

Approcher la qualité textuelle des écrits scolaires : le corpus grenouille à l'essai d'un jugement d'experts

Bernard Slusarczyk
Université P. Mendés France, Grenoble



Synergies Pays Scandinaves n° 5 - 2010 pp. 97-110

Résumé : *Les mesures de qualité de textes peuvent être explorées selon deux grandes perspectives : soit on considère qu'il y a interaction entre le lecteur et le texte, soit on considère le texte comme ayant cristallisé la pensée de l'auteur. Notre étude vérifie si ces deux types de mesures caractérisant chacune de ces deux perspectives peuvent être corrélés sur les copies de récits du corpus Grenouille. Une première mesure de qualité de texte est effectuée, sur l'ensemble des copies d'élèves, par des correcteurs experts, qui utilisent une grille bâtie à partir d'une échelle connue dans la littérature, une seconde est effectuée par un chercheur et vise à quantifier objectivement les contenus factuels et inférés du texte, tels qu'ils sont présentés dans un test standardisé. La comparaison conduite à propos de ces deux types d'évaluation montre que le nombre d'idées générées par les élèves en lien avec la compréhension globale du récit (contenus factuels et inférés) est corrélé au jugement de qualité globale portant sur la structure, la cohérence, la syntaxe et le vocabulaire. Cela renvoie à une question fondamentale de l'enseignement de la rédaction : faut-il apprendre aux élèves à écrire pour autrui ou à réfléchir ?*

Mots-clés : *Qualité de texte, test, évaluation.*

Summary: *Measuring the quality of texts can be explored with two main perspectives: either we focus on an interaction between the reader and the text, or we consider the text as a record of the author's thought. Our study set out to determine whether these two types of measurement could be correlated in composition scripts in the Grenouille corpus. A first measurement of text quality was made on the entire set of pupil scripts by expert graders using a scheme constructed from a literature-endorsed scale. A second measurement was made by a researcher to quantify the factual and inferred text content objectively as presented in a standardised test. Comparison of these measures showed that the number of ideas generated by the pupils linked to the overall understanding of the composition (factual and inferred content) was correlated with the overall judgement of quality in terms of structure, coherence, syntax and vocabulary. This finding points again to a fundamental issue in the teaching of writing skills: should we teach pupils to write for others or to think for themselves?*

Key words: *Writing quality, testing, evaluation.*

1. Introduction

Les modélisations traditionnelles, issues de la rhétorique classique, se sont longtemps attachées à représenter l'activité du rédacteur. Or depuis une trentaine d'années, de nombreux travaux théoriques et empiriques ont abordé le problème de la rédaction en dépassant ces modélisations traditionnelles, qui étaient davantage conçues comme des modèles à imiter (savoir écrire un conte, une nouvelle, un dialogue romanesque ... etc.) que des modèles destinés à rendre compte de l'activité effective du scripteur (Plane, 2006).

Qu'est-ce qui fait qu'un texte est un texte et non pas une suite de mots ? Qu'est-ce qui fait qu'un texte est un texte et non pas une suite de phrases ? Dès que l'on se pose la question de la définition de l'unité linguistique "texte", se pose d'emblée la question de l'unité linguistique plus petite que l'on va utiliser pour définir le tout à partir d'une partie. La psycholinguistique textuelle, qui étudie les mécanismes par lesquels le sujet traite les dispositifs langagiers en vue de produire ou d'interpréter une suite cohérente d'énoncés, a toujours caractérisé le discours comme un enchaînement d'énoncés multiples (Coirier, 1996). Constatant qu'il est en effet rare que nous communiquions avec autrui par le biais d'énoncés isolés, Coirier (1996) pose l'enchaînement d'énoncés comme une caractéristique fondamentale de l'expression, que l'on retrouve dans le discours comme dans le texte. La question posée alors est de savoir ce qui différencie un enchaînement de phrases, dans lequel tout le monde s'accorde à ne voir qu'un enchaînement de phrases, d'un enchaînement de phrases où tout le monde reconnaît un texte. Répondre à cette question n'est pas aisé, et demande à préciser quel type de propriétés l'on cherche à identifier, pour caractériser le texte.

2. Une problématique pédagogique

Lorsque l'on enseigne, peut-on mesurer un texte par ses propriétés caractéristiques ? Les caractéristiques du texte donnent certainement des éléments qui permettent de qualifier, c'est-à-dire, pour reprendre l'étymologie du mot qualifier, de donner l'état du texte tel qu'il est. C'est la mesure des caractéristiques qui ordonne les degrés de qualification du texte. Le terme de qualité de texte ayant des acceptions variées et différentes, nous préférons pour cette raison parler de degrés de qualification pour désigner la qualité textuelle, entendue comme composée de l'ensemble des caractéristiques permettant la qualification.

Le choix du type de propriétés détermine alors et le traitement et la méthode par lesquels le texte sera analysé. Mais, ceci n'est pas sans répercussion sur les modèles implicites qui sont sous-jacentes à la catégorisation employée, ni sur l'exploitation praxéologique que font les enseignants des modélisations actualisées au plan de la recherche. Dans certaines classes, les enseignants accordent une place privilégiée à l'une des quatre propriétés de base d'un texte de qualité définies par les chercheurs (voir ci-après), et considèrent alors de manière prioritaire les indices d'énonciation. Quels sont les modèles utilisés implicitement par les enseignants ?

