L'erreur lexicale dans l'analyse des productions écrites en FLE

Olha Luste-Chaâ

Édition électronique

URL : <http://journals.openedition.org/pratiques/1556>

DOI : 10.4000/pratiques.1556

ISSN : 2425-2042

Éditeur

Centre de recherche sur les médiations (CREM)

Édition imprimée

Date de publication : 15 juin 2010

Pagination : 197-210

Référence électronique

Olha Luste-Chaâ, « L'erreur lexicale dans l'analyse des productions écrites en FLE », *Pratiques* [En ligne], 145-146 | 2010, mis en ligne le 15 juin 2010, consulté le 30 avril 2019. URL : <http://journals.openedition.org/pratiques/1556> ; DOI : 10.4000/pratiques.1556

L'erreur lexicale dans l'analyse des productions écrites en FLE

Olha Luste-Chaâ

Université Paul-Verlaine Metz - CELTED

Depuis les années 1980, la production écrite en langue étrangère et seconde fait l'objet de nombreuses recherches qui s'inscrivent dans des perspectives différentes, qu'elles soient pédagogiques, linguistiques ou psycholinguistiques. Les résultats de ces recherches ont permis de comprendre la complexité des processus cognitifs sollicités et mis en œuvre chez l'apprenant lors de la compréhension et l'expression écrites et, par conséquent, d'amorcer une réflexion didactique, d'améliorer ou de rénover les pratiques d'enseignement en français langue maternelle (désormais FLM), langue étrangère (FLE) et langue seconde (FLS).

Parler de la compréhension et de la production écrites, implique, presque inévitablement, d'évoquer la question des connaissances lexicales et de leur étendue chez l'apprenant. Mais le nombre d'études longitudinales portant sur l'évolution des compétences langagières en général et de la compétence lexicale en particulier reste encore très faible tant en FLM qu'en FLE (C. Garcia-Debanc, M. Fayol, 2002 : 48).

En même temps, la tâche de correction, qui fait partie du quotidien d'un enseignant, amène très souvent à se poser des questions sur les erreurs que l'on trouve dans les textes produits par les apprenants : Pourquoi font-ils des erreurs ? Comment les expliquer et, surtout, peut-on y remédier ? Ces questions résultent de certaines lacunes qui existent encore, malgré la parution de quelques travaux récents sur ce sujet, au niveau de l'appréhension de l'erreur, notamment lexicale, de son rôle et son statut dans l'enseignement / apprentissage du français en général et dans la didactique du lexique en particulier.

Le point de vue que nous adoptons consiste à considérer l'erreur lexicale comme un indicateur de premier ordre permettant d'appréhender les difficultés scripturales et l'évolution de la compétence langagière, et lexicale en particulier, chez l'apprenant, tant par le biais d'une analyse ponctuelle que lors d'une analyse longitudinale.

1. Élaboration d'un corpus d'écrits et recueil d'erreurs lexicales

La méthodologie présentée dans cet article ⁽¹⁾ a été développée initialement pour

(1) Un chaleureux merci à Caroline Masseron (Université de Metz) dont les remarques m'ont permis d'apporter de nombreuses améliorations à la première version de cet article.

les besoins d'une recherche effectuée entre 2006 et 2009 dans le cadre d'un doctorat (O. Luste-Chaâ, 2009). Le but initial visé était d'observer et d'analyser les acquisitions lexicales dans leur corrélation avec les compétences en compréhension et production écrites chez un public non francophone, relevant du FLE / FLS. La démarche consistait à rassembler des travaux écrits qui puissent « refléter » le mieux possible la progression de l'apprentissage et l'évolution des acquisitions lexicales dans l'expression écrite chez le public observé. Cet objectif nous a amenée à recueillir durant neuf mois des productions écrites d'apprenants et à constituer un corpus d'observation représentant un échantillon des acquisitions linguistiques, en l'occurrence des acquisitions lexicales en FLE / FLS chez un public universitaire. Lors de l'analyse du corpus, notre attention s'est focalisée en particulier sur les *erreurs lexicales* dans la production de textes en raison de la complexité de ce phénomène et des perspectives didactiques qu'il pourrait présenter. Ainsi, notre démarche s'est-elle concentrée sur la définition de la nature des erreurs lexicales relevées pour en dégager une typologie opérationnelle qui permette de comprendre la source des erreurs produites et pour les convertir en objets d'enseignement / apprentissage, sans les considérer comme quelque chose d'*a priori* répréhensible.

Lors du recueil des données, nous avons privilégié la constitution d'un corpus *homogène* du point de vue du public (les auteurs de textes relèvent d'une même tranche d'âge et connaissent une situation sociale comparable) et des thèmes abordés (le corpus est divisé en ensembles thématiques), et *restreint* (le corpus inclut uniquement les productions du groupe-témoin). Ces deux premières caractéristiques nous ont permis d'amorcer une analyse qualitative (préférée à un simple relevé statistique) d'erreurs qui peuvent survenir dans le traitement et le choix des unités lexicales par les apprenants FLE / FLS et de proposer une typologie d'erreurs lexicales qui sera présentée plus loin. Notre corpus se caractérise également par une *organisation* basée sur la chronologie de la progression dans l'apprentissage, un *ancrage temporel* (toutes les productions ont été écrites durant une période déterminée), et une certaine *représentativité* et *pertinence* (les faits linguistiques révélés par l'analyse du corpus et les conclusions qui s'en suivent pourraient être appliqués à un autre corpus de productions écrites d'étudiants d'origines différentes).

