Portrait des erreurs lexicales d'élèves de 3^e secondaire en production écrite et proposition de pistes didactiques

Dominic Anctil

Édition électronique

URL : <http://journals.openedition.org/pratiques/3447>

DOI : 10.4000/pratiques.3447

ISSN : 2425-2042

Éditeur

Centre de recherche sur les médiations (CREM)

Édition imprimée

Date de publication : 15 décembre 2012

Pagination : 7-30

Référence électronique

Dominic Anctil, « Portrait des erreurs lexicales d'élèves de 3^e secondaire en production écrite et proposition de pistes didactiques », *Pratiques* [En ligne], 155-156 | 2012, mis en ligne le 18 décembre 2017, consulté le 19 avril 2019. URL : <http://journals.openedition.org/pratiques/3447> ; DOI : 10.4000/pratiques.3447

Portrait des erreurs lexicales d'élèves de 3^e secondaire en production écrite et proposition de pistes didactiques

Dominic Anctil

Université de Montréal, Département de didactique
Observatoire de Linguistique Sens-Texte (OLST)

L'enseignement du lexique est depuis longtemps considéré comme l'un des parents pauvres des cours de français langue maternelle (Bastuji, 1978 ; Simard, 1994 ; Préfontaine, 1998). Rarement objet d'un enseignement systématique et organisé, le vocabulaire est généralement abordé en classe de façon incidente (Vancomelbeke, 2004), et son étude est toujours subordonnée à d'autres apprentissages (Reboul-Touré, 2002). Pourtant, beaucoup d'enseignants continuent de se plaindre de la pauvreté lexicale des élèves (Rocquet, 1998 ; Dreyfus, 2004 ; Vancomelbeke, 2004). Ces commentaires récurrents sur le présumé manque de vocabulaire des jeunes ont constitué le point de départ de notre travail.

En classe, le vocabulaire donne rarement lieu à des évaluations spécifiques; il est presque essentiellement apprécié à travers les productions écrites des élèves, généralement par une évaluation qualitative et globale et à l'aide d'échelles descriptives reposant sur des critères flous comme la richesse, la variété ou le « caractère évocateur » des mots utilisés (Anctil, 2011). Une telle approche d'évaluation nous laisse donc avec un portrait peu précis des lacunes lexicales des élèves et bien peu d'indices concrets pour mesurer leur compétence lexicale réelle.

Dans le cadre de notre recherche, nous avons donc choisi de décrire plus finement la dimension lexicale des textes des élèves, en nous concentrant précisément sur leurs erreurs. Comme le mentionnent Granger et Monfort (1994), l'analyse des erreurs lexicales d'un apprenant ne permet pas à elle seule de rendre compte de sa compétence lexicale, mais constitue une porte d'entrée intéressante pour faire la lumière sur l'une des facettes : la maîtrise qu'a l'apprenant des unités lexicales qu'il utilise spontanément en écriture. Dans le but de mieux éclairer la question lexicale en production écrite dans une

perspective didactique, nous nous sommes aussi intéressé au traitement de l'erreur lexicale par les enseignants lors de la correction de textes.

Nous définirons d'abord la notion d'erreur lexicale, avant de présenter notre méthodologie de recherche et d'exposer notre grille d'analyse des erreurs lexicales. Nous ferons ensuite un tour d'horizon des erreurs relevées dans notre corpus, avant de proposer certaines pistes didactiques pour l'enseignement du lexique et la correction d'erreurs lexicales.

1. L'erreur et l'erreur lexicale : définitions

Reuter, Cohen-Azria, Daunay, Delcambre et Lahanier-Reuter (2007) mentionnent la difficulté, dans une perspective didactique, de bien caractériser le concept d'*erreur* et le fait qu'il a été peu défini dans les travaux sur l'erreur et l'évaluation. On définit souvent l'erreur linguistique comme un écart par rapport à la norme (voir par exemple Debyser, Houis et Rojas, 1967 ; Bailly, 1998), ce qui nous laisse avec le problème, non moins épineux, d'établir ce qu'est la norme, tâche d'autant plus difficile lorsqu'on s'intéresse à la dimension lexicale de la langue (Guilbert, 1972 ; Siouffi et Steuckardt, 2001). En effet, la pluralité des normes (Brent, 1999 ; Verreault, 1999) et l'importance de la variation en matière de lexique rendent difficile l'établissement d'une norme unique, ce qui se manifeste notamment par les tiraillements entre norme descriptive et prescriptions normatives dans les ouvrages lexicographiques (Rey, 1972). Pour ces raisons, nous privilégions, à l'instar de Lennon (1991) et Legendre (2005), une définition de l'erreur qui s'articule autour du concept de *locuteur expert* plutôt que de celui de norme.

