

UN LOGICIEL EN LIGNE POUR AIDER À LA COMPRÉHENSION DES TEXTES DOCUMENTAIRES

De meilleures notes... qu'en bas de page

Brigitte MARIN, Patrick AVEL, Jacques CRINON et Denis LEGROS
Équipe CODITEXTE - IUFM de Créteil (*)

La difficulté des élèves à comprendre un texte documentaire relève de multiples facteurs, en particulier l'ignorance du vocabulaire et la connaissance insuffisante du domaine évoqué. Comment les aider ? Doit-on recourir aux notes de bas de page ? Quel est leur statut ? Des expérimentations sur la compréhension des textes scientifiques, menées dans des classes de CM2 et sixième, a conduit l'équipe Coditexte à réaliser un logiciel d'aide à la lecture. Ouvert à tout type de texte, il repose à la fois sur l'efficacité de l'hypertexte et la nature des aides proposées.

L'équipe de recherche *Coditexte* (Cognition et Didactique du texte) étudie les effets des hypertextes sur la compréhension des textes documentaires scientifiques par des élèves de la fin de l'école élémentaire et du cycle d'observation du collège. Les résultats d'expérimentations menées dans treize classes de la Seine-Saint-Denis ont donné lieu à la conception du logiciel *Aide à la lecture* qui propose aux enseignants des textes à dominante scientifique concernant les programmes de CM2 et de sixième. Ces écrits documentaires, présentés sous plusieurs versions, offrent aux élèves des conditions d'accès aux textes diversifiées et destinées à en faciliter la compréhension. La création du logiciel résulte de recherches exploratoires réalisées à partir d'un protocole expérimental visant à tester et à croiser différents types d'aide à la compréhension de textes documentaires. La difficulté des élèves à comprendre les textes scientifiques relève de facteurs distincts et interdépendants. Le plus facilement identifiable consiste en une connaissance insuffisante du domaine évoqué par le texte ⁽¹⁾ : ignorances liées au vocabulaire et aux objets du monde référés par ce vocabulaire ; manque des connaissances nécessaires

* Cette recherche et le développement du logiciel ont été soutenus par l'IUFM de Créteil.

pour relier entre elles de manière cohérente et logique les informations disponibles dans ces textes afin de combler les « trous sémantiques », c'est-à-dire rétablir, par inférence, les informations implicites, qui ne sont pas directement données à lire. Compenser de telles lacunes ne peut se faire sans l'apport des informations manquantes ni sans l'entraînement à la mobilisation pertinente de ces connaissances, si on veut amener les élèves à construire le « modèle de situation »⁽²⁾. Ce dernier est une représentation mentale constituée à partir des informations présentes dans le texte et de celles qui appartiennent aux connaissances du sujet activées lors de la lecture. Ainsi le lecteur, grâce aux processus d'inférence, peut construire la cohérence de la signification du texte.

DES NOTES AU STATUT DIFFÉRENT

Notre recherche s'est donné pour objectif de mesurer l'efficacité de différents types d'aides à la compréhension. Nous avons supposé qu'en fournissant aux élèves les informations absentes du texte, et pourtant indispensables à son intelligibilité, nous les aiderions à mieux le comprendre. Nous avons fait l'hypothèse que la nature des informations et la façon dont elles sont accessibles aux lecteurs pouvaient les aider à activer les connaissances renvoyant au « modèle de situation » sous-jacent au texte et à s'en fabriquer une image mentale cohérente⁽³⁾. De surcroît, nous avons considéré qu'entraîner les élèves lecteurs de textes scientifiques à relier causalement les informations appréhendées jusqu'à ponctuellement devait leur permettre d'accéder à une compréhension plus fine.

