



HAL
open science

Traitement automatique des interventions des enseignants en CE2 et en CM2 en France et en Italie : entre grammaire et contenu

Sara Mazziotti

► To cite this version:

Sara Mazziotti. Traitement automatique des interventions des enseignants en CE2 et en CM2 en France et en Italie : entre grammaire et contenu. *Bellaterra Journal of Teaching & Learning Language & Literature* , 2020, 13, pp.61 - 76. <10.5565/rev/jtl3.816>. <halshs-04127452>

HAL Id: halshs-04127452

<https://shs.hal.science/halshs-04127452v1>

Submitted on 13 Jun 2023

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.



HAL Authorization



Traitement automatique des interventions des enseignants en CE2 et en CM2 en France et en Italie : entre grammaire et contenu

Automatic processing of teachers' corrections in third grade and fifth grade in France and Italy: between grammar and content

Sara Mazziotti
Université Sorbonne Nouvelle – Paris 3
Università di Bologna



Résumé

La pratique de correction d'un texte écrit a un rôle essentiel dans l'apprentissage de l'écriture de la part de l'élève et elle peut varier d'une façon consistante selon les habitudes de l'enseignant, le type de consigne, le niveau de l'élève et le système linguistique de référence. Mais quels sont les aspects linguistiques et textuels que les enseignants de l'école primaire italiens et français corrigent-ils plus particulièrement ? À partir d'un corpus d'écrits d'élèves de CE2 et de CM2 recueilli en France et en Italie, nous investiguerons les pratiques de correction des enseignants en focalisant ici notre attention sur les méthodes de correction de la grammaire. Les différents types d'interventions des enseignants sont calculés automatiquement à l'aide d'outils informatiques qui les répartiront selon le type d'opération (ajout, suppression, remplacement, commentaire et soulignement). Le but étant d'évaluer, d'une part, le taux d'intervention et la fréquence des différents types de correction chez les enseignants italiens et français et, d'autre part, la prise en compte de toutes ces interventions de la part des élèves.

Mots clés : Pratiques de correction ; Écriture scolaire ; Génétique textuelle ; Transcription

Abstract

Correction practices of a written text have an essential role in the student's learning to write and can vary consistently depending on the teacher's habits, the writing rules, the student's level and the language system of reference. But what linguistic and textual aspects do Italian and French primary school teachers correct in particular? Based on a corpus of writings from third and fifth graders collected in France and Italy, we will investigate teachers' correction practices by focusing our attention here on grammar correction methods. The different types of teacher interventions are automatically calculated using computer tools that will divide them according to the type of operation (addition, deletion, replacement, comment and underlining). The aim is to evaluate, on the one hand, the intervention rate and frequency of the different types of correction among Italian and French teachers and, on the other hand, the consideration of all these interventions by students.

Keywords: Correction practices; School writing; Textual genetics; Transcription

UN INTERET CROISSANT POUR LES CORPUS SCOLAIRES

Cette étude s'inscrit dans une pluralité de recherches menées sur des corpus scolaires et sur leurs transcriptions ayant l'objectif de répondre à des questionnements didactiques et pédagogiques que les linguistes, les enseignants et les chercheurs se posent d'une façon transversale. Si au début la génétique textuelle était appliquée en particulier aux corpus littéraires, pendant ces trente dernières années, elle s'est intéressée aux productions scolaires (Fabre-Cols, 1990) présentant les mêmes opérations d'écriture des manuscrits d'auteurs. Qu'est-ce qui se cache derrière à une rature du point de vue de la conception du texte : est-elle synonyme d'erreur ou pas ? (Penloup, 1994). Plusieurs groupes de recherche en France s'intéressent en effet à différents aspects des textes d'élèves : pour en citer quelques-uns, le type d'erreurs d'orthographe, l'emploi de certains temps verbaux et plus en général, la construction du processus d'écriture, mise en relief par la transcription semi-génétique. Le projet E-calm par exemple, dirigé par Claire Doquet et financé par l'ANR, réunit quatre laboratoires (CIRCEFT-ESCOL, LIDILEM, CLESTHIA et CLLE-ERSS) qui collaborent pour constituer, mettre à disposition de la communauté scientifique et étudier un vaste corpus d'écrits d'élèves et d'étudiants, en vue d'« analyses sociologiquement contextualisées »¹. L'un des objectifs fondamentaux de cette coercion est de mettre au point un protocole de transcriptions et d'annotations de ces écrits qui signale toutes les procédures d'écriture qui ont précédé l'état final du texte. Les spécificités d'un corpus d'écrits scolaires sont en fait nombreuses : il s'agit de productions écrites non normées qui nécessitent une couche supplémentaire d'annotation incluant la correction des erreurs d'orthographe, afin de pouvoir repérer automatiquement, par exemple, les catégories grammaticales employées par l'élève. De plus, ces écrits peuvent porter les traces d'une écriture collective ou d'un premier jet de l'élève qui est ensuite corrigé et modifié par l'enseignant. Il est important alors de pouvoir, d'une part, garder dans les transcriptions l'évolution chronologique des différentes opérations d'écriture et, d'autre part, de signaler qui est le scripteur à l'origine de ces opérations (l'élève-scripteur, un pair ou l'enseignant).