Le public qui a fait l'objet de l'observation est constitué de 17 étudiants chinois âgés de 23-25 ans, qui sont venus faire des études à l'Université de Metz et devaient suivre pendant une année universitaire un programme du français intensif afin de pouvoir intégrer le cursus universitaire ordinaire en vue de la préparation d'un diplôme niveau Licence ou Master dans les domaines différents. Au moment du démarrage des cours en octobre 2005, tous les étudiants avaient suivi au préalable environ 400h de cours de français en Chine et atteint les niveaux A1 / A2 du *CECRL* (2001). Dans l'impossibilité d'obtenir la totalité des productions écrites couvrant l'ensemble du public d'observation, nous avons décidé de former un groupe-témoin en prenant comme critère l'« assiduité optimale » des producteurs des textes. C'est ainsi que s'est constitué un noyau de neuf étudiants, pour lesquels les productions écrites ont, le plus régulièrement possible, été conservées et qui attestent des progrès réalisés en français. Il est à souligner que l'ensemble des étudiants n'étaient pas au courant de l'étude menée sur leurs productions écrites. Nous avons choisi cette option pour ne pas risquer de faire interférer les objectifs et préserver le caractère authentique des productions.

Le corpus final comprend donc 63 productions écrites des 9 étudiants non francophones d'origine chinoise, réalisées au cours des 9 mois d'enseignement / apprentissage intensif du français. L'ensemble des textes se divise en 9 sous-ensembles parmi lesquels on distingue trois évaluations et quelques travaux écrits intermédiaires. Dans les annexes du présent article figure un extrait de notre corpus : les trois

textes produits par la même étudiante et correspondant aux trois évaluations effectuées avant le début des cours (Texte 1), à la fin du premier semestre (Texte 2) et à la fin du cursus en juin (Texte 3). La version dactylographiée des trois textes se conforme strictement aux originaux.

1.1. L'erreur : évolution et typologie

Du point de vue linguistique, l'*erreur*, dans son sens large, peut être définie comme un écart par rapport à la norme. Étant donné la relativité de la notion même de *norme* et la coexistence de plusieurs conceptions, nous avons adopté une conception qui fait de l'erreur un écart de la norme prescriptive du français standard écrit. Cette norme est requise dans le cadre institutionnel de la production verbale écrite en contexte de FLE / FLS, ce qui conduit à juger erroné tout emploi familier. Par ailleurs, nous référant à une conception fonctionnelle (H. Frei, 1929), nous avons relevé comme erronée une unité lexicale jugée inadaptée au contexte de production.

En didactique des langues, le domaine dont relève cette recherche, la conception de l'erreur a subi une évolution assez spectaculaire, depuis le behaviorisme, où l'erreur était considérée comme une anomalie, une faille, voire quelque chose d'inconcevable, jusqu'à la période contemporaine qui fait de l'erreur un moteur de l'apprentissage (la construction d'une compétence complexe par essais et erreurs). À partir des années 1950-1960, l'application des théories psycholinguistiques à l'enseignement / apprentissage des langues a permis de déplacer le débat sur les sources possibles d'erreurs, et d'admettre que l'erreur fait partie de l'apprentissage, qu'elle reflète l'évolution progressive de l'interlangue et qu'elle renvoie à des « hypothèses » émises par l'apprenant au sujet du fonctionnement de la langue cible (P.S. Corder, 1980). Les questions qui occupent désormais les didacticiens des langues portent sur les sources possibles (langue maternelle, langue cible, autre ?) des erreurs relevées dans les productions d'apprenants, et les réponses en termes de consolidation ou de remédiation. Dans cette perspective, plusieurs essais de classement d'erreurs ont vu le jour dans le but de guider le travail de l'enseignant et de le conduire à des explications possibles au sujet de la nature ou de la source d'une erreur. De tels classements ont en commun de chercher avant toute autre chose d'identifier l'aspect de la langue affecté par l'erreur relevée : phonétique, orthographe, morphologie, lexicale, syntaxe. Certaines typologies ont connu un succès particulier, comme notamment la typologie des erreurs en orthographe conçue par N. Catach et son équipe (N. Catach, 1980 ; N. Catach *et al.*, 1980), d'autres ont rapidement été mises de côté comme difficilement applicables dans un contexte scolaire. Si les premières typologies d'erreurs élaborées dans les années 1960-70 ont le plus souvent porté sur les infractions grammaticales et orthographiques, celles des deux dernières décennies se sont de plus en plus souvent consacrées au lexique.

1.2. L'erreur lexicale, un indicateur puissant mais qu'il n'est pas aisé de stabiliser

En retenant les erreurs lexicales comme objet d'étude nous avons considéré qu'elles jouaient un rôle d'indicateur du niveau de langue atteint : mettant en évidence des difficultés scripturales, elles permettent d'identifier des besoins langagiers. Dans ce contexte, nous avons défini l'*erreur lexicale* comme *la manifestation d'un défaut de maîtrise langagière, identifiée par le biais d'une unité, simple ou complexe, matérialisée dans le texte ou même absente de sa surface, qui s'apparente au stock lexical de la langue*.

On peut déjà se douter que l'identification et l'analyse d'erreurs lexicales se ré-

vèlent beaucoup moins aisées que celles d'erreurs d'orthographe par exemple qui sont (en théorie) matérialisées à la surface du texte et qui, en transgressant le code écrit, appellent un jugement univoque. L'exemple suivant illustre combien le constat d'une erreur, peu discutable en orthographe, est hautement discutable dans le « domaine lexical ».

(1) *On peut trouver *beaucoup de choses dans ce site [...].*⁽²⁾

Dans cet énoncé, tiré du Texte 1 (cf. Annexes), **beaucoup* est incontestablement une erreur d'orthographe d'usage. Ce jugement est indépendant du contexte dans lequel l'erreur apparaît et l'analyse se limite à l'unité même. En revanche, la production de (?) *dans ce site* est d'emblée perçue comme plus problématique. Nous y consignons une erreur « lexicale » par facilité : en réalité l'emploi erroné de la préposition *DANS* s'étend à tout le syntagme prépositionnel, impliquant plusieurs facteurs sémantiques qui ne sont pas indépendants de la construction réalisée (*trouver qqch qpart* ; le nom *site* régi par la préposition *dans*).

Une autre difficulté réside dans le fait que les erreurs lexicales peuvent être indirectes. Elles procèdent alors d'une sorte de calcul d'inférence, sur la base de l'absence d'une unité lexicale à la surface du texte. Dans tout discours (oral et écrit), le domaine référentiel sollicité par l'objet dudit discours prédéfini et actualise un certain stock lexical qui est attendu par le locuteur / lecteur et, dans le cas de production écrite, devrait apparaître à la surface du texte. Une unité lexicale requise par l'isotopie dominante du sujet abordé mais absente de la production, peut être indexée au répertoire d'erreurs, dans la mesure où ce dernier a pour objectif principal de recenser les *besoins* langagiers.