Quelles aides à la compréhension de textes documentaires scientifiques fournir et comment les rendre plus efficaces ? Si la première question s'intéresse au dispositif d'appareillage des textes, la seconde concerne directement le support des apprentissages. Pour mesurer et comparer l'impact de différents régimes d'aide, nous avons confronté cent vingt élèves, provenant de classes de cours moyen deuxième année, à la lecture d'un article de journal traitant des dérèglements du climat et des catastrophes naturelles qui en découlent, sur papier ou sur écran d'ordinateur, accompagné dans chacune des deux situations de deux types de notes au statut différent. Nous avons soumis les participants à quatre conditions expérimentales différentes, en faisant varier plusieurs critères et en les



**Parmi les systèmes d'aide possibles,
« l'appareillage » de notes
constitue un ajout d'informations
précieux dont il convient
d'explorer les ressources.**



croisant. Les élèves ont été répartis en quatre groupes, chacun correspondant à une situation particulière afin de comparer leurs performances en fonction des conditions et des aides dont ils bénéficiaient.

Dans le groupe G1, les élèves lisent le texte sur document papier et disposent de six notes de vocabulaire explicitant des informations appartenant à la « base de texte » (contenu sémantique du texte). Ces notes figurent en bas de page.

Dans le groupe G2, les élèves lisent le texte sur écran d'ordinateur et disposent de six notes identiques à celles du groupe G1. Ces notes s'affichent successivement dans la partie inférieure gauche de l'écran lorsque le lecteur clique sur le mot ou le groupe de mots dont il souhaite voir éclairci le sens.

Dans le groupe G3, les élèves lisent le texte sur document papier et disposent de six notes permettant d'activer les connaissances du monde évoquées par le texte (le « modèle de situation »), mais absentes de celui-ci. Ces notes facilitent les inférences et donc la construction d'une représentation cohérente ; elles mettent en évidence les liens logiques qui relient les informations fournies aux éléments du texte concernés.

Dans le groupe G4, les élèves lisent le texte sur écran d'ordinateur et disposent de six notes identiques à celles du groupe G3, placées dans la partie inférieure gauche de l'écran. Ils les activent par lien hypertexte.

Après une tâche distractive, destinée à pallier les effets artificiels de la mémorisation et de la restitution immédiates des informations lues, les élèves ont fait de ce texte documentaire scientifique un rappel écrit, c'est-à-dire qu'on leur a demandé de restituer tout ce

(1) Graesser A.C., Olde B., Pomeroy V., Whitten S., Lu S. & Craig S., *Inferences and Questions in Science Text Comprehension*, In J. Otero, J.A. León & A.C. Graesser (Eds.), 2002 ; *The Psychology of Science Text Comprehension* (pp. 417-436), Mahwah, NJ, Lawrence Erlbaum Associates, 2002.

(2) Kintsch W., *Comprehension, a Paradigm for Cognition*, New York, NY, Cambridge University Press, 1998.

Van Dijk T.A. & Kintsch W., *Strategies of discourse comprehension*, New York, Academic Press, 1983.

(3) Van den Broek P., Virtue S., Everson M.G., Tzeng Y. & Sung Y., *Comprehension and Memory of science texts : inferential processes and construction of mental representation*, In J. Otero, J.A. León & A.C. Graesser (Eds.), 2002 ; *The Psychology of Science Text Comprehension* (pp. 131-154), Mahwah, NJ, Lawrence Erlbaum Associates.

Type d'aide : Explications supplémentaires

Climat

Le dérèglement du climat entraîne une recrudescence de catastrophes naturelles qui se manifestent sous la forme de coups de vent, tempêtes, cyclones, tornades, pluies diluviennes, inondations, éboulements, coulées de boue et sécheresses qui n'épargnent pas les centres urbains. Elles toucheront en particulier les régions équatoriales très peuplées, pauvres, et donc moins aptes à réagir efficacement. L'Europe et la France ne sont pas à l'abri des effets du dérèglement mondial du climat. Les tempêtes de décembre 1999 et les inondations de l'hiver 2000-2001 sont là pour le rappeler. En France, le risque d'inondations touche aujourd'hui plus d'une commune sur quatre et provoque en moyenne plus d'une quinzaine de morts par an.