3. Une double perspective chez les chercheurs

Le texte est parfois considéré comme une entité en soi, ou alternativement plutôt défini par le lien entretenu avec le lecteur ; les propriétés visées sont alors plutôt intrinsèques ou plutôt extrinsèques. Deux orientations, psycholinguistique et discursive, se font jour.

Dans une perspective psycholinguistique, la pertinence du niveau texte apparaît dans le fait que certaines des unités linguistiques le composant, tels les pronoms, les connecteurs et le temps d'emploi des verbes ne peuvent absolument pas s'appréhender au sein seul de la phrase (Fayol, 1995). Les propriétés ici visées sont intrinsèques au texte. Dans cette perspective psycholinguistique, le texte se définit comme une unité de langage en usage à part entière. Le texte constitue le médium langagier spécifique d'une construction de représentations cognitives subjectives visant la transmission de l'information (Coirier, 1996). Dans ce cadre, les représentations textuelles peuvent se résumer à trois bases d'appui : 1) une continuité référentielle, 2) une continuité logique, argumentative et 3) une continuité dite énonciative.

La continuité référentielle est assurée par des dispositifs de cohésion, de structures, de thème, et les systèmes aspectuels et temporels. La continuité logique et argumentative établit la cohérence grâce notamment à l'emploi de structures rhétoriques, repérées depuis longtemps en linguistique. La continuité énonciative opère au plan de l'interlocution pour maintenir la finalité du discours.

Si l'on se place dans une perspective discursive, la question se transforme de la manière suivante : qu'est-ce qui fait que le lecteur dispose des éléments nécessaires qui lui permettent d'interpréter cet ensemble comme un texte pour affecter une signification à ce texte ? Les réponses sont alors directement apportées par tout lecteur potentiel. Plane (2006) donne, selon cette perspective une liste de quatre éléments sur lesquels un lecteur potentiel s'appuie généralement : 1) l'unicité, 2) la socialité, 3) la performativité, et 4) la stabilité référentielle. L'unicité concerne le texte ; dans ce cadre, l'objet textuel est perçu comme un tout par celui qui en prend connaissance. La socialité repose sur le fait que le texte contribue à la constitution de la communauté discursive ; il procède d'un genre reconnu par les initiés. La performativité du texte renseigne, elle, sur l'intention de communication. Le texte s'adresse à un destinataire, il véhicule une intention qui super ordonne les actes de langage dont le texte est l'instrument. Enfin, la stabilité référentielle repose sur le fait que le texte se situe soit dans la réalité soit dans la fiction, jugement nécessairement porté par le lecteur.

Les propriétés ici visées sont toutes extrinsèques au texte. Plane (2006) souligne alors combien la subjectivité du lecteur intervient dans le jugement porté sur un texte. Toutefois, dans cette perspective discursive, Plane (2006) complète les propriétés caractérisant les textes par trois types de propriétés, qu'elle qualifie de propriétés transitionnelles. Il s'agit de 1) la signifiante, 2) l'informativité et 3) la connexité. La signifiante désigne le fait que le texte permet au lecteur de se représenter un référent. L'informativité repose sur le fait que le texte donne des informations qui se cumulent et se complètent, pour aboutir à un ensemble plus

informatif que la simple somme de chacune des informations. En fin, la connexité est envisagée car des marques de segmentation et de jonction guident l'activité cognitive du lecteur.

Pour compléter cet ensemble de dimensions nécessaire à la qualification textuelle, Plane (2006) admet deux propriétés intrinsèques au texte respectivement 1) la cohésion et la topicalité. La cohésion fait que le sens transite au-delà des segments qui le portent, la topicalité, dans ce cas, fait que le texte se donne un objet qu'il construit.

En suivant ces deux perspectives théoriques, qui s'avèrent quasi complémentaires, un texte est un objet structuré, d'une part grâce à des relations internes et d'autre part grâce à l'activation de schémas cognitifs de connaissance autant chez le rédacteur que chez le lecteur. Le texte suit ou subit une double orientation : intrinsèque vs extrinsèque. Au plan des connaissances extrinsèques, la production de texte requiert un ensemble d'opérations, non linguistiques, comme récupérer, analyser, sélectionner les informations pertinentes et prendre en compte le contexte énonciatif, qui mettent à contribution autant le scripteur (connaissances récupérées sur le monde) que le lecteur (orientation donnée à ces connaissances). La production d'un texte présuppose sa compréhension. Le traitement d'un texte doit donc nécessairement analyser l'interaction entre les caractéristiques du texte et celles du contexte, comprenant celles du scripteur et celles du lecteur.