Enfin, un autre facteur de difficulté résulte de la relativité des notions de norme et d'erreur : le jugement d'erreur est discutable et sera discuté par deux correcteurs distincts, par exemple si le jugement d'enfreinte lexicale recourt à une norme, implique un jugement de valeur ou de conformité à un registre de discours ou un genre d'écrit (C. Masseron, O. Luste-Chaâ, 2008 : 519). Dans la plupart des cas, les erreurs lexicales ne permettent pas une commutation stricte d'une unité lexicale, « fausse », par une autre, « juste », dans un énoncé donné.

1.3. Lexies absentes : termes spécifiques et déictiques

Pour illustrer les propos ci-dessus, prenons le Texte 2 qui a été produit lors de l'évaluation intermédiaire, au mois de février, après environ 4,5 mois de cours intensifs de français. Il répond à la consigne suivante :

« Un ami français, spécialisé en informatique, vous a écrit pour vous annoncer qu'il souhaite s'installer en Chine. Vous lui répondez par une lettre en lui donnant des informations et des conseils qui pourraient lui être utiles (démarches administratives, logement, salaire, adresses utiles, façon de présenter un CV...). »

Pour ce qui est de la disponibilité à la surface du Texte 2 des unités lexicales attendues, on peut noter quelques unités lexicales « manquantes » et le fait que l'étudiante recourt aux paraphrases, aux répétitions et aux verbes supports pour combler ses lacunes (ex. 2 *infra*). Il est à noter notamment l'absence de la lexie *LOYER* (et de toute autre lexie désignant le type d'habitation) dans le passage sur le logement

(2) Nous avons consigné en caractères droits le segment identifiant l'erreur lexicale. Quant à l'astérisque (*) placé devant un élément jugé erroné, il sanctionne une graphie ou une forme grammaticale inadéquates.

(ex. 3), ainsi que l'absence du vocabulaire administratif apparemment requis par la consigne (*donner des conseils et des informations sur les démarches administratives*). En témoigne la réécriture (entre crochets) du segment fautif incriminé.

- (2) *Le formulaire de CV est pareil comme celui qui est en France.* < Le modèle de CV en Chine (le formulaire d'un CV chinois ressemble au formulaire français) ressemble à celui qui existe en France.>
- (3) *Il ne faut pas payer beaucoup pour un logement. Comme moi, je paye environ 40 euro par mois.* < Les loyers ne sont pas chers. Moi, par exemple, je paye environ 40 euros de loyer par mois.>

Par ailleurs, notre attention a été attirée par la répétition exagérée du groupe *en Chine* au détriment du déictique *ici* (de notre point de vue, l'unité *absente*) qui aurait peut-être été plus disponible si l'auteure avait effectivement écrit sa réponse depuis la Chine. Ce phénomène rappelle que le contexte institutionnel de la production « désauthentifie » l'écrit, en privant le scripteur des « évidences perceptives » de la situation, que la consigne ne parvient pas à stimuler tout à fait.

Les facteurs susmentionnés rendent sans doute difficile la tâche d'analyse d'erreurs lexicales, mais nous tenons à insister sur l'importance d'associer cette démarche au travail de correction de la part de l'enseignant. L'erreur étant la preuve d'un système de *compétences transitoires* de l'apprenant (P.S. Corder, 1980 : 13), le relevé et l'analyse d'erreurs lexicales permettraient de suivre l'évolution de son interlangue et le déroulement de l'acquisition / apprentissage du lexique en langue étrangère, ainsi que de mieux percevoir les difficultés que l'apprenant peut rencontrer lors de la production mais aussi la réception des discours oraux et écrits. Quant à la correction, son importance vient du fait que l'acte correcteur permet, en adoptant un point de vue linguistique et fonctionnel sur les erreurs lexicales, de localiser avec plus de précision le segment incriminé et d'identifier sa nature, même si ces tâches, et surtout la dernière, sont loin d'être aisées. Dans cette démarche, et pour aboutir à une analyse linguistique cohérente, la correction des erreurs lexicales exigerait du correcteur d'interpréter et d'analyser l'erreur, ainsi que l'influence de celle-ci sur le cotexte, par intermédiaire d'une *réécriture* du segment erroné (M. Monte *et al.*, 1992 ; C. Masseron, 2001, 2005 ; C. Masseron et O. Luste-Chaa, 2008).

2. Les enjeux d'une typologie d'erreurs lexicales

Notre typologie a donc pour objectif d'identifier des types d'erreurs, de chercher à faire formuler par les correcteurs les raisons qu'ils donnent à l'erreur identifiée, si peu évident que ce soit, et d'intégrer ce questionnement à toute démarche didactique.

Deux principes ont été adoptés lors de l'élaboration de ladite typologie. Le premier principe consiste en la réécriture contrôlée du segment erroné, déjà illustré (exemples 2 et 3). Ce procédé permet, à notre avis, de mieux cerner les difficultés scripturales générales aux niveaux de l'organisation du discours, du choix lexical et de la mise en mots. L'analyse concerne alors le lexique théoriquement attendu en fonction de l'isotopie dominante (matérialisé par le biais de la réécriture) et le lexique effectivement utilisé par le scripteur dans son texte, celui-ci réalisant des choix lexicaux en fonction à la fois de son identité et de l'objet de discours à traiter. L'analyse des énoncés obtenus par la réécriture tient ainsi compte des interrelations entre langue et discours dont le lexique est la manifestation. Le deuxième principe prend la forme d'une approche syntagmatique et paradigmatique dans le traitement des erreurs lexicales, fondée sur l'usage et les liens, lexicalement marqués, entre sémantique et syntaxe. Autrement dit, l'identification et l'analyse du segment erroné ne se

limite pas à l'unité lexicale : pour être efficace, une telle analyse doit tenir compte de l'environnement linguistique immédiat (le cotexte) de l'unité lexicale erronée ; solliciter, quand c'est nécessaire, des contextes suffisamment larges qui dépassent le cadre de la phrase et qui nécessitent d'analyser l'ensemble du texte ; et prendre en compte la situation de production du discours donné.

