



HAL
open science

Correction automatique d'écrits d'élèves : rêve ou réalité ?

Sonia Mandin

► **To cite this version:**

| Sonia Mandin. Correction automatique d'écrits d'élèves : rêve ou réalité ?. 2013. hal-04422245

HAL Id: hal-04422245

<https://hal.science/hal-04422245v1>

Submitted on 30 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

L'Agence des usages

Intégrer le numérique dans sa pratique pédagogique



Correction automatique d'écrits d'élèves : rêve ou réalité ?

La correction automatique est aujourd'hui possible grâce à des logiciels d'analyse de textes. Cette correction peut porter sur la qualité générale du texte ou sur certaines caractéristiques uniquement. Une efficacité des logiciels de correction automatique a été démontrée dans six domaines d'application : la correction des réponses à des questions ouvertes, la correction d'essais (rédactions, résumés), la détection du plagiat, l'orthographe et la détection du niveau de difficulté d'un texte.

Quel enseignant n'a pas rêvé de simplement appuyer sur un bouton pour obtenir les évaluations de ses copies ? Quel élève n'a pas rêvé d'en faire autant pour obtenir une correction immédiate de son travail avant de le soumettre ? Le développement de logiciels de correction automatique de textes et l'évolution des recherches en linguistique et informatique ont rendu possibles de telles actions, du moins dans certains contextes d'application.

Cet article décrit les résultats d'études ayant montré une efficacité des logiciels de correction dans six contextes : la correction des réponses à des questions ouvertes, la correction d'essais (rédactions, résumés), la détection du plagiat, l'orthographe et la détection du niveau de difficulté d'un texte. Des pistes d'usage pédagogique sont également proposées aux enseignants.

Réponses à des questions ouvertes

Une étude de 2010, décrite par Mark D. Shermis et Ben Hammer, montre la fiabilité des logiciels d'évaluation automatique dans l'analyse d'écrits (réponses à des questions documentaires et rédactions de textes narratifs, descriptifs ou argumentatifs). Pour cela, ces scientifiques ont utilisé des productions d'élèves de 5^e, de 4^e et de 2nde. Ils les ont réparties aléatoirement dans 3 groupes afin de comparer les notes attribuées par 9 logiciels aux notes attribuées par des évaluateurs humains. Les productions du premier groupe ont servi au calibrage de l'évaluation automatique, celles du second groupe au calcul de leur précision et de leur vitesse, et celles du troisième groupe à la validation des résultats. Pour les deux types d'écrits (réponses à des questions, rédactions), les évaluations automatiques fournissent bien des notes en moyenne équivalentes aux évaluations humaines.

Essais (rédactions, résumés)

L'évaluation automatique permet des retours immédiats. Des logiciels y recourent donc pour seconder les

L'Agence des usages

Intégrer le numérique dans sa pratique pédagogique



enseignants dans leurs corrections ou pour assister les apprenants dans leur travail d'écriture. Deux exemples nous montrent des utilisations possibles de l'évaluation automatique d'essais dans ce dernier cas.

Les chercheurs d'une université taïwanaise, Chi-Fen Emily Chen et Wei-Yuan Eugene Cheng, ont expérimenté un logiciel associant de l'évaluation automatique (évaluation chiffrée sur le maintien du discours autour de la thématique, sur la qualité des idées, sur l'organisation du texte, sur le style et le respect des conventions d'écriture) à d'autres outils destinés à assister le travail d'écriture (éditeur de texte, liste de points à contrôler, thesaurus, etc.). Les chercheurs ont testé trois formes d'utilisation pédagogique du logiciel. Il était ainsi demandé à deux groupes d'étudiants de l'utiliser seulement pour l'élaboration d'un premier brouillon et pour le travail de relecture finale. L'un de ces deux groupes était en outre incité à utiliser davantage les outils d'assistance que l'évaluation automatique. *A contrario*, il était demandé au troisième groupe d'utiliser le logiciel tout au long de la tâche et en particulier la fonction d'évaluation automatique. Les résultats ne sont pas concluants en ce qui concerne la perception par les apprenants de la fiabilité des évaluations et de l'efficacité globale du logiciel à améliorer leur tâche d'écriture. Toutefois, les auteurs expliquent qu'il est mieux perçu quand il s'agit de l'utiliser afin d'aider la construction d'ébauches et d'assister la révision des textes.