Les protocoles de transcription peuvent changer d'une équipe à l'autre sur la base des objectifs de recherche et des spécificités du corpus qui peut inclure ou pas les interventions de l'enseignant. Dans notre étude, nous nous sommes appuyés sur le protocole employé actuellement au sein du projet EcriScol (Doquet, Paris 3)

¹ Description du projet E-CALM financé par l'ANR : <http://e-calm.huma-num.fr/le-projet/>

à partir duquel nous avons annoté manuellement nos écrits d'élèves corrigés. L'évolution génétique du texte est mise en relief par le biais de balises qui délimitent tous les éléments ajoutés, supprimés et remplacés par l'élève et toutes les interventions de l'enseignant apportées sur la première version du texte de l'élève.

OBJECTIFS DU PROJET ET HYPOTHESES DE BASE

L'intérêt de cette recherche porte sur les pratiques de correction des enseignants français et italiens dans le cadre d'une tâche de réécriture. Nous aurons l'occasion ici d'encadrer le début de l'analyse des textes d'élèves recueillis et de proposer quelques traitements informatiques des transcriptions vers l'identification de plusieurs profils d'enseignant-correcteur. Nous rejoignons l'hypothèse confirmée par le linguiste Jean Luc Pilorgé (2010) selon laquelle chaque enseignant pourrait adopter plusieurs « postures de correction » au sein d'une classe ou d'une même copie. Plusieurs facteurs pourraient inciter les enseignants à privilégier un critère de correction plutôt qu'une autre (type de consigne, niveau de l'élève), mais notre objectif principal au moment du recueil était de capturer et d'analyser leurs manières de corriger à partir d'un protocole et d'une consigne uniques. Même s'ils avaient comme seule consigne de corriger la première version du texte selon leurs propres méthodes habituelles, nous sommes conscients du fait qu'il s'agissait pour eux et pour les élèves également d'un exercice proposé par un observateur externe. Nous évoquons en fait toute une série de risques liés à l'analyse de corpus non écologiques recueillis à partir d'un même protocole pour tous les élèves qui pourrait ne pas refléter les pratiques réelles d'écriture et de correction propre à chaque classe. Cependant, notre protocole de recueil unique pour toutes les classes, nous a permis de collecter des productions rédigées pendant la même durée de temps et à partir d'une même consigne, facile à traduire et présentée en classe par l'observatrice.

Les hypothèses à la base de notre recherche consistent en l'idée que les systèmes linguistiques italien et français conditionnent le nombre et le type d'interventions de l'enseignant sur la première version écrite des élèves. Nous avons précédemment constaté à la suite d'une étude menée dans deux classes françaises de CM2 qu'un nombre d'interventions de l'enseignant très élevé et focalisé surtout sur l'aspect orthographique et grammatical du texte inhibait le travail de réécriture de l'élève lors du passage de la première version à la version définitive. Nous avançons donc l'hypothèse qu'un système linguistique ayant une correspondance entre graphème et phonème plus solide, comme dans le cas de la langue italienne, puisse amener les enseignants à proposer davantage des commentaires sur le contenu et sur les modalités de réécriture, vers un texte de meilleure qualité. Un nombre

d'erreurs d'orthographe plus élevé en contexte français amènerait les enseignants à insister sur l'aspect linguistique du texte et à proposer assez rarement de commentaires portant sur d'autres aspects du texte.

Pendant notre analyse, nous rechercherons les cinq « postures de correction » proposées par J-L. Pilorgé (2010) dans notre corpus pour constater si en contexte italien elles varient ou pas et pour en identifier les plus fréquentes parmi nos enseignants. Deux « postures de correction » en particulier seront le plus souvent évoquées ici : le « gardien du code » et le « lecteur-naïf ». Les trois autres postures (« stimulus-réponse, « éditeur » et « critique ») sont prises en compte dans notre étude, mais elles ne seront pas représentées ici. L'enseignant qui adopte la posture du « gardien du code » se focalise sur la langue et ne corrige que l'orthographe, la grammaire, l'emploi des temps verbaux, les répétitions. L'enseignant « lecteur-naïf » en revanche met en question « la représentation du monde » de l'élève et ses « lois du réalisme » (Pilorgé, 2010) à travers des questions et des commentaires qui portent le plus souvent sur la phrase et non pas sur un mot précis.

LE RECUEIL D'ECRITS D'ELEVES EN VUE DES TRANSCRIPTIONS

Le choix d'effectuer le recueil à partir d'une consigne de réécriture est justifié par une volonté d'évaluer le taux de remaniement du texte de la part des élèves et le relier ensuite à la fréquence des différents types de correction des enseignants que nous distinguerons. Des classes de CE2 et des classes de CM2 (troisième et cinquième année d'école primaire en Italie) de quatre écoles, deux italiennes et deux françaises doivent rédiger la consigne « Que feras-tu quand tu seras adulte ? Raconte une de tes journées ».

