génér.	Elabo.	Suppr.	Répét.
4e N=28	7,79 (3,521)	0,54 (0,922)	10,14 (4,327)	9,32 (4,199)	**4e N=27**	6,56 (2,722)	1,26 (1,228)	6,07 (2,702)	7,41 (2,845)
3e N=36	7,14 (4,162)	0,39 (0,645)	11,11 (4,646)	8,06 (3,680)	**3e N=39**	6,13 (3,318)	1,21 (1,128)	6,85 (3,321)	6,90 (3,424)
CAP N=5	6,60 (3,362)	0,40 (0,548)	12,0 (5,612)	9,0 (5,701)	**CAP N=6**	4 (1,897)	0,83 (0,983)	8,33 (1,862)	2,845 (2,503)
2e N=66	6,83 (3,694)	0,52 (0,614)	10,88 (4,562)	8,74 (3,888)	**2e N=85**	6,92 (2,396)	0,94 (0,836)	5,32 (2,042)	7,42 (2,843)
1re N=22	7,32 (2,607)	0,50 (0,673)	10,32 (3,138)	9,14 (4,074)	**1e N=20**	6,65 (2,54)	1,05 (0,945)	4,85 (2,681)	7,95 (2,724)

6 CONCEPTION D'UN TUTEUR D'AIDE AU RESUME DE TEXTE

Nous avons opérationnalisé les macrorègles précédentes pour réaliser un tuteur d'aide à la conception de résumés par les élèves (Mandin, Dessus, Lemaire & Bianco, 2005). Pour couvrir tous les cas, nous avons augmenté le nombre de stratégies employées par les élèves

pour produire un résumé. Tout d'abord, certaines phrases sont réutilisées dans le résumé par l'intermédiaire d'une *paraphrase*. D'autres sont *recopiées* quasiment telles quelles par les élèves. Enfin, certaines sont ajoutées alors qu'elles sont sans rapport avec le TS (*hors-sujet*). Nous obtenons donc cinq catégories de phrases pour le résumé : généralisation, construction, paraphrase, recopie, hors-sujet, et deux pour le TS : supprimée ou non supprimée.

Pour catégoriser chaque phrase du résumé, nous les comparons donc à chacune des phrases du TS. Nous observons alors la distribution de ces similarités sur l'échelle [0, 1] : 0 signale les phrases sans rapport, 1 celles identiques. La figure 3 ci-dessous est un exemple de distribution des 18 phrases de notre texte expositif comparées à la phrase suivante d'un résumé : « Les éléphants sont des mammifères qui mangent parfois de la terre ». Les quatre premières phrases du TS et leur valeur de similarité sont données à titre indicatif. Nous avons besoin de formaliser les notions de « suffisamment proche », « recopiées quasiment telles quelles » ou « sans rapport avec le TS », en utilisant trois seuils : seuilHorsSujet (de valeur 0,10), seuilParaphraseMin (0,35) et seuilParaphraseMax (0,75), qui définissent quatre zones de similarité sémantique (*cf. figure 2 ci-dessous*) :

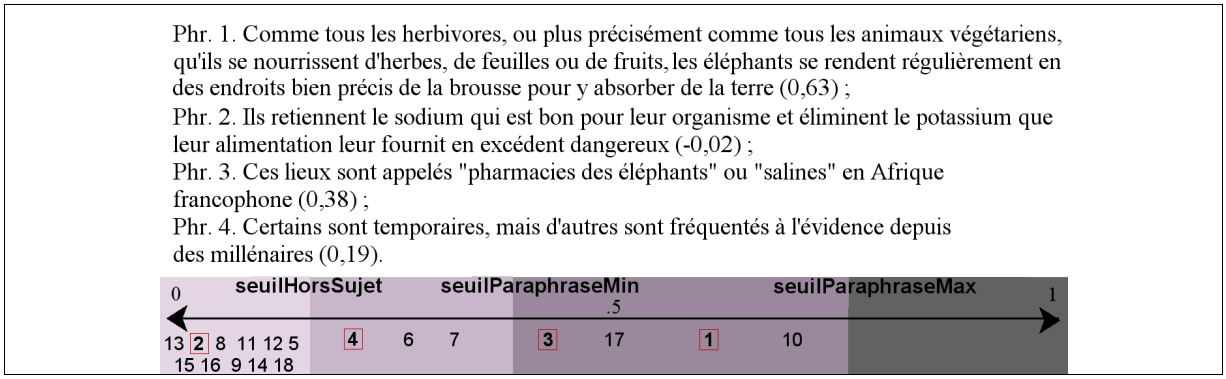

Figure 3 – Représentation graphique de la comparaison entre une phrase du résumé (*cf. texte ci-dessus*) et chaque phrase du TS.