L'augmentation des sécheresses, selon la WWF, entraîne la perte chaque année de 10 millions d'hectares, en particulier dans les régions tropicales qui deviennent plus chaudes et plus sèches, notamment la zone aride qui s'étend de l'Afrique de l'Ouest à l'Indonésie. La multiplication des sécheresses gagne aujourd'hui l'Europe. Les zones arides du pourtour méditerranéen devraient s'étendre sous l'effet de la combinaison température et évaporation l'été, précipitations massives et érosion l'hiver. En Grèce, le débit du plus grand fleuve du pays, l'Acheloos, a baissé de 40% en quatre ans. Mais c'est en Afrique, continent le plus vulnérable, où la désertification ne fait que progresser, que les conséquences pour les populations se feront le plus sentir. Selon l'ONU, 250 millions de personnes ont souffert de famine et de malnutrition en 2000. Et ce chiffre pourrait grimper à 900 millions dans quelques décennies.

| | |
|---|---|
| <p style="text-align: center;">"moins aptes à réagir efficacement"</p> <p>Les régions équatoriales sont très peuplées : les catastrophes naturelles font donc beaucoup de victimes. Pauvres, elles ont peu de moyens pour se protéger des catastrophes.</p> | <p style="text-align: center;">Bloc Notes de "prof" (24/6/2005)</p> <div style="border: 1px solid black; height: 40px; width: 100%;"></div> |
|---|---|

Les élèves du groupe G4, lisant la page documentaire sur écran d'ordinateur, peuvent accéder à des notes renvoyant au « modèle de situation » du texte. Celles-ci consistent, comme dans celles du groupe G3 travaillant également à partir de notes d'inférence mais sur papier, à établir les liens de causalité entre les éléments qui les régissent. Ici, la note hypertextuelle affichée par l'élève est la deuxième dans l'ordre chronologique du texte. Elle permet d'apporter un éclairage à

l'expression « moins aptes à réagir efficacement », appliquée aux régions équatoriales, très peuplées et pauvres. Toutefois, dans la situation « sur écran », la part d'activité de l'élève est déterminante dans le processus de compréhension. En effet, l'activation des notes est liée à son initiative personnelle et à sa volonté de cliquer sur le lien hypertexte pour obtenir une information qu'il juge nécessaire. Ainsi, par exemple, grâce à la note 2, l'expression « très peuplées » est rattachée logi-

quement au nombre important de victimes des catastrophes naturelles, « pauvres » est référé à l'absence de moyens qui permettraient aux populations de se protéger de ces incidents climatiques. Cela n'allait pas de soi à la lecture du seul texte, malgré la présence du connecteur « donc », qui engageait un processus d'inférence difficile à réaliser pour la plupart des élèves.

qu'ils ont retenu du texte. Ensuite, ils ont dû répondre à cinq questions ouvertes permettant de mesurer leur capacité à identifier les relations logiques structurant le texte.

AIDER À TRAITER L'IMPLICITE

Examinons de plus près les variables manipulées dans ces groupes : la présentation (sur écran ou sur papier) et le type de notes (vocabulaire ou inférence). Parmi les systèmes d'aide possibles, « l'appareillage » de notes constitue un ajout d'informations précieux dont il convient d'explorer les ressources. Bon nombre de manuels recourent à ces amplifications explica-

(4) Marin B., Avel P., Crinon J., Legros D., *Aides à la compréhension de textes scientifiques par des élèves de onze ans*, Communication au colloque de l'AIRDF : Le français, discipline singulière ou transversale, Québec, 27-28 août 2004.

tives et informatives censées combler les lacunes rédhitoires pour la construction de la signification. Cependant, il convient de distinguer deux modalités d'aide par les notes au statut bien différent. En effet, il est nécessaire de considérer d'une part les notes de vocabulaire qui renvoient à la « base de texte » et aux informations explicites et, d'autre part, les notes « inférentielles » qui permettent d'accéder au « modèle de situation » du texte, à traiter l'implicite et à produire des informations nouvelles à partir de celles disponibles dans la « base de texte ».

