



HAL
open science

Le correcteur grammatical : un auxiliaire efficace pour l'enseignant ? Quelques éléments de réflexion.

Chantal Charnet, Rachel Panckhurst

► To cite this version:

Chantal Charnet, Rachel Panckhurst. Le correcteur grammatical : un auxiliaire efficace pour l'enseignant ? Quelques éléments de réflexion.. ALSIC - Apprentissage des Langues et Systèmes d'Information et de Communication, 1998, 1 (2), pp.103-114. edutice-00000172

HAL Id: edutice-00000172

<https://edutice.hal.science/edutice-00000172>

Submitted on 5 Nov 2003

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.

Le correcteur grammatical : un auxiliaire efficace pour l'enseignant ? Quelques éléments de réflexion.

Chantal CHARNET, Rachel PANCKHURST

Université Paul Valéry, Montpellier III, France

Résumé : *Comment aider l'enseignant à mieux suivre le parcours acquisitionnel de l'apprenant en langue étrangère ? Que peuvent alors apporter les nouvelles technologies ? Un correcteur grammatical pourrait-il servir d'"auxiliaire" pour identifier les erreurs, voire les lieux d'intervention didactique nécessaires ?*

Par le biais des Industries de la langue, différents outils fournissent un cadre de réflexion pour ce type de problématique. Un logiciel de vérification et de correction grammaticales s'avère utile dans divers contextes d'utilisation. Nous nous proposons d'apporter une contribution pluridisciplinaire, à la fois en linguistique - informatique et en didactique afin de cerner le fonctionnement et l'utilité d'un tel logiciel en situation d'apprentissage du français langue étrangère écrit.

- [1. Introduction](#)
- [2. L'écrit en langue étrangère](#)
- [3. Du corpus d'échantillons à la correction automatisée](#)
- [4. L'espace didactique](#)
- [5. Conclusion](#)
- [Références](#)

1. Introduction

Un des rôles de l'enseignant est d'intervenir dans les processus d'acquisition (soit pour les stimuler soit pour les modifier) quand la construction du système en langue cible par l'apprenant s'avère problématique ; il doit repérer en particulier l'agrammaticalité des productions pour évaluer les points de vue et les règles mis en pratique, afin de reconnaître le niveau de l'apprenant et d'élaborer les apprentissages nécessaires. Cette fonction de correcteur ne doit pas s'arrêter à un simple relevé d'erreurs mais doit être étendue à une identification systématique des erreurs pour que l'enseignant puisse mieux comprendre le parcours acquisitionnel de l'apprenant. En quoi les nouvelles technologies peuvent-elles aider dans ce repérage ? Un correcteur grammatical correspond-il à cet auxiliaire attendu ?



Un "correcteur grammatical" est un logiciel qui permet de vérifier et de corriger un texte en langue naturelle, saisi sur format électronique, par exemple, au sein d'un traitement de texte. Comme son nom l'indique, il doit franchir l'étape de la vérification orthographique, pour atteindre et intégrer des aspects relevant de la vérification grammaticale. La "détection" d'erreurs et l'intervention d'une "correction" s'avèrent cruciales et justifient la raison d'être de ce type d'application.

Les correcteurs orthographiques existent depuis plusieurs décennies, tandis que les correcteurs dits "grammaticaux"^[2] n'arrivent véritablement sur le marché des Industries de la langue que vers les années 1990. Des correcteurs basés sur des approches statistiques (1960-1980), nous sommes passés à ceux incorporant des vérifications dictionnaires (depuis les années 1980), pour aboutir, de nos jours, aux correcteurs d'une "génération nouvelle", qui outre les dictionnaires, incorporent des règles de repérage

"linguistiques".

Dans cet article, nous ne tenterons pas d'apporter une réflexion de type "bancs d'essai comparatifs" de correcteurs grammaticaux divers ou de logiciels incorporant des modules de correction grammaticale (cf. quelques logiciels parmi ceux existant sur le marché des Industries de la Langue : *Sans Faute / Grammaire, Hugo Plus, Correcteur 101, Cordial*, etc.). Ces comparaisons figurent fréquemment dans les magazines spécialisés en informatique (cf. *Univers Mac, Windows Plus, Icônes, SVM, Soft & Micro, PC Expert...*) et constituent, à notre avis, une recherche de nature différente (cf., à titre d'exemple : "Les correcteurs orthographiques et grammaticaux" par J.-L. Brabant, *Icônes*, 56, avril-mai 1996, 53-67). Nous préférons plutôt partir d'un outil donné, quel qu'il soit, et étudier l'utilisation effective qui pourrait être faite de celui-ci, en situation d'apprentissage du français langue étrangère.

Le logiciel choisi est le "Correcteur 101" de l'entreprise québécoise basée à Montréal, [Machina Sapiens](#) (nd). Le choix de ce logiciel face à un autre correcteur grammatical nous importe peu dans la situation présente, bien que cette application ait la réputation de repérer efficacement certains types d'erreurs d'ordre grammatical.

Nous nous intéresserons, via l'intermédiaire de l'ordinateur "vérificateur/correcteur", au fonctionnement du *Correcteur 101* et à son utilité effectifs pour évaluer des productions écrites en français langue étrangère. Plus précisément, lors de l'acquisition de connaissances et de savoir-faire linguistiques et plus particulièrement grammaticaux, dans quelle mesure et par quels moyens le *Correcteur 101* pourra-t-il apporter une analyse qui aidera l'enseignant dans ses tâches d'évaluation et de correction ?

