

Atelier du réseau EVADIDA

Session 1

Atelier A.EVADIDA.1

L'évaluation de la compréhension de l'écrit en français : comment estimer la difficulté des items et des textes proposés aux élèves ? (7393)

*Murielle Roth, Verónica Sánchez Abchi & Jean-François de Pietro
Institut de recherche et de documentation pédagogique, Suisse*

Mots-clés : évaluation de la compréhension de l'écrit, difficulté des items, difficulté des textes

Dans le processus d'évaluation de la compréhension de l'écrit à l'école, il suffit de choisir un « mauvais » texte – un texte trop complexe ou trop simple pour le niveau de connaissance des élèves, un texte avec une structure trop peu conventionnelle... – pour dire qu'un élève comprend beaucoup, un peu, ou rien du tout du texte. De même, les questions qui sont adressées aux élèves peuvent être de complexité variable en fonction notamment de leur formulation, des opérations cognitives que l'élève doit réaliser – repérer, analyser, interpréter... – ou encore du format de questionnement (questions à choix multiples, questions ouvertes demandant une activité d'écriture, etc.). Aux difficultés qui concernent la tâche en soi s'ajoutent également des aspects propres à l'élève comme sa motivation ou ses connaissances préalables qui vont modifier son approche de l'écrit (McNamara, Kintsch, Songer, & Kintsch, 1996). Dans cette contribution, liée à des travaux réalisés pour la Conférence intercantonale de l'instruction publique de la Suisse romande et du Tessin (CIIP), l'accent n'est pas mis sur les aspects relevant de l'élève mais sur les tâches mesurant la compréhension de l'écrit en français langue première, dans le milieu scolaire, ainsi que sur l'estimation de la difficulté des textes à comprendre.

Nous commençons par définir, d'un point de vue méthodologique, ce que nous entendons par « item ». Ensuite, nous présentons une procédure qui permet d'estimer la difficulté des textes et des items, illustrée à partir de tâches d'évaluation de la compréhension de l'écrit provenant d'évaluations ayant pour but d'évaluer les acquisitions des élèves à l'échelle de l'un ou l'autre canton suisse romand pour la 8e année de scolarité (élèves âgés de 11-12 ans).

Pour évaluer les textes, la procédure comprend six critères repris parmi ceux que nous avons rencontrés dans nos lectures et réinterprétés sur la base de nos analyses (McNamara, Graesser, & Louwerse, 2012 ; François, 2009 ; Lafontaine, 2003 ; Kandel & Moles, 1958 ; etc.) : 1) complexité syntaxique, 2) complexité lexicale (Mesnager & Bres, 2008), 3) cohésion textuelle, 4) structure textuelle, 5) familiarité du contenu traité et 6) divers autres éléments du texte qui peuvent fournir des ressources facilitant la lecture et la compréhension (exemples, comparaisons, analogies, reformulations, explications, humour, etc.) ou, au contraire, la rendre plus difficile.

Cependant, même si le choix du texte est essentiel pour évaluer la compréhension, il n'est qu'un facteur parmi d'autres qui interviennent dans la construction des items. Ainsi, pour en estimer la difficulté, il s'avère également nécessaire de prendre en compte la difficulté des questions. Pour ce faire deux critères sont pris en considération : a) le contenu – les objets langagiers concernés et les opérations cognitives que l'élève est censé réaliser –, et b) l'« enveloppe » – la formulation de la consigne, le format de questionnement et ce qu'il demande en plus comme travail pour répondre (questions à choix

multiples, questions ouvertes demandant une activité d'écriture, etc.) ainsi que le matériel complémentaire – les aides fournies pour réussir la tâche telles que les dictionnaires, tableaux de conjugaison, etc.

Pour caractériser l'item (le texte et les questions associées), notre procédure retient au final trois composantes incluant elles-mêmes plusieurs facteurs : le « contenu », l'« enveloppe » et le « matériel langagier » (texte ou corpus) auquel il est attaché (Sánchez Abchi, De Pietro & Roth, 2016). L'ensemble de ces facteurs définit en quelque sorte la manière dont la compréhension du texte est appréhendée et permet de mieux saisir les difficultés potentielles que les élèves vont rencontrer.

Afin de prendre en compte et intégrer ces trois composantes, nous avons alors élaboré un outil qui permet qui permet d'estimer et/ou de calculer la difficulté de chacun des facteurs qu'elles contiennent – sur une échelle comportant trois valeurs : -1 (= facile), 0 (= intermédiaire) et 1 (= difficile) – et d'établir, ensuite, de manière séparée pour chaque composante (contenu, enveloppe et matériel langagier) une moyenne des valeurs que nous convertissons en un indice de difficulté qualitatif pour la composante concernée : ICont, IEnv, IMat. Ainsi, l'indice de difficulté de l'item se présente non pas comme un indice unique mais comme la juxtaposition de trois indices qui devraient pouvoir guider les « concepteur-trices de tests » dans l'analyse et l'estimation de la difficulté des items et des textes, et qui les aident à les interpréter, les élaborer ou les choisir.

Par ailleurs, nous allons évaluer notre procédure en comparant nos estimations à priori de la difficulté des items avec les taux de réussite effectivement obtenus par les élèves à chaque item, dans le but de voir si la difficulté prédite se retrouve dans les résultats des élèves.

Notre présentation se conclut par une mise en évidence des enjeux qui sont sous-jacents à ce travail et qui renvoient à la validité didactique de l'évaluation. A savoir que pour s'assurer d'évaluer ce que nous voulons évaluer, il est nécessaire de bien comprendre, décrire et définir les « outils » employés – soit, en l'occurrence, les questions proposées et les textes utilisés – et de tenir compte de leur difficulté.

Références bibliographiques

- François, T. (2009). Modèles statistiques pour l'estimation automatique de la difficulté de textes de FLE. In Actes de la 11e Rencontre des étudiants chercheurs en informatique pour le traitement automatique des langues (RECITAL), Senlis, 24-26 juin 2009 (pp. 1-10). Paris : Université, Laboratoire d'informatique de Paris-Nord.
- Kandel, L. & Moles, A. (1958). Application de l'indice de Flesch à la langue française. Cahiers études de radio-télévision, 19, 253-274.
- Lafontaine, D. (2003). Comment faciliter, développer et évaluer la compréhension des textes aux différentes étapes de la scolarité primaire (document envoyé au PIREF en vue de la Conférence de consensus sur l'enseignement de la lecture à l'école primaire, les 4 et 5 décembre 2003). Liège : Université.
- Mesnager, J. & Bres, S. (2008). Évaluer la difficulté des textes : cycle 2-3, 6e-5e (CD-Rom). Paris : Nathan.
- McNamara, D.S., Kintsch, E., Songer, N.B. & Kintsch, W. (1996). Are good texts always better ? : interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and instruction*, 14(1), 1-43.
- McNamara, D.S., Graesser, A.C. & Louwerse, M.M. (2012). Sources of text difficulty : across genres and grades. In J.P. Sabatini, E. Albro & T. O'Reilly (Eds), *Measuring up : advances in how we assess reading ability* (pp. 89-116). Lanham : Rowman and Littlefield Publishers.
- Sánchez Abchi, V., De Pietro, J. F., & Roth, M. (2016). Évaluer en français. Neuchâtel : Institut de recherche et de documentation pédagogique. (IRDP).