— zone 1 ($s <$ seuilHorsSujet) : elle présente des phrases du TS sans rapport avec la phrase du résumé (cas de la phrase 14). Si elles sont toutes ici, ce sera le signe que la phrase du résumé considérée est hors-sujet ;

— zone 2 (seuilHorsSujet $\leq s <$ seuilParaphraseMin) : elle présente des phrases du TS ayant une certaine relation avec la phrase du résumé (cas de la phrase 6) ;

— zone 3 (seuilParaphraseMin $\leq s <$ seuilParaphraseMax) : elle présente des phrases proches de la phrase du résumé (cas de la phrase 10) ;

— zone 4 ($s \geq$ seuilparaphraseMax) : elle présente les phrases trop proches de la phrase du résumé (aucune phrase dans notre exemple), signalant probablement une recopie presqu'à l'identique.

Soit $Q_i = (x_1, x_2, x_3, x_4)$ les effectifs de chaque zone pour la phrase *Ri* du résumé. Précédemment, ce quadruplet valait (11, 3, 4, 0), le nombre de phrases du TS valant $x_1+x_2+x_3+x_4$. La répartition des similarités sémantiques entre ces quatre zones permet de catégoriser les phrases du résumé, en posant qu'une variable pouvant prendre une valeur quelconque est représentée par « ? ». Nous dirons que R_i est :

— une *recopie* si $Q_i = (?, ?, ?, N)$, $N \geq 1$ (au moins une phrase du TS très proche de R_i) ;

— une *généralisation* si $Q_i = (?, ?, N, 0)$, $N \geq 2$ (plusieurs phrases du TS proches de R_i) ;

— une *paraphrase* si $Q_i = (?, ?, 1, 0)$ (une seule phrase du TS proche de R_i) ;

— une *construction* si $Q_i = (?, N, 0, 0)$, $N \geq 1$ (aucune phrase du TS proche de R_i, mais au moins une ayant une certaine relation avec R_i). Nous sommes conscients du fait que cette règle, comme pour la généralisation, ne suit pas la définition exacte d'une construction. LSA rend difficilement compte des relations de causalité entre phrases (*cf.* toutefois

Graesser *et al.*, 2000), mais nous pouvons supposer qu'une phrase du résumé généralisée entretiendra une relation sémantique assez proche avec au moins une phrase du TS, puisque faisant état du même fait général ;

— *hors-sujet* si Q_i = (?, 0, 0, 0) (toutes les phrases du TS sans rapport avec R_i).

Cette catégorisation de chacune des phrases du résumé constitue une modélisation des activités cognitives de l'élève qui va nous permettre de le guider. On peut faire la même chose pour le processus de suppression de phrases qui, par définition, n'est pas observable dans le résumé. Nous dirons donc qu'une phrase T_i du TS est :

— *supprimée* s'il n'existe pas de phrase du résumé qui lui soit proche, c'est-à-dire si pour chaque phrase du résumé, Q_i = (?, ?, 0, 0) ;

— ou au contraire *conservée*, si pour chaque phrase du résumé, Q_i = (0, 0, ?, ?).

Nous avons conçu un premier prototype d'EIAH (Environnement informatique d'apprentissage humain) qui combine les informations précédentes (hiérarchie, catégorie) pour établir un diagnostic du résumé de l'élève. Il indique clairement les phrases sur lesquelles l'élève doit porter son attention (*cf.*, figure 4 ci-dessous, une copie d'écran de l'interface montrant les messages produits). Par exemple, il est préférable de l'inciter à utiliser les macrorègles de construction et de généralisation plutôt que les stratégies de recopie, paraphrase ou suppression (Brown *et al.*, 1983), en attribuant à ces opérations des commentaires, mais aussi des poids différents. De plus, ce poids doit dépendre du niveau d'importance de la phrase du TS, tel qu'évalué par le processus de hiérarchisation. Enfin, il est important de prendre en compte la taille des phrases du résumé, les mesures de LSA étant d'autant plus fiables que les phrases comparées sont longues.