Nous ne cherchons absolument pas à repousser le *Correcteur 101* dans ses derniers retranchements pour cerner d'éventuels défauts et limites propres. Si nous exposons certaines limites réelles au cours de l'article, ce ne sera nullement pour montrer du doigt ce logiciel en particulier, mais, simplement pour indiquer quel est, à notre avis, le stade atteint de la technologie actuelle en matière d'avancées en linguistique - informatique, voire d'intelligence artificielle, concernant la correction grammaticale. Par contre, nous voulons explorer l'espace au sein duquel ce type de logiciel pourrait s'avérer utile en nous interrogeant effectivement sur la pertinence et les limites de ce choix d'utilisation en situation didactique. L'approche, double, est motivée par nos domaines de spécialité distincts : l'une s'intéressant de près à l'élaboration d'une réflexion et d'une implémentation éventuelle d'outils en linguistique - informatique, l'autre aux rapports entre l'activité cognitive développée lors des processus d'acquisition et l'apprentissage du français pour des sujets dont cet idiome n'est pas la langue maternelle. Nous tenterons ici de répondre aux diverses questions que notre interaction a suscitées.



2. L'écrit en langue étrangère

2.1. Un lieu privilégié d'intervention dans l'apprentissage ?

Lors de l'apprentissage d'une langue étrangère, un individu est amené à exprimer ses performances en utilisant des canaux différents, l'oral et l'écrit. Ces pratiques sont à différencier car leur utilisation implique des savoir-faire et des connaissances distincts. En effet, comme le précise Goody, l'écriture "n'est pas un simple enregistrement phonographique [...], elle favorise des formes spéciales d'activité linguistique et développe certaines manières de poser et de résoudre les problèmes" (Goody, 1979 : p.267). De plus, contrairement au processus d'interaction à l'oral pendant lequel l'interlocuteur stimule et agit sur le discours du locuteur non natif, à l'écrit, l'apprenant est amené à gérer sa production sans aide externe. Les contraintes qui lui sont posées pour réaliser une production écrite (par exemple, en situation simulée d'examen) sont souvent dictées par des consignes lui indiquant d'agir par écrit en temps limité. L'apprenant doit savoir planifier et organiser son discours pour répondre aux consignes et aux objectifs fixés au préalable.

Avant de parvenir à cette réalisation concrète, selon la conception vygotskyenne du langage écrit, le sujet passe par une phase de réflexion où il élabore intérieurement son texte écrit. Cette période d'élaboration peut d'ailleurs être appréciée dans sa phase de "brouillon" :

La forme écrite aide au déroulement d'un langage relevant de l'activité complexe. C'est là justement ce qui fonde le recours au brouillon. Le cheminement du "brouillon" au "mis au net" est précisément une activité complexe. Mais, même si l'on ne fait pas réellement de

brouillon, l'élément de réflexion est très important dans le langage écrit ; très souvent nous nous disons d'abord pour nous-mêmes ce que nous allons écrire ; il s'agit là d'un brouillon mental. Ce brouillon mental du langage écrit est aussi... un langage intérieur (Vygotsky in Schneuwly 1989 : 113-114).

Ainsi, la production écrite, tant pour l'apprenant que pour l'enseignant, apparaît comme un lieu d'intervention privilégié car elle correspond à "l'actualisation" (cf. Barbéris *et al.*, 1998), au résultat de ce travail intérieur. À la différence de la production orale qui "échappe" à l'apprenant (à moins d'être enregistrée), la production écrite constitue un objet dont l'instabilité et la variabilité fondamentale se concrétisent dans une forme matérialisée, celle-ci permettant une éventuelle intervention ultérieure. Cette intervention pourra être médiatisée par un enseignant voire (partiellement) par un logiciel.

Si nous estimons que la prise de conscience des erreurs par l'apprenant, ainsi que l'assimilation des formes de la langue cible sont pertinentes pour l'acquisition de celle-ci, le retour sur un texte écrit précédemment constitue une technique intéressante pour améliorer la production linguistique. Cette intervention peut être située à deux niveaux : (i) l'évaluation et (ii) la correction voire l'appropriation d'une "compétence linguistique". Pour l'enseignant, le recours à un repérage systématique des erreurs lexicales et grammaticales constitue un regard préliminaire sur la production de l'apprenant, lui permettant d'évaluer à la fois le niveau de ce dernier et les points sur lesquels celui-ci présente des déficits.

2.2. Le correcteur : une pratique pour l'amélioration de l'écrit ?

Selon les perspectives mentionnées en section 2.1, l'utilisation d'un "correcteur" correspond-elle à une pratique pertinente, pour l'enseignant, afin d'évaluer le niveau de l'apprenant et cerner ses déficits linguistiques ? Dans le cadre de cette réflexion, nous proposerons, à partir d'exemples authentiques, d'effectuer un inventaire de certains types d'erreurs repérées et de corrections fournies par le *Correcteur 101*. Une analyse basée sur une expérience dans le domaine de la linguistique - informatique nous guidera ensuite pour comprendre certaines zones de flottement au niveau de la correction proposée (ou le manque de correction) par le logiciel, qui pourraient paraître inexplicables à l'oeil nu.



