



Pratiques

Linguistique, littérature, didactique

195-196 | 2022

Écriture scolaire et universitaire : quelle évaluation pour l'accompagner, quelle formation pour les enseignants ?

Diagnostiquer l'erreur rédactionnelle : contribution d'un outil automatique à la pratique de l'évaluateur

Diagnosing the writing error: contribution of an automatic tool to the evaluator's practice

Marianne Desmets, Laura Noreskal, Sarah De Vogüé et Iris Eshkol-Taravella



Édition électronique

URL : <https://journals.openedition.org/pratiques/12244>

DOI : [10.4000/pratiques.12244](https://doi.org/10.4000/pratiques.12244)

ISSN : 2425-2042

Éditeur

Centre de recherche sur les médiations (Crem Université de Lorraine)

Référence électronique

Marianne Desmets, Laura Noreskal, Sarah De Vogüé et Iris Eshkol-Taravella, « Diagnostiquer l'erreur rédactionnelle : contribution d'un outil automatique à la pratique de l'évaluateur », *Pratiques* [En ligne], 195-196 | 2022, mis en ligne le 18 décembre 2022, consulté le 29 juillet 2024. URL : <http://journals.openedition.org/pratiques/12244> ; DOI : <https://doi.org/10.4000/pratiques.12244>

Ce document a été généré automatiquement le 29 juillet 2024.

Le texte et les autres éléments (illustrations, fichiers annexes importés), sont « Tous droits réservés », sauf mention contraire.

Diagnostiquer l'erreur rédactionnelle : contribution d'un outil automatique à la pratique de l'évaluateur

Diagnosing the writing error: contribution of an automatic tool to the evaluator's practice

Marianne Desmets, Laura Noreskal, Sarah De Vogüé et Iris Eshkol-Taravella

Introduction

- 1 Lors des interactions entre enseignants et étudiants, les deux locuteurs n'ont pas le même rôle (Ricci, 2010) puisque l'enseignant est dans une démarche didactique alors que l'étudiant est dans une démarche d'apprentissage. On sait que ce dialogue est souvent mis à mal par des ruptures de communication (*ibid.*) notamment lors de la correction des productions écrites, lorsque les commentaires portent sur la qualité de l'expression. L'évaluation des erreurs par les enseignants dans les écrits des étudiants peut souvent s'apparenter à des « mal dit » qui ne fournissent pas à l'apprenant les éléments nécessaires pour comprendre l'erreur (Boch, Rinck & Sorba, 2020). Il est donc nécessaire que l'enseignant soit plus clair et explicite afin de rétablir le dialogue. Cependant, comment y parvenir quand la tâche de catégorisation des erreurs est compliquée même pour des experts ? Quelles compétences maîtrise-t-on vraiment quand on n'a pas de problème d'expression écrite à l'université ? Il s'agit vraisemblablement de tâches plus complexes que la pure application de règles apprises dans le cadre scolaire, sans quoi les étudiants n'éprouveraient pas autant de difficultés à l'écrit, et, sans doute s'agit-il aussi de compétences ou de savoir-faire suffisamment inconscients pour que l'enseignant ne soit lui-même pas en mesure de les expliciter facilement sur la copie de l'étudiant.

- 2 Ces questions, qui nourrissent déjà plusieurs travaux menés dans le champ des recherches sur les littératies universitaires (Delcambre & Lahanier-Reuter, 2010 ; Boch & Frier, 2015) et plus récemment dans le cadre du projet écrit¹ (Boch, Rinck & Sorba, 2020 ; De Vogüé, 2022), ont été ravivées à l'occasion du développement d'un outil d'aide à la détection automatique d'erreurs syntaxiques², qui doit figurer parmi les dispositifs de formation du projet. C'est donc dans la perspective d'une recherche en traitement automatique du langage (TAL) que nous souhaitons ici contribuer à la réflexion. Les éléments qui concentrent notre attention portent sur le diagnostic ou la catégorisation de l'erreur rédactionnelle, qui est une étape clé pour la création de l'outil, et *in fine*, pour l'étudiant et/ou l'évaluateur.
- 3 Tout d'abord, afin de situer le cadre dans lequel se place la présente recherche, nous présentons la démarche du projet écrit⁺, nous envisageons quelques caractéristiques qui nous ont semblé définitives des écrits réalisés à l'université et qui déterminent à la fois les attendus rédactionnels et le statut donné à l'erreur dans notre approche ; enfin, nous présentons la constitution du référentiel de compétences rédactionnelles développé par le projet écrit⁺. De là, nous abordons la problématique de la création de l'outil de détection automatique, nous présentons les données recherchées et la méthodologie de constitution du corpus, pour aboutir à la typologie des erreurs. Puis, nous discutons des problèmes qui sont apparus lors du calcul de l'accord inter-annotateur et qui nous ont permis d'établir des degrés de difficulté pour l'évaluateur. Enfin, nous présentons l'analyse quantitative des données recueillies.

Le projet écrit⁺ : démarche pour construire un référentiel de compétences rédactionnelles

- 4 Prenant la mesure des difficultés des étudiants dans la maîtrise du français écrit, des chercheurs en linguistique et des enseignants provenant de différents horizons théoriques et disciplinaires ont mutualisé leurs compétences et leurs ressources au service d'un projet commun écrit⁺. L'objectif est de développer un dispositif national³ public pour la formation, l'évaluation et la certification du français écrit. Le projet se voulant collaboratif, le partage de ressources, de pratiques et de savoirs, permet d'aider non seulement les étudiants mais également les enseignants partenaires du projet. Actif depuis 2018, le projet forme à présent une communauté de recherche.
- 5 Une des particularités de l'approche d'écrit⁺ est de partir des difficultés attestées dans les productions écrites des étudiants, de les analyser et d'en déduire des zones de compétence à acquérir⁴. Ce travail d'envergure a permis de construire une grille de compétences rédactionnelles qui sert de référentiel pour élaborer les ressources et les outils pédagogiques d'écrit⁺. C'est donc une démarche « ascendante » qui est empruntée, et qui permet de repérer une compétence à acquérir grâce à ce qu'on a identifié comme étant une difficulté récurrente⁵. Autrement dit, c'est par le relevé des erreurs situées que le référentiel enregistre ce qui correspond à des compétences rédactionnelles. Les dispositifs qui servent pour la formation, l'évaluation et la certification (en cours ou en ligne) sont conçus à partir de cette grille de compétences.
- 6 Un autre aspect essentiel de la démarche pour construire la grille concerne la définition des attendus, ce qui peut s'apparenter à la question de la langue de référence. En réalité, ce sont des conditions de production particulières à l'écrit

universitaire, que l'étudiant doit apprendre à observer, qui vont déterminer les attendus en termes de production textuelle et de langue. Tout d'abord, l'écrit universitaire peut relever de genres textuels différents et peut s'inscrire dans des types différents (argumentatif, explicatif, descriptif...), mais dans tous les cas il s'agit d'un écrit réflexif – au sens où l'on écrit pour penser et l'on pense pour écrire. Autrement dit, c'est un écrit qui construit du savoir et qui doit se placer explicitement comme tel. En outre, c'est un écrit qui s'adresse à un lecteur, et qui se doit d'être maximale-ment lisible : le texte doit fournir les éléments pour que le lecteur puisse lui-même reconstruire ce savoir. Enfin, l'écrit universitaire s'inscrit dans une tradition, un héritage – il est lié à une sphère sociale – qui connaît des spécificités d'une discipline à l'autre (en termes de forme de texte, de type de registre, de choix de constructions, de formes de langue). C'est donc tout un ensemble d'habiletés linguistiques et de savoir-faire qui se trouvent mobilisés dans le travail de production d'un écrit réflexif, comme le notent au demeurant F. Boch et C. Frier (2015).