4. L'entrée de la psychologie cognitive

L'intérêt d'une approche du texte, en tant qu'entité linguistique unique, achoppe sur une impasse si la déclinaison des différents éléments indiquant la qualité se convertit directement en mesure de qualité au niveau du texte. Effectivement, si la décomposition opère au point que le texte ne soit plus perçu que par la compilation d'autres mesures, il ne s'agit plus vraiment de la qualité du texte, entendu comme aboutissement du processus d'écriture, qu'il s'agisse d'un écrivain ou d'un élève. Les travaux élaborés en psychologie cognitive présentent le texte comme faisant partie intégrante du processus de rédaction de texte. Rédiger un texte est une activité de résolution de problèmes, associant la mise en œuvre d'un ensemble de connaissances langagières au sein de différents processus mentaux (voir Coirier, Gaonac'h & Passerault, 1996, Fayol, 1997, Alamargot & Chanquoy, 2001, 2002, par exemple). Si rédiger revient à traduire une représentation référentielle et personnelle sous forme linguistique, le problème de l'organisation des connaissances et des processus mentaux associés se retrouve au centre. Formaliser l'organisation des connaissances langagières et des processus mentaux à l'œuvre dans la rédaction s'est traduit historiquement par l'élaboration (et la révision) de plusieurs modélisations théoriques visant à rendre compte de l'activité du rédacteur. Les différents modèles mettent alors en perspectives l'activité du scripteur expert et celui du novice. Ces modèles abordent ainsi l'activité rédactionnelle de l'expert, en envisageant les étapes de développement de l'écrit. Ils décrivent les processus rédactionnels et l'activité rédactionnelle dans le cadre d'un système cognitif de traitement de l'information.

Les connaissances et les processus mentaux associés cherchent à rendre compte de la complexité structurelle de l'activité de rédaction du point de vue du rédacteur. Cette complexité explique la difficulté du chercheur à tenter de contrôler ou de manipuler les différentes composantes en même temps. C'est d'ailleurs la raison pour laquelle les différents modèles ont en premier lieu défini des composantes *a priori* afin de pouvoir les étudier (Alamargot & Chanquoy, 2001, 2002). Les recherches ont ainsi progressé pas à pas. Les modèles théoriques proposés au départ étaient descriptifs, schématiques et globaux ; ils intégraient un ensemble d'opérations cognitives lié à l'activité de production. Au plan méthodologique, outre l'étude du produit fini, des études menées en temps réel ont exploité les protocoles verbaux de participants qui expliquaient à voix haute ce qu'ils pensaient être en train de faire, alors même qu'ils rédigeaient. Ces études ont dégagé des éléments importants : c'est sur cette base, en croisant l'analyse du produit fini et la prise d'informations sur le temps réel de l'activité rédactionnelle, que les auteurs se sont accordés sur un découpage en deux composantes centrales (Alamargot & Chanquoy, 2002). Sont distingués et reliés l'environnement physique et l'environnement mental du scripteur. L'environnement physique du rédacteur est défini par les conditions dans lesquelles la tâche d'écriture se déroule. Ces conditions recouvrent la situation de rédaction ; dans le cas de la classe, le dispositif pédagogique, incluant les consignes, les aides permettant au rédacteur d'écrire comme les documents, dictionnaires, etc. L'environnement mental du rédacteur comprend les connaissances stockées en mémoire à long terme, mais aussi les processus cognitifs qui opèrent sur ces connaissances.

Les premières modélisations se sont intéressées à l'activité de l'adulte-expert. Les modélisations ultérieures, considérant les travaux antérieurs concernant l'adulte, se sont davantage centrées sur l'acquisition et le développement de l'activité de rédaction d'enfants ou de jeunes adultes(1). Ces modèles restent, en dépit de leur ambition à rendre compte d'une complexité, essentiellement descriptifs : ils ne permettent alors d'appréhender la production verbale écrite, que comme une activité analysable soit 1) en composantes (vocabulaire, grammaire, orientation, intérêt), soit 2) envisageable comme un tout. Si l'on prend le problème en décomposant un texte en ses composants fonctionnels, on retombe sur les caractéristiques linguistiques connues. Aussi, nous prendrons le chemin inverse, en nous interrogeant d'abord sur la manière de pouvoir considérer un texte comme une totalité linguistique.

5. De l'expert au novice : comment se situer ?

Aussi trivial que cela puisse paraître, la tâche d'évaluation de texte par un enseignant pose problème : l'enseignant, adulte, donc expert, doit déterminer si un novice est expert mais à son niveau d'élève, ou potentiellement expert quand il sera adulte. Comment juger de la qualité des textes produits ? Le corpus grenouille représente dans ce cadre un espace d'exploitation très intéressant : des textes sont produits par des élèves de la classe de CE2 primaire à celle de 4^{ème} collège.

6. Quatre pôles de détermination pour la qualification textuelle

Sur la base des éléments théoriques rapportés plus haut, nous avons déjà, pour déterminer la valeur de textes écrits par des élèves, synthétisé l'apport des auteurs sur cette question (Slusarczyk, 2005, 2010). Il convient, selon nous, pour pouvoir rendre compte d'une qualification des textes écrits par des élèves, de distinguer quatre grands pôles orientant chacun la caractérisation du processus de production et de compréhension de texte. Il s'agit 1) des caractéristiques du système de la langue, 2) des modalités de traitement du domaine de référence, 3) du contexte énonciatif et 4) des structures cognitives en jeu.