Les élèves ont une heure à disposition pour concevoir et écrire leur texte de façon autonome, avec comme seule contrainte l'interdiction de blanco ou de gomme. Notre objectif étant de pouvoir repérer et transcrire à la fin l'aspect génétique des textes et d'annoter tous les éléments ajoutés, supprimés, remplacés, etc. L'observatrice lit à voix haute la consigne, reportée plus haut, aux élèves, même si l'enseignant est lui aussi présent en classe pour pouvoir répondre aux doutes des élèves et pour les aider en cas de blocage. Tous les enseignants ensuite corrigent les textes selon leurs pratiques de correction et les distribuent aux élèves après une semaine environ pour que les élèves regardent les corrections et proposent une version définitive. Nous avons demandé d'assister aux deux heures de rédaction dans chaque classe pour pouvoir suivre le bon déroulement de l'épreuve et pour récolter les métadonnées du devoir (affiches en classe, dictionnaires à disposition des

élèves). Notre corpus est constitué de 158 copies italiennes et de 103 copies françaises qui incluent les premiers jets et les deuxièmes jets.

Le protocole de transcription qui a été adopté est très précis et détaillé et vise à la transcription de dessins, de schémas et de toute intervention de la part de l'enseignant. Comme nous l'avons anticipé, il s'agit du protocole actuellement utilisé par le groupe de recherche EcriScol, qui continue à le perfectionner pour résoudre en particulier des cas d'accumulation d'opérations d'écriture qui sont difficiles à transcrire et qui compliquent le traitement automatique des copies (Doquet, Enouï, Fleury et Mazziotti, 2017). En effet, nous ne nous limitons pas à transcrire le texte avec toutes les procédures d'écriture mises en évidence : nous y ajoutons également la forme normée des erreurs orthographiques des élèves qui permet ensuite d'étudier les différents types d'erreurs (David et Doquet, 2016). Nous proposons les balises principales qui constituent notre protocole :

- //élément ajouté// : par l'élève ou par le Prof //T2P#élément ajouté par le Prof//
- [élément supprimé] : par l'élève ou par le Prof [T2P#élément supprimé par le Prof]
- {T2P#commentaire du Prof}
- {T2Pσ#élément souligné par le Prof}
- £ : fin de ligne, renvoi involontaire
- § : fin de paragraphe, renvoi à la ligne volontaire
- <erreur de l'élève>_<correction du transcripteur>

Avant de procéder avec notre analyse, il ne serait pas inutile de présenter un extrait de notre corpus italien et français. Ici, une production écrite d'un élève de CE2 français (Figure 1) et d'un élève de CE2 italien (Figure 2).

Dans ce premier exemple, nous signalons le premier jet de l'élève au crayon et puis ensuite au stylo bleu, même si nous les considérons comme un seul temps d'écriture, puisque l'enseignant corrige visiblement les deux parties en rouge, donc pendant le temps de correction successif à la première rédaction de l'élève. Nous attirons l'attention sur les suppressions fréquentes apportées par l'enseignant et souvent suivies d'un ajout : dans ce cas, l'opération d'écriture qui apparaîtra dans nos transcriptions est le remplacement d'un seul mot (ligne 4, ®[T2P#se]//T2P#ce//®) ou de plusieurs mots (ligne 4, ®[T2P#J'ai pleurer]//T2P#Je pleurerai//®). Un trait rouge supprime le mot ou les mots qui doivent être corrigés et sans identifier l'erreur à l'aide d'abréviations, l'enseignant propose la forme normée. Il s'agit d'une première méthode de correction de la grammaire qui est

Figure 1. Extrait d'une copie d'un élève de CE2 corrigé par un enseignant français

Production écrite : Quand je serai adulte.

Consigne : Que feras-tu quand tu seras adulte? Imagine et raconte une de tes journées.

Je voudrais faire Maîtreresse de l'École. J'ai 27 ans et 4 h
 serai
 après le travail, je vais à la piscine. Mais au travail,
 le premier jour je fait enfant une petit fille un peut
 méchant rien me taper! J'ai pleurer mes mais ce n'est pas
 pas grave. J'ai écrit à sainte Lint. croix du Mont. L'école
 où je travaille s'appelle t. La route est un peut
 long. Il pour aller à l'école depuis la maison. Un
 petit garçon à venir j'ai appelé sae maman
 au téléphone, sa mamen ne peut pas venir elle
 travaille.

* Un jour, va là ce qui s'est passé.

elle vient taper la maîtresse?

appliquée à toutes les catégories grammaticales d'une manière transversale (nom, verbe, adjectif) et qui parfois cache complètement le premier choix de l'élève : c'est le cas de la surcharge, c'est-à-dire quand l'enseignant rend presque illisible ce que l'élève avait écrit (ligne 7 long@[T2P# t]//T2P#ue//@). En revanche, un commentaire en marge inséré en correspondance d'un segment souligné par l'enseignant ({T2P#elle vient de taper la maîtresse ?}) pourrait être associé à la posture du « lecteur naïf » proposée par Pilorgé, utile pour amener l'élève à réfléchir sur sa propre « représentation du monde » (2010). En bas de page, un ajout de contenu de la part de l'enseignant (ligne 7) que nous resituons dans le texte au moment de la transcription, selon les critères d'une transcription linéarisée semi-diplomatique. Étant le traitement informatique notre objectif principal pendant la conception du protocole et la transcription de ces écrits.