***Erreur linguistique :** forme linguistique ou combinaison de formes linguistiques qui diffère de ce qu'un locuteur expert aurait selon toute probabilité produit dans un même contexte de production.*

Un locuteur expert est un locuteur natif idéalisé possédant une très bonne maîtrise de la langue et un vocabulaire étendu ; privilégiant un emploi standard de la langue, il sait néanmoins s'adapter aux différentes situations de communication (notamment par l'emploi de différents registres de langue) et transgresser certaines règles pour produire des effets de style. Définir l'erreur par rapport au concept de locuteur expert ne nous libère pas du concept de norme (puisque le locuteur expert respecte généralement les prescriptions normatives), mais nous offre une certaine souplesse, particulièrement pertinente pour traiter de problèmes lexicaux, dont plusieurs sont soumis à des normes très floues. Nous verrons dans notre méthodologie de quelle façon nous avons opérationnalisé le concept de locuteur expert.

Une remarque terminologique s'impose par rapport à l'utilisation du terme *erreur*. L'erreur, telle que nous la définissons ici, englobe tous les écarts linguistiques par rapport à ce que produirait un locuteur expert, indépendamment de leur « gravité » ou de leur systématicité. Elle désigne tant les problèmes « condamnables » en vertu d'une norme claire que les emplois simplement marqués par rapport à l'usage normal, qui pourraient être qualifiés

plutôt de *maladresses*. Nous privilégions le terme *erreur* pour désigner tous les problèmes lexicaux parce qu'il est le plus utilisé dans la littérature⁽¹⁾, mais il faut garder en tête que les erreurs présentées dans le présent article ne mériteraient pas toutes la même attention didactique, ni qu'elles devraient être systématiquement sanctionnées lors de la correction par les enseignants.

Maintenant que nous avons défini le terme *erreur*, intéressons-nous plus spécifiquement à définir l'erreur lexicale.

L'erreur lexicale a donné lieu à assez peu de travaux, particulièrement en L1⁽²⁾, et, comme le souligne Agustín Llach (2005), elle est rarement définie clairement dans les travaux s'y intéressant, et souvent sur la base de critères différents. Certains travaux limitent leur conception de l'erreur lexicale aux problèmes sémantiques, alors que d'autres proposent une définition de l'erreur lexicale couvrant un spectre de problèmes beaucoup plus large, notamment en y incluant des erreurs en apparence syntaxiques⁽³⁾. Pour proposer une délimitation précise de l'erreur lexicale, il est indispensable d'avoir une conception claire de ce qu'implique la maîtrise d'un « mot » ; en effet, la connaissance d'une unité lexicale dépasse la simple connaissance de son sens et de son orthographe (Richards, 1976 ; Nation, 2001). Or, les travaux sur l'erreur lexicale semblent rarement reposer sur un cadre théorique précis, qui permettrait de bien circonscrire le concept d'*erreur lexicale*.

Notre conception de l'erreur lexicale s'appuie sur la Théorie Sens-Texte (TST) (Mel'čuk, 1997 ; Milićević, 2006), et plus particulièrement sur la *Lexicologie explicative et combinatoire* (LEC) (Mel'čuk, Clas et Polguère, 1995), branche lexicologique de la TST. L'intérêt de cette théorie linguistique pour notre travail est qu'elle accorde une place de choix à la lexicologie et propose une description exhaustive des unités lexicales, description qui rend bien sûr compte de la forme et du sens des lexies décrites, mais aussi de leur *combinatoire restreinte*, c'est-à-dire « l'ensemble des contraintes propres à [un] signe qui limitent sa capacité de se combiner avec d'autres signes linguistiques et qui ne peuvent être déduites ni de son signifié ni de son signifiant. » (Polguère, 2008 : 40) Le terme *combinatoire* est ainsi entendu dans un sens très large : la combinatoire restreinte d'un mot regroupe l'ensemble des propriétés qui régissent son « comportement » lorsqu'il apparaît en contexte de phrase et limitent les situations de parole dans lesquelles ce mot peut être utilisé ; elles peuvent être d'ordre syntaxique (régime verbal, partie du discours), morphologique (invariabilité, genre nominal), pragmatique (registre de langue) et collocationnel⁽⁴⁾. L'inclusion de ces propriétés de combinatoire dans la description d'une unité lexicale permet de mettre en lumière l'ensemble des

(1) Et beaucoup moins connoté que le terme *faute* utilisé par certains auteurs, dont Debyser (1967), Patris et Vansnick (1992) et Denance (1999).

(2) Parmi les travaux s'y étant intéressés dans une perspective de L2, citons Luste-Chaa (2009), Milićević et Hamel (2007), Agustín Llach (2004), Engber (1995), Arnaud (1984 ; 1992) et Zimmermann (1986).