Au plan des caractéristiques du système de la langue, il s'agit de repérer dans les textes des indices de construction de la structuration textuelle, généralement caractérisée par la présence des connecteurs, et d'indicateur de ponctuation dont la matérialisation des paragraphes. Pour le traitement du domaine de référence, une organisation préalable du domaine de référence existe chez le lecteur et le scripteur, grâce à l'activation des représentations mentales, des schémas cognitifs pour décrire le monde, qui supposent l'exploitation de schémas cognitifs pour analyser le domaine de référence. A ce niveau, se trouvent utilisés des outils cognitifs d'organisation causale, logique et temporelle du monde. En fait, le contexte énonciatif met en adéquation la structuration linguistique du texte avec sa finalité communicative. Les outils sont essentiellement marqués par le jeu sur les personnes, la concordance du temps et des lieux, telle qu'elle s'effectue au niveau de la génération de la syntaxe. C'est le dernier espace qui nous intéresse le plus. Les structures cognitives en jeu sont nécessairement mises en correspondance avec le niveau d'expertise linguistique et textuelle du locuteur. Son degré d'acquisition de lecture-écriture, sa familiarité avec les modèles discursifs, sa capacité à traiter de l'information symbolique, ses ressources attentionnelles, enfin les contraintes liées au stade de développement de sa mémoire importent. Or, l'enseignant, lorsqu'il doit évaluer un texte est nécessairement soumis à un dilemme : construit-il la qualité à son niveau d'expertise ou construit-il un jugement en tenant compte des possibilités liées au niveau de sa classe d'élèves ? Comment savoir comment s'y prennent les enseignants ? Travailler à partir d'un échantillon de textes scolaires paraît intéressant.

7. L'appui sur des jugements d'experts

Spencer & Fitzgerald (1993) ont comparé des mesures de degré de qualification de textes effectuées à partir de productions de texte d'élèves de 6^{ème} primaire. La mesure du degré de qualification des textes était comparée au sein de trois grands domaines généralement associés au niveau textuel. Ces chercheurs ont considéré l'évaluation de la structure, celle de la cohérence et celle de la qualité. Pour qualifier la structure textuelle, la littérature unifiée dans ce domaine définit les éléments constitutifs d'un texte en termes de : lieu, événement, problème, action, résolution ou chute et conclusion. L'évaluation dans ce domaine est consensuelle. Pour appréhender la cohérence textuelle, il n'en va pas de même. Dans une revue de lecture, les auteurs ont dénombré pas moins de 27 définitions différentes pour la notion de cohérence. Ils s'appuient alors sur Bamberg & Hasan (1984) pour présenter la cohérence

comme une mesure au niveau de l'unité du texte, repérable sous 1) l'absence de digressions inutiles, 2) l'identification claire des différents points du texte, et 3) un contexte et des répliques liant le discours. Enfin la qualité textuelle est définie pour être la plus générale possible. Elle se décline au mérite du texte en termes de clarté, de développement de l'idée principale, de suspens, et de construction (architecture d'ensemble) correcte.

Sur chacun de ces trois domaines, deux types de mesures sont explorés. Deux groupes de correcteurs différents opèrent à partir des mêmes productions d'élèves. Une première mesure, conduite par l'un des groupes de correcteur, est basée sur le fait que l'on considère qu'il y a interaction entre le lecteur et le texte. Pour cet aspect, qualifié de constructiviste par Spencer, le point de vue du lecteur, pris en compte, est considérable conformément à la perspective discursive proposée par Plane (2006) : la subjectivité du correcteur intervient dans le jugement posé. Dans l'autre mesure, conduite par le second groupe de correcteur, le texte produit est considéré comme ayant cristallisé la pensée de l'auteur ; dans ce cas, la structure, la cohérence et la qualité sont dans le texte lui-même. Le lecteur n'a qu'à vérifier la présence ou l'absence de chacun des éléments les caractérisant. Par exemple, la cohérence peut être mesurée à partir de l'organisation des relations entre prédicats successifs, assurés notamment par le système des temps verbaux.

Les résultats de l'étude de Spencer & Fitzgerald (1993) montrent des liens positifs, mitigés ou absents entre les mesures selon le domaine considéré : structure, cohérence et qualité. Pour la structure, il existe une corrélation positive entre les jugements intrinsèque (texte) et extrinsèque (lien texte-lecteur). Cette corrélation explique très probablement en quoi la notion de structure fait aujourd'hui partie intégrante de la culture des lecteurs-correcteurs : les professeurs de français s'y appuient régulièrement. En revanche, les résultats n'admettent aucune corrélation entre les deux mesures au niveau de la cohérence. Les résultats sont mitigés pour le domaine de la qualité.