Dans ce deuxième exemple de copie (Figure 2), l'élève rédige son texte au crayon et en lettres capitales, choix partagé par plusieurs élèves de CE2 de notre corpus italien, probablement plus à l'aise qu'avec l'écriture en cursive. Si nous regardons les corrections apportées par l'enseignant italien, il est intéressant d'insister en revanche sur le commentaire global à la fin du texte et associé généralement à

Figure 2. Extrait d'une copie d'un élève de CE2 corrigé par un enseignant italien


Produzione scritta : Quando sarò grande.

Consegna : Cosa farai quando sarai grande? Immagina e racconta una delle tue giornate.

ERA INVERNO E IO AVEVO 35 ANNI, C'ERA IL SOLE E FACEVA MOLTO CALDO, IO ERO A LONDRA E STAVO CURANDO UN CAVALLO. IN FATTI SONO UN ~~VETERINARIO~~ VETERINARIO, E IL MIO CAVALLO AVEVA MAL DI PANCIA E IO DOVEVO ~~CURARLO~~ CURARLO. ERANOTTE ^{DOPO AVERLO CURATO} E IO DOVEVO ANDARE A CENA FUORI CON I MIEI AMICI E AMICHE.

Buona cena,

dopo aver curato il cavallo te la meriti!



un smiley, une combinaison assez récurrente chez cet enseignant. Nous ne pouvons pas considérer ce commentaire comme une appréciation globale du texte du point de vue qualitatif, mais nous pourrions plutôt l'interpréter comme une validation du contenu. L'élève a raconté dans le texte qu'il était vétérinaire et qu'il allait soigner son cheval avant d'aller dîner avec les amis. Nous ne nous attardons pas ici sur l'emploi des temps verbaux surtout au passé qui n'a pas été commenté par l'enseignant, mais nous proposons la traduction du commentaire final pour éclairer le type de stratégie de l'enseignant en vue de la réécriture : « Bon dîner, après avoir soigné le cheval, tu le mérites ! ». En revanche, l'enseignant applique la même méthode de correction des aspects orthographiques à travers l'utilisation de remplacements en rouge : il y a une seule correction orthographique d'une erreur de segmentation (ligne 4, « infatti ») et un seul remplacement d'un autre mot mal segmenté qui n'est pas gardé en tout cas par l'enseignant qui propose une autre formulation (lignes 7 et 8, ®[T2P#eranotte e io]/T2P#Dopo averlo curato//®²). Les enseignants italiens insistent-ils davantage sur le contenu et sur la globalité du texte ou tout simplement corrigent-ils moins d'erreurs d'orthographe, car un nombre inférieur d'erreurs est présent dans les copies italiennes ? Nous n'aurions pas ici l'occasion de présenter l'analyse des interventions de tous nos enseignants, notre

² « En effet » a été remplacé par « Après l'avoir soigné » (le cheval).

traitement des données n'étant qu'au début, mais nous montrerons comment répondre à ce type de questions à travers l'exploitation informatique des annotations.

LE DEGRE D'INTERVENTION DES ENSEIGNANTS PENDANT LA CORRECTION

Nombreuses sont les différentes étapes de notre analyse qui vise à prendre en compte également les métadonnées de la classe (spécificités du manuel de langue, nombre d'enfants étrangers) et du devoir (affiches en classe, outils à disposition pendant la rédaction) que nous avons recueillies. Nombreux également les questionnements que nous pourrions nous poser à partir d'un corpus d'écrits scolaires corrigés par des enseignants italiens et français et auxquels l'on pourrait répondre à l'aide d'outils informatiques. Nous faisons ici alors le choix de mettre en avant, d'une part, l'un des scripts générés pour exploiter notre corpus et qui permet d'extraire et de classer automatiquement toutes les interventions des enseignants ; et, d'autre part, l'un des logiciels qui nous donne la possibilité de mettre en comparaison les deux versions de l'élève pour évaluer le taux de remaniement du texte après l'intervention de l'enseignant.

Pour pouvoir étudier les corrections apportées par les enseignants italiens et français et en particulier les corrections sur la grammaire, nous nécessitons d'abord l'extraction automatique de toutes les interventions et leur mise en relation avec le nombre total de copies corrigées, afin de calculer le nombre moyen d'interventions par copie. Voici les catégories dans lesquelles notre script classe les interventions et qui correspondent aux opérations annoncées dans notre protocole :

- Nombre d'interventions total de l'enseignant ;
- Nombre de remplacements ;
- Nombre de soulignements ;
- Nombre de commentaires ;
- Nombre de suppressions ;
- Nombre d'ajouts.