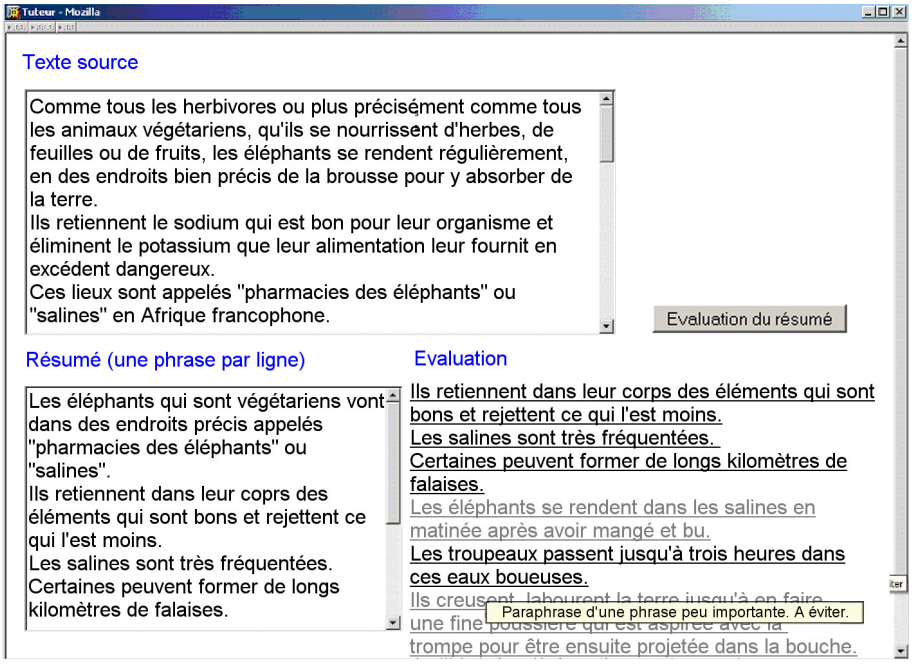

Figure 4 – Copie d'écran de notre environnement.

7 TRAVAUX FUTURS

7.1. L'activité de correction experte de résumés de textes

Le modèle d'utilisation de macrorègles précédemment décrit pourrait être utilisé pour évaluer automatiquement des résumés de textes. Afin de vérifier cette hypothèse et éventuellement

fournir des données empiriques permettant d'ajuster le modèle, nous envisageons de solliciter des enseignants pour corriger à voix haute les résumés de textes que nous avons faits réaliser à des élèves de collèges et lycées. A l'heure actuelle, un enseignant s'est déjà prêté à cet exercice, sur deux résumés de deux TS différents. Le traitement de son discours montre que les résumés sont effectivement jugés sur les critères de pertinence des informations sélectionnées et des règles appliquées pour passer du TS au résumé (la généralisation est préférée à la copie et à la paraphrase). Nous nous apercevons aussi que dans le cas du résumé d'un texte narratif, l'utilisation correcte du discours indirect est particulièrement appréciée par le correcteur.

7.2. Validation du modèle d'utilisation des macrorègles

Jusque-là, nous avons supposé que la façon dont on utilisait LSA dans le modèle nous permettait de détecter des copies, des paraphrases, des généralisations, des élaborations et des hors-sujets réellement réalisés par les producteurs. Une comparaison à des jugements humains est toutefois encore nécessaire pour pouvoir l'affirmer. Un protocole a donc été mis en place pour comparer les règles détectées par des experts (étudiants) dans des résumés aux règles détectées avec notre modèle. Nous espérons également que nous pourrons ainsi affiner les seuils utilisés par LSA.

7.3. Conception et développement d'un tuteur électronique d'aide au résumé de texte

Ce tuteur, tel qu'il a été conçu pour l'instant, a pour objectif d'aider les apprenants à améliorer leur utilisation des macrorègles et leur compréhension du TS. Nous prévoyons de tester deux facteurs d'apprentissage. Le premier concerne l'effet d'une négociation. Nous envisageons de gérer dans le tuteur deux types d'évaluations de l'utilisation des macrorègles : celle de l'utilisateur et celle du système. Celle de l'utilisateur se fera par étapes guidées et celle du système se fera par l'application du modèle opérationnel utilisant LSA. Nous espérons que la confrontation de l'apprenant aux accords et divergences qui existent entre ces deux évaluations lui permettront de réfléchir à sa pratique et de l'améliorer. Un second facteur doit aussi être testé. Il s'agit de l'exposition à des exemples : soit des exemples d'autres résumés de textes, soit des exemples d'applications de macrorègles. L'activité de résumé de textes est complexe et difficilement explicite. Nous faisons donc l'hypothèse que l'exposition à des exemples permet aux apprenants de converger vers un modèle de l'activité tout aussi bien, sinon mieux, que d'essayer de lui enseigner classiquement les règles pour résumer.