(3) Plusieurs auteurs ont d'ailleurs souligné la difficulté à départager erreurs lexicales et erreurs syntaxiques (Patris et Vansnick, 1992 ; Denance, 1999).

(4) La LEC accorde d'ailleurs une place de choix à la description des collocations liées à chaque unité lexicale, à l'aide d'un formalisme sophistiqué appelé *fonctions lexicales* (pour une introduction, voir Mel'čuk *et coll.*, 1995).

connaissances liées à sa maîtrise, de clarifier les interactions entre lexique et syntaxe et, par conséquent, de mieux circonscrire les problèmes qui découlent de la méconnaissance d'une unité lexicale.

Voici, sur la base de cette conception de l'unité lexicale, la définition que nous proposons de l'erreur lexicale, définition qui se rapproche de celles de Granger et Monfort (1994) et de Hamel et Milićević (2007).

***Erreur lexicale :** erreur linguistique qui concerne une propriété inhérente à l'unité lexicale, que cette propriété soit d'ordre formel, sémantique, lexicogrammatical, collocationnel ou pragmatique.*

En d'autres mots, nous pourrions dire que nous sommes confrontés à une erreur lexicale chaque fois que l'explication de cette erreur consiste à pointer une propriété spécifique à une unité lexicale, peu importe la nature de cette propriété. C'est sur la base de cette définition de l'erreur lexicale que nous avons développé la typologie descriptive des erreurs lexicales qui nous a permis de classer les erreurs relevées dans notre corpus.

2. Méthodologie

2.1. Le corpus analysé

Le corpus de textes analysé comprend 300 rédactions d'élèves francophones⁽⁵⁾ de 3^e secondaire (14 ou 15 ans) réalisées en classe à l'automne 2007 ou au printemps 2008 dans le cadre d'une évaluation sommative de fin d'étape ou de fin d'année scolaire. Les productions proviennent de quatre écoles de Montréal et des alentours et ont une longueur moyenne de 385 mots. Nous avons choisi de nous concentrer sur les élèves francophones dans le but d'assurer une certaine homogénéité au corpus d'erreurs recueilli.

Les textes recueillis correspondent aux types à l'étude en 3^e secondaire : environ les deux tiers des textes étaient donc à dominante narrative, alors que le tiers restant étaient à tendance descriptive/explicative. Le thème abordé dans les rédactions varie d'un établissement à l'autre. Environ la moitié des élèves (148) devaient écrire un conte merveilleux (texte narratif) à partir d'éléments proposés dans la consigne de rédaction (personnage, lieu, événement déclencheur). Près d'un sixième des rédactions (48) consistaient en la description d'une publicité et la présentation des techniques publicitaires qui y sont utilisées (texte descriptif/explicatif). Finalement, le tiers des productions écrites du corpus (104) portaient sur le thème du travail d'été ; les élèves étaient invités soit à exposer les avantages et les inconvénients d'un emploi d'été (texte explicatif), soit à parler de leur première expérience en tant que jeune occupant un tel emploi (texte narratif). Un travail de préparation à l'écriture a été réalisé en classe pour l'ensemble des groupes et nous pouvons supposer que ce travail préalable a fourni aux élèves un certain vocabulaire de base pour traiter le thème

(5) Nous avons considéré comme francophone un élève ayant été scolarisé en français depuis la maternelle et parlant le français à la maison avec au moins un membre de sa famille.

imposé (par exemple, vocabulaire lié au monde de la publicité ou à l'univers du conte merveilleux). Lors de la rédaction, le recours au dictionnaire, à la grammaire et à un guide de conjugaison était permis.

Comme la plupart des productions écrites avaient déjà été évaluées au moment de notre collecte de données, nous avons pu obtenir 224 rédactions corrigées par les enseignants, ce qui représente 75 % du corpus. Ces copies nous ont permis d'observer de quelle façon les enseignants traitent l'erreur lexicale dans leur correction. Ces traces des pratiques de correction ont été complétées par une série de huit entretiens semi-dirigés menés auprès d'enseignants de français de 3^e secondaire, entretiens au cours desquels nous avons fait émerger les conceptions des enseignants relativement à l'erreur lexicale et avons recueilli leurs pratiques déclarées de correction de textes.

2.2. La correction des textes et le codage des erreurs

Une fois les rédactions recueillies et anonymisées, chaque texte a fait l'objet d'une première lecture, dans le but de bien en comprendre le sens, puis d'une seconde, au cours de laquelle les erreurs lexicales ont été relevées, puis retranscrites dans notre logiciel d'analyse (*Excel*), avec un contexte suffisant pour permettre de bien comprendre le problème⁽⁶⁾. Une correction était ensuite systématiquement proposée, avec un souci de rester le plus près possible de la formulation de l'élève, ce que Milićević et Hamel (2007) appellent une *correction minimale*.