Par ailleurs, il convient d'analyser la potentialité de correction de ce logiciel pour des productions en français langue étrangère, sachant que celui-ci n'a pas été spécialement conçu pour corriger des productions écrites par des sujets en cours d'acquisition d'une langue étrangère, mais plutôt pour le grand public (professionnel ou non) souhaitant écrire des textes "irréprochables" en français ("Toute personne rédigeant des textes en français sur un ordinateur peut bénéficier des services de 101", *Manuel de l'utilisateur du Correcteur 101* : 10).

Plusieurs scénarios de mise en place de l'espace didactique s'offrent à nous, et les types de questions que nous nous posons sont les suivantes : Le *Correcteur 101* est-il adapté à une recherche des erreurs produites (lexicales et/ou grammaticales) en français langue étrangère ? Lors de l'évaluation, le *Correcteur 101* permet-il une appréciation plus rapide du niveau acquis de l'apprenant ? Lors de la correction, le *Correcteur 101* permet-il de cerner les "lieux de travail" (Coïaniz, 1996:187) nécessaires, les points de langue sur lesquels il faudrait intervenir pour une acquisition améliorée de la langue cible ? En somme, l'utilisation de ce moyen automatisé permet-elle de repérer les lieux d'intervention de manière plus saillante que lorsque la méthode utilisée est plus classique (*i.e.* par repérage manuel) ?

3. Du corpus d'échantillons à la correction automatisée

3.1. Le corpus

Nous avons recueilli 20 copies manuscrites anonymes (que nous avons saisies ensuite, en prenant garde de ne pas générer d'autres erreurs typographiques !) d'étudiants étrangers, qui passaient, respectivement l'examen terminal du niveau 1 (apprenants débutants, ayant suivi environ 200h d'enseignement du français -- 10 copies), et l'examen terminal du niveau 4 (apprenants avancés, ayant suivi environ 800h d'enseignement du français -- 10 copies), de l'Institut des Étudiants Étrangers (I.E.E.) de l'Université Paul Valéry - Montpellier III [3]. Pour les étudiants du niveau 1, il s'agissait de rédiger un texte court (sujets 1 et 2 suivants au choix), suite à un texte annonçant un mariage :

Sujet 1 :

Imaginez la réponse écrite de Patricia qui accepte l'invitation.

Sujet 2 :

Vous êtes allé(e) au mariage de Patricia et Pierre. Racontez cet événement à un ami.

Pour les étudiants du niveau 4, il convenait de faire une synthèse en 250 mots de deux articles parus dans le *Figaro* et l'*Événement du Jeudi* à propos du vieillissement. Ce corpus très restreint d'échantillons ne doit servir qu'à titre d'exemple de mise en pratique en situation didactique ; il ne peut évidemment pas permettre un examen exhaustif de toutes les possibilités de correction permises par le logiciel.



3.2. Spécificités du Correcteur 101

Le *Correcteur 101*^[4] a été commercialisé pour la première fois en 1992 ; nous avons pu apprécier son évolution jusqu'à la version v3.2 (juillet 1997) dont nous nous sommes servis pour cette étude. Son module linguistique n'est pas basé sur une théorie linguistique particulière, mais s'inspire du *Bon Usage* et de la *Nouvelle grammaire française* de Grevisse et Goose, pour l'élaboration des règles grammaticales. Le *Correcteur 101* permet plusieurs niveaux de correction, en allant de la correction typographique (apport de la version 3), à la correction syntaxique, en passant par le "balayage lexical". La correction syntaxique est proposée de manière "assistée" en interaction avec l'utilisateur (phrase par phrase) ou de manière "automatique/autonome" (travail en arrière-fond et sauvegarde des corrections). Le but du logiciel est de "redonner confiance en leur écriture à ceux qui craignent de méconnaître les règles, et d'aider les autres à réparer leurs errements" (*Manuel de l'utilisateur du Correcteur 101* : 11).

3.3. Utilisation du logiciel

Afin de cerner rapidement les erreurs, dites lexicales, on fera fonctionner "l'assistant lexical" du *Correcteur 101*. La [figure 1](#) montre le début du résultat de ce balayage. Le lecteur remarquera que certains éléments (noms propres et noms communs dans l'exemple) sont répétés ; ceci est nécessaire dans la mesure où le texte est balayé de manière linéaire et chaque erreur est repérée, et ce, à chaque occurrence erronée. Ainsi, l'utilisateur a le choix d'ajouter les noms propres à un dictionnaire personnel, par exemple, avant de lancer la correction grammaticale proprement dite.

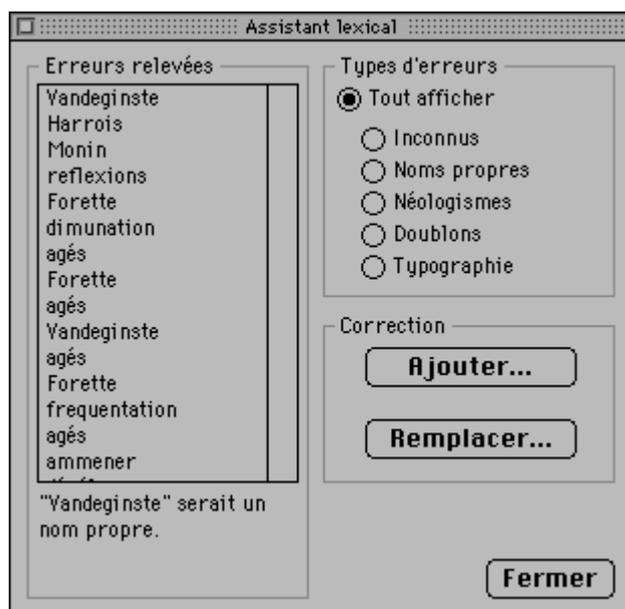


Figure 1 : L'assistant lexical

Cela étant, la correction peut être lancée sans que l'on ait procédé au balayage lexical précédemment. À partir de la [figure 2](#), "Extraits de phrases", les phrases, contenant des erreurs d'accentuation, seront

analysées et corrigées, si l'utilisateur le souhaite, comme le montre la [figure 3](#).