8 CONCLUSION

En résumé, voici les principaux résultats de ces études.

Tout d'abord, il semble qu'on puisse dégager des profils d'étudiants résumeurs en comparant leur résumé au texte-source (étude 1). Ensuite, il apparaît que l'espace sémantique dans lequel se font les comparaisons a une grande importance : l'espace sémantique « adulte » permet d'obtenir systématiquement des performances plus proches de celles des élèves. En revanche, les comparaisons interclasses n'ont pas donné des résultats conformes à nos hypothèses : ce ne sont pas toujours les élèves des niveaux supérieurs (1re) qui se distinguent des autres (*i.e.*, qui ont des résultats plus proches de certaines modélisations que ceux des autres classes).

Ensuite, les performances des différents modèles testés (études 2 et 3) semblent être sensibles à la nature des textes traités (narratif *vs* expositif). La méthode de hiérarchisation par comparaison avec le texte-source donne de bien meilleurs résultats avec le texte expositif ; la méthode CI/LSA, elle, est moins sensible au type de texte, tout en donnant des résultats moins proches de ceux des participants. La méthode basée sur la cohérence donne des résultats

inférieurs aux deux autres. Ces différences peuvent s'expliquer par le fait que CI/LSA traite mieux la structure du texte, et donne donc de meilleurs résultats pour des textes narratifs. En revanche, pour des textes expositifs, la simple sommation des vecteurs des mots rend suffisamment compte de la signification globale du texte.

9 PUBLICATIONS CONSECUTIVES A CES RECHERCHES

Lemaire, B., Mandin, S., Dessus, P. & Denhière, G. (2005). Computational cognitive models of summarization assessment skills. In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society (CogSci' 2005)* (pp. 1266-1271). Mahwah : Erlbaum.

Mandin, S. (2005). Exemple d'un tuteur d'aide au résumé dans l'utilisation d'une analyse sémantique latente. *2nd séminaire sur l'Analyse des Discours en Éducation (ADEE 2005).* Grenoble : IUFM, 5-6 juillet.

Mandin, S., Dessus, P., Lemaire, B. & Bianco, M. (2005). Un EIAH d'aide à la production de résumés de textes. In P. Tchounikine, M. Joab, & L. Trouche (Eds.), *Actes de la conférence EIAH 2005* (pp. 69-80). Paris : I.N.R.P.

Mandin, S., Dessus, P. & Lemaire, B. (à paraître). Résumer pour apprendre, apprendre pour résumer. In P. Dessus & E. Gentaz (Eds.), *Comprendre les apprentissages* (T. 2). Paris : Dunod.

10 REFERENCES

Brewer, W. F. (1980). Literary theory, rhetoric, and stylistics: Implications for psychology. In R. J. Spiro, B. C. Bruce & W. F. Brewer (Eds.), *Theoretical Issues in Reading Comprehension* (pp. 221-239). Hillsdale: Erlbaum.

Brown, A. L., Day, J. D., & Jones, R. S. (1983). The development of plans for summarizing texts. *Child Development, 54*, 968-979.

Denhière, G. & Lemaire, B. (2004a). Representing children's semantic knowledge from a multisource corpus. *Proceedings of the 14th Annual Meeting of the Society for Text and Discourse*, Chicago.

Denhière G. & Lemaire, B. (2004b). A Computational Model of Children's Semantic Memory, in *Proceedings of the 26th Annual Meeting of the Cognitive Science Society (CogSci'2004)*, 297-302.

Fayol, M. (1985). Analyser et résumer des textes : Une revue des études développementales. *Etudes de Linguistique Appliquée*, *59*, 54-64.

Foltz, P. W. (1996). Latent semantic analysis for text-based research. *Behavior Research Methods, Instruments and Computers, 28*(2), 197-202.

Foltz, P. W., Kintsch, W., & Landauer, T. K. (1998). The measurement of textual coherence with Latent Semantic Analysis. *Discourse Processes, 25*(2-3), 285-307.

Graesser, A., Karnavat, A., Pomeroy, V., & Wiemer-Hastings, K. (2000). Latent Semantic Analysis captures causal, goal-oriented, and taxonomic structures. *Proc. Int. Conf. CogSci 00*, Philadelphie, 2000.