Une description linguistique de l'erreur à l'aide de plusieurs critères constituait le cœur de notre codage : classe de mots de l'unité lexicale erronée, type d'unité lexicale en jeu (lexème ou locution), classement de l'erreur à l'aide de la typologie des erreurs lexicales et évaluation de sa gravité sur la base de trois critères (position des dictionnaires retenus pour l'étude, impact sur la compréhension du message, perception de l'intégration de l'erreur à l'usage). Finalement, une hypothèse sur la (ou les) source(s) possible(s) de l'erreur était avancée dans le cas où une telle hypothèse nous apparaissait naturellement. Les résultats présentés ici concernent notre premier niveau d'analyse, soit la description linguistique des erreurs ; la gravité des erreurs lexicales et leur explication ne seront pas abordées.

2.3. La typologie descriptive des erreurs lexicales

Vous trouverez à la page 13 la typologie descriptive des erreurs lexicales développée à partir de nos travaux antérieurs (Ancil, 2005) et des classements proposés par d'autres auteurs (Arnaud, 1984 ; Zimmermann, 1986 ; Patris et Vansnick, 1992 ; Granger et Monfort, 1994 ; Engber, 1995 ; Agustín Llach, 2004 ; Milićević et Hamel, 2007). Notre classement regroupe les différents types d'erreurs lexicales selon la dimension de l'unité lexicale dont elles relèvent.

(6) Ce contexte minimal permettant de détecter l'erreur correspond à la notion de *domaine de l'erreur* chez Lennon (1991).

Nous spécifions que notre typologie est descriptive, parce nous avons voulu éviter le chevauchement entre la description des erreurs et leur explication, typique de plusieurs travaux sur l'analyse d'erreurs (Porquier, 1977 ; James, 1998 ; Granger, 2003). Ainsi, les sous-classes d'erreurs présentes dans notre grille demeurent résolument descriptives et ne visent qu'une caractérisation linguistique des problèmes relevés ; l'explication de la source des erreurs constituait, comme nous l'avons mentionné plus haut, une étape subséquente dans notre démarche d'analyse.

2.4. Opérationnalisation du concept de locuteur expert

Nous avons souligné le caractère flou de la norme lexicale, ce qui nous a conduit à définir l'erreur lexicale comme un écart à ce qu'aurait produit un locuteur expert dans un même contexte de production. Si ce choix nous permet en quelque sorte de contourner la notion de norme, il ne l'occulte pas complètement, puisqu'un locuteur expert est quelqu'un dont l'usage correspond à la norme. Il nous a donc fallu trouver un moyen d'opérationnaliser cette norme lexicale, afin d'éviter de se poser, en tant que chercheur locuteur expert, comme seul détenteur de la norme. Le dictionnaire demeurant la meilleure illustration de la norme lexicale (Guilbert, 1972), nous avons retenu trois ouvrages comme balises dans notre analyse : *Le Petit Robert électronique*, ainsi que deux dictionnaires québécois ⁽⁷⁾, *Le Multidictionnaire* (de Villers, 2009) et le logiciel *Antidote HD* (2009).

Nous avons considéré comme correct tout emploi lexical apparaissant dans l'un ou l'autre de ces dictionnaires, pourvu bien sûr que cet emploi soit approprié dans le contexte dans lequel il était utilisé (*cf.* registres de langue). De la même façon, tout emploi condamné explicitement par l'ensemble des ouvrages a été considéré comme une erreur ; dans le cas où les ouvrages avaient des positions différentes relativement à certains usages, nous avons adopté celle du dictionnaire le plus permissif. Ce n'est que dans les cas où les ouvrages ne faisaient aucunement mention du cas problématique analysé que nous avons dû nous fier à nos intuitions de locuteur, que nous avons systématiquement validées à l'aide d'Internet.

3. Résultats

La correction des 300 rédactions a révélé un total de 1144 erreurs lexicales ⁽⁸⁾, ce qui correspond à une moyenne de 3,5 erreurs par 350 mots ⁽⁹⁾. La *figure 1* présente la répartition des erreurs des différents types.

(7) Comme les textes analysés ont été produits en contexte québécois, il était primordial que les ouvrages sur lesquels repose notre analyse reflètent bien la norme au Québec. Ces ouvrages fournissent une description fine et nuancée du français québécois et contiennent beaucoup d'indications relatives aux registres de langue, qui permettent de bien distinguer ce qui relève du français québécois standard de ce qui appartient au registre familier.