Il est important de préciser que dans la catégorie « remplacements » deux opérations sont calculées : la suppression et l'ajout. En revanche, dans les catégories « suppressions » et « ajouts » nous calculons les occurrences de ces deux opérations qui n'ont pas été déjà calculées dans les remplacements, c'est-à-dire des suppressions qui ne sont pas suivies d'un ajout ou des ajouts qui ne sont pas précédés d'une suppression. Comme nous l'avons déjà précisé, pour pouvoir les extraire, ces opérations ont été précédemment annotées manuellement selon les modalités

brièvement présentées plus haut. À partir du nombre de corrections moyen par copie retrouvées dans notre corpus, nous avons définis trois niveaux d'intervention, en essayant de constituer trois groupes ainsi répartis :

- Bas : moins de 18,2 interventions par copie (cinq enseignants, dont un seul français) ;
- Moyen : entre 18,2 et 34,9 interventions par copie (quatre enseignants, dont un seul français) ;
- Élevé : plus de 34,9 interventions par copie (six enseignants, dont un seul italien).

Figure 3. Capture d'écran du script utilisé pour calculer le nombre d'interventions d'un enseignant français de CE2 avec un taux d'intervention moyen.

```
Nombre d'interventions enseignant: 349
Nombre remplacements: 111
Nombre de soulignements: 19
Nombre de commentaires: 7
Nombre de suppressions: 26
Nombre d'ajouts: 75
```

Dans la Figure 3, nous présentons une capture d'écran du script qui vient d'être exécuté : il affiche le nombre d'interventions d'un enseignant français de CE2 qui a corrigé les textes de dix élèves, parmi lesquels nous retrouvons le jeune scripteur du texte en Figure 1. Le nombre moyen d'interventions est de 34,9, c'est-à-dire un taux moyen au seuil d'un taux élevé selon notre classement. Le remplacement est l'opération la plus fréquente avec 111 occurrences, suivi par l'ajout qui compte 75 occurrences. Les commentaires sont en revanche 7 sur 349 opérations totales, donc comme nous l'avons vu dans le manuscrit, il est possible que cet enseignant ait fourni quelques commentaires en marge ou en bas de page qui portent plutôt sur la globalité du texte. En ce qui concerne l'enseignant italien de CE2 qui a corrigé la copie présentée dans la Figure 2, nous obtenons les résultats suivants à l'aide du même script :

Nombre d'interventions enseignant : 491

Nombre remplacements : 111

Nombre de soulignements : 29

Nombre de commentaires : 45

Nombre de suppressions : 58

Nombre d'ajouts : 134

Le nombre d'interventions total, c'est-à-dire le nombre total d'occurrences de l'indication « T2P », est 491 sur 21 copies, ce qui correspond à un nombre moyen d'interventions par copie de 23,3. Cet enseignant aussi a donc un taux d'intervention moyen, mais avec un chiffre beaucoup plus proche au niveau bas que l'enseignant français. Le nombre de remplacements est le même, mais le nombre de commentaires est plus élevé, ainsi que le nombre d'ajouts. Une attention particulière à¹ l'aspect linguistique est donc attendue chez cet enseignant également, malgré son habitude à apporter un commentaire global à la fin de chaque copie de ses élèves.

Sur la base de notre classement en trois groupes, nous constatons toutefois que la quasi-totalité des enseignants français ont un taux d'intervention moyen ou élevé, alors que presque la moitié des enseignants italiens a un taux d'intervention bas et un seulement a un taux élevé. Dans les deux contextes linguistiques, les interventions les plus fréquentes apparaissent en correspondance de l'erreur, donc elles ont une portée locale (Pilorgé, 2008 ; Doquet, 2011). La catégorie « nombre de commentaires » est la seule qui pourrait inclure des interventions du « lecteur-naïf », puisque le « gardien du code » agit localement sur le mot ou la lettre écrits de manière erronée. Cependant, nous avons annoté en tant que commentaires également les abréviations du type (orth, gr, MD, lex, etc...³), qui apparaissent souvent associées au soulignement de l'erreur. Il demeure alors important de vérifier combien parmi tous les commentaires sont des appréciations globales qui portent sur tout le texte, et combien sont en revanche des abréviations, qui identifient l'erreur et ne se réfèrent qu'à un endroit précis de la phrase ou du texte. Cette distinction ne pourra pas être effectuée automatiquement, puisque comme nous l'avons précisé, dans notre protocole nous n'avons pas signalé de manière différente les commentaires de type « abréviation » et les commentaires sur la globalité du texte. Nous avons alors extrait et divisé tous les commentaires d'une façon manuelle dans deux groupes : les commentaires avec portée locale et les commentaires avec portée globale. Dans le premier groupe, nous avons classé toutes les abréviations qui indiquent à l'élève le type d'erreurs (orth, gr, etc...) et les commentaires du type « mal dit » ou « cela est hors sujet » généralement dans la marge latérale de la feuille. Les commentaires qui portent en revanche sur l'intégralité du texte seront classés dans le deuxième groupe : nous nous référons ici aux appréciations globales qui valident le contenu du texte, qui donnent des conseils à l'élève pour améliorer la production

³ Orthographe, grammaire, mal dit, lexique.

écrite ou qui reprennent les éléments auxquels il devrait faire attention pendant la réécriture.