- 7 Dans ce contexte, ce que nous appelons « erreur » a un statut particulier puisque cela correspond à une mauvaise adéquation à – au moins – l'un des trois critères qui sont, à grands traits : la construction d'un savoir, la compréhensibilité optimale et le respect d'une tradition. C'est donc plus largement un problème de *savoir-faire* rédactionnels, puisqu'à bien des égards les étudiants sont de très bons locuteurs⁶. Pourtant, les problèmes rencontrés touchent toutes les composantes de l'activité linguistique, depuis la sélection et la construction de mots jusqu'à la structuration de textes et le développement de discours, en passant par le déploiement de phrases qui soient bien construites. De ce point de vue, à la suite des travaux de P. Corder (1980), nous considérons que l'erreur est un problème de compétence et non de performance. Autrement dit, dans l'opposition erreur/faute⁷ proposée par P. Corder (*ibid.*), nous considérons bien avoir affaire à des erreurs et non à des fautes, les fautes relevant de la performance et « n'[étant] pas significatives par rapport au processus d'apprentissage d'une langue », alors que les erreurs relèvent de la compétence. Les erreurs « ont une triple signification » (*ibid.*, p. 13) : elles permettent à l'enseignant mais aussi à l'apprenant de vérifier les compétences acquises par l'apprenant, et elles permettent aux chercheurs de comprendre les processus en action lors de l'apprentissage d'une langue. Nous adoptons également la vision de M. Larruy (2003, p. 120 ; voir aussi Demirtas & Gümüs, 2009, p. 128) selon qui, l'erreur relève « d'une méconnaissance de la règle de fonctionnement » alors que la faute se réaliserait lors d'erreurs « de type (lapsus) inattention/fatigue que l'apprenant peut corriger ». D'après l'auteure, ce n'est pas l'analyse des fautes qui apporte des informations sur l'apprentissage de la langue mais bien celle des erreurs.
- 8 La question de l'analyse ou du diagnostic de l'erreur est centrale, particulièrement dans le travail de création de l'outil automatique, qui vise à proposer un jeu d'étiquettes explicites à l'utilisateur (étudiant ou enseignant). Dans l'approche qui a été suivie, la réflexion s'est appuyée sur différents travaux de référence, tant pour l'analyse des erreurs (Frei, 1929 ; Perdue & Porquier, 1980 ; Reichler-Béguelin, 1992) que pour la constitution de la grille d'évaluation (Boivin & Pinsonneault, 2020 ; Boch, Rinck & Sorba, 2020). Nous avons veillé à répondre aux observations de F. Boch, F. Rinck & J. Sorba (2020) qui analysent les erreurs et maladroites des étudiants grâce à une base de données d'écrits accompagnés des commentaires d'enseignants. En effet, les

auteures notent trois problèmes qui se posent lors de l'analyse des erreurs par l'évaluateur :

- Le passage du global au local. Il est nécessaire de diagnostiquer précisément l'erreur ;
 - Le passage de l'intuition à la catégorisation. Il faut réussir à regrouper les erreurs semblables à partir des commentaires des enseignants ;
 - Le passage d'une objectivité illusoire à une intersubjectivité assumée. Celui-ci nécessite de trancher sur la nature d'une erreur ou d'une maladresse en adoptant, selon les auteures, une approche variationniste.
- 9 Face à ces considérations, l'outil que l'on propose doit fournir :
- Un diagnostic situé (par exemple, par un surlignage de la zone problématique) ;
 - Une catégorisation explicite de l'erreur (par un message invitant à vérifier une ou deux catégories d'erreur détectées).
- 10 En ce qui concerne le point 3, les catégories d'erreur sont le fruit d'un travail collectif correspondant à une étape technique nécessaire à tout outil de détection automatique, ce que l'on nomme « l'accord inter-annotateur », et cet accord donne lieu à une mesure (voir *infra* dans la partie « Difficultés dans le calcul de l'accord inter-annotateur »).
- 11 Les catégories d'erreur que l'outil sera capable de repérer doivent être en correspondance avec des items de la grille générale du projet écrit+. On l'a dit, la constitution de cette grille vise à définir un référentiel de compétences rédactionnelles, et s'inscrit à ce titre dans une démarche particulière. Ainsi que le notent M.-C. Boivin et R. Pinsonneault (2020), les grilles existantes posent un certain nombre de problèmes, en particulier, elles se contentent souvent de catégories très générales comme « Syntaxe », « Orthographe » ou « Lexique » sans pour autant préciser ou détailler les sous-types d'erreurs. Les auteures suggèrent qu'une grille devrait être conçue en respectant quatre critères d'évaluation, qui, selon elles, « assurent que la grille de codage offre une vision globale organisée de la langue » (*ibid.*, 92). Ces critères sont l'adéquation à un modèle théorique, l'adéquation empirique, la complétude et la comparabilité. La grille du projet écrit+ répond implicitement à ces critères épistémologiques, tout en prenant le parti de rendre compte des habiletés linguistiques et des savoir-faire mobilisés lors de la production d'écrits réflexifs. Cela a conduit à la constitution d'un référentiel de compétences dédié à ce type d'écrits, qui se trouve être assez différent à ce titre des référentiels existants (ceux du CECRL notamment). Quatre champs très généraux ont d'abord été distingués : les champs du contenu, du texte, des mots et des phrases. Cependant, ils sont à concevoir en interaction, la production d'un écrit réflexif impliquant nécessairement que contenus, structures et moyens se déterminent mutuellement. On a alors procédé de manière récursive, jusqu'à arriver à des habiletés qui soient à la fois situées et testables (Perrenoud, 2000). Ainsi a-t-on identifié dans chaque champ les compétences qu'il recouvre, puis pour chaque compétence les capacités qu'elle mobilise, et pour chaque capacité l'ensemble des habiletés qu'elle convoque. Ces habiletés sont notamment celles repérées à l'occasion du recueil de données, entendues donc comme des savoir-faire et comme des savoir-faire situés. S'y ajoutent d'autres, relatives à de la compréhension (parce que rédaction et compréhension sont indissociables ; voir Bishop & Rouxel, 2007), mais aussi aux différents moments du travail rédactionnel que sont, à côté de la mise en texte proprement dite, la planification et la révision (Hayes & Flowers, 1980). Ainsi, ce sont quelques 850 habiletés différentes qui ont pu être distinguées, les unes élémentaires,

les autres expertes, le référentiel visant à rendre compte de profils de compétences rédactionnelles à la fois diversifiés et évolutifs.

- 12 À présent, nous abordons la problématique de la création de l'outil de détection automatique, nous présentons les données recherchées et la méthodologie de constitution du corpus, puis la typologie des erreurs que nous avons obtenue.

Objectif de l'outil, données recherchées et constitution du corpus

- 13 Parmi les dispositifs créés pour aider les étudiants à améliorer leurs habiletés rédactionnelles, nous avons souhaité développer un outil capable de détecter automatiquement certains problèmes de construction syntaxique. Parmi les copies que nous avons collectées, ou les extraits que les enseignants nous ont signalés, figuraient un certain nombre de séquences « mal réussies » – typiquement des séquences que l'enseignant commente par « mal dit », ou parfois « mal construit ». Nous avons pu y relever des éléments linguistiques réguliers : ce sont des séquences phrastiques longues comportant des coordinations ou des juxtapositions, parfois enchâssées, avec ou sans ellipses ; voici deux exemples caractéristiques :

(1) Avant de visionner la comédie musicale, il faudra étudier avec les élèves la période révolutionnaire pour comprendre les raisons de la Révolution et de rendre cette activité ludique mais pédagogique.

(2) La partie EMC du SCCC du cycle 3 illustre cette idée d'engagement « L'esprit de coopération doit être encouragé, la responsabilité vis-à-vis d'autrui mise à l'épreuve des faits » et de travailler sur les connaissances suivantes « S'estimer et être capable d'écoute et d'empathie ».

- 14 De façon générale, dans les copies, les problèmes de construction qui sont en cause ici sont moins saillants que les erreurs d'orthographe lexicale ou grammaticale, mais ils renseignent sur le niveau d'habileté de l'étudiant. Ils mettent en jeu la capacité à planifier l'organisation des propositions tout en maintenant la cohérence des relations logiques, et la maîtrise de la valence des lexèmes par-delà la coordination, ou la distance. Ces problèmes de construction mettent également l'évaluateur en difficulté puisque plusieurs facteurs linguistiques sont à l'œuvre et l'analyse précise du ou des facteurs déficients est très coûteuse. En bref, l'évaluateur a le sentiment d'un « mauvais » équilibre dans l'énoncé mais n'identifie pas vraiment d'où il provient. Nous avons donc cherché à construire un outil automatique capable d'appeler le scripteur à la vigilance en signalant les zones sensibles de son texte – des séquences propices aux erreurs –, tout en aiguillant sa recherche en lui proposant de vérifier un, deux ou trois éléments pour chaque zone identifiée.
- 15 La première étape de création de l'outil est celle de la constitution du corpus – l'outil devant être entraîné sur un corpus de référence pour « apprendre » à repérer les zones sensibles. Nous avons opté pour un corpus dit « écologique »⁸ reflétant réellement les difficultés des étudiants afin d'éviter les biais d'analyse, et nous avons procédé à une fouille exhaustive des phrases contenant des coordinations. Le corpus utilisé est composé d'écrits réflexifs, ce sont des rédactions évaluatives telles que les mémoires, les rapports de stage, les exercices et les devoirs maison. Bien qu'ils soient différents par leur temps de rédaction et leurs conditions de réalisation, ces quatre types de rédactions permettent d'observer les productions des étudiants lorsqu'ils sont évalués⁹.

Tableau 1. Composition du corpus d'écrits¹⁰.