Une première recherche papier/crayon⁽⁴⁾ nous avait permis de comparer l'effet du type de notes sur la compréhension du même texte. De façon très fortement significative, nous

avons pu alors observer que les élèves disposant des notes d'inférence produisaient des textes plus longs et plus riches en informations reliées par des liens de causalité que ceux ayant eu accès à des notes de vocabulaire. Leur compréhension du texte, jugée à l'aune de la qualité de la restitution, s'avérait donc plus cohérente et plus approfondie.

Il convenait ensuite de comparer l'impact des aides sur document papier, qu'il s'agisse des notes de vocabulaire (groupe G1) ou des notes d'inférence (groupe G3) à des aides sur support informatique. C'est le but de la recherche présentée ici. Notre hypothèse de travail attribuait au groupe bénéficiant de l'effet croisé des liens hypertextes et des notes d'inférence (groupe G4) les meilleurs scores de compréhension de l'ensemble des situations expérimentales proposées. Nous nous attendions à des performances moindres de la part des participants au groupe G2 pouvant activer des liens hypertextes renvoyant à la seule « base de texte ».

MIEUX COMPRENDRE LA CAUSALITÉ

L'analyse qualitative des réponses à la question correspondant à la note 5 relative à l'érosion (« Quand la terre est emportée par les eaux, rien ne pousse. Les paysans ne peuvent plus rien cultiver. ») fournit des éléments d'appréciation sur l'effet croisé des aides proposées. Nous avons sélectionné les réponses d'élèves présentant des niveaux de compréhension en lecture analogues, afin d'établir une comparaison, toutes choses égales par ailleurs.

Question : Pour quelles raisons une terre ne peut plus être cultivée ?

Réponses d'élèves du groupe G3
(notes d'inférence sur support papier)

Yanis : « Car il pleut beaucoup et car il fait chaud ; donc elle est soit humide, soit sèche. »

Anastassiya : « Car si il fait très chaud les légumes vont brûler et si il fait très froid ils vont mourir de froid. »

Réponses d'élèves du groupe G4
(notes d'inférence sur support informatique)

Coumba : « La terre ne peut pas être cultivée parce que l'eau peut emporter la terre en cas de dérèglement. »

Brahim : « La raison est : parce que il ne pleut pas et il y a moins de terre pour cultiver. »

LES DEUX TYPES DE NOTES

Notes renvoyant au « modèle de situation » du texte Notes proposées aux groupes G3 et G4

(1) Le dérèglement du climat

Le climat se modifie. Quand ces modifications sont trop importantes, il y a dérèglement du climat. Cela entraîne des catastrophes naturelles.

(2) moins aptes à réagir efficacement

Les régions équatoriales sont très peuplées : les catastrophes naturelles font donc beaucoup de victimes. Pauvres, elles ont peu de moyens pour se protéger des catastrophes.

(3) la perte chaque année de 10 millions d'hectares

Les sécheresses ont des conséquences sur l'agriculture, car il est impossible de faire pousser quelque chose sur des terres qui ne reçoivent pas de pluie.

(4) précipitations massives

Les fortes précipitations ont aussi des conséquences graves sur l'agriculture, car rien ne pousse quand le sol a été emporté par les eaux.

(5) érosion

Quand la terre est emportée par les eaux, rien ne pousse. Les paysans ne peuvent plus rien cultiver.

(6) les conséquences pour les populations

L'Afrique a de moins en moins de terres cultivables et c'est un continent pauvre. Cela entraîne le manque de nourriture.

Notes de vocabulaire concernant la « base de texte » Notes proposées aux groupes G1 et G2

(1) dérèglement du climat

Le climat regroupe les caractéristiques d'une région, comme la température, le vent, la pluie. Quand le climat se dérègle, les caractéristiques changent.