Sur 91 commentaires totales des enseignants italiens, 23 sont commentaires locaux alors que 68 portent sur la globalité du texte. En revanche, chez les enseignants français, sur 630 commentaires totales, 583 ont une portée locale et uniquement 47 commentaires portent sur le texte entier. Nous en déduisons que les enseignants français utilisent majoritairement la procédure du commentaire pour signaler des erreurs de grammaire et d'orthographe, alors que les enseignants italiens privilégient ce type d'intervention pour commenter le texte, pour valider son contenu ou pour suggérer des pistes d'amélioration du texte. Dans le tableau suivant quelques exemples de commentaires tirés du corpus français et du corpus italien (table 1).

Table 1. Tableau présentant quelques exemples de commentaires avec portée locale et avec portée globale chez les enseignants français et italiens

Enseignants	Portée locale	Portée globale
Italiens	{T2P#sei passata al tempo passato remoto meglio continuare con il presente}* {T2P#?}	{T2P#tanti elenchi, qualche errore, ma nel complesso benino!}** {T2P#Devi usare il futuro perché é...quando sarò grande} {T2P#Che inseguimento!}
	{T2P#Maj} {T2P#connecteur} {T2P#inutile et cela n'est pas logique} {T2P#mal dit}	{T2P#Parle un peu plus de cette équipe (tes amis?), Contre qui feras-tu des matchs? Combien de matchs? Il y aura du public? Explique un peu plus tes soirées.}

* Traduction du commentaire italien avec portée locale : {T2P#Tu es passée au temps passé simple, il vaut mieux continuer avec le présent}

** Traduction des trois commentaires italiens avec portée globale : {T2P#Beaucoup de listes, quelques erreurs mais bien dans l'ensemble} ; {T2P#Tu dois utiliser le futur car c'est...quand je serais grand} ;{T2P#Quelle poursuite !}

FOCUS SUR DEUX POSTURES DE CORRECTION : LE « GARDIEN DU CODE » ET LE « LECTEUR-NAÏF »

Nous nous concentrons maintenant sur ces deux postures de correction avancées par Jean-Luc Pilorgé qui montrent deux méthodes complètement différentes d'intervention dans le texte. La posture du « gardien du code » nous permettra d'étudier les pratiques de correction de la grammaire : selon Pilorgé, l'enseignant qui adopte cette dernière posture « se projette dans un rôle uniquement technique [...]. Il corrige le texte en privilégiant le niveau de surface (ponctuation, mise en paragraphes, orthographe principalement) pour le rendre conforme à un modèle qui n'est pas envisagé comme devant être explicité » (2010, p. 297). Ce n'est qu'à l'intérieur de cette posture que nous devons alors rechercher toutes les différentes

méthodes de correction de la grammaire : biffure suivie d'un ajout ou simple biffure ? Soulignement du mot mal orthographié ou soulignement suivi de l'ajout de la forme normée ? Pour n'en citer que certaines parmi les nombreuses opérations associées au « gardien du code ». Sur la base de l'étude menée précédemment et que nous avons évoquée plus haut, les effets d'une intervention massive du « gardien du code » auprès des élèves pendant la phase de réécriture peuvent être de deux types : l'abandon du premier jet très massivement corrigé par l'enseignant en faveur de la rédaction d'un texte *ex novo* ou la tentative d'une mise au propre de la première version pour accepter toutes les corrections de l'enseignant.

Le « lecteur-naïf » est en revanche déjà considéré tout d'abord un lecteur : il s'éloigne de la langue pour s'intéresser au contenu à travers ces commentaires. La « représentation du monde » de l'élève est mise en question pour lui signaler une incongruence souvent sous forme de question (Pilorgé, 2010, p. 300). Dans la Figure 4, un extrait de copie d'élève de CE2 corrigé par l'enseignant italien précédemment étudié. En marge, un commentaire qui succède un soulignement dans le texte de la part de l'enseignant et qui insiste sur l'idée que depuis New York il est compliqué d'arriver à Rio De Janeiro en train : un premier smiley triste y est associé, mais toujours sans biffer ni critiquer le choix de l'élève, l'enseignant renforce son message en bas de page, lui propose de prendre l'avion et y ajoute un smiley

Figure 4. Extrait d'une copie d'un élève de CE2 corrigé par un enseignant italien

HO PRESO L'AEREO E SONO ANDATO A NEW YORK
DOVE HO VISTO LA STATUA DELLA LIBERTA'
~~HO ANCHE VISTO UN TAXI~~ ^{IL PAESANO} WHO
CHIAMATO E SONO ANDATO ALLA STAZIONE
HO PRESO IL TRENO E SONO ANDATO
A RIO DE JANEIRO DOVE HO COMPRATO UN
UCCELLO TROPICALE, SONO ANDATO
IN SPIAGGIA POI SONO SALITO SU UNA
MONTAGNA E HO USATO IL DELTA PLANO
E DOPO AVER VOLATO HO MANGIATO FRUTTA
TROPICALE. HO PRESO LA BARCA NAVE
SONO TORNATO IN ITALIA HO CHIAMATO IL
TAXI SONO TORNATO A CASA HO CENATO
MI SONO MESSO IN PIGIAME E SONO ANDATO A DORMIRE.