Dans une synthèse de la littérature, Sonia Mandin décrit plusieurs logiciels d'entraînement ou d'assistance au résumé, fondés sur de l'évaluation automatique de textes. L'évaluation automatique sert alors à générer un retour lié au choix des idées reportées dans les résumés et à la façon de les présenter. Parmi ces logiciels, l'un a pour objectif d'améliorer la maîtrise du résumé chez des élèves de seconde en proposant à ces derniers 3 tâches : identifier les phrases les plus importantes d'un texte, le résumer et déterminer les stratégies utilisées pour construire chaque phrase ainsi produites. Une évaluation automatique de la première et de la dernière tâche est utilisée pour inciter un questionnement des élèves sur leur travail. L'expérimentation du logiciel mettait en scène un groupe d'élèves qui utilisait le logiciel original et un autre qui utilisait une version modifiée dans laquelle la première et la dernière tâches étaient substituées par d'autres ne pouvant pas influencer sur l'activité de résumer. Après 10 utilisations du logiciel, la même chercheuse, Sonia Mandin, a constaté que les compétences à résumer des élèves s'amélioraient : meilleure reconnaissance des stratégies utilisées et meilleure utilisation de ces stratégies pour résumer les textes aux idées les plus saillantes.

Détection du plagiat

Une autre utilisation fréquente de l'évaluation automatique d'écrits vise la détection du plagiat. à ce propos, Lucie Audet fait état d'un logiciel utilisé en réseau dans les écoles publiques d'Ontario. Le logiciel compare les productions rentrées par l'enseignant à d'autres productions d'élèves ainsi qu'à des documents accessibles sur le Web. Pour l'auteure, ce type de logiciel présente principalement un effet dissuasif et formatif. Il permet aussi d'alléger les corrections et d'être plus équitable dans le traitement des copies. En

L'Agence des usages

Intégrer le numérique dans sa pratique pédagogique



2010, le rapport de la mission parlementaire Fourgous, intitulé « Réussir l'école numérique », souligne qu'un grand nombre d'universités sont déjà équipées de tels logiciels et incite les lycées à en faire de même pour former les élèves à l'exploitation des données trouvées sur Internet. Depuis, des établissements du secondaire en ont fait l'expérience en France.

C'est par exemple le cas dans l'académie de Bordeaux. Katrine Delage, interlocutrice académique TICE et Documentation décrit ainsi, dans un [compte rendu](#) publié sur le site d'Eduscol, l'utilisation d'un logiciel anti-plagiat dans l'un des lycées du secteur. Le logiciel a été présenté en cours de travaux pratiques encadrés (TPE) de façon à apprendre aux élèves comment, à partir des résultats d'analyse affichés, réfléchir au contenu de leurs productions écrites et des références utilisées. L'auteure note une attention supérieure des élèves portée sur le contenu de leur travail (organisation, formulation et citations) et sur la pertinence des sources.

Orthographe

De nombreux correcteurs orthographiques existent aussi. La plupart analysent automatiquement le document saisi pour fournir des corrections. Plus rares sont ceux qui ont une réelle visée pédagogique comme le logiciel de correction orthographique et grammaticale expérimenté par la chercheuse Samra Bensalem. Cette dernière a mené une étude sur l'utilisation en classe du logiciel dans le but d'améliorer, chez des élèves algériens apprenant le français, la relecture de leurs productions. L'outil utilisé, accessible en ligne, a la particularité de ne pas corriger automatiquement les erreurs détectées mais seulement de les signaler en rappelant la règle correspondante. L'étude a impliqué 20 élèves en équivalence de 1^{ère} littéraire. Elle a tout d'abord permis une analyse typologique des fautes d'orthographe encore omises par les élèves après une relecture non assistée, dans des lettres rédigées par ces derniers. Puis, l'analyse a été réitérée après que les mêmes élèves eurent révisé leur lettre à l'aide du correcteur orthographique. L'étude montre que les élèves corrigent ainsi davantage d'erreurs. Hormis les fautes engendrées par la saisie informatique ou laissées par manque de familiarisation avec l'outil, il ne reste plus dans les lettres que celles qu'ils ne savent pas corriger et celles qui n'ont pas été détectées par l'ordinateur.

Difficulté d'un texte

Enfin, notons que l'utilisation des logiciels d'évaluation automatique ne s'arrête pas à l'évaluation des discours produits par les apprenants. D'autres sont utilisés pour permettre aux enseignants de proposer des textes adaptés au niveau de leur public. De tels logiciels proposent par exemple des alternatives au vocabulaire difficile ou une mise en évidence des passages complexes. Mark D. Shermis et ses collègues de l'université d'Akron pré-cités mentionnent les études conduites par deux universités américaines dans le but d'évaluer l'effet de ces logiciels sur la conscience que les enseignants ont de l'impact des caractéristiques linguistiques d'un texte sur la compréhension. Soixante-dix enseignants d'anglais

L'Agence des usages

Intégrer le numérique dans sa pratique pédagogique



participent à l'étude dans le cadre de cours en ligne pour le développement professionnel. Ils sont formés à l'adaptation de textes à l'aide d'un logiciel pour la moitié d'entre eux et sans, pour les autres. Les résultats montrent que cette activité permet à tous les enseignants participant à l'expérience d'améliorer leurs connaissances sur les caractéristiques linguistiques. Cependant, les enseignants utilisant le logiciel se montrent plus performants dans l'adaptation des textes destinés à des apprenants étrangers.