Au travers cet exemple, on voit à quel point les référents théoriques d'un domaine peuvent plus ou moins infuser dans le domaine des praticiens, avec des conséquences plus ou moins importantes. Pour le domaine de la structure, l'adéquation entre un relevé exhaustif d'éléments caractéristiques du texte et le jugement constructiviste du correcteur ne porte aucun préjudice à la compétence mesurée pour déterminer la qualité du processus de production de texte des élèves (Spencer & Fitzgerald, 1993). En revanche, quand cette adéquation n'est pas vérifiée (cohérence) ou peu évidente (qualité, cf. Spencer & Fitzgerald, 1993), au plan de la recherche scientifique, l'idée même que l'on peut se faire de la cohérence ou de la qualité d'un texte n'active pas plus de consensus dans la culture des praticiens. Les enseignants divergent vis-à-vis de ces notions autant que différent les théoriciens du domaine. Au-delà du jugement sur la structure, qui est consensuel, faut-il pour autant retomber dans une mesure des parties composant le texte pour juger le texte ? Doit-on, lorsque l'on évalue un texte, abandonner l'idée de son entité linguistique ? C'est ce que nous avons tenté de mettre en évidence sur le corpus grenouille, selon une méthodologie déjà rôdée.

8. Le corpus grenouille

Nous avons exploité, dans le corpus d'ensemble, les seconds jets de récits des élèves de CE2, 6^{ème} et 4^{ème}, soit neuf textes. S'appuyant sur les travaux de Spencer & Fitzgerald (1993), nous avons exploré les liens entre deux types de mesures pour approcher un degré de qualification des textes. L'une, basée sur le fait que l'on considère qu'il y a interaction entre le lecteur et le texte. L'autre, basée sur le fait que le texte produit est considéré comme ayant cristallisé la pensée de l'auteur.

Concernant le premier type de mesure, le degré de qualification des textes correspond à une évaluation faite par des enseignants, qui considérés comme experts, pratiquent un jugement tel qu'il pourrait être pratiqué dans une classe. Nous avons cependant formé les enseignants à l'utilisation d'une grille de correction permettant une harmonisation en rendant l'évaluation du texte comparable d'un correcteur à l'autre. Nous avons élaboré cette grille à partir de the holistic scoring method de Shell et al. (1989). Celle-ci est basée sur la méthode décrite par Cooper (1985, voir Rankin, Bruning, Timme & Katkanant, 1993). Quatre enseignants ont cinq points à évaluer : 1) la créativité, 2) la cohérence et l'organisation du texte, 3) la densité du texte (nombre d'idées différentes), 4) la clarté du texte et la qualité du vocabulaire et 5) la syntaxe pour les neuf textes de récits. Les enseignants disposaient des copies manuscrites photocopées, la classe de l'élève était indiquée. Le jugement porté sur ces différents points est subjectif ; une échelle de satisfaction en 20 points était proposée pour graduer le jugement. Le point de vue de l'enseignant lecteur est au centre du jugement produit sur le texte. Il rend compte de l'interaction texte-lecteur.

La seconde mesure est effectuée par un chercheur. Elle vise à quantifier objectivement deux types d'éléments. Nous avons réutilisés des éléments repérables et quantifiables dans le test standardisé de Khomsi, Nanty, Pasquet, & Parbeau-Guéno (2005) destiné à évaluer des compétences scolaires des élèves au plan de la maîtrise de la langue. Dans le test, l'objectivité du jugement est assurée grâce à l'étude préalable des images soumis aux élèves pour construire un récit. Les images renferment des contenus factuels (on voit -ou non- un garçon) et des éléments inférés (on déduit au passage d'une image à l'autre ce que le garçon a fait).

8.1 L'environnement physique : analyse du support de récit du corpus grenouille

Les six images imposées aux élèves pour le corpus grenouille présentent l'avantage d'une analyse préalable équivalente. Nous avons étudié le contenu des images et repéré pour chaque image des contenus factuels (exemple : *on voit le petit garçon en train de manger*) et des contenus inférés (exemple : *on doit deviner que les parents sont fâchés*). Le texte des élèves est jugé en fonction d'une pondération entre contenus factuels et inférés, comme dans le test de Khomsi & al. Nous avons défini neuf items factuels *a priori*. Ces éléments reposent sur trois types d'éléments, les personnages présents, les lieux et les actions successives des personnages. Ainsi, 1/Personnages : (au moins deux) enfant, grenouille, 2/une action, un premier lieu : l'enfant se change devant la

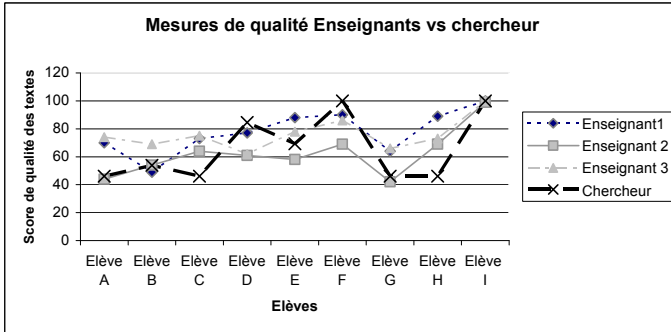
glace, dans sa chambre... 3/un deuxième lieu : au restaurant avec ses parents, 4/une deuxième action : la grenouille saute dans le trombone, 5/une troisième action : la grenouille saute dans la salade, 6/une quatrième action : la grenouille saute dans le verre d'un client du restaurant, 7/une cinquième action : le serveur se débarrasse de la grenouille, 8/une sixième action : /le garçon récupère sa grenouille, 9/une septième action : la famille rentre à la maison en voiture. Nous avons isolés quatre items inférés *a priori*: 1/le garçon se fait beau pour aller au restaurant (l'évènement est important) ; 2/la grenouille s'était cachée dans la poche de la veste, le garçon l'avait prise en cachette (explication de sa présence) ; 3/la grenouille sème la panique dans le restaurant (conséquence logique) ; 4/ La soirée est gâchée, la sœur et la mère sont en colère (rancœur contre le personnage principal ou conséquence pour l'ensemble de la famille).