(8) À noter que nous avons exclu les 1426 erreurs d'orthographe d'usage de notre corpus.

(9) Si nous décidons ici d'établir la moyenne pondérée pour 350 mots, c'est qu'il s'agit de la longueur de texte minimale spécifiée dans les consignes d'écriture.

CLASSES D'ERREURS			
FORME	Erreur d'orthographe lexicale ex. <i>chemain, dedant, endomagé</i>		
	Barbarisme (signifiant déformé) ex. <i>désambuler</i> (pour DEAMBULER), <i>suptil</i> (→ SUBTIL)		
	Erreur de flexion ex. <i>travails, armes à feux</i>		
	Forme analytique ex. <i>un poison qui fait dormir</i> (→ SOMNIFERE)		
	Unité lexicale inexistante ex. <i>un village sombré</i> (au sens de 'plongé dans le noir')		
SENS	Utilisation inappropriée d'un quasi-synonyme ex. <i>il n'y avait pas d'issue pour entrer</i> (→ OUVERTURE)		
	Utilisation inappropriée d'une unité lexicale de sens proche ex. <i>Il passait tous ses passe-temps sur son projet.</i> (→ TEMPS LIBRES)		
	Impropriété	Sens fictif ex. <i>Ce logo fait appel à un arbre en fleurs.</i> (→ EVOQUER)	
		Anglicisme sémantique ex. <i>Elle pouvait voir dans le futur.</i> (→ AVENIR)	
	Non-respect des contraintes sémantiques pesant sur les actants d'une unité lexicale ex. <i>L'arbre s'affala sur le ruisseau.</i> (→ TOMBER)		
	Pléonasme ex. <i>Le soleil éblouissait ses yeux.</i> (→ l'éblouissait)		
COMBINATOIRE RESTREINTE	GRAMMATICALE	SYNTAXIQUE (REGIME)	Absence d'un complément obligatoire de l'unité lexicale ex. <i>Une nuit, Mathieu partit s'aventurer.</i> (→ s'aventurer dans la ville)
			Emploi transitif indirect d'un verbe transitif direct ex. <i>La peur avait gagné sur Amélie.</i> (→ avait gagné Amélie)
			Utilisation d'une préposition/conjonction erronée ex. <i>Le chef s'empressa à aller chercher la formule.</i> (→ d'aller)
			Préposition régie absente ex. <i>Le jour vint succéder la nuit</i> (→ succéder à la nuit)
			Emploi transitif d'un verbe intransitif ex. <i>Ils se sont enfuis du loup-garou.</i> (→ sauvés du loup-garou)
		MORPHOLOGIQUE	Construction erronée dans un complément ex. <i>C'est une technique pour attirer les gens à acheter.</i> (→ INCITER)
			Non-respect de la classe de mots de l'unité lexicale ex. <i>Elle lui planta l'arme direct dans le cœur.</i> (→ directement)
			Non-respect du genre du nom ex. <i>Une gigantesque éclair déchira la ciel.</i>
			Non-respect du caractère invariable de l'unité lexicale ex. <i>des scènes d'actions</i> (→ d'action), <i>prendre une vacance</i>
			Utilisation d'un verbe auxiliaire erroné ex. <i>La nuit avait déjà tombée.</i>
	LEXICALE	Choix d'un collocatif inapproprié (collocation inexistante) ex. <i>Ces publicités portent une mauvaise influence sur les jeunes.</i> (→ ont)	
		Choix d'un collocatif valable, mais mauvaise fonction lexicale ex. <i>À bout de souffle, il vit une bicyclette dans un coin et fit du vélo pour se sauver du meurtrier.</i> (→ l'utilisa, continua à vélo)	
		Non-respect des contraintes d'emploi du collocatif ex. <i>Soudain, un monstre fit apparition.</i> (→ fit son apparition)	
	PRAGMATIQUE	Emploi d'une unité lexicale appartenant à un registre de langue inapproprié ex. <i>Il a lâché son emploi.</i> (→ QUITTER)	
		Emploi d'une unité lexicale véhiculant un contenu connotationnel incompatible avec le contexte ex. <i>Cette aide à domicile devrait permettre à plusieurs vieux de demeurer plus longtemps à la maison.</i> (→ aînés, personnes âgées)	
		Emploi d'une unité lexicale appartenant à une variété dialectale incompatible avec le contexte ex. <i>Utilisation de NONANTE et SEPTANTE en contexte québécois.</i>	

Tableau 1. Typologie descriptive des erreurs lexicales (Anctil, 2011)

Figure 1. Répartition des 1144 erreurs lexicales du corpus

Les erreurs sémantiques occupent le premier rang (342 cas), suivies par les problèmes formels (241 cas), puis par les erreurs lexicopragmatiques (195 cas) ; les problèmes lexicosyntaxiques, lexicomorphologiques et de combinatoire lexicale sont présents dans des proportions presque identiques (111 à 128 cas). Il est important de souligner que près de la moitié des erreurs (49 %) sont liées à la combinatoire restreinte des unités lexicales, ce qui réaffirme la pertinence de considérer ces propriétés dans une définition du concept d'*erreur lexicale*. Observons maintenant la répartition des différents types d'erreurs à l'intérieur de chaque grande classe à l'aide d'exemples d'erreurs tirés des rédactions ⁽¹⁰⁾.

