

L'application du diagnostic cognitif au domaine de la rédaction est examinée en rapport avec la spécification de la structure interne des réponses et la définition de la tâche. Un modèle stratifié des niveaux de représentation du discours peut être utilisé pour identifier la structure interne du texte. Des méthodes de traçage, développées dans les recherches sur la résolution de problèmes, peuvent être utilisées afin d'obtenir des données sur la définition de la tâche par le rédacteur. A cause du caractère mal défini de la tâche de rédaction, des méthodes d'analyse du discours doivent être appliquées pour analyser les traces de façon à définir une tâche de rédaction particulière. Ces caractéristiques suggèrent que l'application du diagnostic cognitif à la rédaction aura un impact d'abord sur le rédacteur, plutôt que sur l'enseignant uniquement.

LE DIAGNOSTIC COGNITIF DANS LA REDACTION

par R.J. Bracewell et A. Breuleux

R.J. BRACEWELL
et A. BREULEUX
Centre de Science Cognitive
de l'Université McGill
Laboratoire de Science
Cognitive Appliquée
3700, rue Mc Tavish
Montréal, Canada
H3A 1Y2

La préparation de ce manuscrit a été rendue possible grâce à des subventions du Conseil de recherche en sciences humaines du Canada et du Ministère de l'enseignement supérieur et des sciences du Québec.

* Ohlsson, 1987.

Cet article a pour objectif d'examiner comment le diagnostic cognitif peut être appliqué au domaine de la rédaction. Cet examen porte sur les structures et les processus qui engendrent le discours, et non sur les structures et processus plus élémentaires impliqués dans l'orthographe et la grammaire. Nous allons brièvement examiner le diagnostic cognitif tel qu'il est appliqué dans d'autres domaines, puis donner un aperçu des caractéristiques des théories de la rédaction pouvant supporter la pratique du diagnostic, et finalement spéculer sur les modifications devant être apportées au diagnostic cognitif afin qu'il puisse être appliqué avec succès à la rédaction.

Le terme diagnostic cognitif réfère au processus par lequel les états cognitifs d'une personne sont inférés à partir de la performance de cette personne à une tâche*. Défini de façon aussi générale, le diagnostic cognitif se distingue difficilement d'autres activités de recherche et d'évaluation cognitives. Néanmoins, la nature du diagnostic cognitif est davantage caractérisée par les domaines de connaissances dans lesquels il a été appliqué et par l'usage de modèles très spécifiques des processus et des connaissances qui déterminent la performance.

En général, les domaines d'application du diagnostic cognitif ont été des domaines dans lesquels les connaissances sont très structurées, comme l'arithmétique ou les problèmes verbaux simples*. Ce haut degré de structuration permet au diagnostiqueur d'utiliser les réponses qu'un sujet donne à un problème afin d'inférer les principaux aspects des processus appliqués par le sujet pour produire la réponse.

De façon plus spécifique, Ohlsson et Langley (1987) ont proposé que les domaines pour lesquels on prévoit de faire du diagnostic cognitif devraient avoir les caractéristiques suivantes :

- 1) Les réponses ou les productions ont une structure interne (par opposition aux réponses «oui-non») ;
- 2) La réponse, ou le but à atteindre, n'est pas fourni dans l'énoncé du problème ;
- 3) La réponse ou le résultat provient d'un ensemble très grand de réponses possibles.

Ces caractéristiques assurent que la réponse produite par le sujet fournit des informations pertinentes sur le traitement cognitif qui l'engendre. La première caractéristique procure une heuristique afin de faire des inférences à propos du traitement (c'est-à-dire,

* voir, par exemple, Brown et Burton, 1978 ; VanLehn, 1983.

la structure observée dans la réponse est le résultat de processus spécifiques), les deux dernières caractéristiques empêchant un processus de réponse banal (c'est-à-dire, que la réponse ait été produite en devinant).

Le domaine particulier que Ohlsson et Langley visaient est celui des problèmes de soustraction en arithmétique. On voit aisément, toutefois, que les trois caractéristiques s'appliquent tout aussi bien au domaine de la rédaction. Le principal obstacle à l'usage du diagnostic cognitif dans la rédaction a été un manque de méthodes permettant de spécifier la structure interne dont il est question dans la première caractéristique. Or, certains développements théoriques récents permettent de déterminer la structure interne des compositions en spécifiant les niveaux de représentations sémantiques et langagiers ainsi que les liens entre ces niveaux*. Ces développements, de même que les validations de cette approche théorique sont décrits dans les deux sections suivantes.

Le diagnostic cognitif est également caractérisé par l'utilisation de modèles détaillés et spécifiques des processus et des connaissances déterminant la structure rencontrée dans les réponses des sujets. Ces modèles expliquent un ensemble d'observations empiriques qui est très vaste en comparaison avec les théories psychologiques habituelles. Un modèle adéquat doit non seulement expliquer les types de réponses observées (c'est-à-dire, réponse correcte-incorrecte), il doit également prédire quels types de réponses incorrectes ne se trouveront pas parmi les réponses des sujets*. Comme VanLehn (1983) l'a démontré, le développement d'un modèle ayant une telle portée implique une interaction entre la spécification des processus qui agissent sur les connaissances et la façon dont les connaissances sont représentées. Dans la mesure où le modèle prédit à la fois les réponses qui apparaissent et celles qui n'apparaissent pas, il constitue une théorie des processus psychologiques et des connaissances utilisées par les sujets. Ainsi, un des résultats du développement d'un tel modèle est qu'il fournit une méthode pour valider les types de représentations qui sont utilisées pour caractériser la structure interne des réponses.