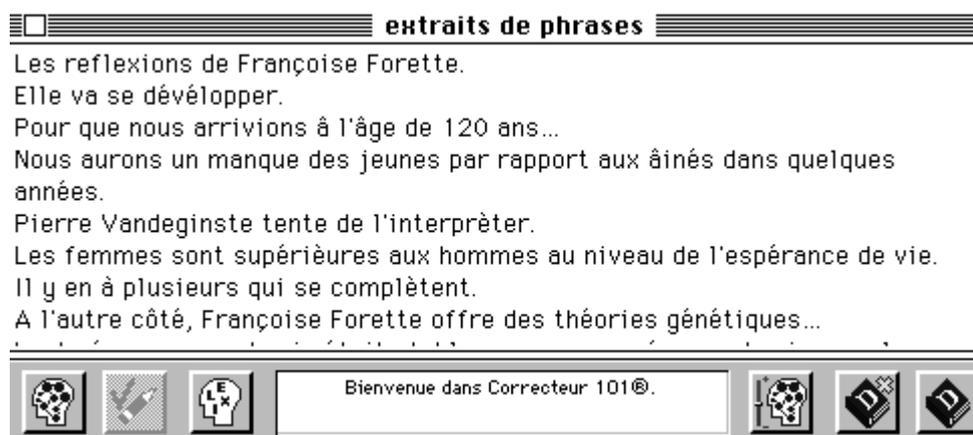


Figure 2 : Exemple de phrases avec des erreurs d'accentuation. La [figure 3](#) montre les corrections proposées.

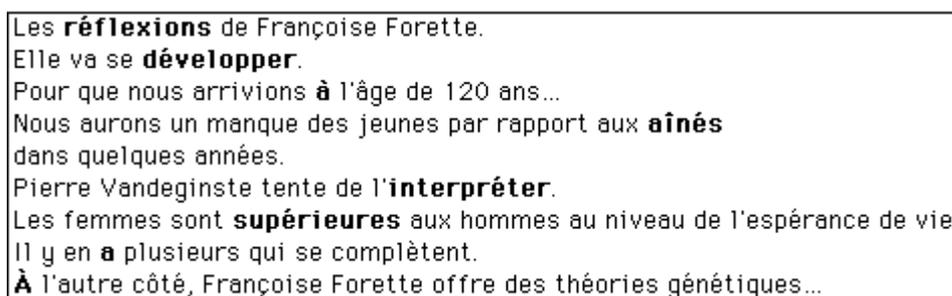


Figure 3 : correction de problèmes d'accentuation (cf. [figure 2](#)).

3.4. Inventaire et typologie des corrections

On peut classer les erreurs repérées et effectivement corrigées par le *Correcteur 101*, pour notre corpus, en trois catégories majeures[5].

Orthographe, signes diacritiques, typographie

Dans une première catégorie, pourront être corrigés : l'orthographe "pure", les signes diacritiques et la typographie. Le [tableau 1](#) donne des exemples de telles erreurs. Ces corrections appartiennent à une correction "locale", limitée, en contexte restreint. Une base de données dictionnaire peut suffire, où la comparaison de type mot à mot sera employée pour la vérification et des règles classiques (cf. section 3.5) serviront pour la correction. Ici, l'aspect syntaxique n'a pas lieu d'intervenir.



(1a) Orthographe "pure"		
Lettre manquante : origins, rallongment, shéma, developpemen...	Lettre superflue, doublée ou non : ammener, millieu, troupp...	Lettre erronée : rube, veudrais, muriage...
(1b) Signes diacritiques		
Accent manquant : reflexion, frequentation...	Accent erroné : â l'âge, interpréter...	Accent mal placé : âinés...

Tableau 1 : Erreurs d'orthographe, signes diacritiques, typographie.

Entre lexicale et syntaxe

Cependant, une erreur d'accentuation, par exemple, ne peut pas être systématiquement corrigée par ce type de moyen, tout simplement parce que le contexte doit être élargi pour qu'elle puisse être repérée. Un des exemples mentionnés dans la [figure 3](#) ("Il y en à plusieurs qui se complètent") est intéressant à cet égard. Un simple logiciel vérificateur orthographique ne pourra évidemment pas offrir une possibilité de correction ici ; il faut que le contexte soit élargi, d'où la prise en compte nécessaire de règles relevant de la syntaxe, que ce soit au niveau du groupement structurel des syntagmes ou au niveau global de la phrase. De même, deux exemples tirés de notre corpus : "Les français", et "Nous allons peut être..." ne pourront pas être corrigés efficacement par un logiciel non doté d'un module d'analyse morfo-lexico-syntaxique.