3.1. Les erreurs sémantiques

Une partie importante des erreurs sémantiques observées (43,6 %) mettent en jeu une unité lexicale erronée dont le sens se rapproche celui de l'unité lexicale proposée comme correction et fidèle au sens visé par l'élève ⁽¹¹⁾. Dans certains cas, il s'agit de termes (quasi)synonymes (12,3 %), comme dans l'exemple (1) ⁽¹²⁾.

- (1) [L'héroïne du récit part à la chasse d'un dragon pour délivrer son village.] *Après une longue **promenade** [marche] d'environ une heure, elle dut passer par la Vallée verte.* [112]

(10) Toutes les sous-classes d'erreurs de la typologie ne pourront être présentées, mais les exemples fournis illustreront les types les plus fréquents.

(11) Il serait en fait plus prudent de dire « *présumée fidèle au sens visé par l'élève* ». Le problème de restitution du sens visé par l'élève lors de la proposition d'une correction est particulièrement sensible dans le cas d'erreurs sémantiques (Corder, 1981; Marquilló Larruy, 2003). Soulignons tout de même que la très grande majorité des contextes nous permettaient de proposer une correction avec assurance.

(12) Les phrases tirées des rédactions ont été recopiées telles quelles, c'est-à-dire que nous n'y avons corrigé aucune erreur, ceci dans le but de fournir des extraits représentatifs du niveau de maîtrise linguistique des élèves. L'unité lexicale erronée apparaît en gras et la correction proposée, entre crochets. Le nombre entre crochets à la suite de l'exemple correspond quant à lui au numéro de la rédaction analysée. Comme toutes les phrases présentées contiennent des erreurs, nous avons omis l'astérisque marquant l'agrammaticalité.

L'élève qui a produit la phrase (1) ne semble pas conscient du caractère récréatif de la promenade, caractère plutôt incompatible avec la chasse au dragon relatée dans son récit ! Dans la plupart des cas mettant en jeu des termes synonymes, l'élève utilise un terme erroné sémantiquement plus riche que l'unité lexicale visée, ce qui témoigne peut-être d'un effort d'utilisation d'un vocabulaire sophistiqué, mais pas totalement maîtrisé. D'autres exemples de ce type de confusion seraient PARADE au sens de 'défilé' [147], FILAMENT au sens de 'fil' [27], FRACTURER au sens de 'briser' [80] ou ISSUE pour 'ouverture' [107].

(2) *Il passait tous ses **passes-temps** [son temps libre] sur celui-ci (= son projet). [31]*

En (2), le lien sémantique entre PASSE-TEMPS et TEMPS LIBRE est évident, puisque le sens de la seconde unité lexicale fait partie de celui de la première, un passe-temps étant une 'activité agréable que l'on pratique dans son temps libre'. Ce type de confusion entre des termes sémantiquement proches sans être synonymes était très fréquent dans notre corpus (31,3 % des erreurs sémantiques).

(3) *Tous les objets de la cuisine se déplaçaient, ce qui causait tout un **charabia** [désordre]. [18]*

(4) ***Dernièrement** [Finalement, En dernier lieu], un slogan apparaît également au bas de la page. [159]*

Les impropriétés, comme en (3) et (4), représentent plus du tiers (35,4 %) des erreurs sémantiques. Il s'agit pour la plupart de cas où la forme utilisée ne véhicule pas du tout le sens visé. Dans de tels cas, le lien sémantique entre l'unité lexicale utilisée et l'item visé est très ténu – par exemple DÉSORDRE et CHARABIA partagent une vague idée de 'confusion' – ou inexistant – par exemple entre DERNIÈREMENT et FINALEMENT. Ce dernier cas nous permet d'illustrer la distinction entre description et explication de l'erreur. D'un point de vue strictement descriptif, DERNIÈREMENT signifie 'ces derniers temps' et ne partage aucun lien de sens avec FINALEMENT : il s'agit donc d'une erreur sémantique d'utiliser DERNIÈREMENT pour exprimer 'finalement'. Bien entendu, nous pouvons supposer que l'élève a construit DERNIÈREMENT par une dérivation abusive, en s'appuyant sur le modèle de PREMIÈREMENT, DEUXIÈMEMENT, TROISIÈMEMENT ; cette hypothèse relève par contre de l'explication de la source de l'erreur et intervient à un autre moment de notre analyse.