Im treno
fai
molto
fatica!
☹☹

Mamma mia! Un viaggio di mesi!
Forse ti conviene usare l'aereo. 😊

souriant. La « naïveté » de ce lecteur-correcteur demeure dans le fait qu'il ne pointe pas le doigt contre le scripteur du texte, mais l'invite à réfléchir sur sa propre représentation du monde. Comme nous l'avons constaté dans les Figures 1 et 5, il ne s'agit pas vraiment d'une posture qui peut être adoptée d'une façon exclusive, puisque dans les deux cas, les enseignants choisissent de corriger également la grammaire, la ponctuation, l'orthographe.

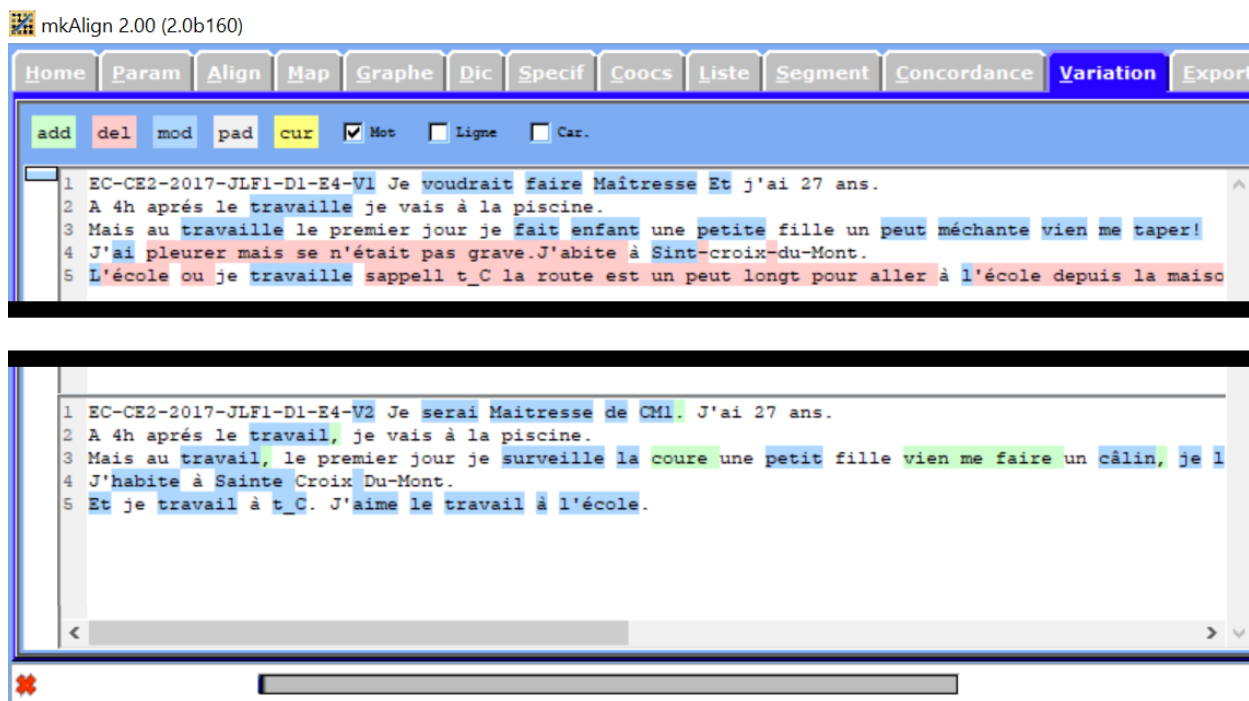
LA PRISE EN COMPTE DES INTERVENTIONS DES ENSEIGNANTS

La deuxième phase de notre analyse, qui a l'objectif d'évaluer la prise en compte des interventions des enseignants lors du passage de la première version au texte définitif, nous permet d'identifier les postures de correction les plus favorables à la réécriture. Quelles sont les postures de correction qui favorisent le plus le remaniement du texte de la part de l'élève et qui l'amène à modifier non seulement la grammaire, mais aussi le contenu ? Une simple mise au propre orthographique ne permet pas d'améliorer le texte du point de vue de la structure ou de la cohérence par exemple, et parfois certains problèmes liés notamment aux chaînes référentielles ne sont pas résolus, en particulier si l'élève n'a plus envie de relire son premier jet, bien caché derrière les corrections de l'enseignant.

Nous avons choisi de comparer les deux versions des productions écrites de chaque élève à l'aide du logiciel MkAlign (Fleury) pour évaluer la prise en compte des corrections des enseignants dans le texte final. Premièrement, un nouveau script a éliminé automatiquement toutes les balises et les annotations des enseignants que logiciel aurait sinon signalées comme absentes dans la version finale. En effet, le logiciel signale en vert les éléments nouveaux ajoutés dans la deuxième version, en rouge les éléments supprimés qui apparaissent uniquement dans la première version et en bleu tous les éléments modifiés. La Figure 5 présente une capture d'écran de la mise en comparaison de la copie française proposée en Figure 1 et de sa version définitive rédigée par l'élève.