8.2 L'environnement mental de l'élève : analyse des récits d'élèves

Dans le test ECL (Khomsi & al. 2005), le nombre d'idées factuelles et inférées est considéré comme indicateur du niveau de compétences scolaires des élèves. Plus l'élève génère d'idées factuelles et d'idées inférées indispensables à la compréhension, par le lecteur de l'histoire, sans le support imagé, meilleur est le récit. Le jugement du chercheur est alors proche de celui demandé aux enseignants car c'est une interaction texte-lecteur qui est là mesurée. Dans une étude précédente Slusarczyk, (2005), étudiant la corrélation entre des mesures de qualité comparées entre enseignants-experts et chercheur, obtenait une corrélation significative. Preuve que le jugement de qualité globale d'un texte, composée sur la base de multiples critères était identique à la mesure composée d'une étude du texte portant uniquement sur la génération d'idées (factuelle et inférées). Le processus de reprise d'idées factuelles, complétés de génération d'inférences indispensables au lecteur est-elle au cœur de l'expertise rédactionnelle ? Si un élève génère suffisamment d'idées, est-il de facto capable de produire un texte jugé de qualité ? L'étude précédente (Slusarczyk, 2005) nous amenait à constater que la mesure de qualité par un enseignant expert intégrant une définition la plus générale possible est corrélée à une mesure du nombre attendu d'idées générées en conformité avec un listing d'idées établi *a priori*. Si tel est le cas dans le corpus grenouille, cela signifierait que vocabulaire, grammaire, syntaxe, cohérence, structure sont des éléments dépendants finalement d'un indicateur central.

Le corpus grenouille ne permet pas d'établir une corrélation entre les deux mesures, eu égard au peu de copies exploitées. Nous mettons seulement en évidence (voir figure 1) les points de recouvrements entre les jugements du chercheur et ceux des trois enseignants. La proximité des courbes va dans le sens des corrélations établies dans notre étude antérieure : il n'y a pas d'écart manifeste entre les deux formes de jugement.

Mesure du degré de qualification des textes du corpus Grenouille Enseignants vs Chercheur



Légende : A, B, C (CE2), D, E, F (6^{ème}), G, H, I (4^{ème}) respectivement placés selon leur niveau : faible, moyen et fort.

Toutefois, on remarque, sur ce graphique, que les évaluations du chercheur atteignent pour deux élèves les 100% : ce sera rarement le cas en classe. On note, parmi ces scores impressionnants, que l'élève de 4^{ème}, I, remporte un score de 100% avec un accord inter-juges de 100% ! Le texte atteint la qualité maximale. On pourrait penser que les enseignants n'arriveraient pas à une saturation de leur notation : pourtant un bon élève de 4^{ème} parvient à rédiger un récit représentant un niveau d'expert. Une phrase extraite, pour illustrer ce fait, du récit de cet élève est éloquente : « Peter, bien qu'il soit jeune, se fit plus beau que jamais devant son miroir ». Le héros est nommé d'un nom anglais pour traduire la distinction du personnage; la syntaxe est travaillée ; le passé simple est maîtrisé ; le lexique choisi est recherché. La phrase finale, « le trajet du retour se passa dans un silence pesant, qui ne présageait rien de bon pour Peter et sa grenouille », manifeste des mécanismes inférentiels non seulement présents mais associés à une maîtrise du langage autant que de la langue : l'élève manie fort bien l'implicite. Pareille copie est exemplaire...

Deux scores attribués par le chercheur, aux élèves G et H, respectivement jugés faible et moyen en classe, sont situés sur des niveaux qui correspondent à ceux d'élèves de CE2, et sont mieux jugés par les enseignants. Ces élèves, G et H, n'ont en fait actualisé dans leur texte aucune information inférentielle et réalisé un faible score (6/9) sur les informations factuelles. Ils sont ainsi assimilables à ce que peut produire un élève de CE2 : rappeler des informations factuelles correspondant aux images les plus faciles à commenter, ce qui correspond au score attribué par le chercheur aux trois copies de CE2. De fait l'élève de CE2 qui a produit une information inférée est un bon élève, qui a conclut son texte en envisageant les conséquences de la soirée. Les enseignants, face à ces copies, G et H, d'élèves de 4^{ème} collège, sont plus indulgents et donnent des avis qui reflètent pratiquement la hiérarchie des niveaux scolaires (non fournis) : comment ont-ils examiné ces copies ? On voit clairement, à ce propos, poindre des tendances d'évaluation reflétant trois degrés d'indulgence croissant entre les enseignants 2, 3 et 1 qui donnent une idée de la difficulté du jugement, et interroge sur la perspective des jugements de textes dans le cadre scolaire.