Durant le développement d'un modèle, particulièrement pour la spécification des processus, il est important de considérer les connaissances que le chercheur a de la tâche effectuée par le sujet. Afin de comprendre les réponses, le chercheur doit avoir certaines informations sur comment le sujet perçoit la tâche*. Par exemple, une réponse erronée à un problème de soustraction qu'un sujet interprète comme un problème d'addition révèle très peu de choses concernant les processus cognitifs qu'il utilise pour la soustraction. La détermination de la tâche effectuée par le sujet ne pose pas de problème dans des domaines tels que l'arithmétique, où les connaissances sont très structurées. Par contre, la détermination de la tâche devient une préoccupation centrale dans des domaines tels que la rédaction, où les connaissances peuvent être structurées de manière plus variable et complexe. Les implications de ce degré de complexité plus grand sont examinées dans la dernière section de cet article.

Les méthodes de représentation du discours

L'application efficace du diagnostic cognitif dans le domaine de la rédaction requiert une analyse adéquate de la structure interne d'une composition. Une telle analyse dépend en grande partie de la spécification des concepts, relations et opérateurs qu'un sujet utilise pour effectuer une tâche donnée. Dans les recherches antérieures sur le diagnostic cognitif dans des domaines structurés, une spécification adéquate de ces entités (et par conséquent de la structure interne d'une réponse) est effectuée par une analyse informelle de la tâche de la part du chercheur*. Dans un domaine comme la rédaction, où les connaissances sont plus complexes, il est nécessaire d'utiliser des méthodes plus formelles pour l'analyse de la structure des réponses. Un certain nombre de critères peuvent être proposés en ce qui concerne ces méthodes.

Premièrement, les méthodes doivent avoir une *portée suffisante* pour produire une spécification intégrée de la structure de la composition. En d'autres mots, l'analyse du discours ne peut porter sur un seul dispositif langagier (p. ex. : la structure cohésive) ; elle doit plutôt porter sur le texte et intégrer les structures linguistiques, sémantiques et pragmatiques de la composition.

* Bracewell, Frederiksen et Frederiksen, 1982 ; Frederiksen, Bracewell, Breuleux et Renaud, sous presse ; Frederiksen, Frederiksen et Bracewell, 1987.

* VanLehn, 1983 ; VanLehn, Brown et Greeno, 1984.

* Ohlsson et Langley, 1987.

* Ericsson et Simon, 1984 ; Ohlsson et Langley, 1987.

Deuxièmement, les méthodes doivent être *flexibles* afin de traiter la grande variabilité des structures de texte d'une composition à l'autre. La question de savoir comment traiter ces variations de structure est particulièrement aiguë dans l'investigation et l'évaluation de la rédaction, à cause de la nature essentiellement constructive de la tâche. Même si d'autres tâches impliquant le langage, comme la compréhension et la traduction, sont également constructives, le contrôle exercé par le chercheur sur la structure du texte initial que les sujets doivent lire et/ou traduire réduit significativement les variations entre différents textes produits. Ce contrôle permet, premièrement, l'utilisation de procédures d'analyse plus rigides pour coder ce qui est commun à travers les textes produits et, deuxièmement, permet de traiter les variations résiduelles à titre d'erreurs. Le problème pour l'analyse des productions en rédaction est que le traitement des variations comme erreurs risque d'occulter la structure commune à travers les compositions. La façon de traiter la variation devient une caractéristique importante des méthodes d'analyse elles-mêmes.

Troisièmement, les méthodes doivent être *précises* afin de produire une analyse détaillée de la structure des compositions. Ce n'est qu'avec une analyse qui fournit beaucoup de détails sur la structure du texte qu'une quantité suffisante d'informations peut être extraite des réponses afin d'inférer les conduites cognitives complexes concourant à la réalisation du texte.

Dans nos recherches sur la production de textes, nous tentons de rencontrer ces trois exigences pour les méthodes d'analyse du discours, en incorporant deux caractéristiques générales dans nos modèles. Premièrement, nos modèles de la structure du langage naturel adoptent une *approche stratifiée* de la représentation du discours, dans laquelle la structure d'un texte est représentée à plusieurs niveaux allant du niveau sémantique au niveau syntaxique. Deuxièmement, l'analyse d'un niveau de représentation est déterminé par un système de *règles de formation* définissant les objets propres à un niveau ainsi que les relations admissibles entre ces objets*.

* Frederiksen, Bracewell, Breuleux et Renaud, sous presse.

L'utilisation d'un modèle stratifié permet que l'analyse ait une portée englobante. En effet, des niveaux de représentation sont spécifiés pour les structures linguistiques, sémantiques, et pragmatiques d'un texte, et l'intégration entre ces niveaux est accomplie en spécifiant les relations de transformation entre les objets à différents niveaux. Par ailleurs, l'utilisation d'un système de règles pour définir des représentations aide à rencontrer les exigences de flexibilité et de précision. L'application d'un système de règles à une composition produit un arbre décrivant la structure du texte ; la flexibilité dans la structure est atteinte simplement en émondant l'arbre de façon à ce qu'uniquement les structures représentées dans le texte soient codées. Ainsi, la variabilité des structures de composition est traitée à l'intérieur des méthodes d'analyse. L'utilisation d'un système de règles permet également un haut degré de précision dans l'analyse, étant donné que les composantes du système peuvent en principe être décomposées en ajoutant des règles additionnelles. Ainsi, il est possible d'effectuer des analyses de plus en plus détaillées de la représentation du discours.

Le tableau 1 présente les principaux éléments de notre modèle de la représentation du discours. Des structures discursives sont définies à trois niveaux : (a) les réseaux conceptuels, qui représentent la structure d'ensemble d'un texte, (b) les structures propositionnelles, qui fournissent des représentations sémantiques détaillées pour les concepts des réseaux conceptuels, et (c) les structures thématiques, qui créent une structure entre les phrases du texte. Au sein de chacun de ces niveaux, des types qualitativement différents de représentations sont définis. Les réseaux conceptuels sont composés de cadres narratif, descriptif, et procédural, chaque cadre se distinguant par les types de relations et de concepts qui le composent*. Les structures propositionnelles sont composées de représentations pour les événements, les états, les relations algébriques et les relations de dépendance, définies par des relations de cas et d'attribut, ainsi que de valeurs de temps, de modalité et de vérité*. Les structures thématiques sont composées de séquences d'informations thèmes ou rhèmes dans la structure de surface du texte. Les structures thématiques à l'intérieur des phrases sont définies par des relations syntaxiques ; les structures thématiques entre phrases sont définies par des relations de cohésion*.