Rédactions	Nombre	Types
Devoirs maison	188	Préparés
Exercices	167	Non-préparés
Mémoire	27	Préparé
Rapport de stage	47	Préparé

- 16 Le corpus de phrases coordonnées a été constitué à partir de 429 rédactions dont 188 devoirs maison, 167 exercices, 27 mémoires et 47 rapports de stage. Les différentes rédactions nous ont été transmises par des enseignants mais aussi parfois par des étudiants. Certaines rédactions ont également pu être consultées *via* les archives de l'Université de Paris Nanterre.
- 17 Les rédactions ont été séparées en deux : les rédactions préparées rédigées sur une longue période et avec l'aide d'un tiers, et les rédactions non-préparées rédigées dans une courte limite de temps et obligeant l'étudiant à travailler seul. Dans Noreskal, Eshkol-Taravella & Desmets (2020), cette séparation a permis de vérifier l'hypothèse selon laquelle il y a plus d'erreurs dans les rédactions non-préparées que dans les rédactions préparées.
- 18 À partir de ces rédactions, il a été possible d'extraire les phrases coordonnées grâce au logiciel Unitex (Paumier, 2011). Cet outil permet de rechercher et d'extraire des constructions syntaxiques d'un texte par l'intermédiaire de graphes. Le corpus est ainsi composé de 6 645 phrases contenant des coordinations.

Tableau 2. Nombre de phrases coordonnées extraites par type de rédactions.

Types	Nombre	Total phrases	Phrases coordonnées	Fréquence des phrases coordonnées dans le type d'écrits
Devoirs maison	139	7 635	1 467	19 %
Exercices	167	2 502	593	24 %
Mémoire	27	13 244	1 890	14 %
Rapports	47	16 311	2 695	17 %
Total	380	39 692	6 645	16 %

- 19 Le tableau 2 présente le nombre de phrases coordonnées extraites du corpus de productions. Ainsi, 1 467 phrases ont été extraites des devoirs maison, 593 des exercices, 1 890 des mémoires et 2 695 des rapports. Les phrases coordonnées représentent 19 % du total de phrases pour les devoirs maison, 24 % pour les exercices, 14 % pour les mémoires et 17 % pour les rapports.

Annotation des constructions problématiques et typologie des erreurs

- 20 Une fois le corpus de référence constitué, il doit être préparé pour devenir ce qu'on appelle un corpus « d'apprentissage » afin de pouvoir entraîner l'outil de détection à repérer les constructions recherchées. Pour ce faire, on procède à une tâche technique appelée *annotation du corpus*¹¹ qui doit permettre de réaliser des statistiques sur les données pour contrôler et interpréter l'efficacité d'un outil. L'annotation est un moyen d'ajouter des informations pertinentes sur les données du corpus, et dans le cas qui nous intéresse, elle doit pouvoir renseigner sur deux niveaux :
- (1) Phrase correcte/erronée ;
 - (2) Le type d'erreur.
- 21 Nous avons procédé à l'annotation systématique de toutes les phrases du corpus avec des étiquettes (1) « correcte/erronée »¹², et ce faisant nous avons cherché les régularités dans les types d'erreurs afin d'aboutir à (2) une typologie et un jeu d'étiquettes correspondant.
- 22 De cette première étape d'annotation, nous avons tiré deux constats. D'une part, certains types d'erreurs sont plus fréquents que d'autres, d'autre part, beaucoup d'énoncés problématiques présentent simultanément des défections à plusieurs niveaux, au niveau syntaxique et au niveau sémantique. Au moins dans un premier temps, l'outil n'arriverait pas à détecter tous les types d'erreurs, il a donc fallu faire le choix de se limiter à la détection des éléments les plus identifiables, les erreurs syntaxiques, et les éléments les plus fréquents. Dans les séquences analysées, ce sont les constructions phrastiques longues avec des coordinations, et le plus souvent des ellipses, qui constituaient les contextes linguistiques concentrant le plus de problèmes. Dans ces séquences le scripteur mobilise plusieurs mécanismes simultanément sur plus d'une proposition, sur un temps long. On sait que les constructions coordonnées sont soumises à de fortes contraintes de parallélisme syntaxique et sémantique (Abeillé, 2005, p. 42-66 ; Mouret, 2003 ; 2005) et qu'elles sollicitent une maîtrise robuste du cadre de sous-catégorisation des lexèmes mis en facteur commun, maîtrise qui est d'autant plus sollicitée en cas de distance entre les conjoints. Dans les données que nous étudions, l'édifice se rompt le plus souvent par la violation des contraintes de parallélisme, entraînant l'ambiguïté voire l'incohérence de l'énoncé final, sans que ce qui constitue la source même de l'erreur ne soit facilement repérable. Ainsi, dans l'exemple (1) reproduit ici en (3), la séquence coordonnée soulignée doit pouvoir être mise en parallèle avec un syntagme correspondant du contexte antérieur :
- (3) Avant de visionner la comédie musicale, il faudra étudier avec les élèves la période révolutionnaire pour comprendre les raisons de la Révolution et de rendre cette activité ludique mais pédagogique.
- 23 En (3) le meilleur candidat est le syntagme infinitif « étudier avec les élèves la période révolutionnaire pour comprendre les raisons de la Révolution » en mettant en facteur commun le verbe *falloir*. Le problème vient de la forme du syntagme conjoint souligné, un syntagme prépositionnel en *de*, qui ne permet pas de construire syntaxiquement, ni surtout sémantiquement, le parallélisme qui était prévu (« il faudra étudier la période révolutionnaire et rendre cette activité ludique mais pédagogique »).
- 24 Nous avons donc cherché à obtenir une typologie des erreurs qui soit « optimale » pour l'outil, c'est-à-dire comportant le moins d'étiquettes possibles mais les étiquettes les

plus pertinentes. Tous les types d'erreurs identifiés ont fait l'objet d'un enregistrement sous la forme d'une « habileté » dans le référentiel général d'écrit+. Ce premier travail a donc permis d'enrichir la grille générale. Seules quelques étiquettes ont été retenues pour former le guide pour l'annotation du corpus de l'outil. La typologie est la suivante : elle comporte quatre grands types, les erreurs relatives aux *prépositions*, à la *punctuation*, aux *groupes conjoints* et à l'*accord sujet-verbe*. Ces catégories sont à leur tour divisées en sous-types, pour un total de dix sous-types que nous détaillons ci-après.

Les prépositions

- 25 Le non-respect des valences des verbes, des noms ou des adjectifs dans les constructions coordonnées engendre des erreurs impliquant très souvent des prépositions. Voici les trois grands cas de figure repérés.

Absence d'une préposition (PREP ABS)

(4) Il s'agit de permettre aux plus âgés de s'ouvrir à la diversité et la comprendre.

- 26 Ici, la catégorie du complément attendu par le verbe « permettre » (permettre de faire quelque chose) est un syntagme prépositionnel indirect introduit par la préposition *de*, or le second conjoint est un syntagme infinitif direct (« la comprendre »). La construction syntaxique attendue est « Il s'agit de permettre aux plus âgés de s'ouvrir à la diversité et de la comprendre ».

Ajout d'une préposition non-attendue (PREP ADD)

(5) Les personnes qui ont recours à la GPA peuvent être des couples femme/femme-homme en cas d'infertilité féminine, ou à des hommes célibataires ou un couple homosexuel.

- 27 En (5), le second conjoint correspond au complément du verbe « être », or la préposition *à* ne fait pas partie de la sous-catégorisation attendue par le verbe « (pouvoir) être ».

Remplacement d'une préposition (PREP REMP)

(6) Effectivement, j'ai bien aimé la contrainte qui consistait, à la fois, de trouver une suite cohérente à l'incipit proposé mais aussi de terminer le récit par une phrase bien précise.

- 28 L'exemple (6) illustre un problème de choix de préposition. Le verbe « consister » attend un syntagme prépositionnel en « à » ou « en », cependant qu'ici le syntagme prépositionnel réalisé est introduit par « de ».

La punctuation

- 29 Pour comprendre les attendus du rôle joué par la punctuation dans les écrits réflexifs, il nous semble important de rappeler ici les trois grands critères d'évaluation de ces textes : la construction d'un savoir, la lisibilité, l'adhésion aux formes héritées. Il est normal que les énoncés soient longs pour construire un raisonnement, qu'ils comportent plusieurs propositions. La punctuation joue alors pleinement le rôle de marqueur de la structuration syntaxique permettant d'interpréter les liens de discours

et les liens logiques entre les propositions. Deux sous-types de difficultés relevant de ce domaine ont été répertoriés :

- (a) des séquences comportant beaucoup de propositions imbriquées les unes dans les autres où la ponctuation ne permet pas la lisibilité attendue – nous les étiquetons « structures lourdes » (sans considération stylistique) ;
- (b) des séquences où la « ponctuation autour des incidents » n'est pas observée.

Structure lourde (SL)

(7) L'auteur dénonce les contrats sociaux très importants entre les pays du nord et les pays du sud plus précisément, la faim à l'échelle mondiale et il va donc accentuer ses disparités à travers une caricature qui à une visée humoristique mais également critique puisque elle met en relief les pays du Nord, riches et industrialisés, dominé par la Triade qui regroupe les grands pôles économiques du monde et les pays du Sud, pauvres et désigné comme le « Tiers-monde » qui réfléchissent à une solution pour lutter contre la famine autour d'un repas.