Ainsi, nous proposons une deuxième catégorie, intermédiaire entre lexicale et syntaxe, que l'on pourrait intituler "orthographe exigeant une connaissance morfo-lexico-syntaxique supplémentaire" (cf. [tableau 2](#)).

"que à cause", "que autour"... (élision)
"à les"... (amalgames[6])
"Il y en à..."
"Cet phénomène..."
"Les français"
"Nous allons peut être..."

Tableau 2 : Erreurs entre lexicale et syntaxe.



Syntaxe

Bien entendu, le logiciel doit intervenir efficacement au niveau syntaxique s'il est intitulé "correcteur grammatical". Pour ce qui concerne notre corpus, nous avons surtout relevé des erreurs d'accord (que ce soit au sein d'un même syntagme, ou au niveau phrastique), corrigées effectivement par le logiciel. Le [tableau 3](#) n'inventorie que quelques types d'erreurs repérées par le *Correcteur 101*. Il est capable d'intervenir également dans les accords de temps, les règles de participes passés sophistiquées avec des référents éloignés, au sein du groupe verbal, etc.[7]

Groupe nominal (genre) :	un majorité, le non influence du vieillesse, un formule...
Groupe nominal (nombre) :	l'actions, aux être, des raison principale, un ver ou un champignons infirme...
Groupe nominal contenant un adjectif :	une basse taux, la nouvelle problème, les côtés positives, les progrès médicaux...
Groupe nominal et groupe verbal (participe passé) :	Nous pourrons en l'avenir être surveillé, ... une qualité de vie pour les gens augmenté.

Tableau 3 : Erreurs de syntaxe repérés par le logiciel.

3.5. Blocages

Parfois, le logiciel bloque, et l'utilisateur peut être étonné par un manque de correction, qui serait pourtant, à première vue, simple. Plusieurs raisons peuvent être à l'origine du problème, mais l'enseignant doit être plus averti que le simple utilisateur : il devra comprendre le logiciel d'une manière plus approfondie pour ne pas tomber dans le jugement de valeur exprimé par certains étudiants francophones lors d'une première

expérimentation avec le *Correcteur 101* dans le cadre d'un de nos cours : "C'est nul, ce logiciel !"

En effet, plus il y a d'erreurs contenues dans une séquence, plus le logiciel prendra du temps pour la correction et se trouvera éventuellement en difficulté. Que l'on songe un instant aux méthodes de correction d'antan dans lesquelles on créait un algorithme basé sur une hypothèse qu'il n'existait qu'une seule erreur par mot. Lorsqu'un mot n'était pas reconnu dans un dictionnaire, 4 opérations de réparation intervenaient : interversion, substitution, insertion, suppression. Les exemples qui suivent figurent dans les copies manuscrites sélectionnées, à l'exception du premier, concernant l'interversion, qui relève davantage d'une erreur de saisie sur machine :

interversion : *lingusitique* -> *linguistique*
substitution : *vaconce* -> *vacance*
insertion : *fair* -> *faire*
suppression : *enttendre* -> *entendre*

Après chaque tentative le mot était de nouveau recherché dans le dictionnaire. À l'étape de la substitution, un mot de 8 lettres impliquerait $8 * 26$ (lettres de l'alphabet) = 208 vérifications dans le dictionnaire. Mais, est-il pertinent d'utiliser des algorithmes conçus à l'origine pour corriger des erreurs de saisie au clavier et non pas, *a priori*, pour corriger des erreurs liées à une méconnaissance morpho-lexicale dans le cadre de l'apprentissage des langues ?



Page 109

Prenons un exemple de "blocage" par le *Correcteur 101*. Soit la séquence suivante, figurant dans une des copies de notre corpus d'échantillons (niveau 4) :

Vandeginste explique qu'on vieillisse d'un raison pathologique : les vieux ont plus de maladies et un phsygologique concernant une diminution de sécretions hormonales.

Si on la soumet telle quelle au *Correcteur 101*, on obtiendra le résultat de la [figure 4](#). On pourrait aboutir à une conclusion hâtive et (partiellement) erronée, si on connaissait mal le logiciel :

- les noms communs *physiologie*, *diminution*, *sécrétions* et/ou l'adjectif *physiologique* n'existeraient pas dans le dictionnaire ;
- le genre des noms communs semblerait inconnu ;
- le subjonctif ne serait pas remplacé par un indicatif.

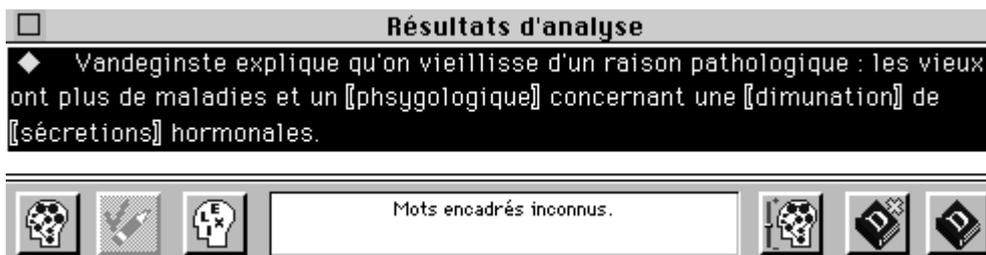


Figure 4 : un exemple de "blocage"