Conformément à ces deux principes (réécriture et contexte), notre typologie des erreurs lexicales distingue trois niveaux d'analyse allant du plus « global » (la sélection lexicale en fonction de l'acte de parole et de l'objet du discours) au plus « local » (le choix d'une unité lexicale et la formation du lexique), en passant par les relations sémantiques et syntaxiques de l'unité lexicale avec son cotexte (le lexique-grammaire). Les paragraphes qui suivent présenteront les caractéristiques des trois niveaux de classement illustrés d'exemples tirés des trois textes figurant dans les Annexes.⁽³⁾ Comme dans les exemples *supra*, nous avons consigné en italique des fragments qui nous ont paru problématiques ; le segment central, identifiant l'erreur, figure en droits. Entre crochets, est proposée une réécriture du segment incriminé, ce qui permet de caractériser l'erreur en termes sémantiques et énonciatifs.

2.1. Appréhension du niveau global (l'explication contextuelle)

Le premier niveau de notre typologie concerne les erreurs relevées au niveau de la mise en mots d'une intention signifiante globale (*sens intenté*) et de la disponibilité lexicale. La sélection lexicale est dépendante des caractéristiques personnelles de l'apprenant ainsi que des paramètres énonciatifs et thématiques de la production. À ce niveau sont recensés les erreurs de planification, d'énonciation et de modalités énonciatives, les erreurs de connecteurs et marqueurs de portée, d'orientation et point de vue du procès, et les erreurs de disponibilité de l'unité lexicale à la surface du texte en fonction du domaine référentiel identifié (isotopie).

- (4) *Selon les experts dans le secteur médical, on trouve que les médecines douces sont adaptées à beaucoup de maladies [...].* (Texte 3) <Selon les experts du secteur médical, les médecines douces seraient plus adaptées à beaucoup de maladies [...]. La sémantique et la pragmatique des temps verbaux. L'expression du point de vue du procès.
- (5) *J'ai appris comme conduire la voiture, il était un peu difficile pour moi. Donc, j'ai discuté avec l'autre personne qui a conduit bien. Puis, j'ai fais des exercices.* (Texte 1) <J'ai appris comment conduire une voiture. Mais c'était un peu difficile pour moi. J'ai discuté alors avec une autre personne qui conduisait bien.> L'emploi et la valeur sémantique des connecteurs.
- (6) *Après l'avoir préparé, tu pourras l'envoyer dans des entreprises que tu t'intéresses beaucoup.* (Texte 2) <[...] envoyer aux entreprises qui t'intéressent le plus.> Orientation et point de vue du procès.
- (7) *C'est vrai qu'on n'a pas encore de capacité de manifester les valeurs des médecines douces. C'est-à-dire que les recherches n'ont pas arriver à trouver les influences des médecines douces.* (Texte 3) <C'est vrai qu'on ne peut pas encore prouver les bienfaits des médecines douces car les résultats de recherches n'ont pas réussi à confirmer les effets des médecines douces.> La disponibilité des unités lexicales attendues à la surface du texte.

Même si les erreurs de ce niveau ne sont pas « purement » lexicales, nous avons

(3) Pour une description plus ample de la typologie d'erreurs lexicales, cf. O. Luste-Chaâ (op. cit.).

maintenu cette première catégorie comme la manifestation des liens qui unissent le lexique à deux autres versants linguistiques, ceux de la sémantique et de la pragmatique.

2.2. Collocations et cotexte étroit

Au deuxième niveau de la typologie, il s'agit d'observer l'influence réciproque de l'erreur et de son environnement immédiat (le cotexte). Sont regroupées ici les erreurs qui concernent essentiellement les relations qu'entretient une unité lexicale avec d'autres éléments de l'énoncé (propriétés de combinatoire grammaticale et lexicale). Se trouvent ainsi associées la représentation sémantique et l'organisation syntaxique : des erreurs d'emploi du passif, des quantificateurs, indéfinis et évaluatifs, des problèmes de forme et d'orientation d'énoncé, et des erreurs de collocations.

- (8) *Ici, tous les chose sont excellentes.* (Texte 2) <Ici, tout est excellent.> Les tours « indéfinis » en usage.
- (9) *Votre vie seront assez bien.* (Texte 2) <Ce sera la belle vie. / Tu aimeras ta vie. / Ta vie sera très bien.>. L'emploi des évaluatifs et des adverbess intensifs.
- (10) *Mais avec un marché de plus en plus florissant, il existe toujours des problèmes. Dans un premier temps, il durera normalement longtemps, si on choisit les médecines douces.* (Texte 3) <[...] Dans un premier temps, il faudra attendre longtemps avant que les gens ne recourent davantage aux médecines douces.> La forme et l'orientation de l'énoncé.
- (11) *En France, il n'y a pas de remboursement pour le choix des médecine douces.* (Texte 3) <En France, les médicaments homéopathiques ne sont pas remboursés. / Le choix de la médecine douce va de pair en France avec le non remboursement des médicaments.> La valeur du passif et l'implication à valeur hypothétique (si p, q).
- (12) *Alors, j'ai gagné des progress.* (Texte 1) <J'ai fait des progrès.> Collocation et expression verbale à partir du support *faire*.
- (13) *D'abord, il n'y a pas d'effet seconde.* (Texte 3) <D'abord, il n'y a pas d'effets secondaires.> Collocation, expression figée (dérivation erronée).

L'échantillon d'erreurs en contexte étroit (cotexte), ainsi que le classement auquel nous avons soumis les occurrences erronées, attestent que les problèmes linguistiques rencontrés dans le corpus sont loin d'être homogènes. Les erreurs lexicales recensées à ce niveau de classement dépendent certes moins du contexte énonciatif de la production, mais elles ne sont pas pour autant plus faciles à identifier, à analyser et surtout à classer, car elles font toujours ressortir les corrélations complexes du lexique avec la syntaxe et la sémantique, même dans un environnement étroit. Elles confirment que le lexique joue le rôle d'un indicateur fonctionnel de l'interface *forme-sens*.