Haxaire, M. (1989). *Réussir le résumé.* Lacoste.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*, 163-182.

Kintsch, W. (1998). *Comprehension, a Paradigm for Cognition*. Cambridge: Cambridge University Press.

Kintsch, W. (2001). Predication. *Cognitive Science, 25*(4), 173-202.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*(5), 363-394.

Landauer, T. K. (2002). On the computational basis of learning and cognition: Arguments from LSA. *The Psychology of Learning and Motivation, 41*, 43-84.

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*, 211-240.

Langston, M., & Trabasso, T. (1999). Modeling causal integration and availability of information during comprehension of narrative texts. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representation during reading* (pp. 29-69). Mahwah: Erlbaum.

Lemaire B., Bianco M. (2003). Contextual Effects on Metaphor Comprehension: Experiment and Simulation. In *Proceedings of the 5th International Conference on Cognitive Modelling (ICCM'2003)*, 153-158.

Lemaire, B., Denhière, G., Bellissens, C., Jhean-Larose, S. (à paraître). A Computational Model for Simulating Text Comprehension. *Behavior Research Methods, Instruments, and Computers.*

Lemaire, B., Mandin, S., Dessus, P. & Denhière, G. (2005). Computational cognitive models of summarization assessment skills. In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society (CogSci' 2005)* (pp. 1266-1271). Mahwah : Erlbaum.

Mandin, S., Dessus, P., Lemaire, B. & Bianco, M. (2005). Un EIAH d'aide à la production de résumés de textes. In P. Tchounikine, M. Joab, & L. Trouche (Eds.), *Actes de la conférence EIAH 2005* (pp. 69-80). Paris : I.N.R.P.

Pfeffer, P. (1989). Les pharmacies des éléphants. In *Vie et mort d'un géant.* Paris : Flammarion.

Trabasso, T. & Sperry, L. L. (1985). Causal relatedness and importance of story events. *Journal of Memory and Language*, *24*, 595-611.

Vidal, N. (1984). *Miguel de la faim.* Paris : Rageot.

Wade-Stein, D., & Kintsch, E. (2004). Summary Street: Interactive Computer Support for Writing. *Cognition and Instruction, 22*(3), 333-362.

Yeh, J.-Y., Ke, H.-R., Yang, W.-P., & Meng, I.-H. (2005). Text summarization using a trainable summarizer and latent semantic analysis. *Information Processing & Management, 41*(1), 75-95.

11 ANNEXES

11.1. Annexe 1 – Le fonctionnement de CI/LSA par l'exemple

Prenons un exemple analysé avec une mémoire sémantique issu du corpus Textenfants (Denhière & Lemaire, 2004b). Soit le texte suivant :

Le jardinier cultive ses roses. Soudain, un chat miaule. Le jardinier lui jette une fleur.

La première proposition est *cultiver(jardinier,roses)*. Les termes qui sont des voisins de *cultiver* mais également proches de *jardinier* et *roses* sont : *légume*, *légumes*, *radis*. Les voisins de *jardinier* sont : *jardin*, *bordure*, *potager*. Les voisins de *roses* sont : *fleurs*, *bouquet*, *violettes*. Les similarités sémantiques entre toutes ces paires de mots sont calculées par LSA. Après la phase d'intégration, la mémoire de travail contient les termes suivants :

cultiver(jardinier,roses) (1.00)
cultiver (.850)
jardinier (.841)
bordure (.767)
jardin (.743)
roses (.740)
fleurs (.726)

Les termes qui ne sont pas reliés au contexte ont été supprimés (comme *radis*).

La deuxième proposition est *miauler(chat)*. Aucun terme n'est récupéré de la mémoire épisodique car la seconde phrase n'est pas reliée sémantiquement à la première. Les termes qui sont des voisins de *miauler* et également associés à *chat* sont : *miaule*, *miaou*, *ronronner*. Les voisins de *chat* sont : *miauler*, *miaule* et *miaou*. Ils sont ajoutés à la mémoire de travail. Les similarités sémantiques entre toutes ces paires de mots sont calculées par LSA. Après la phase d'intégration, la mémoire de travail est la suivante :

miauler(chat) (1.00)
cultiver(jardinier,roses) (.967)
chat (.795)
miaule (.776)

Tous les termes associés à la première phrase (bordure, jardin, ...) ont disparus parce que la seconde phrase n'est pas reliée à la première. Cependant, la proposition entière est restée.