9. Les jugements sur la copie de l'élève moyen de 4^{ème}

La copie de l'élève moyen a fait l'objet d'un jugement par l'enseignant le moins indulgent, 2, d'une sensibilité au fait que le texte manquait d'organisation et de densité : ce sont les deux catégories qui sont moins bien notées que les autres. Seul cet enseignant paraît ainsi traduire dans l'attribution de ses notes par catégories fournies une forme d'adéquation à ce que révèle le test du chercheur, et être ainsi sensible au fait que le texte n'est pas complet, dense et mal organisé : c'est le score d'organisation qui est le plus faiblement jugé parmi l'ensemble des catégories à renseigner, en atteignant cependant une note de 12/20. Les deux autres enseignants ont évalué cette catégorie respectivement à 15/20 ou 18/20. Le texte de l'élève, telle en témoigne l'une des phrases extraite, est dense en informations selon un mode d'organisation des idées sous forme d'empilement, « Il avait tout d'abord un chien qu'il avait eu en cadeau quand il était petit, une tortue et il y a quelques temps il recueillit une grenouille qu'il appréciait beaucoup et la grenouille appréciait aussi le jeune homme », qui traduirait ce que les psychologues nomment une planification au pas à pas, plus qu'une organisation d'ensemble. Toutefois, les enseignants repèrent des éléments d'organisation internes aux paragraphes qui leur indiquent que le rédacteur dispose d'une compréhension de la situation. L'enseignant médian est toutefois alerté pour sa part par la syntaxe de cet élève : il la juge avec un écart de 3 à 4 points de moins que les autres items. Deux phrases composées comme la suivante « Mais la grenouille n'avait pas fini de se promener et sauta directement sur un plateau où (ou) était disposé(es) une assiette de salade et (est) un verre rempli d'eau » ont sans doute conduit à ce jugement.

9.1 Les jugements sur la copie de l'élève faible de 4^{ème}

La copie de l'élève faible est évaluée par les enseignants d'une manière homogène quant au niveau syntaxique ; les trois enseignants ont attribué un score inférieur à la moyenne identique de 8/20. L'enseignant le moins indulgent juge l'ensemble du texte entre 8 et 9 sur toutes les catégories. Les deux enseignants les plus indulgents en revanche attribuent leur plus forte note à la catégorie d'organisation de texte de 15/20. La copie de cet élève faible met bien en évidence la difficulté d'évaluer des écrits scolaires. Le jugement est soit bas soit déséquilibré. Le chercheur, quant à lui, repère dans son test un niveau faible touchant à des mécanismes qui se situent en deçà de la performance verbale et ne tiennent pas compte d'aspects positifs présents dans le récit. Tous les enseignants sont entravés dans la lecture par une syntaxe qui déroute ; voici une phrase extraite de la copie, même s'il est délicat d'illustrer sur un seul extrait ce qui paraît évident à la lecture d'ensemble : « Arrivé au restaurant le jeune garçon regardait le menu et pendant ce temps la grenouille en profita pour sortir de la poche du jeune garçon en sautant allant se propager dans l'orchestre qui jouait dans les instruments 'tambours' ». De fait, avec peu de retouches, quelques virgules, le texte parviendrait très vite à un seuil de cohérence d'ensemble. Le vocabulaire employé présentant un double registre (se propager, parler au responsable, sceptique / sa famille, parents énervés, récupérer) manifeste une volonté d'adaptation au lecteur « le jeune garçon rentra chez lui avec sa grenouille, mais ses parents ainsi que sa sœur n'étaient guère contents, plutôt sceptiques » sans toutefois atteindre le but escompté.

Le corpus Grenouille, à travers cette comparaison des jugements, illustre que les enseignants, au sein d'évaluation corrélée au test du chercheur, prouvent leur sensibilité à des éléments importants dans les copies d'élèves. Ces jugements seront-ils utiles aux élèves ? Nous pensons qu'ils pourraient l'être à la condition de manifester aux élèves le double point de vue suivant : leurs idées sont bonnes et on peut les développer sur leur propre texte (inférences, curiosité, vocabulaire, changement de tournures). Dans ce cadre l'enseignement de la syntaxe ne doit ni obérer la volonté d'écrire des élèves ni trahir une absence d'ambition pour les plus faibles. Il faut faire passer le rédacteur au statut de lecteur critique de sa propre écriture.