2.3. Erreurs lexicales au sens strict

Le troisième niveau de notre typologie regroupe des erreurs que l'on pourrait qualifier d'« exclusivement » lexicales. Le segment erroné se limite souvent à une unité lexicale ou, tout au plus, au syntagme auquel celle-ci appartient. Le dernier exemple cité (13, *supra*) en offrait déjà une première illustration : *seconde* et *secondaire* relèvent, comme nous l'avons indiqué, des règles morphologiques de dérivation.

Les erreurs recensées dans ce dernier groupe relèvent en effet du niveau formel de la langue, notamment sur ses versants morphosyntaxique et morphosémantique (dérivation, construction, composition, figures de mots, relations lexicales, emploi

des prépositions), ainsi que des relations sémantiques entre les unités lexicales (hyponymie, hypéronymie, synonymie, polysémie, combinatoire sémantique, etc.). En voici quelques exemples :

- (14) *progress ; *exemple ; la *music ; *sourir (Texte 1). L'orthographe lexicale de la langue française.
- (15) *Chaque année, de plus en plus* d'animal sont disparaissent. (Texte 1) <Chaque année, de plus en plus d'animaux disparaissent.> La construction interne des unités lexicales, la morphologie flexionnelle.
- (16) *Maintenant, je travail comme ingénieur informatique dans une grande* entreprise logicielle. (Texte 2) <Maintenant, je travaille comme ingénieur informatique dans une grande entreprise de logiciels.> La composition lexicale (dérivation ou unité « construite »).
- (17) *Je crois que tu auras un travail* que tu rêves [...]. (Texte 2) <Je pense que tu auras un travail dont tu rêves [...].> La construction externe, l'emploi des prépositions régies.
- (18) *On peut trouver beaucoup de choses dans ce site (Par exemple : la music, le sportif, le politique etc.).* (Texte 1) <On peut y trouver beaucoup de choses, par exemple de la musique, du sport, de la politique, etc.> L'emploi d'un dérivé spécifique (*sportif*) dans le sens d'un nom générique (*sport*).
- (19) *J'ai gagné le permis de conduire.* (Texte 1) <J'ai obtenu (eu) mon permis>. Les propriétés de la combinatoire sémantique et leur application aux locutions verbales (*gagner, obtenir, avoir qqch*).
- (20) *On utilise de plus en plus de nouvelles manières pour soigner les malades, par exemple, la méthode de l'homéopathie.* (Texte 3) <On utilise de plus en plus de nouvelles méthodes pour soigner les malades, comme par exemple l'homéopathie>. La sélection d'un terme synonymique (les *manières*) ou d'un autre (les *méthodes*) en contexte (médecine).
- (21) *Mais elles ne pourront pas remplacer la place des médecines traditionnelles.* (Texte 3) <Mais elles ne pourront pas remplacer la médecine traditionnelle>. Pléonasm.

Dans cette dernière famille d'erreurs, l'unité repérée comme déviante est à part entière une forme-sens relevant d'une description morphosémantique. Les erreurs de ce niveau sont plus facilement repérables à la surface du texte, elles sont *marquées*, mais ne se soumettent pas pour autant plus facilement à un classement définitif.

Le travail sur la typologie d'erreurs lexicales que nous venons de présenter ci-dessus montre que décidément le lexique ne répond pas facilement à un travail classificatoire homogène.

2.4. Les limites d'une typologie lexicale

Plusieurs arguments nous paraissent pouvoir expliquer les limites d'une telle tentative. Tout d'abord, nous avons vérifié dans un corpus pourtant restreint que l'absence d'unité lexicale à la surface du texte est aussi essentielle que la présence d'une unité erronée. Ensuite, il nous est apparu que l'unité lexicale, associant une forme à un sens, n'était pas dissociable des données sémantiques et syntaxiques. Par conséquent, le classement esquissé doit être considéré comme un analyseur de ces interrelations et du caractère composite, incomplètement réglé du lexique d'une langue. Nous avons en outre été confortée dans la nécessité de dépasser, en contexte didactique, la séparation trop tranchée entre langue et discours. Les plans d'énon-

ciation, l'aspect, la référence, la quantification, les processus rédactionnels sont autant de structurations qui pèsent sur la sélection lexicale.

Enfin, nous restons consciente d'une certaine relativité de notre classement. En effet, d'un correcteur à l'autre, le jugement dépendant notamment de la norme de référence, il n'y aura pas toujours de consensus sur le degré de gravité et d'acceptabilité, ni sur la nature des erreurs que nous avons relevées, ni sur nos propositions de réécriture. D'autre part, il est aussi difficile de garantir l'étanchéité des catégories de classement, surtout quand il s'agit du lexique. En effet, il s'avère souvent impossible de déterminer catégoriquement de quel domaine relève une erreur lexicale donnée car certaines erreurs sont à la frontière entre tel ou tel domaine. Rares sont finalement les erreurs qui ne se situent qu'à un seul niveau et qui n'auraient pas d'influence (même indirecte) sur d'autres niveaux.

3. L'erreur lexicale dans une analyse longitudinale des productions écrites en FLE

Selon la vision *constructiviste* que nous avons adoptée pour aborder l'enseignement / apprentissage du FLS, l'erreur n'est plus un fait répréhensible, mais se révèle être « porteuse de sens » pour l'enseignant. Elle lui permet de juger, d'une part, de l'état de la compétence lexicale acquise par l'apprenant et, d'autre part, de sa progression en acquisition / apprentissage du français. Ainsi envisagée pour procéder à l'analyse longitudinale des textes produits par les apprenants FLE, l'erreur acquiert le statut d'indicateur et d'outil d'évaluation de leurs acquisitions lexicales. Quant à la typologie, elle permet de statuer notamment sur le type, le niveau d'infraction et la fréquence des erreurs recensées.