- 30 En (7), la ponctuation ne permet pas d'interpréter « les pays du Sud » et « les pays du Nord » comme deux compléments de « met en relief ». L'énoncé demande un coût d'interprétation trop important au regard du critère de lisibilité.

Ponctuation autour des incidents (INCIDENT PONC)

- 31 L'incident est un segment de la phrase qui permet au locuteur d'exprimer un commentaire en arrière-plan de la prédication principale (Marandin, 1999). Graphiquement, la ponctuation marque la délimitation du commentaire par un encadrement du segment entre virgules.

(8) Il en est de même aux Etats-Unis où la Cour suprême qui est le sommet du pouvoir judiciaire aux États-Unis et le tribunal de dernier ressort tranche des litiges et rend des décisions de justice qui en font des règles de droit.

- 32 Dans l'exemple (8), la relative incidente « qui est le sommet du pouvoir judiciaire aux États-Unis et le tribunal de dernier ressort » n'est pas entourée de virgules, ce qui entame la compréhension de l'énoncé (problème de lisibilité).

Les groupes conjoints

- 33 Deux problèmes ont été observés entre les groupes coordonnés :

- (a) la mauvaise cohérence des groupes syntaxiques ;
- (b) la grande distance entre conjoints.

Mauvaise cohérence des groupes syntaxiques (MCGS)

- 34 Sous cette étiquette, nous avons dû réunir des séquences répondant à plusieurs cas de figure, afin de limiter le nombre de catégories pour l'outil. On trouve des erreurs liées à un problème de parallélisme comme en (9), et des cas de constructions de phrases elliptiques problématiques comme en (10) et (11).

(9) Il ne me parle pas d'un divorce en particulier, mais que l'action de divorcer était interdite.

- 35 Dans la phrase (9), la forme du second conjoint ne respecte pas la valence du verbe « parler » mis en facteur commun (*« parler que quelque chose »). Dans l'absolu, des syntagmes de catégories distinctes coordonnés sont possibles à condition que leurs

fonctions respectives soient identiques et qu'elles respectent les attentes du lexème mis en facteur commun, ce qui n'est pas le cas en (9) (COI *mais* *COD). Ici, c'est probablement le verbe « dire », autre verbe de parole, qui est entré en concurrence lexicale avec le verbe « parler » au moment de la production écrite.

- 36 La construction d'une phrase elliptique peut être défectueuse pour d'autres raisons. Nous avons observé des cas de mauvaise résolution anaphorique comme en (10), mais aussi de nombreux cas où les conditions qui permettent à la séquence omise d'être restituée (interprétée) dans la phrase elliptique ne sont pas respectées (11).

(10) Les maires et les présidents des régions ne dépendent pas du ministre alors que les recteurs et DASEN le sont.

- 37 En (10), la résolution de l'ellipse par proforme verbale pose problème, il est nécessaire de répéter le verbe à l'identique (« en dépendent »). On sait qu'en français, les circonstancielles sont un contexte possible pour les ellipses polaires, ou modales¹³, mais la liste des verbes utilisés est très limitée et le verbe « être » (« les recteurs et DASEN le sont ») n'en fait pas partie (Busquets & Denis, 2001 ; Dagnac, 2008).

(11) Laurent (2004) explique les avantages de la méthode inductive ; avec celle-ci, les enfants retiennent mieux car ils construisent et de ce fait rend actif l'apprenant.

- 38 En (11), la construction de l'ellipse est défectueuse parce que le syntagme nominal « *la méthode inductive* » ne peut être mis en facteur commun pour être interprété comme le sujet du dernier conjoint « et de ce fait [la méthode inductive] rend actif l'apprenant ».

Grande distance entre conjoints (DIST CONJOINT)

(12) L'entreprise ONDULED est spécialisée dans la vente de batteries à condensateurs appelées aussi des régulateurs varométriques ou des onduleurs tout simplement ainsi que des équipements d'éclairage en LED.

- 39 Dans certaines phrases, les groupes qui doivent être interprétés comme des conjoints se trouvent éloignés l'un de l'autre par la présence d'un autre constituant – en (12) « de batteries à condensateurs et des équipements d'éclairage en LED ». Ici, il s'agit d'un constituant épithète (apposé, ou incident) comprenant lui-même une coordination « appelées aussi des régulateurs varométriques ou des onduleurs tout simplement ». Ainsi, il devient difficile d'interpréter l'énoncé, car le rattachement du deuxième conjoint est trop coûteux.

Accord Sujet-verbe

- 40 Plusieurs phrases présentent des problèmes d'accord sujet-verbe. La forme verbale peut être non finie (gérondif, infinitif), ou bien mal fléchie :

(13) Cette phrase renvoie au moment où le Sénat et le peuple romain donna à Octave le surnom augustus en 27 av. J.-C.

- 41 En (13), alors que les deux syntagmes nominaux coordonnés le « Sénat » et « le peuple romain » forment le sujet de la phrase, le verbe « donna » ne porte pas l'accord pluriel. On peut supposer que le verbe a été conjugué avec le conjoint le plus proche.

Difficultés dans le calcul de l'accord inter-annotateur

- 42 La typologie adoptée, utilisée pour l'outil, doit être compréhensible et applicable par des experts tels que les linguistes et enseignants. Pour vérifier sa compréhensibilité et sa bonne application, il est nécessaire de calculer ce qu'on nomme un « accord inter-annotateur ». Cette mesure, propre au TAL, calcule la concordance entre les différentes annotations des experts afin d'évaluer la pertinence de la typologie. En d'autres termes, grâce au calcul de l'accord inter-annotateur, il est possible de savoir si notre typologie permet d'évaluer les erreurs repérées. Plusieurs mesures existent pour calculer l'accord inter-annotateur, tels que le kappa de J. Cohen¹⁴ (1960) qui permet de comparer les annotations de deux annotateurs ou le kappa de J. L. Fleiss (1975) qui permet de comparer les annotations de plus de deux annotateurs. L'évaluation de l'accord peut aller de 0 (total désaccord) à 1 (accord parfait). Par conséquent, plus l'accord s'approche de 1, meilleur est l'accord entre les différents annotateurs.
- 43 Dans notre cas, le groupe d'experts, composé de deux chercheuses linguistes, trois étudiantes en TAL et une professeure des écoles, a annoté 200 phrases du corpus. L'hétérogénéité relative des experts est méthodologiquement nécessaire, elle permet d'observer les convergences et divergences selon le niveau des annotateurs. Les annotateurs disposent du guide d'annotation réunissant les différents types d'erreurs recherchés avec les explications et les exemples détaillés. Ils ont également un fichier Excel dans lequel ils doivent remplir 4 colonnes pour chaque phrase (figure 1).

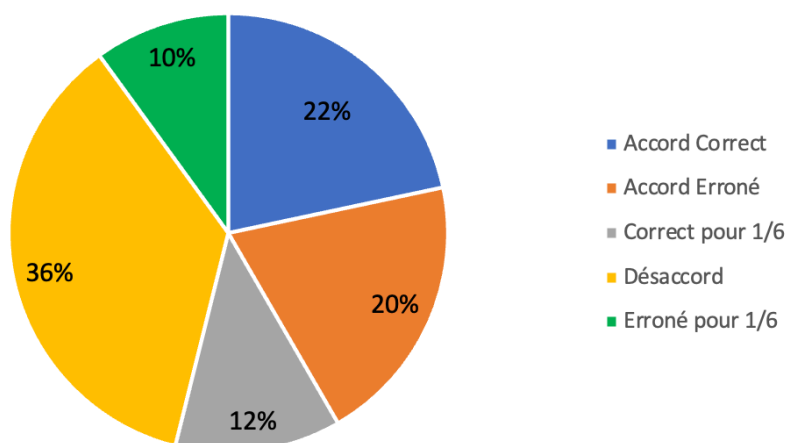
Figure 1. Exemple de document Excel pour l'accord inter-annotateur.

ID	Phrases	ERREUR (OUI/NON)	Type Erreur	Autres types d'erreur	Explications
id-Dem447	À chaque fin de trimestre, il me faudra les rencontrer individuellement pour leur donner en main propre les livrets d'évaluation et les aider à les comprendre car explicités, leur lecture sera constructive et évitera un rejet éventuel de leur contenu.				
id-Mem540	À droite et à gauche de la composition se trouvent deux étoiles.				
id-Mem528	À grands renforts de courbes, de cercles et d'arabesques, le cadre se fraye un chemin et devient le fond et le décor.				
id-Mem1682	À l'adolescence en revanche, elle se décrit comme introvertie.				
id-Mem510	À l'intérieur de ces cercles, comme des arrière plans, se déploient à gauche un paysage de nuit, avec la lune et six étoiles, et à droite, un soleil qui rayonne, et prend toute la place à l'intérieur de ce cercle, avec quelques nuages.				
id-Rap889	À plusieurs niveaux, étant en moyenne montagne, il y a des fonds de vallée avec des terrains plus humides, de coteaux plus sec et moins abrités voir des plateaux secs de types landes.				