En fait, puisque les noms existent effectivement dans le dictionnaire, pourquoi le *Correcteur 101* ne les trouve-t-il pas ? La réponse réside dans le fait que les trois mots contiennent des degrés de complexité distincts : *sécrétion* ne contient qu'une erreur minimale d'accentuation ; *diminution* contient deux erreurs exigeant une substitution de deux lettres *u->i*, *a->u* ; *physyologique* est fort complexe, mais si une recherche dictionnaire est effectuée à partir de *physyologique*, en ôtant le *g* initial, la substitution *y->i* est retrouvée sans encombre. En effectuant la correction manuelle (à l'aide du dictionnaire en ligne toutefois), l'orthographe correcte de *sécrétion* est automatiquement retrouvée. Par ailleurs, une fois levés les problèmes lexicaux, l'accord en genre du déterminant est corrigé par rapport au nom commun *raison*. Reste le subjonctif qui n'est effectivement pas remplacé par un indicatif, car la sous-catégorisation verbale (types de compléments et temps verbaux requis) est un problème complexe qui est loin d'être résolu au sein des logiciels de correction grammaticale de manière générale.

Nous espérons que cet exemple aura servi d'illustration quant à la complexité du processus de correction

automatique ou automatisé[8]. Ainsi, même si le *Correcteur 101* semble "se bloquer" au niveau de l'analyse syntaxique dans certains cas (par exemple, lorsque la phrase contient trois erreurs ou plus[9]), l'enseignant se devra de rester vigilant pour ne pas aboutir à des conclusions hâtives concernant les corrections proposées par le *Correcteur 101*. Bien entendu, cette vigilance n'est possible que si les didacticiens lui expliquent les limites de fonctionnement d'un logiciel de ce type en situation didactique.



4. L'espace didactique

Le lecteur l'aura sans doute remarqué dans le tableau récapitulatif des erreurs corrigées ; le *Correcteur 101* ne peut repérer des erreurs du point de vue de l'interprétation sémantique. Dans le dernier exemple apparaissant dans la [figure 3](#), le *Correcteur 101* corrige l'accent manquant sur la majuscule dans "A l'autre côté...", mais il sera incapable de remplacer cette production par d'autres possibilités : "De l'autre côté" vs. par exemple, "À l'autre extrémité". Les exemples suivants (copies de niveau 1) témoignent de cette absence de correction, bien que certains homonymes soient effectivement signalés (Dernière **foie** que je t'ai vu etait à Paris) :

Il y avait un **loin** temps que j'ai eu des nouvelles de toi.
Je parte en Suède une semaine pour **regarde** ma famille.
J'espère un **bien** temps en Suède.
C'est **pré** d'ici pour moi.

Son analyse, comme pour tous les logiciels de ce type existant à l'heure actuelle, s'arrête à un niveau essentiellement syntaxique ; la correction sémantique, voire discursive, en est exclue. Certaines connaissances, qui dépassent le cadre syntaxique *stricto sensu*, i.e. la concordance des temps, figurent effectivement au sein du logiciel (*Si tu serais venu, j'aurais fait... -> Si tu étais venu...*), mais d'autres, concernant par exemple l'analyse verbale, (*J'ai commencé pleurer*), les pronoms clitiques (*Je vais téléphoner à toi*), etc. sont méconnues.

Cependant, bien que les détections/corrections d'erreurs concernant les constructions stylistiques des phrases, les niveaux syntaxico-sémantiques, sémantiques, voire pragmatiques soient exclues à ce stade, le lecteur ne doit pas s'en étonner ; celles-ci reflètent tout simplement l'état actuel des recherches en linguistique - informatique. Il n'en demeure pas moins que le *Correcteur 101* propose des corrections convaincantes pour les catégories mentionnées dans la section 3.4, et l'enseignant pourra aisément concevoir des exercices en fonction des "lieux de travail" repérés au préalable.

4.1. Correction automatisée de tests : pour ou contre ?

L'enseignant pourrait-il se servir du *Correcteur 101* pour faire une correction préliminaire (lexicale et syntaxique) de tests avant d'intervenir (manuellement) pour la partie sémantico-pragmatique ? L'outil pourrait effectivement permettre de cerner les problèmes lexicaux puisqu'une liste des mots non reconnus est fournie via l'assistant lexical (*cf.* section 3.3).

Ensuite, les zones d'ombre syntaxiques sont soit corrigées (en gras) soit annotées (en italique). Bien entendu, il sera primordial de faire attention aux aspects soulignés en section 3.5. D'où un travail certain de vérification manuelle de la part de l'enseignant. Mais, celui-ci, dégagé qu'il serait des problèmes morfo-lexico-syntaxiques, pourrait concentrer tous ses efforts sur l'intervention stylistique-sémantico-pragmatique.

Par ailleurs, la saisie des copies sur format électronique par les étudiants participant à l'examen n'est pas un aspect anodin. En effet, d'autres types d'erreurs peuvent être introduites à ce stade. Par exemple, comment distinguer l'orthographe méconnue d'un mot d'une saisie erronée dans laquelle a été introduite une substitution ou une interversion de lettres ? Ou bien, au niveau grammatical, un mauvais accord au participe passé dû davantage à une surcharge cognitive via l'utilisation de l'ordinateur en tant que médiateur, qu'à une incompréhension d'une règle grammaticale ? Certains étudiants pourraient, selon le type d'exercice proposé en examen, se plaindre, à juste titre, de leur lenteur de saisie qui les désavantagerait par rapport à d'autres ayant suivi des cours de dactylographie !