* Frederiksen, 1986 ; Frederiksen et Breuleux, sous presse.

* Frederiksen, 1975.

* Clements, 1976 ; Halliday, 1985 ; Halliday et Hasan, 1976.

Tableau 1 : Les niveaux de représentation du discours et les processus s'y rattachant*(d'après Frederiksen, Bracewell, Breuleux et Renaud, sous presse)*

Structures discursives	Processus
Réseaux conceptuels	
Descriptif (relations catégorielles)	Génération de structures de cadre
Narratif (relations temporelles)	Recouvrement en mémoire
Procédural (relations de conditions et de buts)	Intégration de types de cadre
	Séquençement des informations
	Sélection d'informations
Structures propositionnelles	
Événements (relations de cas)	Génération de propositions
États (relations d'attributs)	Recouvrement de propositions
Relations algébriques (p. ex. : ordre, équivalence, proximité)	Évaluation de la cohérence sémantique
Relations de dépendance (p. ex. : cause, condition)	Segmentation des représentations
Structures thématiques	
Propositions syntaxiques (relations syntagmatiques)	Génération de structures syntaxiques
Chaînage cohésif (relations référentielles et conjonctives)	Décisions lexicales
	Évaluation de la cohérence de surface

A chacun de ces niveaux de représentation sont associés des processus cognitifs qui opèrent sur les représentations ou accomplissent une mise en correspondance entre les niveaux de représentation. Les systèmes de règles mentionnés plus haut sont des modèles des processus qui génèrent et contraignent les structures à un niveau particulier de représentation. D'autres processus contraignent les choix parmi les structures sémantiques ou syntaxiques possibles en opérant à travers différents niveaux. Ainsi, des processus de sélection s'appliquent aux structures sémantiques afin de déterminer ce qui doit être exprimé spécifiquement dans le discours et ce qui peut être laissé implicite et déterminé par le lecteur.

Un nombre croissant de recherches indiquent que ces niveaux de représentation correspondent aux structures que les individus engendrent lorsqu'ils utilisent la langue naturelle pour la production, la compréhension, la traduction ou l'acquisition de connaissances*. En ce qui concerne la rédaction, les recherches ont visé principalement à démontrer la validité des différents types de représentation dans des tâches de composition. Des preuves ont été apportées concernant l'utilisation des cadres narratif et procédural*, les structures thématiques**, et la mise en correspondance systématique entre les réseaux conceptuels et les structures thématiques***. La concentration de la recherche sur les représentations au détriment des processus a amené un certain degré d'incertitude dans l'interprétation des observations. Ainsi, les différences dans le degré de développement, par exemple, du cadre narratif peuvent être attribuées à des différences soit dans les connaissances du rédacteur concernant les événements impliqués, soit dans le processus de génération produisant le cadre narratif du texte. Bracewell (1983) a présenté différentes méthodes pour différencier dans la rédaction les effets basés sur les connaissances de ceux basés sur les processus.

Dans l'ensemble, ces résultats suggèrent qu'un modèle stratifié de la représentation du discours commence à ouvrir le domaine de la rédaction au diagnostic cognitif. En particulier, l'application du modèle à la détermination de la structure du discours rencontre les exigences initiales de Ohlsson et Langley pour le diagnostic cognitif, à l'effet que les réponses doivent avoir une structure interne. Les niveaux de représentation présentés plus haut concernent principalement les structures et processus sémantiques (réseaux et propositions sémantiques) et linguistiques (thématisation) impliqués dans la rédaction.

Les méthodes de représentation des tâches de rédaction

Le modèle stratifié présenté ici peut être considéré comme un modèle de base des connaissances et des processus impliqués dans la rédaction*. Un tel modèle est nécessaire à la spécification de la structure interne des réponses dans une tâche de rédaction ; toutefois, certaines extensions au modèle sont requises afin de permettre le diagnostic cognitif. En effet, un traitement de la tâche effectuée par le rédacteur doit être ajouté afin que le modèle puisse prédire la structure des réponses et, plus particulièrement, définir

* Bracewell et al, 1982 ; Frederiksen et Breuleux, sous presse ; Frederiksen et al, 1986 ; Frederiksen et al, (sous presse).

* Bracewell et al, 1982 ; DeRemer et Bracewell, 1989 ; Frederiksen et al, 1986 ; Senecal, 1989.

** Bracewell, 1986, 1987.

*** Bracewell, 1986, 1987 ; Frederiksen et al, 1986.

* Frederiksen et al, 1986, p. 285.

les processus appliqués aux représentations.

Dans les domaines hautement structurés dans lesquels le diagnostic cognitif est actuellement employé, une définition adéquate de la tâche peut initialement être obtenue par des experts du domaine en effectuant une analyse informelle de la tâche*. L'adéquation de la définition de la tâche est utilisée ici en référence avec deux activités : la détermination de la tâche exacte effectuée par le sujet à partir d'un ensemble de réponses, et le développement d'un modèle de diagnostic cognitif. Ainsi, dans une tâche d'arithmétique, la structure des entiers en base dix contraint les représentations utilisées pour les connaissances, et les opérations concernant l'addition et la soustraction des nombres contraint la définition des processus cognitifs applicables à ces représentations. Par la suite, la spécification détaillée des représentations et opérations d'un modèle n'est évidemment pas subjective, et dépend des formes particulières de structure interne rencontrées dans les réponses des sujets.