Ce type d'analyse n'a pas encore été effectué sur la totalité des textes qui seront ensuite relationnés aux profils d'enseignant-correcteur que nous aurons dessinés avant d'en tirer des conclusions. Cependant, nous remarquons déjà que, dans les cas où le logiciel signale en bleu plusieurs items dans les deux versions (Figure 5), cela indique que l'élève a modifié non pas des phrases entières mais plutôt l'orthographe de mots sans doute soulignés et/ou corrigés par l'enseignant dans le premier jet. Probablement, l'enseignant avait apporté des corrections locales orthographiques ou grammaticales que l'élève a ensuite adoptées. Cette méthode de

Figure 5. Capture d'écran du logiciel MkAlign utilisé pour la mise en comparaison du premier jet (à gauche) et du deuxième jet (à droite) de l'élève en Figure 1



correction nous rappelle le type d'intervention qui caractérise le « gardien du code » : cette posture risque en effet d'inhiber le remaniement du contenu de la part de l'élève et de favoriser en revanche une mise au propre au niveau orthographique.

PREMIERS RESULTATS

L'outil de transcription nous permet de rétablir l'ordre chronologique des opérations d'écriture de l'élève et de l'enseignant et d'aboutir à des transcriptions linéarisées semi-diplomatiques exploitables à l'aide d'outils informatiques. En particulier, nous pouvons générer des scripts qui analysent automatiquement tous les aspects précédemment annotés dans les transcriptions de copies d'élève et nous pouvons affiner les résultats à l'aide de logiciels de textométrie.

Tout d'abord, les interventions des enseignants français en CE2 et en CM2 sont beaucoup plus nombreuses en moyenne que les interventions des enseignants italiens, cela dû au nombre plus élevé d'erreurs d'orthographe qui est limité en contexte italien par un système linguistique bénéficiant d'une correspondance entre graphème et phonème plus consistante. D'ailleurs, le classement des commentaires avec portée locale et avec portée globale, nous a permis de constater que les enseignants italiens apportent davantage des commentaires qui portent sur la globalité

du texte, alors que les enseignants français utilisent le commentaire pour identifier des erreurs ponctuelles, le plus souvent orthographiques.

De toute manière, dans les deux contextes linguistiques, c'est la posture du « gardien du code » parmi les cinq postures proposées par Pilorgé qui est associée à la correction de la grammaire et qui se concrétise en une intervention locale par le biais de procédures d'écriture hétérogènes : remplacements, suppressions, ajouts et soulignements. L'exclusivité de cette posture inhibe le travail de réécriture de la part de l'élève qui est interprété trop souvent par l'enseignant aussi comme une étape de mise au propre du premier jet. L'évaluation successive du taux de remaniement du premier jet nous aidera alors à constater les effets des différents types d'interventions locales et globales chez les jeunes scripteurs italiens et français pendant la réécriture.

REFERENCES

- David, J., et Doquet, C. (2016). Les écrits d'élèves : Un corpus de référence pour le français contemporain. Dans F. Neveu, G. Bergounioux, M.-H. Côté, J.-M. Fournier, L. Hriba et S. Prévost (Eds.), *Actes du 5^e Congrès Mondial de Linguistique Française. SHS Web of Conferences*, 27, 11001. <https://doi.org/10.1051/shsconf/20162711001>
- Doquet, C. (2011). *L'écriture débutante. Pratiques scripturales à l'école élémentaire*. Rennes : Presses universitaires de Rennes.
- Doquet, C., Enouï, V., Fleury, S., et Mazziotti, S (2017). Problèmes posés par la transcription et l'annotation d'écrits d'élèves. *Corpus*, 16. <https://journals.openedition.org/corpus/2776>
- Fabre-Cols, C. (1990). *Les Brouillons d'écoliers ou l'entrée dans l'écriture*. Grenoble : Ceditel / L'atelier du texte.
- Penloup, M.-C. (1994). *La rature n'est pas un raté : Plaidoyer pour le brouillon*. Rouen : Académie de Rouen.
- Pilorgé, J.-L. (2010). Un lieu de tension entre posture de lecteur et posture de correcteur : Les traces des enseignants de français sur les copies des élèves. *Pratiques*, 145-146, 85-103. <https://doi.org/10.4000/pratiques.1513>

**SARA MAZZIOTTI**

Doctorante contractuelle en Sciences du langage à l'Université Sorbonne Nouvelle - Paris 3 et en cotutelle avec le Département de Sciences de l'éducation de l'Université de Bologne. Membre du groupe de recherche EcriScol dirigé par Claire Doquet au sein du laboratoire CLESTHIA à Paris 3.

sara.mazziotti@studio.unibo.it



Mazziotti, S. (2020). Traitement automatique des interventions des enseignants en CE2 et en CM2 en France et en Italie : entre grammaire et contenu. *Bellaterra Journal of Teaching & Learning Language & Literature*, 13(1), 61-76. <https://doi.org/10.5565/rev/jtl3.816>



Rebut / Recibido / Received / Reçu : 04-09-2019

Acceptat / Aceptado / Accepted / Accepté : 19-12-2019

<https://revistes.uab.cat/jtl3/>