Légende : « Erreur (Oui/ Non) » : la phrase contient-elle une erreur ou non ? ; « Type d'erreur » : dans le cas où la phrase est erronée, quel est le type d'erreur ? ; « Autres types d'erreurs » : présence de plusieurs erreurs simultanément ? ; « Explications » : expliquer sa décision.

- 44 À la fin de la tâche, l'accord inter-annotateur a été calculé en utilisant le kappa de J. L. Fleiss. Pour calculer l'accord nous avons pris en compte la colonne « Erreur : Oui/ Non ». Le score était de 0,4, ce qui signifie qu'il y a un accord modéré entre les six annotateurs selon l'échelle de J. R. Landis (& Koch, 1977). Cependant, selon R. Artstein & M. Poesio (2008), les valeurs inférieures à 0.8 sont le signe de désaccords et non pas celui d'un accord modéré. Ainsi, afin de comprendre ce résultat, une analyse des convergences et des divergences entre les annotateurs a été réalisée. Les résultats sont présentés dans la figure 2.

Figure 2. Pourcentages d'annotations pour l'accord inter-annotateur.



- 45 Parmi les 200 phrases données pour calculer l'accord, les annotateurs ont été d'accord pour 42 % des phrases et en désaccord pour 58 % des phrases. Parmi les 42 % de phrases pour lesquelles les annotateurs ont été d'accord, 20 % concernent l'annotation erronée et 22 % l'annotation correcte. En ce qui concerne les désaccords, 12 % des phrases ont été annotées comme erronées par 5 personnes sur 6 et 10 % des phrases ont été annotées comme correctes par 5 personnes sur 6.
- 46 Ces résultats montrent que la tâche d'annotation d'erreurs est complexe, et cela même pour des experts informés par le guide d'annotation, puisqu'ils n'ont pas toujours été d'accord sur la présence d'une erreur ou non. Résultat plus intéressant encore, l'accord inter-annotateur s'avère dépendre du « type d'erreur » en jeu. Nous en avons déduit que toutes les erreurs ne présentent pas le même degré de difficulté pour l'évaluateur. Grâce à l'analyse des convergences et des divergences nous avons pu établir un classement avec deux polarités : d'un côté, un ensemble de séquences pour lesquelles l'évaluation d'erreur est unanime, de l'autre, un ensemble de séquences pour lesquelles il y a le plus de désaccords. Regardons dans le détail les types d'erreurs concernés.

Avis unanimes pour l'annotation erronée

- 47 Les avis unanimes portent sur les erreurs de type « manque de cohérence des groupes syntaxiques », les erreurs de « structure lourde », les erreurs de « remplacement d'une préposition par une autre » et les erreurs d'« ajout de préposition non-attendue ». Les constructions en cause sont celles où le parallélisme n'est pas respecté, celles où la construction de l'ellipse est défectueuse¹⁵, celles dont l'enchaînement des propositions n'est pas fluide en raison de mauvais liens logiques, de problèmes de ponctuation et/ou de rattachements de syntagmes, ou encore celles pour lesquelles le complément du lexème mis en facteur commun est un syntagme prépositionnel erroné.
- 48 Si la présence d'erreurs dans ces séquences est unanimement reconnue par le groupe d'experts, on peut estimer que pour l'évaluateur dont le niveau d'expertise est médian, les séquences concernées sont facilement repérables comme « erronées ». Cependant, s'il y a accord sur la présence d'une erreur, il n'y a pas nécessairement accord sur le type d'erreur concerné. Dans une séquence donnée, l'analyse fine du type d'erreur en

cause demeure une tâche complexe. Autrement dit, la compétence inconsciente du locuteur-annotateur produit un sentiment clair sur le fait que la phrase n'est pas « bonne » ; le même locuteur-annotateur n'est en revanche pas facilement capable de dire pourquoi cette phrase est erronée. Nous avons observé beaucoup d'indécision. Par conséquent, la tâche qui vise à expliciter la source de l'erreur relève d'un niveau d'expertise poussé, demande une pratique didactique, un entraînement, pour les formateurs sans doute, mais aussi pour les annotateurs.

Désaccords maximaux pour l'annotation correcte/erronée

49 Paradoxalement, pour les séquences qui concentrent le taux de désaccords le plus important sur l'étiquette « correcte/erronée », les pistes d'explications sont plus claires. Les erreurs en cause sont les « absences de préposition », comme en (4) et (18) ci-dessous :

50 (18) Un orthophoniste aborde enfin tout ce qui est lié à la neuropsychologie, les sciences du langage, la linguistique et pleins d'autres domaines.

51 Ici, on se situe dans le domaine de la variation, du choix du locuteur, du possible de langue. Ainsi en (18), certains experts acceptent de mettre en facteur commun la préposition à et réalisent le calcul de la coordination avec les quatre syntagmes « la neuropsychologie, les sciences du langage, la linguistique et pleins d'autres domaines ». D'autres ont besoin de distribuer la préposition à sur chacun des syntagmes nominaux conjoints « lié à la neuropsychologie, aux sciences du langage, à la linguistique et à plein d'autres domaines ». On rencontre cette partition entre experts chaque fois que les deux chemins de calcul sémantique sont possibles, comme en (19) où 67 % des annotateurs trouvent la phrase erronée et 33 % la trouvent correcte.

(19) En effet, selon une étude menée en 2005-2006, pour la majorité des élèves, l'histoire et la géographie consistent principalement à identifier et décrire les phénomènes historiques ou géographiques et non les classer, les hiérarchiser, les structurer etc., ou encore rédiger une argumentation.

52 Ce constat nous permet de décider que ce que nous avons initialement repéré comme une erreur ne doit pas être traité comme tel. Nous avons donc relancé le travail d'annotation en éliminant tous les cas relevant de la catégorie « absences de préposition ». Le score alors obtenu a été de 0,88, calculé avec le kappa de J. Cohen. Le fait de dépasser 0,8 est caractérisé comme un excellent accord selon l'échelle de J. R. Landis (Landis & Koch, 1977) et signe d'accord selon R. Artstein & M. Poesio (2008). Ce nouveau résultat a ainsi permis de valider la typologie et d'annoter tout le corpus en type d'erreurs.

53 Avant de passer à l'analyse des résultats, nous devons préciser quelques points. Tout d'abord, même si l'étape d'analyse des erreurs (réalisée à partir du corpus des rédactions) est cruciale pour le travail des chercheurs en TAL et pour celui du formateur, les perspectives dans lesquelles ceux-ci s'inscrivent sont distinctes. Les premiers ont pour objectif de trouver la typologie la plus efficace pour l'outil, ce qui entraîne la recherche de la typologie optimale : le moins d'étiquettes possible et les étiquettes les plus pertinentes pour mener à bien le repérage automatique des séquences erronées. Les seconds ont pour objectif de développer des dispositifs pédagogiques à partir de bases empiriques solides, attestées, pour entraîner les apprenants sur des difficultés ciblées. Par conséquent, les « grains » de compétences

rédactionnelles qui figurent dans le référentiel général ont vocation à être plus nombreux que les étiquettes utilisées pour entraîner l'outil automatique.

54 À présent, nous présentons les résultats de l'analyse quantitative en détails, classés par type de rédaction.

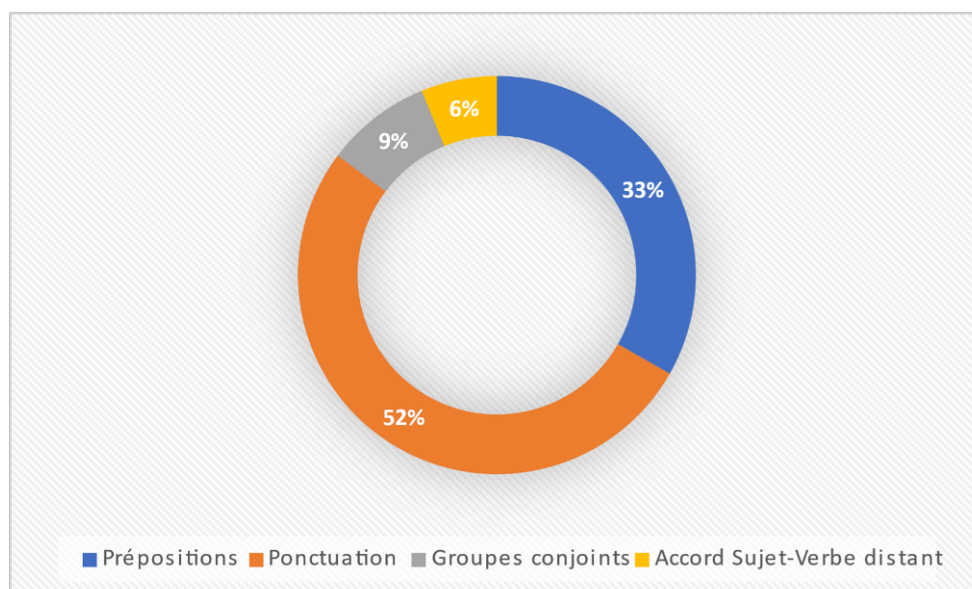
Analyse quantitative

55 Après validation, une partie du corpus a été annotée selon la typologie. La composition du corpus est présentée dans la figure 3.