* Ohlsson et Langley, 1986 ; Van-Lehn, Brown et Greeno, 1984.

Dans le domaine de la rédaction, une définition adéquate des pré-requis pour une tâche nécessite de traiter ce qui, dans une analyse linguistique générale, serait appelé la pragmatique de la composition, c'est-à-dire l'utilisation faite par le rédacteur des représentations sémantiques et discursives à certaines fins de communication dans différentes situations, pour des auditoires particuliers. Les informations pertinentes à la pragmatique de la composition sont souvent regroupées sous la catégorie de «contexte situationnel»*, et comprennent les présupposés du rédacteur à propos de ce que l'auditoire connaît du domaine abordé, le rapport entre le rédacteur et l'auditoire (p. ex. : enseignant/étudiant), le genre et le registre que le rédacteur croit approprié en fonction de sa représentation de l'auditoire et de son rapport avec lui, et les objectifs que le rédacteur compte atteindre au moyen du texte ainsi que les répercussions anticipées sur l'auditoire*.

* Firth, 1957.

* voir Halliday, 1978 ; Hymes, 1967.

Ce sont les connaissances du contexte situationnel du rédacteur qui définissent une tâche particulière de rédaction. Les représentations et les processus qui constituent ces connaissances et les relient au texte rédigé sont présentés sommairement au tableau 2. La forme de ces représentations est la même que celle des cadres de connaissances pour le contenu sémantique du texte. L'auditoire et les caractéristiques langagières sont représentés par un cadre descriptif qui est engendré par un processus de recouvrement en mémoire et de génération de cadre à partir d'une grammaire descriptive. La fonction d'un texte est définie par un ensemble de buts engendrés suivant une grammaire procédurale. Le plan du texte fournit les liens avec le texte lui-même, servant de structure de contrôle pour la génération du texte.

Tableau 2 : Les représentations et processus cognitifs pour la définition de tâches

Représentations	Processus
Caractéristiques de l'auditoire	
Connaissances présumées	Recouvrement en mémoire
Rapport à l'auteur	Génération d'un cadre descriptif
Caractéristiques langagières	
Genre	Recouvrement en mémoire
Registre	Génération d'un cadre descriptif
Fonctions du texte	
Buts finaux	Spécification de buts
Plan du texte	
	Spécification et organisation de buts
	Evaluation de plans

Un examen du contenu présenté au tableau 2 commence à révéler la nature du problème de définition de la tâche, et indique le rôle et les limites de l'analyse informelle de la tâche dans le domaine de la rédaction.

Premièrement, du point de vue du chercheur (ou du diagnostiqueur), les composantes de la définition de la tâche sont semblables à des paramètres libres qui peuvent prendre un ensemble assez grand de valeurs différentes. Autrement dit, les caractéristiques de l'auditoire, le rapport rédacteur/auditoire, les buts, et le genre définissent la tâche, mais ces valeurs sont puisées parmi un grand nombre de valeurs possibles et correspondent à un ensemble de possibilités combinatoires.

De plus, le chercheur s'est intéressé à expliquer comment ces caractéristiques sont effectivement organisées et utilisées par le rédacteur pour contrôler la production du

texte. En l'absence d'informations sur les processus de contrôle, il est difficile d'identifier quels processus sont responsables d'une réponse ayant une structure «erronée». Etant donné que la définition de la tâche ne se manifeste habituellement que de manière indirecte dans le texte lui-même, il est nécessaire d'avoir une méthode pour déterminer les processus de contrôle utilisés par un rédacteur pour une composition particulière.

Deuxièmement, du point de vue du rédacteur, les caractéristiques définissant une tâche sont connues, mais étant donné que la tâche implique une communication, cette définition de la tâche est dynamique. En rédigeant, l'auteur s'adresse à une communauté discursive*, et bien qu'il ne soit pas en relation immédiate avec cette communauté, l'auteur doit, pour la production de textes complets, anticiper le processus de réponse par la communauté au moment de la lecture du texte. Ainsi, la définition de la tâche évolue durant la production du texte, au fur et à mesure que le rédacteur considère les effets possibles des parties du texte et les révisé pour satisfaire les besoins de l'auditoire. On peut dire que ces aspects de la tâche ne peuvent habituellement pas être déterminés au préalable de manière fiable par le rédacteur. Des méthodes permettant d'enregistrer cette évolution de la définition de la tâche sont nécessaires pour le diagnostic cognitif.

* Faigley, 1984.

Troisièmement, la relation entre la définition de la tâche et la structure éventuelle du texte est essentiellement heuristique, au sens où le rédacteur ne peut anticiper avec précision plusieurs aspects importants de la structure du texte*. Cela semble être dû à la définition dynamique de la tâche et aux limites du système humain de traitement de l'information. Bien que le contenu pragmatique joue un rôle important dans la détermination des opérations appliquées aux représentations des cadres, des propositions, et des thématisations en vue de produire un texte, ces opérations dépendent pour une large part des évaluations comparatives du texte produit et des intentions. Le diagnostic cognitif nécessite d'enregistrer comment les heuristiques comme la planification sont utilisées par le rédacteur pour produire un texte.

* Simon, 1978.

Ces facteurs suggèrent qu'une variété de techniques devront être utilisées afin de définir des tâches de rédaction suffisamment en détail pour permettre le diagnostic cognitif. La nature de ces techniques peut être illustrée en considérant deux situations extrêmes : les tâches de rédaction déterminées de l'extérieur, dont la définition est extérieure au rédacteur, et les tâches de rédaction déterminées par le rédacteur.