Conclusions

L'ensemble de ces recherches donne un aperçu de la grande diversité qui caractérise les utilisations de l'évaluation automatique d'écrits. Ces logiciels concernent de nombreux domaines, et sont destinés tout autant aux acteurs de la formation qu'aux apprenants. Le panel d'évaluations ainsi proposé est large : des notes, des suggestions et des corrections peuvent être générées à partir de l'analyse automatique des textes soumis. Les évaluations peuvent porter sur la qualité générale du texte ou sur certaines caractéristiques uniquement. Toutefois, chaque logiciel est prévu pour une utilisation bien définie. C'est donc essentiellement l'adéquation entre les besoins des utilisateurs potentiels (apprenants ou formateurs) et les fonctionnalités proposées qui conduit à l'intégration des logiciels dans l'enseignement. Un grand nombre de besoins sont encore aujourd'hui non satisfaits mais les recherches avancent à grands pas dans le domaine. En attendant, rappelons comme le chercheur David M. Williamson et ses collaborateurs, que les logiciels d'évaluation automatique peuvent être utilisés pour fournir une évaluation unique et rapide ou, *a contrario*, pour doubler une évaluation humaine ou compléter les évaluations automatiques d'autres logiciels.

Sonia Mandin - titulaire d'un doctorat en sciences de l'éducation, chercheuse associée au Laboratoire des sciences de l'éducation (Université Pierre-Mendès-France, Grenoble)

date de publication : 30/01/2013

Recommandations

- Lors de l'utilisation des logiciels de correction automatique, il est préférable que les productions soient typographiées directement par les élèves.
- Ces logiciels, s'ils manquent de fiabilité ou ne mesurent pas toutes les dimensions attendues, peuvent être doublés par d'autres évaluations humaines ou automatiques.
- Le choix du logiciel d'évaluation automatique doit être guidé par le type d'évaluation attendue. Par exemple, si l'évaluation doit être formative, préférer des logiciels d'évaluation automatique qui fournissent un feedback guidant les apprenants vers ce qu'ils doivent faire pour remédier aux erreurs détectées.

L'Agence des usages

Intégrer le numérique dans sa pratique pédagogique



- Le choix du logiciel d'évaluation automatique doit être guidé par les caractéristiques des discours évalués (type de tâche demandée, population ciblée, etc.) et par la vitesse du logiciel (vitesse de connexion, vitesse de calcul).

Références bibliographiques

- Audet L. (2011), *Les pratiques et défis de l'évaluation en ligne. Rapport du réseau d'enseignement francophone à distance du Canada*.
- Bensalem S. (2011), « Le correcteur orthographique : quel apport pour l'amélioration de la compétence orthographique ? », *Synergies Algérie*, 12, 233-240.
- Chen C.-F. E. & Cheng W.-Y. E. (2008), "Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes", *Language Learning & Technology*, 12(2), 94-112. <http://llt.msu.edu/vol12num2/chencheng.pdf>
- Delage K. (2012), *Compilatio, outil de détection de plagiat*, compte rendu d'expérimentation, Eduscol. <http://eduscol.education.fr/cdi/anim/reunion-des-interlocuteurs-academiques/reunion-iatice-doc-2012/compilatio>.
- Dessus P., Lemaire B., Loiseau M., Mandin S., Villiot-Leclercq E. & Zampa V. (2011), "Automated free-text assessment: Some lessons learned. International Journal of Continuing", *Engineering Education and Life-Long Learning*, 21(2/3), 140-154.
- Fourgous J.-M. (2010), Réussir l'école numérique. Rapport de la mission parlementaire. <http://www.ladocumentationfrancaise.fr/rapports-publics/104000080/index.shtml> - book sample
- Loiseau M., Dupré D. et Dessus P. (2011), « Pensum, un logiciel d'aide à la compréhension de cours à distance », *In M. Bétrancourt, C. Depover, V. Luengo, B. De Lièvre et G. Temperman (coord.), Actes de la conférence EIAH 2011*, Mons : Presses de l'université de Mons, p. 287-299.
- Shermis M. D., Burstein J., Higgins D., & Zechner K. (2010), "Automated essay scoring: Writing assessment and instruction", *In E. Baker, B. McGaw & N. S. Petersen (Eds.), International Encyclopedia of Education* (Vol. 4, pp. 20-26), Oxford, UK: Elsevier. http://www.mkzechner.net/AES_IEE09.pdf
- Shermis M. D., Hammer B. (2012), *Contrasting State-of-the-Art Automated Scoring of Essays: Analysis*.
- Williamson D. M., Bennett R. E., Lazer S., Bernstein J., Foltz P. W., Rubin D. P., Way W. D., Sweeney K. (2010), *Automated Scoring for the Assessment of Common Core Standards*.

Article en ligne : <https://www.reseau-canope.fr/agence-des-usages/correction-automatique-decrits-deleves-reve-ou-realite.html>