(2) recrudescence

Les catastrophes naturelles sont de plus en plus graves et nombreuses. On dit qu'elles sont en recrudescence ou en augmentation.

(3) pluie diluvienne

Une pluie diluvienne est une pluie très abondante et torrentielle. En peu de temps, il tombe beaucoup d'eau.

(4) zone aride

Une zone aride est une partie du monde où il y a peu de végétaux. En général une zone aride est très sèche.

(5) érosion

L'érosion est l'usure de tout ce qu'il y a à la surface de la Terre. Tout ce qui est à la surface de la Terre s'abîme, comme les montagnes, les falaises, les bâtiments.

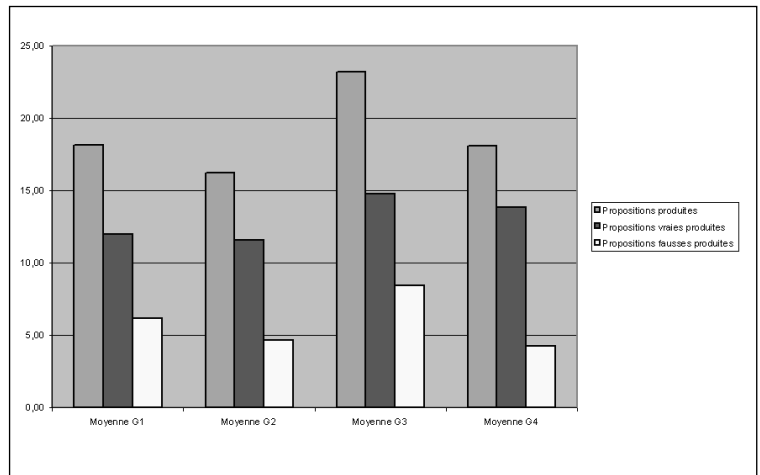
(6) vulnérable

Une personne vulnérable est une personne fragile et exposée au danger. De même, un continent vulnérable est exposé au danger.

Nous observons, avec des notes identiques mot à mot, la supériorité de compréhension du texte chez les élèves du groupe G4, alors que seul diffèrait le support de ces notes. Les quatre réponses utilisent des liens logiques abondants qui témoignent d'une prise en compte de la causalité induite par les notes. Toutefois, un élément essentiel de la réponse attendue (l'érosion), explicité par la note 5, est absent des réponses des élèves du groupe travaillant sur papier. En revanche, le phénomène de l'érosion fait l'objet d'une mention implicite chez les deux élèves issus du groupe utilisant les notes d'inférence dans un environnement logiciel. Certes, le terme « érosion » ne figure pas, mais cette absence terminologique n'est pas le gage d'un défaut de compréhension dans la mesure où le processus de l'érosion se trouve défini. Au contraire, la reformulation témoigne d'un traitement cognitif de l'information. Être capable de reformuler une information en en modifiant la forme linguistique révèle une appropriation relevant d'une bonne compréhension.