Rédactions	Phrases erronées	Phrases correctes	Phrases annotées	Nombre total de phrases coordonnées
Devoirs maison	151	242	393	1467
Exercices	313	279	592	592
Mémoires	69	222	291	1890
Rapports	127	281	408	2695
Total	660	1024	1684	6644

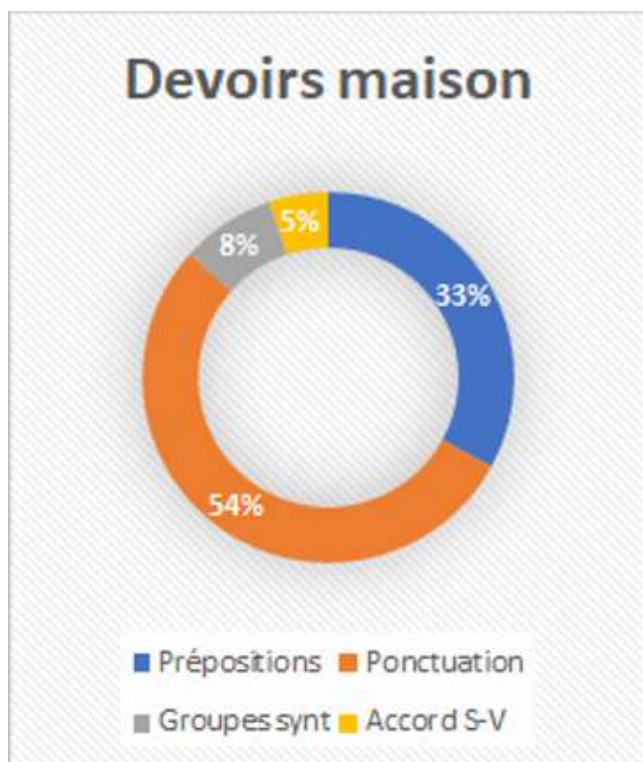
56 Dans le corpus de phrases, nous avons annoté 393 phrases provenant de devoirs maison, 592 d'exercices, 291 de mémoires et 408 de rapports, soit 1 684 phrases au total. Parmi les phrases annotées, nous avons 1 024 phrases correctes et 660 phrases erronées. Les rédactions préparées (devoirs maison, mémoires et rapports) contiennent moins de phrases erronées que de phrases correctes alors qu'on observe le schéma inverse pour les rédactions non-préparées (exercices).

Figure 3. Répartition des erreurs dans le corpus.



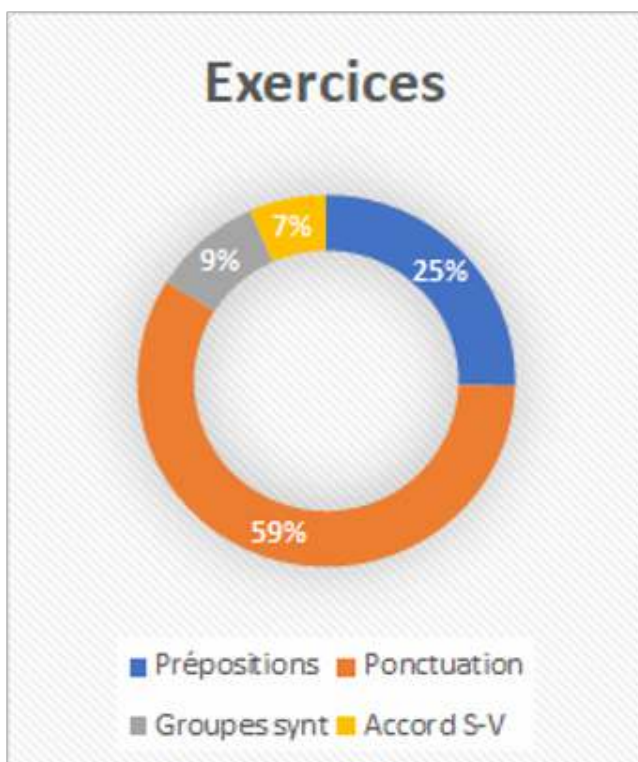
- 57 La catégorie d'erreurs la plus répandue est la ponctuation » avec 52 % d'apparition dans le corpus. Les erreurs de « prépositions » sont en deuxième position avec 33 %, puis suivent les « groupes conjoints » et les « accords Sujet-Verbe » avec respectivement 9 % et 6 %. Les erreurs de « ponctuation » possèdent le plus de sous-catégories, donc il peut sembler normal qu'il y ait plus d'erreurs de ce type dans le corpus.

Figure 4. Répartition des erreurs dans les devoirs maison.



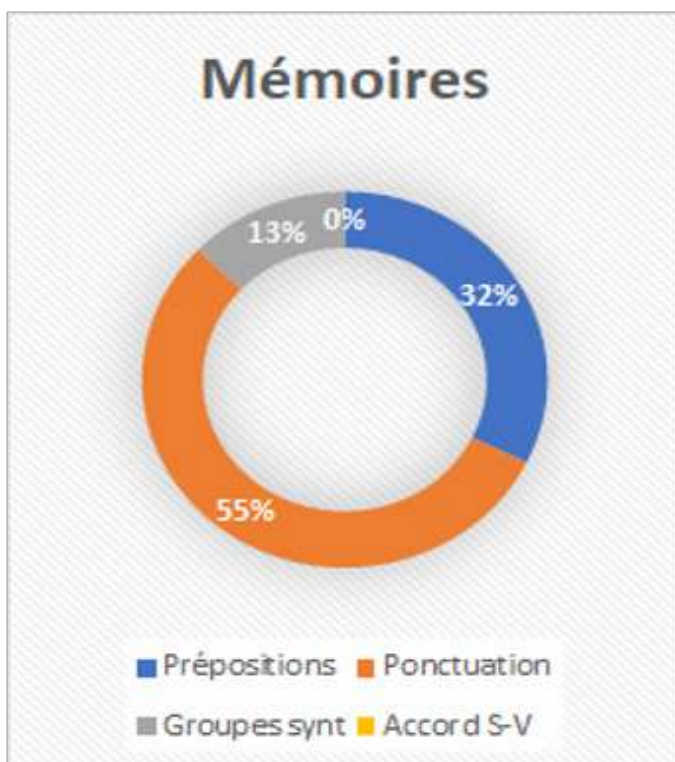
- 58 Dans les devoirs maison, les problèmes les plus courants sont les erreurs de « ponctuation » puisqu'elles représentent 54 % des erreurs. La catégorie « prépositions » est deuxième avec 33 %. Les catégories « groupes conjoints » et « accord Sujet-Verbe » sont moins présentes dans le corpus de devoirs maison puisqu'elles représentent respectivement 8 % et 5 %. Bien que plus courts que les mémoires ou les rapports de stage, les devoirs maison sont des écrits réflexifs par excellence qui concentrent les habiletés à mobiliser pour construire un savoir « lisible ». Ces tâches demandent de planifier la pensée, d'établir clairement les propositions et les relations logiques pour construire les structures phrastiques, c'est une véritable difficulté pour les étudiants.

Figure 5. Répartition des erreurs dans les exercices.



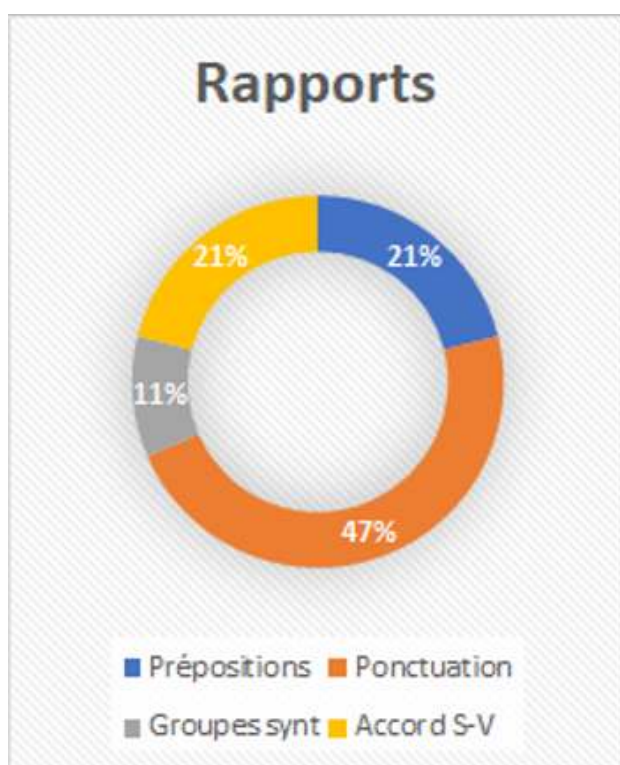
- 59 Les difficultés les plus courantes dans les exercices sont celles liées à la « ponctuation ». Celles-ci représentent 59 % des erreurs. Les erreurs de « prépositions » sont moins présentes que dans les autres écrits puisqu'elles ne représentent que 25 % des erreurs. Dans les exercices, il y a de nombreuses phrases longues sans ponctuation, les séquences montrent des problèmes de lisibilité. Ici, plusieurs pistes d'explication peuvent être avancées. D'une part, le facteur du temps limité s'ajoute à la contrainte rédactionnelle – il y a manifestement peu ou pas de place laissée à la relecture. D'autre part, on peut se demander si la nature procédurale de l'exercice n'induit pas une relative complexité du contenu à développer, ce qui expliquerait à la fois la longueur des phrases et les problèmes de ponctuation.