Les tâches de rédaction déterminées de l'extérieur



Les tâches de rédaction déterminées de l'extérieur sont les tâches pour lesquelles l'objectif, l'auditoire, le genre, et le contenu sont spécifiés pour le rédacteur. Les tâches de ce type sont courantes dans les contextes de recherche (où les variations pragmatiques sont utilisées comme traitement), en éducation (souvent au désespoir des pédagogues qui désirent favoriser un éventail complet d'habiletés de rédaction), et dans l'entreprise et les organismes gouvernementaux (où des procédures, des procédés, et des systèmes doivent être documentés). La spécification des informations pragmatiques, de même que les éléments du modèle stratifié de la représentation et du traitement discursifs, fournissent les informations nécessaires à la définition de tâches par une analyse informelle. Pour ce type de tâche, le diagnostic cognitif s'effectue comme pour les tâches d'arithmétique et d'algèbre. La principale différence avec les domaines très structurés est la nécessité de déterminer si le rédacteur a assimilé les spécifications externes et en tient compte durant la rédaction.

Le rôle crucial de ces informations pour le diagnostic peut être illustré à l'aide de l'exemple suivant. Le tableau 3 présente un extrait de deux phrases de la version préliminaire d'une thèse soumise par un de nos étudiants.

Tableau 3 : Extrait d'un texte d'un étudiant

Phrases initiales :

La variable indépendante dans cette étude est le type de révision que l'élève est amené à effectuer, soit ajouter soit effacer un élément du texte. Les évaluations concernant lequel des textes original et révisé est meilleur, telles que fournies a) l'élève, b) en comparant les structures de cadres, et c) par un juge indépendant, sont les variables dépendantes.

Phrases révisées :

La variable indépendante dans cette étude est le type de révision que l'élève est amené à effectuer, soit ajouter soit effacer un élément du texte. Les variables dépendantes sont les évaluations concernant lequel des textes original et révisé est meilleur, telles que fournies a) par l'élève, b) en comparant les structures de cadres, et c) par un juge indépendant.

Ces phrases proviennent du début de la section sur les résultats et servent à résumer les variables dépendantes et indépendantes de la recherche. La version originale est présentée au haut du tableau, la version révisée dans la partie inférieure.

Le problème de la première version est que la thématization des deux phrases ne signale pas adéquatement le cadre descriptif qui structure les informations concernant les variables utilisées pour la recherche : dans la première phrase, un concept de catégorie (*La variable indépendante*) est thématized et la spécification des variables (*le type de révision...*) est mise en position de rhème. Dans la seconde phrase, cette structure est inversée, la spécification des variables (*Les évaluations concernant...*) étant thématized, et le concept de catégorie (*les variables dépendantes*) étant placé en position rhème. Or, on comprend aisément, à la simple lecture du texte, que le cadre descriptif est composé de la désignation de deux contenus similaires et symétriques : la notion de variable indépendante et son contenu, et la notion de variable dépendante et son contenu. Un des principaux objectifs de la présentation de ces informations est d'en faciliter le décodage par le lecteur à partir du texte. Etant donné cet objectif, une signalisation appropriée de la similitude sémantique dans la structure de surface du texte consiste à thématiser les unités d'information d'un même type. La structure préférable est la thématization de l'information au sujet de la catégorie, avec la spécification du contenu de la catégorie en position de rhème. (Cette structure est préférable parce qu'elle correspond à l'ordre des informations données et nouvelles - les lecteurs de textes décrivant une recherche expérimentale s'attendent à trouver la mention de variables dépendantes et indépendantes, l'information nouvelle étant ce que sont ces variables dans la recherche particulière qui est décrite). Les phrases révisées, qui montrent la structure thématique préférable, sont présentées au bas du tableau 3.

Ce court extrait et sa révision sont particulièrement intéressants parce que nous avons également des informations concernant la définition de la tâche par le rédacteur au moment de produire la structure thématique initiale. Questionné au sujet de cette structure, le rédacteur explique que l'alternance dans la structure thématique des deux phrases est délibérée, et élaborée afin que le texte soit plus intéressant pour le lecteur. L'analyse présentée plus haut interprète le problème comme étant un problème d'asymétrie entre la structure sémantique et la structure thématique en fonction d'un but particulier. Toutefois, à la lueur des justifications fournies par le rédacteur, on peut voir que l'asymétrie est en fait motivée par le but de produire un texte plus intéressant. De façon plus générale, le rédacteur de ce texte avait défini la tâche davantage en fonction du genre littéraire que du genre scientifique.

L'erreur commise, s'il en est, ne réside pas tant dans la structure du texte lui-même, mais bien dans l'inadéquation entre la tâche que se définit le rédacteur et celle que le diagnostiqueur définit. Autrement dit, la spécification et l'organisation de buts est également une tâche, et les plans élaborés en fonction de cette tâche ont eux-mêmes une structure. Pour que le diagnostic cognitif soit efficace dans les tâches de production du discours, il faut avoir accès à ces réponses de haut niveau et pouvoir les analyser aussi bien que la structure du texte lui-même. Cela justifie l'utilisation de questions (p. ex. : le rappel des instructions pour la tâche) en supplément à l'analyse des textes, les réponses aux questions et le texte constituant la matière à diagnostiquer.

Les tâches de rédaction déterminées par le rédacteur

A l'autre extrême, les tâches de rédaction dont la définition est déterminée par le rédacteur nécessitent un large éventail de techniques pour le diagnostic cognitif. La plus représentative de ce type de tâches est sans doute la rédaction d'un roman. La décision de rédiger un roman ne fournit qu'une spécification très générale du contenu pragmatique, l'écrivain devant par la suite puiser un agencement particulier de caractéristiques de l'auditoire, de caractéristiques langagières, et d'objectifs parmi l'ensemble assez grand de valeurs possibles. L'agencement choisi par le rédacteur définit essentiellement la tâche, et les connaissances de ces informations sont nécessaires au diagnostic cognitif.