Les anglicismes constituent une autre sous-classe d'impropriétés et sont surtout présents dans notre corpus sous la forme d'anglicismes sémantiques comme BATTERIE au sens de 'pile' [48], [l'e] FUTUR au sens d' 'avenir' [297], DÉFINITIVEMENT au sens d' 'assurément' [275] et ACADÉMIQUE au sens de 'scolaire' [234]. Nous retrouvons aussi certains calques, comme PRENDRE POUR ACQUIS [257] et PÂTE À DENTS [136], ainsi que des anglicismes formels tels que COACHER_v [49] et FAN au sens d' 'admiratrice' [282]. Ces anglicismes sont tout à fait intégrés à la langue, ce qui nous laisse croire que les élèves les ayant produits ne sont pas passés par l'anglais, mais ont simplement utilisé un item lexical de leur vocabulaire sans être conscients de la position de la norme prescriptive

relativement à ces usages ; c'est pour cette raison que nous incluons *anglicisme* dans notre typologie descriptive des erreurs.

- (5) *L'arbre s'affala [tomba] sur le ruisseau. [8]*
- (6) *Toutpetit **empoigna** [saisit] la poudre de Papif et en souffla dans la pièce puis les insectes devinrent de joli papillons. [74]*

Un peu plus d'une erreur sémantique sur dix peut être expliquée par le non-respect des contraintes pesant sur les actants sémantiques de l'unité lexicale utilisée, comme en (5) et (6). C'est ici la cooccurrence du verbe avec son sujet ou son complément qui permet de constater que l'élève ne maîtrise pas tout à fait les contraintes sémantiques qui pèsent sur les actants de ces verbes. Dans le cas d'**AFFALER**, le type d'actant sémantique pouvant remplir le rôle de sujet est clairement décrit dans *Antidote HD*, où on définit ce verbe comme 'se laisser tomber lourdement, s'affaïsser, **en parlant d'une personne**'. Pour l'exemple (6), par contre, il n'est pas aussi aisé d'expliquer l'erreur, puisque les contraintes pesant sur les actants sémantiques du verbe en jeu ne sont pas énoncées explicitement dans les ouvrages. *Le Robert* définit ainsi **EMPOIGNER** : 'prendre en serrant dans la main'. Comment expliquer qu'on ne puisse dire **empoigner de la poudre*, comme il aurait été étrange de dire **empoigner une feuille* ou *une ficelle* ? Sans doute existe-t-il certaines contraintes sur les unités lexicales pouvant jouer le rôle de complément du verbe **EMPOIGNER** (rigidité ? grosseur ?), mais elles demeurent difficiles à cerner, même si la plupart des locuteurs s'entendent pour dire qu'**empoigner de la poudre* est carrément bizarre. Les erreurs de typage des actants sont donc souvent difficiles à diagnostiquer, puisque le recours aux ouvrages de référence n'aide pas toujours ; on peut donc imaginer qu'un enseignant devant expliquer ce genre d'erreur puisse être un peu dépourvu.

3.2. Les erreurs formelles

Les erreurs liées au signifiant des unités lexicales représentent 21 % des erreurs observées ; rappelons que nous avons exclu de notre corpus les 1426 erreurs d'orthographe lexicale pour ne conserver que celles mettant en jeu une déformation du signifiant affectant la prononciation : les barbarismes. Ces barbarismes, qui représentent 86,7 % des erreurs formelles, découlent la plupart du temps d'une confusion entre deux phonèmes (**suptile* [84] pour *subtil*, **tournade* [170] pour *tornade*, **tribucha* [197] pour *trébucha*, **mignature* [89] pour *miniature*) ou de l'omission d'un phonème (**acquesça* [187], **ostination* [209], **abrute* [31]).

Il est à noter que plus du tiers des barbarismes de notre corpus concernent des locutions. Voici deux cas dans lesquels l'un des éléments de la locution a été remplacé par un autre, ce qui va à l'encontre du caractère figé de ce type d'unité lexicale, même si certaines locutions sont plus malléables que d'autres.

- (7) *... l'un de vos chevaliers l'affrontera **au périple de sa vie** [au péril de sa vie]. [60]*
- (8) ***Peu d'instant**s après [Peu de temps après, quelques instants plus tard], Eric sortit sa cantine de son sac pour ensuite pouvoir dîner. [177]*

Il existe d'autres cas, comme les suivants, où le figement de l'expression est transgressé de manière plus subtile, ce qui produit malgré tout des phrases douteuses.