Enfin, si un logiciel de traitement de texte est utilisé à l'examen, on ne peut écarter, à l'insu de l'enseignant examinateur, le recours de l'étudiant à une vérification orthographique et semi-grammaticale par le logiciel utilisé !

4.2. "Lieux de travail"

Nous sommes convaincues que certains lieux de travail ou lieux d'intervention didactique peuvent être repérés de manière synthétique et automatisée par le *Correcteur 101*. En effet, le logiciel permet un repérage systématique de différents types de problèmes :

- orthographe, signes diacritiques, typographie...
- morphologie, lexicque, syntaxe...

De prime abord, le "balayage lexical" fourni par le *Correcteur 101* permet de faire afficher les mots inconnus, les noms propres, les néologismes, les doublons, la typographie (cf. schéma 1). L'enseignant pourrait alors élaborer certains exercices ciblés, en exigeant de la part de l'apprenant une participation active et une vérification systématique par le logiciel. Ensuite, étant donné l'accès possible au dictionnaire, des exercices de type lexical et lexicographique pourraient être proposés à l'apprenant. Aussi, les tables de conjugaison qui figurent au sein du *Correcteur 101* permettraient-elles un certain travail de révision. Du point de vue morphosyntaxique, l'enseignant pourrait orienter un premier travail en direction des problèmes de flexion (genre, nombre, personne...), avant de vérifier des connaissances plus directement liées à la structure globale d'une phrase.

5. Conclusion

Au départ, nous avons posé le problème de l'utilité d'un correcteur grammatical par rapport à l'acquisition du français langue étrangère écrit, à l'aide de l'utilisation d'un logiciel de correction grammaticale. À condition que l'enseignant maîtrise le logiciel afin d'intervenir dans la "détection" et la "correction", celui-ci peut être un auxiliaire fort utile. Par ailleurs, si un exercice peut être effectivement élaboré et proposé au niveau morpho-lexico-syntaxique, ce ne sera pas le cas pour une éventuelle stimulation de la production sémantico-pragmatique.

L'identification systématique des erreurs des apprenants peut permettre à l'enseignant d'avoir une vision plus globale de leur compétence linguistique. Tout d'abord, elle permet à l'enseignant d'intervenir sur une modification de la structure cognitive des apprenants ; ensuite, elle lui fait prendre conscience des erreurs dues à l'inférence du système de la langue première (quand un groupe issu d'une même communauté commet les mêmes erreurs) et de celles dues à une construction plus idiolectale du nouveau système. Mais, dans cette perspective, une vision globale, structurée, des erreurs commises, serait plus appréciée pour déterminer le stade d'acquisition en production écrite, et pour permettre l'élaboration d'applications pédagogiques ultérieures :

Les correcteurs pourraient supporter la tâche de l'enseignant, et par conséquent l'apprentissage des élèves, s'ils pouvaient établir un portrait des erreurs commises par chaque élève en utilisant une typologie d'erreurs adaptée à la situation pédagogique. (Désilets, 1998, p. 22).



Ce relevé, fourni par le logiciel, permettrait-il de vérifier de manière efficace l'acquisition des connaissances, sous forme de "bilan" grammatical à la fin de chaque niveau en classe de langue ? Pour s'adapter à un public de français langue étrangère, le correcteur devrait prendre en considération le contexte pédagogique explicite. C'est seulement à partir d'une identification exhaustive des erreurs que l'enseignant trouverait dans cet outil un "auxiliaire" efficace.

Dans une prochaine étude, nous présenterons une recherche de terrain (actuellement en cours), concernant l'utilisation effective en classe de langue d'un logiciel comme le *Correcteur 101* par des apprenants étrangers. Celle-ci nous permettra de fournir une réponse effective quant aux pertinences et aux limites du

point de vue de l'apprenant, et de confirmer (ou d'infirmer !) l'hypothèse selon laquelle ce type de logiciel, bien utilisé, peut être un outil didactique précieux dans la prise de conscience du système de la langue cible. Enfin, il serait intéressant de fournir quelques pistes sur les principales fonctionnalités d'un logiciel correcteur à vocation didactique.

En attendant, même si les correcteurs de nos jours ont largement progressé dans le domaine syntaxique, la citation de Sabah (1988:184) demeure d'actualité :

La correction d'erreurs syntaxiques et sémantiques, comme l'émission de messages intelligents, constituent encore des sujets de recherche à long terme. Ils sont fondés sur le développement de mécanismes de représentation des connaissances (en particulier la notion d'univers de croyances propre à chaque interlocuteur) et l'utilisation des règles pragmatiques de la conversation. Ils nécessitent donc le développement de l'ensemble des recherches sur la compréhension des langues en général.