Dans ce dernier cas, le diagnostic cognitif nécessite une trace de l'évolution des spécifications pragmatiques par lesquelles le rédacteur définit graduellement sa tâche au fur et à mesure de la composition. Des méthodes pour obtenir de telles traces ont été

développées dans les recherches sur la résolution de problèmes, en particulier pour les domaines de connaissances hautement structurés dans lesquels les objectifs (ou les réponses) sont connus d'avance mais où le cheminement entre l'énoncé du problème et l'objectif n'est pas connu. Le type même de ce genre de tâche est le puzzle de la tour de Hanoi*. Etant donné que l'effort principal dans ce genre de tâche consiste à découvrir le chemin qui mène de la position de départ à la solution, les méthodes de cueillette de traces utilisant l'enregistrement vidéo, les verbalisations des pensées, et les questions durant la performance se révèlent être des techniques privilégiées de cueillette de données.

* Simon, 1975.

La méthode principalement utilisée dans les recherches sur la résolution de problèmes est basée sur les pensées verbalisées - le compte-rendu verbal des pensées et des actions produit par le sujet durant l'exécution d'une tâche. Les protocoles obtenus durant ces verbalisations constituent une banque d'informations ; le recouvrement des données à partir des protocoles dépend de la spécification par le chercheur d'une procédure de codage, ou d'analyse, des verbalisations. Dans les domaines de connaissance hautement structurés, les procédures de codage des protocoles de pensées à voix haute peuvent être obtenues à partir d'une analyse informelle de la tâche, étant donné que l'uniformité de la tâche à travers les sujets et les occurrences de la tâche permet d'inférer de façon fiable les concepts et les opérations que les sujets utilisent ou effectuent. Bien qu'il existe certaines controverses concernant les biais engendrés par la verbalisation simultanée*, lorsqu'elle est utilisée correctement la méthode fournit un résumé précis (au moins) des stratégies impliquées dans la résolution d'un problème*.

* Nisbett et Wilson, 1977.

* voir Ericsson et Simon, 1984, pages 63 à 172.

* Hayes et Flower, 1983.

Des méthodes de traçage, et en particulier la méthode des pensées à voix haute, peuvent être appliquées à la rédaction*. La principale difficulté de cet usage réside dans la mise au point d'une procédure de codage des protocoles adéquate. Ces difficultés sont à la fois théoriques et méthodologiques. En ce qui concerne la théorie, les domaines tels que la rédaction ne se prêtent pas à une analyse informelle de la tâche, qui ne procure pas une sémantique suffisamment complète pour le codage des concepts et des opérations que le sujet verbalise concernant la composition. En ce qui a trait à la méthodologie, le langage contenu dans les protocoles de pensées à voix haute est en langage naturel, et manifeste la grande variation lexicale et syntaxique qui est le propre des productions en langue naturelle. Pour ces raisons, il est difficile de prédire avec une précision suffisante pour le développement d'une procédure de codage rigoureuse, ce que le rédacteur va dire à propos de la tâche qu'il effectue. Autrement dit, les insuffisances de l'analyse informelle de la tâche pour le diagnostic cognitif s'appliquent également aux protocoles de pensées à voix haute.

Cet obstacle au codage des protocoles de pensées à voix haute peut sembler insurmontable ; toutefois, il peut être contourné en appliquant le modèle stratifié d'analyse du discours présenté plus haut. Breuleux* a développé une procédure de codage permettant d'identifier à partir de protocoles de pensées à voix haute les processus de planification et d'évaluation utilisés par un rédacteur durant la production d'un texte. Cette méthode porte essentiellement sur la détermination des buts spécifiés par le rédacteur et sur l'organisation de ces buts pour former un plan pour la production du texte. Les buts sont définis en termes de caractéristiques de contenus et de relations sémantiques au niveau des propositions ; un exemple de définition d'un but étant une configuration sémantique comportant un agent à la première personne et une action dont le temps est marqué comme étant futur (p. ex. : «Je vais écrire un article sur...»). L'organisation des buts entre eux est également définie au niveau propositionnel, plus spécifiquement en termes de relations temporelles et de relations de dépendances (p. ex. : «Premièrement, je veux parler des perspectives... et ensuite expliquer pourquoi elles sont ainsi...»). Finalement, les énoncés d'évaluation sont définis au niveau propositionnel en termes de relations d'attributs (p. ex. : «c'est bon»), qui marquent la vérification de l'atteinte d'un état souhaité ou d'un but.

* 1989 ; sous presse.

Nous travaillons actuellement à étendre cette approche pour couvrir à la fois les informations descriptives qui composent le reste du contenu pragmatique de la définition de la tâche (les caractéristiques de l'auditoire, etc.), et le codage des relations entre le plan et la structure du texte lui-même. Les informations pragmatiques descriptives semblent se manifester habituellement comme contenu des buts ou dans les relations de dépendances entre les buts (p. ex. : «Je vais faire cela dans un style éditorialiste»). Les relations entre le plan et le texte se manifestent de deux façons : premièrement à travers la spécification

par le rédacteur de buts au niveau détaillé de la structure du texte (p. ex. : «utilisons un autre mot que soviétiques...»), deuxièmement par la présence d'énoncés d'évaluation qui signalent une comparaison entre le texte et un but (p. ex. : «ce n'est pas vraiment bien dit, non, ça ne marche pas, il faut que je reformule cela»).

Cette approche résout à la fois les problèmes théoriques et les problèmes méthodologiques du codage des protocoles de pensées à voix haute dans des domaines complexes. La sémantique appropriée pour le codage est fournie par des listes d'attributs sémantiques motivés par la théorie (p. ex. : la référence pronominale ou le temps futur) et de relations entre concepts (p. ex. : l'agence). Pour ce qui est de la méthodologie, la spécification des structures sémantiques réduit considérablement les variations rencontrées dans les structures de surface en langue naturelle en un nombre fini de types.

De plus, l'approche permet la mise au point de batteries de questions pour le rédacteur motivées théoriquement, le contenu des questions étant déterminé par les types de représentations utilisées par le rédacteur pour définir la tâche. De cette façon, les réponses aux questions viennent compléter les pensées à voix haute afin de définir la tâche.