EFFICACITÉ DE L'HYPERTEXTE

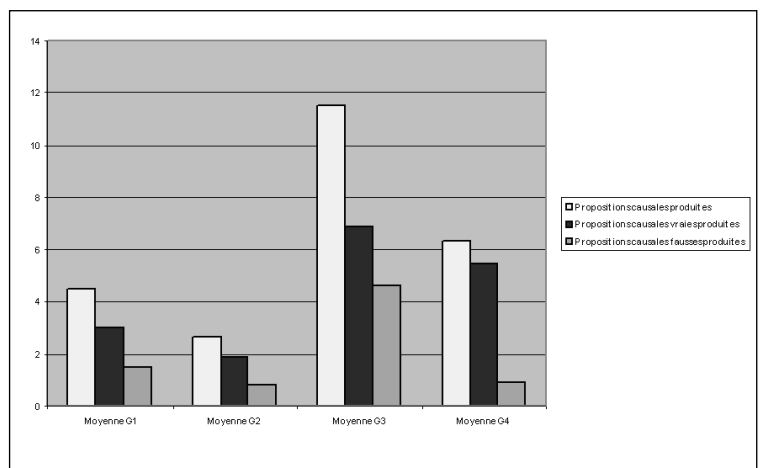
Afin de quantifier les informations produites dans les « rappels » des élèves et de rendre compte de leur pertinence, nous les avons codées et comptabilisées en unités minimales, les propositions sémantiques. Les premiers résultats issus de ces analyses ne répondent pas à notre hypothèse⁽⁵⁾. En effet, si les « rappels » sont riches en relations causales dans le groupe G4, ils s'avèrent aussi plus courts lorsque l'on considère l'ensemble des textes produits. Comment interpréter cette tendance à des écrits laconiques dans le groupe qui bénéficiait des conditions les plus favorables d'accès à la compréhension ? À une moindre mémorisation des informations ? À un souci du détail moins prégnant ? En fait, il n'en est rien. Le déficit en informations remarqué est directement imputable à la sélection avisée des propositions sémantiques formulées et à la qualité des informations produites lors de ce « rappel ». En effet, l'examen du « rappel » montre le poids des informations erronées dans les productions plus longues des participants au groupe G3, ayant lu le texte sur papier. L'observation des informations relevant de rapports de causalité vient appuyer et confirmer cette explication. En effet, les informations causales produites par le groupe G4 (liens hypertextes et notes d'inférence) montrent la pertinence de leurs rappels. Ces écrits se distinguent par une



Les informations produites par les élèves dans leurs « rappels » sont comptabilisées en propositions sémantiques.

Le graphique ci-dessus montre que les élèves des groupes G3 et G4 (qui bénéficient de notes d'inférence) produisent davantage de propositions que ceux des groupes G1 et G2 (qui bénéficient de notes de vocabulaire). Les élèves du groupe G4 commettent moins d'erreurs que leurs camarades des autres groupes.

Le graphique ci-dessous montre que les élèves des groupes G3 et G4 produisent nettement plus de propositions causales que leurs camarades des groupes G1 et G2. Le petit nombre de propositions causales fausses produites par les élèves du groupe G4 traduit l'efficacité de l'hypertexte comme appareillage du texte.



concision qui privilégie les informations causalement reliées, et font état de peu d'erreurs dans la mise en relation des processus et événements mentionnés.

Parmi les pistes interprétatives, il nous paraît vraisemblable d'expliquer l'effet catalyseur de l'aide informatisée sur la sélection pertinente des informations par le processus cognitif lié à l'activation de l'hypertexte⁽⁶⁾. En effet, chaque note est alors mobilisée à l'initiative de l'élève et dans une temporalité qui rend son contenu congruent par rapport à une situation de résolution de problème.

De surcroît, le plus faible taux d'informations causales fausses produites dans les rappels de ce groupe, comparativement à ce qui est observé dans le groupe disposant des mêmes notes accumulées en bas de page (telles qu'elles ont été présentées dans les encadrés), conduit à une interprétation complémentaire. La somme des notes souvent lues de manière enchaînée, les unes à la suite des autres permet certes une familiarisation avec les processus d'inférence, et conduit l'élève à établir des relations de causalité entre les différentes informations qu'il mentionne. Cependant, les lecteurs semblent alors exercer un moindre contrôle sur la cohérence des informations reliées.

UN LOGICIEL OUVERT

Le logiciel *Aide à la lecture*, qui nous a permis de recueillir les données exposées plus haut, fournit à tout enseignant la possibilité d'entrer dans la base de données ses propres textes, quels que soient leur type et leur discipline d'appartenance. Par exemple, des textes documentaires historiques, des textes de littérature de jeunesse, y trouvent tout autant leur place que les documentaires scientifiques.