Références

Bibliographie

- Barbérís, J.-M., Bres, J. & Siblot, P. (dirs.) (1998). *De l'actualisation*. Paris : CNRS-Éditions.
- Brabant, J.-L. (1996). "Les correcteurs orthographiques et grammaticaux", *Icônes*, 56, avril-mai. pp 53-67.
- Carré et al. (1991). *Langage humain et machine*. Paris : Presses du CNRS.
- Charmet, C (1995). "Ah ça y est c'est le désert". Apprendre en classe de langue : une pratique interactive". *LIDIL*, 12, septembre, "L'interaction en questions" coordonné par V. de Nuchèze & J.-M. Coletta. pp 135-148.
- Coñaniz, A. (1996). *Faute et itinéraires d'apprentissage en classe de français langue seconde*, Série Travaux de Didactique. Presses de l'Université Paul Valéry - Montpellier III.
- Désilets, M. (1998). "Les correcteurs d'orthographe et l'apprentissage", *Cahiers pédagogiques*, no 362, mars. pp 22-23.
- Goody, J. (1979). *La raison graphique*. Paris : Minuit.
- Leeman-Bouix, D. (1994). *Les fautes de français existent-elles ?*. Paris : Seuil.
- Manuel de l'Utilisateur, Le Correcteur 101* (1992). Machina Sapiens : Montréal.
- Panckhurst, R. (1998). "Marques typiques et ratages en communication médiée par ordinateur". *Actes du colloque CIDE 98*, INPT, Rabat, 15-17/04/98. Paris : Europa Productions. pp 31-43.
- Peterson, J. (1980). "Computer programs for detecting and correcting errors". *Communications of the ACM*, 23, 12.
- Pollock, J. & Zamora, A. (1984). "Automatic spelling correction in scientific and scholarly text". *Communications of the ACM*, 27, 4.
- Riegel et al. (1994). *Grammaire méthodique du français*. Paris : P.U.F.
- Sabah, G. (1988). *L'intelligence artificielle et le langage*, 2 vol.. Paris : Hermès.
- Sabatier, P. (1987). *Contribution au développement d'interfaces en langage naturel*. Thèse d'État en Informatique, Université Paris VII.
- Schneuwly, B. (1989). "La conception vygotkienne du langage écrit", *Études de linguistique Appliquée*, 73. pp 107-117.



Véronis, J. (1988). *Contribution à l'étude de l'erreur dans le dialogue homme-machine en langage naturel*. Thèse de l'Université Aix-Marseille III.

Site internet

Machina Sapiens (nd). *Site de la société diffusant le "Correcteur 101"*. Consulté en novembre 1998 : <http://www.machinasapiens.com>.

Notes

[1] Nous remercions les relecteurs anonymes pour leurs remarques et critiques très pertinentes.

[2] Faute de place, nous ne retracerons pas tout l'historique des correcteurs orthographiques - grammaticaux. Entre autres : pour un survol historique, cf. Carré et al. (1991) ; Véronis (1988) traite le problème de l'erreur dans le dialogue homme - machine en langage naturel ; Sabah (1988) expose les différents "niveaux" de traitement quant aux "non attendus", qui correspondent aux données à traiter qui n'ont pas été "prévues".

[3] Nous remercions H. Boyer, C. Santodomingo et Y. Boisson d'avoir gracieusement mis à notre disposition les copies citées.

[4] L'entreprise québécoise Machina Sapiens a choisi ce nom pour son logiciel de correction du français, puisqu'il correspond à la loi 101, qui protège l'utilisation du français au Québec.

[5] Ces catégories ne sont qu'indicatives, et ne constituent pas un classement formel et définitif. Bien entendu, les erreurs répertoriées peuvent parfois appartenir à plusieurs catégories.

[6] Nous préférons "amalgames" à "articles contractés" : "Ces amalgames avec les prépositions *de* et *à* sont traditionnellement nommés, de façon très contestable, "articles contractés", Riegel *et al.* (1994:154).

[7] Voici quelques exemples supplémentaires de correction, dont certains sont proposés dans la base d'exemples du *Correcteur 101* :

Si je viendrais, je dirais tout -> Si je venais, je dirais tout.
Viendrais-tu s'il serait là ? -> Viendrais-tu s'il était là ?
Les journaux que tu as dit à Marie que tu avais acheté hier ont été livrés ce matin. -> Les journaux que tu as dit à Marie que tu avais achetés hier ont été livrés ce matin.

[8] Le linguiste - informaticien utilisateur de ce type de logiciel ne peut qu'élaborer des hypothèses d'interprétation quant au contenu réel du code informatique, et aux éventuelles zones de flottement dans les corrections fournies, puisque l'accès direct à celui-ci demeure nécessairement "opaque", et ce, essentiellement pour des raisons de commercialisation tout à fait compréhensibles.

[9] Cette information a été vérifiée auprès de Machina Sapiens pour la version utilisée, soit la v3.2. Les versions ultérieures corrigeront sans doute ce type de problème.

A propos des auteurs

Chantal CHARNET : maître de conférences (linguistique) à l'Université Montpellier III. Ses réflexions portent sur les fonctionnements discursifs en communication exolingue et l'acquisition des langues secondes.

Courriel : charnet@danaid.univ-montp3.fr

Adresse : Praxiling - ESA 5475, Université Paul Valéry - Montpellier III, Route de Mende, 34199 Montpellier Cedex 5, France.

Rachel PANCKHURST : maître de conférences (linguistique - informatique) à l'Université Montpellier III. Ses centres d'intérêt actuels sont : la communication médiée par ordinateur, l'évaluation de logiciels, la sémantique lexicale.

Courriel : rachel@bred.univ-montp3.fr

Adresse : Praxiling - ESA 5475, Université Paul Valéry - Montpellier III, Route de Mende, 34199 Montpellier Cedex 5, France.



[ALSIC](#) | [Sommaire](#) | [Consignes aux auteurs](#) | [Comité de rédaction](#) | [Inscription](#)

© *Apprentissage des Langues et Systèmes d'Information et de Communication*, 15 décembre 1998