Figure 6. Répartition des erreurs dans les mémoires.



- 60 Dans les mémoires, 55 % des erreurs sont dues à la « ponctuation ». En revanche, nous n'avons pas rencontré de problèmes d'« accord sujet-verbe » dans les mémoires. Cependant, les erreurs de « prépositions » sont bien présentes avec 32 %, suivies par les erreurs de « groupes conjoints » avec 13 %. Dans les mémoires, la plupart des séquences problématiques sont des phrases longues pour lesquelles la compréhension est ambiguë, en raison de ponctuations mal placées ou de groupes conjoints distants.

Figure 7. Répartition des erreurs dans les rapports



- 61 Contrairement aux autres écrits, les rapports de stage contiennent les quatre types d'erreurs en nombres considérables : 47 % des erreurs sont dues à la *ponctuation*, 21 % aux « prépositions », 21 % aux « accords sujet-verbe distant » et 11 % aux « groupes conjoints ». Dans cet exercice rédactionnel, l'étudiant décrit les conditions du stage ainsi que les différentes tâches effectuées. Il est courant de trouver des phrases longues comparant différentes techniques ou tâches, où les erreurs de ponctuation et les erreurs d'accords sujet-verbe sont nombreuses.

Conclusion

- 62 Les compétences à l'écrit sont multi-dimensionnelles, elles sont le fruit de l'exécution de tâches complexes tant pour le scripteur que pour l'évaluateur. Dans cet article, nous avons cherché à montrer comment le travail de développement d'un outil automatique de détection d'erreurs syntaxiques pouvaient contribuer à mieux comprendre les difficultés de l'évaluateur à diagnostiquer l'erreur dans certains types de séquences. Pour cette recherche, menée dans le cadre du projet *écri+*, nous nous sommes concentrées sur les séquences phrastiques longues comportant des coordinations, avec parfois des ellipses, et le plus souvent des structures syntaxiques complexes. Nous pensons qu'elles sont assez caractéristiques des « écrits réflexifs » demandés à l'université, où il s'agit de produire des textes qui construisent un savoir, dont la forme doit être lisible pour que le lecteur puisse à son tour construire ce savoir. Mais ce sont des zones rédactionnelles difficiles, propices aux erreurs, que l'enseignant a souvent du mal à analyser explicitement. Ces séquences mobilisent en effet simultanément des habiletés linguistiques nombreuses. En l'occurrence nous avons pu identifier des difficultés au niveau de la valence des lexèmes mis en facteur commun pour construire

les coordinations, des problèmes dans la construction des phrases elliptiques, des difficultés de ponctuation ou encore des accords sujet-verbe problématiques. En réponse à la difficulté de l'évaluateur à commenter ou diagnostiquer l'erreur, nous avons montré que la typologie utilisée pour l'outil permet de proposer un premier repérage plus explicite. Nous avons également essayé d'apporter des éclairages à partir de l'analyse de l'accord inter-annotateur, reflétant les éventuels accords ou désaccords pouvant exister entre évaluateurs. Nous avons identifié :

- Un ensemble de structures pour lesquelles l'indécision est linguistiquement justifiée car elle tient compte de variations possibles entre locuteurs ;
 - Un ensemble de structures que les locuteurs identifient nettement comme erronées, mais pour lesquelles le repérage du type d'erreur est difficile.
- 63 Pour ces structures, il est important que le formateur soit entraîné à discrétiser les contraintes linguistiques en jeu afin d'être capable de les rendre explicites auprès de l'apprenant. Elles dépassent le cadre des compétences travaillées dans le parcours scolaire et sont propres à la rédaction d'écrits réflexifs. Les compétences fines qui sont en jeu sont traduites dans la grille du référentiel d'écrit+ que nous avons pu enrichir grâce au travail présenté ici (voir également Noreskal, Eshkol-Taravella & Desmets, 2021). Une fois implémenté, l'outil de détection automatique figurera parmi les modules de remédiation du projet. À l'issue d'un exercice de production écrite, il permettra d'attirer l'attention de l'étudiant sur les structures possiblement erronées de son texte avec des messages de retours immédiats (feedback), et l'invitera à consulter les ressources écrit+ correspondantes.

BIBLIOGRAPHIE

- ABEILLÉ, A. (2005). « Les syntagmes conjoints et leurs fonctions syntaxiques ». *Langages* 160 (4), p. 42-66.
- ARTSTEIN, R. & POESIO, M. (2008). « Inter-coder agreement for computational linguistics ». *Computational linguistics* 34 (4), p. 555-596.
- BISHOP, M.-F. & ROUXEL, A. (2007). « Présentation : Sujet lecteur, sujet scripteur, quels enjeux pour la didactique ? ». *Le français aujourd'hui* 157 (2), p. 3-8. En ligne : <https://www.cairn.info/revue-le-francais-aujourd-hui-2007-2.htm>.
- BOCH, F. & FRIER, C. (éds) (2015). *Écrire dans l'enseignement supérieur. Des apports de la recherche aux outils pédagogiques*. Grenoble : Éditions littéraires et linguistiques de l'Université de Grenoble.
- BOCH, F., RINCK, F. & SORBA, J. (2020). « Les acquisitions tardives en français écrit : une base de données sur les erreurs et maladroites à un niveau avancé », 7^e *Congrès Mondial de Linguistique Française*. En ligne : <https://hal.archives-ouvertes.fr/hal-02931262/document>.
- BOIVIN, M.-C. & PINSONNEAULT, R. (2020). « La catégorisation des erreurs linguistiques : Une grille de codage fondée sur la grammaire moderne ». *Le français aujourd'hui* 209 (2), p. 89-116. En ligne : <https://doi.org/10.3917/lfa.209.0089>.

- BUSQUETS, J. & DENIS, P. (2001). « L'ellipse modale en français : le cas de *devoir* et *pouvoir* ». *Cahiers de grammaire* 26, p. 55-74. En ligne : <http://w3.erss.univ-tlse2.fr/textes/publications/CDG/26/CG26-4-Busquets.pdf>.
- CHOMSKY, N. (1965). *Aspects of the theory of syntax*. Cambridge, The M.I.T. Press.
- COHEN, J. (1960). « A coefficient of agreement for nominal scales ». *Educational and Psychological Measurement* 20 (1), p. 37-46.
- CORDER, S. P. (1980). « Que signifient les erreurs des apprenants ? ». *Langages* 57, p. 9-15. En ligne : <https://doi.org/10.3406/lgge.1980.1833>.
- DAGNAC, A. (2008). « L'Ellipse modale en français : arguments pour une ellipse du TP ». *8^e Congrès mondial de linguistique française*, p. 2453-2465. En ligne : <https://www.linguistiquefrancaise.org/articles/cmlf/pdf/2008/01/cmlf08211.pdf>.
- DE VOGÜÉ, S. (2022). « Orthographe et compétence linguistique : ce que dit le *e de crie* ». *SHS Web of Conferences* 143. En ligne : https://www.shs-conferences.org/articles/shsconf/pdf/2022/13/shsconf_onea2022_01002.pdf.
- DELCAMBRE, I. & LAHANIER-REUTER, D. (2010). « Les littéracies universitaires. Influence des disciplines et du niveau d'étude dans les pratiques de l'écrit ». *Diptyque* 18, p. 11-42. En ligne : https://www.forumlecture.ch/myUploadData/files/2010_3_Delcambre_Lahanier.pdf.
- DEMIRTAS, L. & GÜMÜS, H. (2009). « De la faute à l'erreur : une pédagogie alternative pour améliorer la production écrite en FLE ». *Synergies* 2, p. 125-138. En ligne : <https://gerflint.fr/Base/Turquie2/lokman.pdf>.
- FLEISS, J. L. (1975). « Measuring agreement between two judges on the presence or absence of a trait ». *Biometrics* 31 (3), p. 651-659.
- FREI, H. (1971) [1929]. *La Grammaire des Fautes*. Genève : Slatkine.
- HAYES, J. R. & FLOWER, L. (1980). « Identifying the Organization of Writing Processes ». In : Gregg, L. W. & Steinberg, E. R. (éds), *Cognitive Processes in Writing: An Interdisciplinary Approach*. Hillsdale, Lawrence Erlbaum, p. 3-30.
- LANDIS, J. R. & KOCH, G. G. (1977). « The Measurement of Observer Agreement for Categorical Data ». *Biometrics* 33 (1), p. 159-174.
- MARANDIN, J. M. (1999). *Grammaire de l'incidence*. En ligne : http://www.llf.cnrs.fr/sites/llf.cnrs.fr/files/u63/grammaire_incidence.pdf.
- MARQUILLÓ LARRUY, M. (2003). *L'Interprétation de l'erreur*. Paris : Éd. Clé International.
- MOURET, F. (2003). « Conjonction de coordination et constituance en français ». *6^{es} Rencontres de l'École Doctorale de Paris III*. Paris, 24 mai. En ligne : <http://www.llf.cnrs.fr/Gens/Mouret/Mouret-RJC-2003.pdf>.
- MOURET, F. (2005). « La syntaxe des coordinations corrélatives du français ». *Langages* 160, p. 67-92. En ligne : <https://doi.org/10.3917/lang.160.0067>.
- NORESKAL, L., ESHKOL-TARAVELLA, I. & DESMETS, M. (2020). « Approche méthodologique pour l'étude des constructions elliptiques dans les rédactions étudiantes ». *SHS Web of Conferences* 78. En ligne : https://www.linguistiquefrancaise.org/articles/shsconf/pdf/2020/06/shsconf_cmlf2020_06007.pdf.