La troisième proposition est jeter(homme, fleur)[1]. *Fleurs, bouquet*, *roses* et *violettes* sont récupérés en mémoire épisodique et placés en mémoire de travail parce qu'ils sont proches de l'argument *fleurs*. Les termes qui sont des voisins de *jeter* et aussi proches de *homme* et *fleur* sont : *ordonner*, *envoyait*, *valet*. *Homme* est trop fréquent pour produire de bons voisins ce qui fait que le programme ne le retient pas. Les voisins de *fleur* sont : *pétales*, *pollen*, *tulipe*. Les similarités sémantiques entre toutes ces paires de mots sont calculées par LSA. Après la phase d'intégration, la mémoire de travail est la suivante :

fleur (1.00)
fleurs (.979)
pétales (.978)
cultiver(jardin,roses) (.975)
roses (.974)
violettes (.932)
bouquet (.929)

[1] Cette proposition est en réalité jeter(homme, fleur, chat), mais le système n'ayant aucun moyen de résoudre l'anaphore, il ignore le troisième argument.

jeter(homme,fleur) (.917)
tulipe (.843)
pollen (.832)

Cet exemple montre que (1) le modèle est capable de fournir des associés pertinents à la proposition courante, grâce à l'espace sémantique construit par LSA et (2) la mémoire épisodique permet la récupération de concepts (voire de propositions) qui apparaissaient antérieurement dans le texte mais ont disparus de la mémoire de travail en raison d'un changement de thème local dans le texte.

11.2. Annexe 2 – Protocole de l'étude 2 et 3

EXERCICE 1

Durée : 30 minutes

Exercice : rédigez, sur la page suivante, un résumé qui raconte la même histoire que le texte ci-dessous en un nombre de mots plus petit. La longueur de votre résumé ne doit pas dépasser les lignes pré imprimées sur la copie.

Vous ne pouvez utiliser pour cet exercice que les copies qui vous sont fournies. Vous n'avez donc droit ni à des feuilles de brouillon ni à d'autres documents (dictionnaire, cours, etc.). En revanche vous pouvez barrer sur votre copie et annoter ou souligner le texte ci-dessous.

Le village dormait, la maison était noire et paisible.

Miguel s'efforçait d'être calme et de ne pas écouter les douleurs qui lui fouettaient le ventre. Il était tendu à force d'essayer de marcher sans bruit.

Arrivé près de l'arbre, il posa son argent plus une médaille bénite pour le cas où le prix des poulets excéderait la somme qu'il pouvait y mettre.

Puis, avec son couteau, il coupa doucement la ficelle qui tenait au tronc, s'avança, les deux mains prêtes à saisir le cou des volailles en même temps, afin de les étrangler avant qu'elles ne crient, et frappa.

Hélas ! Il était trop épuisé pour faire du bon travail. Les poules crièrent et battirent des ailes comme des forcenées. Un homme sortit sur le pas de la porte avec son fusil.

- Ne tirez pas, capitaine ! s'écria Miguel. Ce n'est que moi… je venais prendre les poulets que j'ai achetés.

- Que tu as achetés, voleur ! dit l'homme, sur un ton de colère.

Et il leva son fusil.

- Pose ces poulets.

- Je les pose, mais je les ai achetés, dit Miguel, désespéré, sans lâcher son bien. J'ai mis par terre de l'argent et une médaille bénite.

Le fait était si étrange et l'accent de l'enfant si sincère que l'homme s'approcha. Dix-sept pièces d'or et une médaille luisaient doucement dans le clair de lune.

Le paysan s'accroupit pour les regarder de près, dans une sorte d'étonnement amusé.

- Il n'y a là même pas assez pour un poulet, dit-il d'une voix rude destinée à effrayer l'enfant.

- Oui, mais la médaille n'a pas de prix, elle a été bénite par le cardinal de Recife, ainsi me l'a assuré Federico qui me l'a donnée. Il l'avait achetée pour son fils à Juan Dias, le porteur d'eau…, son fils est mort, malgré le lait de la chèvre. Il faut dire qu'après un si long voyage, la chèvre ne donnait pas beaucoup.

Miguel parlait avec difficulté, son ventre le torturait au point qu'il avait un brouillard devant les yeux. A rester ainsi debout sans pouvoir s'appuyer, il perdit ce qui lui restait de force et s'évanouit.

Soulignez directement sur cette feuille entre 3 et 5 phrases qui vous paraissent les plus importantes dans le texte ci-dessous.