Ainsi, il devient possible d'analyser en détail la définition de la tâche par le rédacteur et, en conjonction avec la structure du texte produit qui est révélée par l'analyse du discours, de satisfaire les conditions d'exercice du diagnostic cognitif.

Les implications pour le diagnostic cognitif dans la rédaction

La discussion précédente présente les méthodes et les justifications indiquant que le diagnostic peut être utilisé pour évaluer les structures et les processus impliqués dans la production du discours écrit. Toutefois, l'implantation effective de ces méthodes soulève certains problèmes, à la fois d'ordre pratique et théorique.

Le problème pratique concerne les procédures pour obtenir la définition de la tâche. Tel que mentionné plus haut, la définition spécifique d'une tâche varie d'un texte à l'autre, et même d'un moment à l'autre durant la production d'un même texte.

Dans de nombreuses tâches de rédaction, afin d'obtenir des informations suffisamment détaillées sur la tâche que le rédacteur effectue il est nécessaire d'avoir une trace de la définition de la tâche. L'obtention d'une telle trace implique que le rédacteur ait comme compagnon un analyste du discours pouvant coder les verbalisations du rédacteur au sujet de la tâche, soit les pensées à voix haute soit les réponses à des questions. Actuellement, et jusqu'à ce que les logiciels capables d'interpréter le langage naturel soient développés, ce compagnon doit être un humain. Ainsi, pour n'importe quelle application pratique, comme dans l'enseignement ou l'aide à la performance, ces conditions limitent l'exercice du diagnostic cognitif à des situations de tutorat.

Le problème théorique concerne la localisation des «erreurs» dans les représentations et les processus impliqués dans la rédaction. Comme dans le cas du problème pratique présenté plus haut, c'est la définition de la tâche elle-même (en particulier les buts et leur structuration) qui est l'élément principal du problème. Dans son expression la plus simple, la question d'identifier la source des erreurs concerne la difficulté d'attribuer l'erreur à des buts non optimaux pour la tâche (et réalisés correctement dans le texte), ou à un texte réalisant incorrectement des buts appropriés. Le texte présenté au tableau 3 illustre ce dilemme : étant donné le but d'écrire un texte scientifique, la structure originale est erronée ; mais étant donné le but d'écrire un texte littéraire (ce qui en soit n'est pas adapté à la situation), la structure du texte est correcte.

A cause de ces problèmes théoriques et pratiques, la réalisation du diagnostic cognitif dans l'enseignement de la rédaction aura une forme différente de celle dans les domaines plus structurés. Le diagnostic cognitif dans les domaines structurés a eu un impact principalement sur l'enseignement, en ce qu'il a modifié la conception des erreurs chez les enseignants (elles sont conçues comme étant motivées plutôt qu'aléatoires), et a mené à un contenu pédagogique mieux circonscrit (l'enseignant sait à quoi il faut remédier et comment y remédier). La nécessité d'obtenir des traces dynamiques de la définition de la tâche aura pour effet d'orienter les techniques de diagnostic cognitif vers le rédacteur plutôt que vers l'enseignant.

Pour les tâches de rédaction définies de manière externe, le diagnostic cognitif aura

la forme qu'il a dans les domaines structurés. Cela est dû au fait que le diagnostiqueur peut obtenir des informations au sujet de la définition de la tâche à partir de sources autres que le rédacteur. La principale différence avec les techniques pré-existantes consiste à déterminer si le rédacteur utilise effectivement les spécifications externes pour élaborer une définition de la tâche et à déterminer l'efficacité avec laquelle ces informations sont utilisées pour contrôler la production du texte. Cette différence implique des systèmes pédagogiques utilisant des procédures de questionnement et de rapports rétrospectifs (en particulier avec les sujets plus habiles) afin d'obtenir une définition de la tâche permettant le diagnostic cognitif. La rédaction et les révisions présentées au tableau 3 illustrent ce type de procédure de diagnostic.

Pour les tâches de rédaction définies par le rédacteur, la principale ressource disponible pour le diagnostic cognitif est la personne qui rédige. L'implantation du diagnostic cognitif pour ces tâches se concentrera probablement sur l'entraînement du rédacteur à analyser et spécifier sa tâche, à représenter de façon plus explicite et plus détaillée les divers niveaux de représentation élaborés pour un texte particulier, et à rendre explicites les relations de correspondance entre ces divers niveaux. Les systèmes pédagogiques refléteront cette emphase sur le diagnostic cognitif par le rédacteur, et on peut prévoir au moins deux façons complémentaires d'assister les conduites du rédacteur. Premièrement, il s'agit de développer des environnements informatiques d'aide à la rédaction permettant au rédacteur de représenter explicitement les structures à différents niveaux ainsi que leurs relations. En particulier, l'utilisation d'interfaces à base d'objets graphiques permettrait au rédacteur de construire et de manipuler directement des structures de niveau sémantique (plutôt qu'indirectement par le biais de la structure de surface du texte). Deuxièmement, il s'agit d'utiliser des méthodes de questionnement des pairs pour obtenir rétroactivement des informations et des commentaires sur la structure du texte et sur ses effets. Dans les situations où les interactions et les intérêts au sein d'un groupe sont suffisants pour un type de rédaction, il devrait être possible de former une communauté discursive procurant le feedback nécessaire à l'acquisition d'habiletés de rédaction, particulièrement si les membres de la communauté disposent des outils appropriés concernant la nature et la dénomination des unités d'analyse décrites dans cet article. Pour ces deux méthodes, toutefois, le terme de *diagnostic cognitif* semble inadéquat. Puisque ces activités visent à assister le déroulement des opérations, soit par le rédacteur soit par la communauté, il s'agit plutôt de *monitoring cognitif*.