Comme nous l'avons déjà mentionné, les trois textes qui figurent dans les Annexes sont de la même auteure. Le premier répond à un « questionnaire de goût et d'opinion » tandis que les deux autres sont des discours élaborés (une incitation épistolaire à se rendre en Chine pour y travailler et une discussion argumentative sur les médecines douces). Correspondant aux trois étapes clés du cursus (les trois évaluations principales), ces écrits devraient nous permettre de voir l'évolution de la compétence langagière de l'étudiante en neuf mois d'enseignement / apprentissage du français et de faire ressortir les principales étapes du développement de sa compétence lexicale.

Très globalement, on estime que des progrès dans le domaine lexical ont été réalisés. C'est superficiellement le jugement que l'on peut formuler en constatant que les erreurs de dérivation (*disparaissent* au lieu de *disparaissent*, par exemple) et les anglicismes (*progress*) ont diminué et que la compétence en matière d'élaboration textuelle s'est renforcée.

Rappelons succinctement les trois niveaux de structuration envisagés pour classer les erreurs :

- Niveau supérieur de structuration (pragmatique) : les modalités d'énonciation et le contexte au sens large,
- Les unités lexicales en contexte d'énoncé (le lexique-grammaire) : par exemple l'emploi du passif, la gestion des quantifieurs, la forme et l'orientation des énoncés, les collocations et les expressions idiomatiques,
- La formation et l'organisation du lexique : morphologie lexicale, structure actancielle des verbes, relations sémantiques et combinatoire sémantique des unités lexicales.

Nous reprenons ce classement dans l'analyse longitudinale qui suit.

3.1. Les indicateurs lexicaux du niveau supérieur de la structuration

Concernant la planification et les modalités énonciatives, trois catégories présentent des difficultés au scripteur : l'emploi des connecteurs, l'emploi des temps verbaux et l'actualisation des unités lexicales attendues en fonction de l'objet du discours. Pour ce qui est de l'enchaînement des informations, le lecteur peut déplore un manque de « fluidité ». Mais cette impression de lecture n'est pas forcément imputable à un défaut d'organisation ou des marqueurs d'organisation du discours qui seraient absents ou défaillants. Au contraire. La réponse 5 du Texte 1 et le Texte 2 sont jalonnés de connecteurs. À cet égard on aura noté une planification assez « scolaire » qui obéit à un principe chronologique et causal (*d'abord, puis, enfin, donc*) ou bien qui abuse d'énoncés généraux dont le propos est quelque peu banal ou naïf (*D'ailleurs tu vas connaître une nouvelle culture*, Texte 2). Mais cette impression de lecture heurtée tient peut-être davantage à la formation des énoncés, perçus comme « peu naturels ». D'une part, les expressions idiomatiques sont absentes ou fautives – voir notre point suivant –, d'autre part le choix des verbes modaux peut occasionnellement priver le discours du liant énonciatif et aspectuel dont il a besoin, mais enfin et surtout la cohésion référentielle s'avère délicate.

Par exemple, concernant la cohésion nominale et les enchaînements interphrastiques, on repère des difficultés dans la maîtrise des liens anaphoriques. Le Texte 2 offre un aperçu de telles difficultés. Est ici (22) en cause la distance qui sépare le pronom de son antécédent référentiel :

- (22) *Parce que on utilise souvent l'anglais quand on travaille. Après l'avoir préparé, tu pourras l'envoyer dans des entreprises que tu t'intéresses beaucoup. <Après avoir préparé ton CV, tu pourras l'envoyer [...]>.*

Sans doute la rupture de l'enchaînement passerait-elle inaperçue à l'oral, au moins y serait-elle moins gênante. Mais l'écrit nécessite un accès référentiel explicite.

3.2. Collocations artificielles et expressions idiomatiques

Le réseau des connexions lexicales n'étant pas aussi performant que chez un locuteur natif, l'apprenant non francophone ne parvient pas toujours à actualiser le vocabulaire requis par l'isotopie dominante et recourt à des stratégies compensatoires différentes pour combler ses lacunes : du vocabulaire générique, répétitif et passe-partout aux périphrases en passant par le recours aux unités lexicales (traduites littéralement ou laissées inchangées) qui sont empruntées à d'autres langues étrangères, notamment l'anglais. Si l'emploi du vocabulaire non traduit caractérise plutôt le premier texte, dans le dernier, on constate le recours à des stratégies plus élaborées, avec les moyens disponibles en langue cible (périphrases, emploi des synonymes). La disponibilité d'une unité lexicale dans le lexique mental de l'apprenant dépend aussi de la qualité et le nombre de connexions qui la relie à d'autres unités, ceci dépendant à son tour des occurrences de l'unité lexicale donnée auxquelles a été exposé l'apprenant. Autrement dit, si le vocabulaire de l'étudiante en question est tout à fait « acceptable » si l'on considère qu'il permet de comprendre le sens voulu (surtout dans les Textes 2 et 3), il gâgnerait beaucoup à être « plus précis ».

Les collocations sont le domaine central des acquisitions lexicales, celui qui requiert à la fois une attention particulière de la part de l'enseignant et un effort cognitif de mémorisation de la part de l'apprenant. En effet, les liens qui unissent les constituants d'une collocation ne sont pas déductibles à partir du sens de l'unité lexicale qui est la base de collocation, mais sont en même temps soumis aux con-

traintes des propriétés de combinatoire de ladite unité. Si pour le locuteur natif, la plupart des collocations sont prédictibles grâce à son intuition, le locuteur non francophone ne bénéficie pas de cet avantage, d'où les erreurs fréquentes de collocations. C'est pourquoi il n'est pas étonnant de constater que les erreurs de collocations sont présentes dans toutes les trois productions : par exemple, *arrêter la situation*, *gagner des progrès* (Texte 1), *effet seconde*, *être mal à la tête* (Texte 3), etc.