Vous ne pouvez utiliser pour cet exercice que les copies qui vous sont fournies. Vous n'avez donc droit ni à des feuilles de brouillon ni à d'autres documents (dictionnaire, cours, etc.).

Comme tous les herbivores, ou plus précisément comme tous les animaux végétariens, qu'ils se nourrissent d'herbes, de feuilles ou de fruits, les éléphants se rendent régulièrement en des endroits bien précis de la brousse pour y absorber de la terre. Ils retiennent le sodium qui est bon pour leur organisme et éliminent le potassium que leur alimentation leur fournit en excédent dangereux. Ces lieux sont appelés "pharmacies des éléphants" ou "salines" en Afrique francophone. Certains sont temporaires, mais d'autres sont fréquentés à l'évidence depuis des millénaires. De leurs défenses, de leurs sabots et de leurs cornes, les éléphants et les autres animaux ont creusé de véritables cavernes et mêmes des couloirs les liant les unes aux autres ou se terminant en cul-de-sac. Les célèbres salines de Gazao, en Centrafrique, forment sur des kilomètres de véritables falaises, aussi déchiquetées que celles d'Etretat, mais qui sont les résultats du travail des animaux et surtout des éléphants dont les défenses ont laissé leur marque sur leurs parois d'ocre.

Dans les régions où il reste des éléphants et où ils bénéficient d'une tranquillité relative, c'est en général au milieu de la matinée, après s'être nourris et avant d'aller à l'eau, que les éléphants s'adonnent à cette cure de terre que les scientifiques désignent par le terme de "géophagie". Les solitaires restent moins longtemps sur les salines que les troupeaux qui y passent facilement 2 ou 3 heures. Ils y effectuent un travail épuisant, stupéfiant, qui explique le relief lunaire de certaines d'entre elles. Utilisant aussi bien leurs pieds de devant que leur trompe et surtout leurs défenses, ils creusent, labourent et broient la terre en une fine poussière qui est aspirée, toujours avec la trompe, et projetée dans la bouche.

Les éléphants fréquentent leurs "pharmacies" aussi bien en saison sèche qu'au moment des pluies. Au lieu de manger de la poussière, ils se gavent alors de boue, avec autant d'avidité et de gloutonnerie. La saline, comme le point d'eau, est habituellement un lieu de repos, pour les herbivores au moins. La plupart d'entre eux s'en vont pourtant quand arrive un groupe d'éléphants. Il est probable que ce n'est pas par crainte, mais à cause du manque de discrétion et du sans-gêne de ces encombrants et remuants visiteurs. En revanche, j'ai vu un éléphant reculer dans une saline devant un rhinocéros, tout en trahissant son inquiétude par des claquements d'oreilles sonores et des tortillements de trompe du plus haut comique. Le rhinocéros ne montrait pourtant pas la moindre agressivité, mais, avec son air obtus, manœuvrait toujours pour manger la terre sous le nez de l'éléphant qui finit par lui céder la place.

A l'exception des lions dont ils ne peuvent supporter la présence et qu'ils chargent avec des barrissements aigus, jusqu'à ce qu'ils dégagent les lieux en rampant et en se fouettant rageusement les flancs avec la queue, les éléphants ont avec la plupart des animaux des rapports de bon voisinage, c'est-à-dire d'indifférence totale.

11.3. Annexe 3 – Résultats Bruts de l'étude 2

Résultats du traitement du texte narratif « *Miguel* ». Valeurs moyennes en pourcentages des soulignements, par phrase.