NORESKAL, L., ESHKOL-TARAVELLA, I. & DESMETS, M. (2021). « Erroneous Coordinated Sentences Detection in French Students' Writings ». *Communications in Computer and Information Science* 1463, p. 586-596.

PAUMIER, S. (2011). *Unitex. Manuel d'utilisation*. Trad. de l'anglais par C. Martineau. En ligne : <https://hal.archives-ouvertes.fr/hal-00639621/document>.

PERDUE, C. & PORQUIER, R. (1980). « L'analyse des erreurs : un bilan pratique ». *Langages* 57, p. 87-94. En ligne : <https://doi.org/10.3406/lgge.1980.1840>.

PERRENOUD, P. (2000). « L'approche par compétences, une réponse à l'échec scolaire ? ». AQPC *Réussir au collégial*. En ligne : http://www.unige.ch/fapse/SSE/teachers/perrenoud/php_main/php_2000/2000_22.html.

REICHLER-BÉGUELIN, M.J. (1992). « L'approche des "anomalies" argumentatives ». *Pratiques* 73 (1), p. 51-78. En ligne : <https://doi.org/10.3406/prati.1992.1656>.

RICCI, L. (1996). « Le dialogue interrogatif ou les deux faces d'une liberté didactique ». *Les Carnets du Cediscor* 4, p. 131-151. En ligne : <https://doi.org/10.4000/cediscor.412>.

Université Ouverte des Humanités et Université de Strasbourg, s. d., *Écrit+. Évaluer - former, certifier. Français*. En ligne : <https://ecriplus.fr/>.

Université Ouverte des Humanités et Université de Strasbourg, s. d., *Écrit+ Test*. En ligne : <https://app.tests.ecriplus.fr/>.

Université Paul-Valéry Montpellier 3 et Université Paris Nanterre, s. d., *J'améliore mon français écrit*. En ligne : http://uoh.univ-montp3.fr/j_ameliore_ma_maitrise_du_francais/portail/.

VESLIN, O. & VESLIN, J. (1992). *Corriger des copies : évaluer pour former*. Paris : Hachette.

NOTES

1. Le projet écrit+ (ANR-17-NCUN-0015) est un dispositif mis en place pour aider les étudiant·es à développer leurs compétence à l'écrit. Plus d'information à l'adresse suivante : <https://ecriplus.fr/>.
2. Thèse en cours de Laura Noreskal sur la détection automatique d'erreurs dans les constructions coordonnées. L'outil a pour vocation d'alerter l'étudiant lorsque son texte contient des séquences potentiellement erronées.
3. Les dispositifs de formation s'appliquent à un enseignement en présentiel et à distance, on peut citer, entre autres outils, une trentaine de modules de formation *J'améliore mon français écrit*, des MOOC, des tests de positionnement, une plateforme d'auto-apprentissage écrit+test (plus de 600 questions) dont les résultats cumulatifs permettent la certification progressive de l'apprenant.
4. Une plate-forme permettant le recueil des écrits étudiants a aussi été développée afin d'identifier les erreurs récurrentes dans les différents champs disciplinaires (Boch, Rinck & Sorba, 2020).
5. Ceci diffère d'une démarche « descendante » où l'on passerait d'abord par l'apprentissage de la règle pour travailler hors sol des fondamentaux de langue, comme c'est parfois le cas dans le cadre scolaire.

6. Pour la plupart, les étudiants n'ont pas de difficultés majeures lorsqu'ils doivent réaliser une phrase simple (et non un texte), et à l'oral ils montrent les compétences requises pour des conversations ordinaires.
 7. Les notions d'erreur et de faute sont souvent au centre des débats. Pour Pit Corder (1980), la dichotomie erreur/faute se rapporte à la dichotomie performance/compétence de Chomsky (1965).
 8. Un corpus écologique est un corpus naturel qui n'a pas été créé dans un but de recherche. Par conséquent, les données sont brutes, les erreurs n'ont pas été corrigées.
 9. La méthodologie de constitution du corpus a été expliquée dans l'article de L. Noreksal, I. Eshkol-Taravella & M. Desmets (2020), nous ne la développons pas avec exhaustivité ici.
 10. Pour chaque écrit, le niveau d'étude (L1 à M2) ainsi que le domaine d'étude (Sciences du langage, Histoire, Littérature, Psychologie, master Meef) ont été gardés en tant que métadonnées pour de futures analyses.
 11. L'annotation décrite ici est une tâche propre au traitement automatique du langage naturel qui ne doit pas être confondue avec « les annotations » effectuée par l'évaluateur lors de la correction d'une copie d'élève ou d'étudiant (au sens de Veslin & Veslin, 1992).
 12. Pour rappel, nous précisons bien qu'un couple comme *correcte/erronée* correspond à un jeu d'étiquettes utilisé par l'outil, qui a vocation de repérer dans un texte les zones potentiellement problématiques. Ces étiquettes ne correspondent pas directement aux messages visibles par l'utilisateur.
 13. Une ellipse polaire correspond à l'omission du syntagme verbal dans le second conjoint (verbe principal et compléments). Les éléments omis sont remplacés par un adverbe de polarité, par exemple : « Pierre viendra mais pas Marie ». L'ellipse modale est une résolution par proforme verbale ou par un auxiliaire modal sans argument, par exemple « Il a fait tout ce qu'il a pu ».
 14. Aussi désigné « κ de Cohen ».
 15. Pour des raisons multiples, la forme du conjoint n'est pas celle qui est attendue.
-

RÉSUMÉS

Parmi les dispositifs mis en œuvre par le projet *écri+* pour la maîtrise du français écrit à l'université, nous développons un outil d'aide à la formation permettant d'identifier automatiquement des erreurs de constructions syntaxiques. Les données que nous analysons, issues de rédactions d'étudiants, sont plus spécifiquement les séquences phrastiques longues comportant des coordinations, avec ou sans ellipses. Confrontés à la problématique de l'évaluation de l'erreur, nous montrons comment une approche en traitement automatique du langage peut contribuer à identifier les compétences rédactionnelles mobilisées et à mieux comprendre les différentes postures de l'évaluateur selon les types d'erreurs.

Among the devices developed by the *écri+* project for the improvement of the written French skills at university, we are currently implementing a training aid tool that automatically detects errors in syntactic constructions. The data that we analyze, from student writings, are complex and long sentences including coordinations, with or without ellipses. Faced with the issue of error evaluation, we show how an automatic language processing approach can help identify the drafting skills mobilized and better understand the different postures of the evaluator according to the types of errors.

INDEX

Keywords : Coordinated constructions, syntactic errors, students' writings, evaluation, annotation, corpus linguistics, NLP

Mots-clés : Constructions coordonnées, erreurs syntaxiques, rédactions universitaires, évaluation, annotation, linguistique de corpus, TAL

AUTEURS

MARIANNE DESMETS

Université Paris Nanterre, LLF, F- 92000 Nanterre, France

LAURA NORESKAL

Université Paris Nanterre, Modyco, F- 92000 Nanterre, France

SARAH DE VOGÜÉ

Université Paris Nanterre, Modyco, F- 92000 Nanterre, France

IRIS ESHKOL-TARAVELLA

Université Paris Nanterre, Modyco, F- 92000 Nanterre, France