3.3. Morphologie et formation lexicale

Quant à la formation des unités lexicales, c'est le Texte 1 qui se distingue par le nombre d'erreurs de ce type qui y sont très nombreuses. Il s'agit notamment des erreurs d'orthographe lexicale (*progress, *exemple, la *music, *sourir, etc.), de morphologie (par exemple, *de plus en plus* *d'animal sont disparaissent ; *j'ai fais des exercices, etc.) et de morphosyntaxe (*de plus en plus* *d'animal ; *La plus* *important pour l'humanité est la nature ; *le situation). La production suivante (Texte 2) offre une image à peine meilleure de ce point de vue. Ce n'est que dans le dernier texte que l'on peut constater une amélioration, due probablement au fait que l'étudiante se sentait déjà suffisamment à l'aise en français pour se concentrer simultanément sur le traitement des éléments de bas et de haut niveaux.

Si le nombre de certaines erreurs décroît avec la progression dans l'acquisition / apprentissage de la langue (par exemple, les erreurs de morphologie et d'orthographe lexicales, d'orientation de l'énoncé, d'emploi des temps verbaux), d'autres persistent et relèvent des aspects linguistiques apparemment les plus difficiles à acquérir, à savoir les structures actancielles des verbes, les collocations et les propriétés de la combinatoire sémantique des unités, la synonymie et la polysémie, ainsi que la valeur du passif et l'emploi des prépositions régissantes. Nous noterons aussi la présence à ce stade des erreurs dues à l'influence du sens voulu sur le choix des unités lexicales, leur forme grammaticale et le cotexte, ce qui implique une analyse multidimensionnelle de ce type d'erreurs car le classement n'est pas du tout aisé.

L'analyse longitudinale des productions montre l'évolution des compétences de l'auteure qui passe d'un vocabulaire approximatif, générique ou répétitif dans le Texte 1 à des emplois lexicaux plus riches et plus précis dans la dernière production, entraînant des erreurs d'un type nouveau, d'ordre plus sémantique.

Conclusion

Le corpus d'observation constitué à partir des productions écrites d'un groupe-témoin a permis de resserrer la problématique de la recherche autour de la notion d'erreur lexicale et de nous concentrer essentiellement sur son statut et son rôle comme indicateur des difficultés scripturales et du niveau de maîtrise lexicale. Ce premier travail a permis de comprendre l'intérêt d'une analyse ponctuelle et longitudinale d'erreurs lexicales combinant la réflexion linguistique de la part du correcteur (enseignant) et une visée didactique. À la suite de l'application de notre typologie d'erreurs dans l'analyse longitudinale des productions écrites d'apprenants, il nous est devenu possible de mieux orienter la réflexion sur les activités didactiques portant sur les domaines les plus problématiques pour les apprenants (les collocations, la combinatoire sémantique, l'emploi et la valeur du passif, le choix des prépositions, etc.). Les pistes qu'offre une typologie d'erreurs lexicales sont certes nombreuses, mais demandent des investigations complémentaires, plus spécifiquement didactiques, qui visent des apprentissages identifiés et circonscrivent des objets d'enseignement ; la typologie, évolutive, n'a d'autre but que d'outiller cette démarche.

Bibliographie

- CATACH, N. (1980) : *L'orthographe française : traité théorique et pratique avec des travaux d'application et leurs corrigés*, Paris, Nathan Université.
- CATACH, N., DUPREZ D., LEGRIS M. (1980) : *L'enseignement de l'orthographe : l'alphabet phonétique international, la typologie des fautes, la typologie des exercices*, Paris, Nathan.
- CONSEIL DE L'EUROPE (2001) : *Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer*, Paris, Didier.
- CORDER, P.S. (1980) : « Que signifient les erreurs des apprenants ? », *Langages*, 57, p. 9-15.
- FREI, H. ([1929] 1971) : *La grammaire des fautes*, Genève, Slatkine Reprints.
- GARCIA-DEBANC, C., FAYOL, M. (2002) : « Apports et limites des modèles du processus rédactionnel pour la didactique de la production écrite. Dialogue entre psycholinguistes et didacticiens », *Pratiques*, 115 / 116, pp. 37-50.
- LUSTE-CHAAÛ, O. (2009) : *Les acquisitions lexicales en français langue seconde : conceptions et applications*, Thèse de doctorat, 2 volumes, Université Paul Verlaine – Metz.
- MASSERON, C. (2001) : « Du projet de discours à la langue du discours produit : nature et enjeux des erreurs scripturales », *Pratiques*, 109-110, pp. 207-247.
- (2005) : « Indicateurs langagiers et stratégies scripturales – Du discours à la langue », *Pratiques*, 125-126, pp. 205-249.
- MASSERON, C., LUSTE-CHAAÛ, O. (2008) : « Typologie d'erreurs lexicales : difficultés et enjeux », in J. Durand, B. Habert, B. Laks (resp.), *Actes du 1^{er} Congrès mondial de linguistique française*, Paris, 9-12 juillet 2008, p. 62, 519-531. Disponible en ligne : [<http://www.linguistiquefrancaise.org/index.php?option=article&access=doi&doi=10.1051/cmlf08230>], consulté le 23.06.2009.
- MONTE, M., TOUCHARD, Y., SANTACROCE, M., VÉRONIQUE, D., VION, R. (1992) : « Quand des enseignants réécrivent des textes d'élèves... », *Pratiques*, 73, pp. 96-118.
- SCHLYTER, S. (2003) : « Stades de développement en français L2. Exemples d'apprenants suédophones, guidés et non-guidés, du "corpus Lund" ». Disponible en ligne : [luur.lub.lu.se/luur?func=downloadFile&fileOID=624010], consulté le 09.06.2009.

Annexes

Texte 1 (octobre 2005)

- « 1. Quel est votre site favori sur internet ? Décrivez-le.
2. Qu'est-ce qui vous étonne le plus dans le monde actuel ?
3. Quelle est pour vous la découverte la plus importante pour l'humanité ?
4. Que vous inspire un beau ciel étoilé ?
5. Quel a été votre dernier sujet d'énervement ? Racontez.
6. Quel est l'objet qui vous est indispensable ? Pourquoi ?
7. Quel animal vous fascine ? Pourquoi ?
8. Que changeriez-vous dans le monde ?
9. Quelle est l'époque à laquelle vous auriez aimé vivre ? Pourquoi ?
10. Qu'appréciez-vous et que détestez-vous chez les Français ? »

1. Mon site favori sur internet est www.sina.com.cn. Il est le site le plus grand en Chine. On peut trouver beaucoup de choses dans ce site (Par exemple : la music, le sportif, le politique etc.) Je l'utilise pour bavarder avec mes amis qui sont à l'étranger. On peut utiliser aussi e-mail dans ce site.