10. Conclusion

Binet, en plaisantant, définissait l'intelligence comme étant ce que mesurait son test ! Mais l'instrument de mesure ne se confond pas simplement avec l'objet qu'il mesure : quelle que soit la perspective choisie pour évaluer un texte, celle-ci ne donnera qu'un éclairage partiel de la complexité portée par celui-ci. La prudence, dans l'évaluation, doit donc être de mise. Évaluer un texte, sur le seul repérage des contenus factuels et inférés comme indicateurs de qualité, permet probablement une économie évaluative du lecteur « évaluateur » préoccupé par la recherche d'indicateurs à la fois simples et informatifs : là n'est pas l'essentiel. Ce qui est particulièrement intéressant ici, c'est que la comparaison entre la compréhension globale d'un récit au travers de son contenu factuel et inféré et le jugement de qualité globale portant sur sa structure, sa cohérence, sa syntaxe et son vocabulaire, convoque le lien entre la compréhension du rédacteur et le produit de sa compréhension pour autrui (le texte). L'enseignant doit-il chercher à évaluer le produit ou doit-il chercher à évaluer le niveau de compréhension du rédacteur porté par le texte ? Quoiqu'il en soit, la compréhension est toujours implicitement évaluée. Poser la question de l'évaluation de la compréhension débouche alors sur une question fondamentale pour enseigner la rédaction. Comment faut-il apprendre aux élèves à rédiger ? Faut-il leur apprendre à écrire pour autrui ou faut-il leur apprendre à réfléchir ?

Notes

¹ Alamargot & Chanquoy (2002) donnent un aperçu précis de l'apport des différentes modélisations auxquelles se réfère un ensemble de travaux très importants : Hayes et Flower (1980) ; Hayes (1996) ; Bereiter et Scardamalia (1987) ; Berninger et Swanson (1994) et Berninger & al. (2003, 2006).

Références

- Alamargot, D. & Chanquoy, L. 2001. *Through the Models of Writing*, Dordrecht, Boston, London : Kluwer Academic Publishers.
- Alamargot, D. & Chanquoy, L. 2002. Les modèles de rédaction de texte. In M. Fayol (Ed.), *Production du langage* (pp.45-65). Paris : Hermès Lavoisier.
- Bereiter, C., & Scardamalia, M., 1987. *The psychology of written composition*. Hillsdale, NJ: Erlbaum.

- Berninger, W., Fuller, F., & Whitaker, D. 1996. A Process Model of Writing Development Across the Life Span, *Educational Psychology Review*, 8(3), 193-218.
- Chanquoy, L., & Alamargot, D. 2002. Mémoire de travail et rédaction de textes: Evolution des modèles et bilan des premiers travaux. *L'Année Psychologique*, 102, 363-398.
- Coirier, P., Gaonac'h, D. & Passerault, J.M. 1996. *Psycholinguistique textuelle, approche cognitive de la compréhension et la production des textes*, Paris : Armand Colin/Masson.
- Fayol, M. (1997). *Des idées au texte. Psychologie cognitive de la production verbale, orale et écrite*. Paris : PUF.
- Hayes, J.R. & Flower, L.S. 1980. Identifying the organisation of writing processes. Dans L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* Hillsdale, NJ: Lawrence Erlbaum, 3-30.
- Khomsi, A., Nanty, I., Pasquet, F., Parbeau-Guéno, A. 2005. *Evaluation des compétences linguistiques au collège*. Paris : E.C.P.A.
- Mac Cutchen, D. 2006. Cognitive factors in the development of children's writing. Dans C.A. Mac Arthur (ed.), S. Graham & J. Fitzgerald, *Handbook of writing research* (pp. 115-130). New York: The Guilford Press.
- Plane, S. 2006. Singularités et constantes de la production d'écrit, l'écriture comme traitement de contraintes. In J. Laffont-Terranova & D. Colin (Eds.), *Didactique de l'écrit, La construction des savoirs et le sujet-écrivain* (pp.33-54). Presses Universitaires de Namur.
- Rankin, J.L, Bruning, R.H., Timme, V.L. & Katkanant, C. 1993. Is writing affected by spelling performance and beliefs about spelling? *Applied cognitive psychology* . vol. 7, no2, pp. 155-169.
- Slusarczyk, B. 2010. *L'orthographe dans la production écrite du CE2 à l'entrée au collège*. Thèse de Doctorat en Sciences de l'Education, Grenoble : Université Pierre Mendès France.
- Slusarczyk, B., Bressoux, P. & Fayol, M. 2006. *Les performances en production orthographique influent-elles sur la production verbale écrite ?* Congrès scientifique Approche Cognitive de l'apprentissage de la langue écrite, Rennes 19-20 octobre 2006.
- Spencer, S.L., & Fitzgerald, J. 1993. Validity and Structure, Coherence, and Quality Measures in Writing. *Journal of Reading Behavior*, 25(2), 209-231.

Annexe

Grille distribuée aux enseignants

Code	Catégorie	Note /20
real	<u>Le texte produit est un récit</u> L'élève a fait preuve de créativité. Le texte est original.	/20
orga	<u>Le texte est organisé</u> La consigne est respectée. Le texte est cohérent. Les informations progressent dans la logique du texte. Il y a une introduction. Il y a une conclusion.	/20
dens	<u>Le texte est dense</u> Il comporte des idées distinctes en nombre.	/20
clari	<u>Le texte est riche et clair</u> Le vocabulaire utilisé est riche et varié. Le vocabulaire est adapté Le vocabulaire contribue à l'organisation du texte	/20
lang	<u>La syntaxe est correcte</u> Les phrases utilisées sont syntaxiquement correctes.	/20
qual	<u>Qualité globale du texte</u>	/100