Références



- BRACEWELL R.J. (1983). Investigating the control writing skills. In P. Mosenthal, S. Walmsley et L. Tamor (Eds.), *Research in writing : Principles and methods*. London : Longman Press.
- BRACEWELL R.J. (1986, Mai). *Signalling semantic representation : The role of topical patterning in text*. Article présenté au Fourth Annual International Conference on the Teaching of English, Ottawa.
- BRACEWELL R.J. (1987, Avril). *Semantic and textual constraints students use in revising their writing*. Article présenté à la rencontre annuelle de l'American Educational Research Association, Washington.
- BRACEWELL R.J., FREDERIKSEN C. H., et FREDERIKSEN J.D. (1982). Cognitive knowledge and processes as a guide to composing and comprehending discourse. *Educational Psychologist*, 17, 146-164.
- BREULEUX A. (1989). *L'élaboration et l'exécution de plans dans une tâche de rédaction*. Thèse de doctorat. Université de Montréal, Montréal.
- BREULEUX A. (sous presse). The analysis of writers' think aloud protocols : Developing a principled coding scheme for ill-structured tasks. In G. Denhière et J.-P. Rossi (Eds), *Texts and text processing*. Amsterdam : North-Holland.
- BROWN J.S. et BURTON R.B. (1978). Diagnostic models for procedural bugs in basic mathematical skills. *Cognitive Science*, 2, 155-192.
- CLEMENTS P. (1976). The effects of staging on recall from prose. In R. Freedle (Ed.), *Advances in discourse processing : Vol. 2* Norwood, NJ : Ablex.
- DEREMER M. et BRACEWELL R.J. (1989, March). *Students' use of semantic structure in revising their writing*. Article présenté à la rencontre annuelle de l'American Educational Research Association, San Francisco.
- ERICSSON K.A. et SIMON H.A. (1984). *Protocol analysis : Verbal reports as data*. Cambridge, MA : MIT Press.
- FAIGLEY L. (1984). Nonacademic writing : The social perspective. In L. Odell et D. Goswami (Eds), *Writing in nonacademic settings*. New York : Guilford Press.
- FIRTH J.R. (1957). Personality and language in society. In J.R. Firth (Ed.), *Papers in linguistics 1934-*

1951. London : Oxford University Press.

– FREDERIKSEN C.H. (1975). Representing logical and semantic structure of knowledge acquired from discourse. *Cognitive Psychology*, 7, 371-348.

– FREDERIKSEN C.H. (1986). Cognitive models and discourse analysis. In C.R. Cooper et S. Greenbaum (Eds), *Written communication annual : An international survey of research and theory : Vol 1 : Linguistic approaches*. Beverly Hills, CA : Sage.

– FREDERIKSEN C.H. et BREULEUX A. (sous presse). Monitoring cognitive processing in semantically complex domains. In N. Frederiksen, R. Glaser, A. Lesgold et M. Shafto (Eds), *Diagnostic monitoring of skill and knowledge acquisition*. Hillsdale, NJ : L. Erlbaum Associates.

– FREDERIKSEN C.H., FREDERIKSEN J.D. et BRACEWELL R.J. (1987). Discourse analysis of children's text production. In A. Matsubashi (Ed.), *Writing in real time : Modelling production processes*. Norwood, N.J. : Ablex.

– FREDERIKSEN C.H., BRACEWELL R.J., BREULEUX A. et RENAUD A. (sous presse). The cognitive representation and processing of discourse : Function and dysfunction. In Y. Joannette et H. Brownell (Eds), *Discourse ability and brain damage : Theoretical and empirical perspectives*. Amsterdam : Springer Verlag.

– HALLIDAY M.A.K. (1985). *An introduction to functional grammar*. London : Edward Arnold.

– HALLIDAY M.A.K. (1978). Language as social semiotic : *The social interpretation of language and meaning*. London : Edward Arnold.

– HALLIDAY M.A.K. et HASAN R. (1976). *Cohesion in English*. London : Longman Press.

– HAYES J.R. et FLOWER L.S.J. (1983). Uncovering cognitive processes in writing : An introduction to protocol analysis. In P. Mosenthal, S. Walmsley et L. Tamor (Eds), *Research in writing : Principles and methods*. London : Longman.

– HYMES D. (1967). Models of the interaction of language and social setting. *Journal of Social Issues*, 23, 8-28.

– NISBETT R.E. et WILSON T.D. (1977). Telling more than we can know : Verbal reports on mental processes. *Psychological Review*, 84, 231-259.

– OHLSSON S. (1987). Some principles of intelligent tutoring. In R.W. Lawler et M. Yazdani (Eds), *Artificial intelligence and education : Vol. 1. Learning environments and tutoring systems*. Norwood, NJ : Ablex.

– OHLSSON S., LANGLEY P. (1987). Psychological evaluation of path hypothesis in cognitive diagnosis. In H. Mandl et A. Lesgold (Eds), *Learning issues for intelligent tutoring systems*. New York : Springer Verlag.

– SENECAI (1989, Avril). *Elementary students' first-language narrative writing : A longitudinal study*. Article présenté à la Seventh Annual National Testing Network Conference on Writing Assessment. Montréal.

– SIMON H.A. (1975). The functional equivalence of problem solving skills. *Cognitive Psychology*, 7, 268-288.

– SIMON H.A. (1978). Information-processing theory of human problem solving. In W.K. Estes (Ed.), *Handbook of learning and cognitive processes : Vol. 5 Human information processing* (pp. 271-295). Hillsdale, NJ : Erlbaum Associates.

– van LEHN K. (1983). On the representation of procedures in repair theory. In H.P. Ginsberg (Ed.), *The development of mathematical thinking*. New York : Academic.

– van LEHN K., BROWN J.S. et GREENO J.G. (1984). Competitive argumentation in computational theories of cognition. In W. Kintsch, J. Miller et P. Polson (Eds), *Methods and tactics in cognitive science*. Hillsdale, NJ : L. Erlbaum Associates.