2. Il y a encore la guerre dans le monde actuel.

3. La plus important pour l'humanité est la nature. Chaque année, de plus en plus d'animal sont disparaissent. il est très dangereux pour l'humanité. Il faut arrêter le situation.

4. Je vais sourir et Je vais me sens très bien. Quand je marche dans la rue, je vais chanter.

5. Mon dernier sujet d'énervement s'est passé en Chine. J'ai appris comme conduire la voiture il était un peu difficile pour moi. Donc, j'ai discuté avec l'autre personne qui a conduit bien. Puis, j'ai fais des exercices. Alors, j'ai gagné des progress. J'étais très heureuse. En fin, j'ai passé l'examen et j'ai gagné le permis de conduire.

6. C'est ma famille. J'aime mon père et ma mère

7. Le chien. Parce que ils sont très aimables et ils sont très utiles pour les gens.

Texte 2 (février 2006)

« Un ami français, spécialisé en informatique, vous a écrit pour vous annoncer qu'il souhaite s'installer en Chine. Vous lui répondez par une lettre en lui donnant des informations et des conseils qui pourraient lui être utiles (démarches administratives, logement, salaire, adresses utiles, façon de présenter un CV...). »

Nanling, 10-02-2006

Chère Laura,

Quels temps passe vite ! Après avoir obtenir mon diplôme français, il y a deux mois que j'ai retourné en Chine. Maintenant, je travail comme ingénieur informatique dans une grande entreprise logicielle.

Ici, tous les chose sont excellentes. J'ai déjà sais que tu avais eu beaucoup de difficultés pour trouver un job en France. Dans ce cas, je suis très contente que tu puisses changer ton avis. Travailler en Chine, ça sera un bon choix. En Chine, on a besoin de beaucoup de gens qui ont les techniques avancées grâce au développement de l'économie.

Il y a des milliers d'emplois disponibles, surtout dans le domaine d'informatique. Si tu es en Chine, tu auras plus de chances. D'ailleurs, tu vas connaître une nouvelle culture. Tu pourras essayer de parler le chinois qui est de plus en plus utile dans le monde entier.

Pour la réussite de tes projets, d'abord, il faut préparer

Un CV. le formule de CV est pareil comme celui qui est en France.

C'est important que tu puisses parler l'anglais et le français. Parce que on utilise souvent l'anglais quand on travaille. Après l'avoir préparé, tu pourras l'envoyer dans des entreprises que tu t'inté-

resses beaucoup. Je te donne les sites, comme www.sina.com.cn etc. Tu obtiendras des informations utiles sur ces sites. Je crois que tu auras un travail que tu rêves grâce à ton diplôme de Master en informatique. Tu as aussi eu un bon stage en France pendant six mois. Donc, tu auras un avenir superbe.

Ici, les consommations ne sont pas chères. Il ne faut pas payer beaucoup pour un logement. Comme moi, je paie environ 40 euro par mois. Mais, tu vas gagner l'argent beaucoup. Tu seras satisfait à son salaire. Donc tu pourras vivre en Chine. On pourra partir en voyage régulièrement. Votre vie sera assez bien. Si tu as des problèmes en Chine, je vais t'aider directement.

J'espère que tu iras en Chine le plus tôt possible.

Je t'embrasse.

Texte 3 (juin 2006)

« Rédigez un texte argumenté d'au moins 25 lignes, présentant les avantages et les limites des médecines douces. Votre texte devra comporter une introduction, deux parties distinctes présentant des arguments illustrés d'exemples et une conclusion. »

Les médecines douces

De nos jours, les médecines douces comme les médecines naturelles, se développent fortement. On utilise de plus en plus de nouvelles manières pour soigner les malades, par exemple, la méthode de l'homéopathie.

Avec une série des avantages, de plus en plus de gens préfèrent les médecines douces. D'abord, il n'y a pas d'effet secondaire. C'est un bon choix pour les gens âgés, ainsi que les enfants et les femmes.

Puis, les prix sont relativement plus bas grâce à l'utilisation des produits naturels. Ensuite, les médecines douces sont plus efficaces que les médecines traditionnelles. Je vous donne un exemple qui s'est passé chez moi. Après avoir fait une piqûre chez un médecin doux, ma mère n'a plus de problèmes pour l'endormir maintenant, alors qu'elle était toujours mal à la tête toute la nuit après avoir pris des médicaments sans effets un an plus tôt. Selon les experts dans le secteur médical, on trouve que les médecines douces sont adaptées à beaucoup de maladies, surtout les maladies chroniques. Contrairement, dans les pays occidentaux, il faut dépenser beaucoup d'argent pour les recherches des nouveaux médicaments. Le chiffre d'affaire a déjà baissé depuis longtemps dans les industries médicales.

Mais avec un marché de plus en plus florissant, il existe toujours des problèmes. Dans un premier temps, il durera normalement longtemps, si on choisit les médecines douces. En France, il n'y a pas de remboursement pour le choix des médecines douces. Donc, les pauvres ne peuvent pas bénéficier des nouvelles médecines. Dans un second temps, une part des gens ne croit pas à l'effet des médecines douces.

D'après eux, les médecines douces, comme les placebos, sont inutiles.

C'est vrai qu'on n'a pas encore de capacité de manifester les valeurs des médecines douces. C'est-à-dire que les recherches n'ont pas réussi à trouver les influences des médecines douces.

Je crois que les chercheurs vont réussir à nous démontrer les valeurs des médecines douces. L'avenir des médecines douces sera plus attirant. La plupart des patients ne pourront plus souffrir terriblement d'une opération. La place des médecines douces restera la plus importante dans le domaine médical. Mais, ils ne pourront pas remplacer la place des médecines traditionnelles.