L'enseignant a également la possibilité de créer lui-même les aides dont il souhaite faire bénéficier ses élèves, en fonction de quatre catégories, dont celles envisagées lors de cette étude : notes de vocabulaire et notes d'inférence. Les autres possibilités consistent en un résumé et une version du texte simplifiée du point de vue de la syntaxe et du vocabulaire.

L'enseignant peut ainsi, via une interface Web, gérer la base de textes (ajout / suppression de textes et de notes) et celle des utilisateurs (entrer des classes et des élèves, visualiser la classe entière ou l'ensemble des classes d'une école et les « logs » de chaque utilisateur). L'écran *Élève* permet d'afficher un texte, à partir d'une liste cliquable, d'accéder à la consultation de différents types d'aide et à la prise de notes personnelles consultables à tout moment.

Chacun des types d'aide proposés par le logiciel a son utilité. Leur complémentarité, organisée et conjuguée en fonction des objectifs pédagogiques des enseignants, multiplie les voies d'accès à la compréhension des textes. Cette application s'utilise en ligne à partir d'un navigateur, en se connectant à <http://coditexte.creteil.iufm.fr>, et en choisissant ensuite *Travail en ligne* dans le menu. Développé en php et utilisant une base de données MySQL, *Aide à la lecture* est placé sous licence GPL. Il est ainsi libre de droits,

son code source peut être distribué, utilisé ou modifié par ses utilisateurs.

Les ressources liées à l'utilisation du logiciel *Aide à la lecture* sont multiples et génèrent des possibilités d'exploitation pédagogique adaptées à la diversité des publics scolaires.

D'une part, elles offrent aux enseignants susceptibles de s'emparer de l'outil un moyen efficace de rendre les élèves plus autonomes en leur laissant, le cas échéant, l'initiative de déterminer eux-mêmes le type d'aide dont ils ont besoin. De surcroît, en fonction de leurs compétences dans le domaine évoqué par le texte et de leur compréhension, les élèves peuvent utiliser le bloc-notes pour y noter les informations qu'ils jugent importantes. Ils créent ainsi une « mémoire externe » qui les aidera à hiérarchiser et à synthétiser le contenu du texte lu.

La variété des fonctionnalités offertes par le logiciel peut trouver des applications didactiques dans la gestion de l'hétérogénéité des classes. Elle est de nature à pallier les lacunes particulières de certains élèves, si les enseignants proposent une utilisation guidée et conçoivent des notes qui mettent en jeu des savoirs prérequis non encore assimilés dans certains cas.

D'autre part, la base de données textuelles que l'on peut constituer à partir du logiciel *Aide à la lecture* présente un support qui invite au travail coopératif entre enseignants. En effet, la mutualisation et l'accroissement des ressources pédagogiques sont facilités par la possibilité de partager la somme des ressources créées avec les autres enseignants utilisateurs.

À partir de ce corpus de textes, littéraires, documentaires scientifiques, historiques ou appartenant à un champ disciplinaire différent, il appartient à chaque enseignant d'inventer les scénarios pédagogiques dans lesquels s'inscrira la lecture de ces textes. ■

(5) Des résultats statistiques plus complets des données fournies par cette recherche sont disponibles dans Crinon, J., Avel P., Legros, D., Marin B. (2005), *Vers des aides logicielles à la compréhension de textes documentaires scientifiques pour les élèves de cycle 3 et de collège*, Actes de la conférence EIAH, Montpellier, 25-27 mai 2005, INRP, 2005, pp. 21-32.

(6) À propos des effets de l'hypertexte sur la lecture, voir Rouet J.-F. & Levonen J.-J., « Studying and learning with hypertexts : empirical studies and their implications », In J.-F. Rouet, J.J. Levonen, A.P. Dillon, R. J. Spiro (Eds.), *Hypertexts and cognition*, Mahwah, NJ, Lawrence Erlbaum Associates, 1996, pp. 9-23.