	Moyennes de % de sélection des élèves									Valeurs LSA/Corpus	
Phrase	4e sp.	4e	3e	CAP	2e BEP	2e	1re	Tothtech	Total	Enfants	Adulte
1	40,0	8,0	20,5	0,0	61,5	18,2	10,5	15,9	19,6	0,241	0,219
2	20,0	28,0	41,0	42,9	30,8	27,3	52,6	34,8	34,4	0,194	0,163
3	20,0	8,0	10,3	14,3	15,4	1,8	0,0	5,1	6,7	0,391	0,344
4	60,0	60,0	48,7	42,9	38,5	52,7	52,6	52,9	51,5	0,450	0,448
5	40,0	20,0	23,1	14,3	38,5	30,9	36,8	27,5	28,2	0,531	0,400
6	20,0	32,0	28,2	14,3	30,8	10,9	15,8	20,3	20,9	0,308	0,259
7	0,0	24,0	35,9	0,0	30,8	34,5	26,3	31,9	29,4	0,026	0,119
8	0,0	48,0	41,0	14,3	69,2	60,0	47,4	50,7	49,1	0,462	0,331
9	20,0	12,0	15,4	14,3	7,7	5,5	10,5	10,1	10,4	0,208	0,173
10	0,0	24,0	25,6	0,0	0,0	25,5	21,1	24,6	20,9	0,324	0,246
11	0,0	20,0	12,8	42,9	23,1	23,6	15,8	18,8	19,6	0,363	0,301
12	20,0	4,0	2,6	0,0	0,0	1,8	0,0	2,2	2,5	0,424	0,287
13	0,0	4,0	0,0	0,0	0,0	5,5	5,3	3,6	3,1	0,135	0,066
14	20,0	32,0	20,5	14,3	15,4	27,3	21,1	25,4	23,9	0,425	0,346
15	0,0	20,0	28,2	0,0	30,8	34,5	15,8	27,5	25,8	0,320	0,321
16	40,0	16,0	17,9	14,3	15,4	18,2	36,8	20,3	20,2	0,462	0,386
17	20,0	12,0	20,5	14,3	7,7	5,5	5,3	10,9	11,0	0,201	0,150
18	40,0	4,0	12,8	0,0	0,0	1,8	5,3	5,8	6,1	0,427	0,258
19	20,0	28,0	25,6	42,9	15,4	32,7	31,6	29,7	28,8	0,326	0,323
20	0,0	28,0	41,0	42,9	30,8	36,4	36,8	36,2	35,0	0,163	0,241
21	20,0	0,0	10,3	28,6	0,0	1,8	0,0	3,6	4,9	0,486	0,335
22	0,0	0,0	5,1	42,9	0,0	0,0	5,3	2,2	3,7	0,267	0,221
23	0,0	20,0	17,9	28,6	0,0	21,8	15,8	19,6	17,8	0,434	0,317
24	20,0	48,0	51,3	42,9	38,5	56,4	63,2	54,3	51,5	0,549	0,283

Légende : Total htech : sans les classes « techniques »

Résultats détaillés à propos du texte « Les pharmacies des éléphants ».

Eléphants	Moyennes de % de sélection des élèves									Valeurs LSA	
Phrase	4e sp.	4e	3e	CAP	2e BEP	2e	1re	Tothtech	Total	enfants	Adulte
1	0,0	52,0	59,0	50,0	67,7	81,5	86,4	70,7	68,3	0,613	0,532
2	33,3	36,0	20,5	33,3	48,4	51,9	31,8	37,1	38,9	0,254	0,280
3	66,7	36,0	48,7	0,0	54,8	37,0	45,5	41,4	42,8	0,422	0,309
4	33,3	4,0	7,7	16,7	3,2	1,9	0,0	3,6	4,4	0,505	0,387
5	33,3	44,0	30,8	16,7	22,6	38,9	31,8	36,4	33,3	0,586	0,451
6	0,0	16,0	30,8	33,3	12,9	9,3	13,6	17,1	16,7	0,650	0,512
7	33,3	36,0	61,5	50,0	58,1	59,3	59,1	55,7	55,6	0,585	0,528
8	100,0	32,0	7,7	50,0	22,6	1,9	0,0	8,6	13,9	0,323	0,234
9	0,0	4,0	7,7	33,3	22,6	7,4	0,0	5,7	9,4	0,444	0,295
10	0,0	52,0	30,8	33,3	22,6	37,0	45,5	39,3	35,6	0,482	0,437
11	100,0	16,0	41,0	16,7	51,6	37,0	50,0	36,4	39,4	0,338	0,243
12	33,3	28,0	17,9	0,0	12,9	11,1	13,6	16,4	15,6	0,337	0,260
13	33,3	20,0	25,6	0,0	22,6	11,1	27,3	19,3	19,4	0,308	0,324
14	33,3	8,0	2,6	0,0	0,0	5,6	0,0	4,3	3,9	0,364	0,193
15	0,0	0,0	2,6	33,3	3,2	1,9	0,0	1,4	2,8	0,258	0,191
16	0,0	0,0	10,3	16,7	12,9	7,4	4,5	6,4	7,8	0,292	0,252
17	0,0	4,0	12,8	16,7	3,2	0,0	0,0	4,3	4,4	0,464	0,403
18	33,3	36,0	59,0	50,0	41,9	64,8	50,0	55,7	52,8	0,668	0,511

Légende : Total htech : sans les classes « techniques »