- (9) *Ils ont **fait une pierre deux coups** [fait d'une pierre deux coups], car ça sensibilise les conducteurs présents sur les routes, mais aussi les adolescents. [139]*
- (10) *Je les (= les canards) nettoyé tout doucement avec **un coeur gros** [le coeur gros] tellement ils faisaient pitié. [230]*

Le second type d'erreurs formelles de notre corpus est constitué de formes complètement inexistantes. Dans la plupart des cas rencontrés, il est aisé de comprendre le sens de l'unité lexicale erronée, puisque la forme produite par l'élève partage une grande parenté formelle avec une unité lexicale existante⁽¹³⁾. Voici des exemples de ces formes fictives sémantiquement transparentes : **apeurant* [57] au sens d' 'inquiétant', **sombré* [78] au sens de 'plongé dans le noir', **exagèrement* [134] au sens d' 'exagération', **incomprénable* [171] au sens d' 'incompréhensible', **infinissable* [280] au sens d' 'interminable', **ennuyée* [282] au sens de 'qui s'ennuie'. Certains cas, aussi facilement compréhensibles, étaient tout simplement des formes empruntées directement à l'anglais : **dunk* [86] au sens de 'smash' et **drum* [290] au sens de 'tambour'.

3.3. Les erreurs pragmatiques

La quasi totalité des erreurs pragmatiques (98,5 %) concernaient l'emploi d'unités lexicales appartenant à un registre de langue familier dans un contexte demandant l'utilisation d'une langue standard. Ces erreurs affectent toutes les classes de mots, comme en témoignent les exemples suivants :

Noms : JOB [39, 43, 47, 230, 236, 245, 263, 266, 275, 279, 288], PUB⁽¹⁴⁾ [53, 81, 126] au sens de 'publicité', GARS [29, 208, 271], CHUM au sens de 'petit ami' [256], BLONDE au sens de 'petite amie' [250], PLACE au sens d' 'endroit' [157, 158, 213, 245, 258, 264, 270].

Verbes : SE CHICANER [49], LÂCHER [un emploi] [43, 44], DÉCAMPER [57] au sens de 's'enfuir', PARTIR [à] au sens de 'se mettre à' [169, 173].

Adverbes : TOUT D'UN COUP [48] au sens de 'tout à coup', POUR DE VRAI [132] au sens de 'réellement', PAREIL [48, 49] au sens de 'quand même' (ex. *Il n'était pas trop convaincu, mais James essaya **pareil**.*)

Adjectifs : BÊTE [260] au sens d' 'antipathique', TANNÉ [291] au sens de 'qui en a assez', BARAQUÉ [261], BOUSILLÉ [292].

(13) On remarque que ces formes erronées semblent avoir été produites par l'application de règles de dérivation par ailleurs productives dans la langue. Encore une fois, cette hypothèse relève plutôt de l'explication des erreurs et c'est pourquoi, d'un point de vue descriptif, nous nous limitons à dire qu'il s'agit de barbarismes.

(14) PUB et JOB constituent les unités lexicales familières les plus fréquentes dans notre corpus. Cette surreprésentation est assurément liée aux thèmes de préparation proposés aux élèves et au fait que ces unités lexicales familières apparaissaient sans doute dans les textes de préparation consultés par les élèves, ce qui légitime en quelque sorte leur utilisation. D'ailleurs, nous avons remarqué que les enseignants ont peu relevé ou pénalisé ces « erreurs ».

Prépositions : À pour exprimer la possession [59, 61, 166] (ex. *la maman à Christopher*), CÔTÉ [300] au sens de ‘en matière de’ (ex. *Nous nous entendons bien côté goûts musicaux.*), APRÈS [251, 278] au sens de ‘à’, ‘sur’ ou ‘contre’ [21, 118] (ex. ... *le colla après un poteau.*).

Déterminants : PLEIN DE [57, 61, 73, 111, 217, 245, 252, 254, 262, 266, 277] au sens de ‘beaucoup de, plusieurs’, GRAND [267] comme déterminant en contexte négatif (ex. *Je ne mis pas grand temps à regretter d’avoir voulu cet emploi.*).

Conjonctions : À CAUSE QUE [48, 134, 203, 266] au sens de ‘parce que’.

Si beaucoup des termes familiers relevés dans le corpus appartiennent au français québécois, on note tout de même la présence de plusieurs termes propres au français de France, comme BOULOT [12, 229, 288], BALÈZE [224], BOUSILLÉ [292] ou EN AVOIR MARRE [227]. Nous pouvons supposer que les élèves qui utilisent ces termes à « consonance européenne » ont l’impression d’utiliser un vocabulaire sophistiqué, alors qu’ils emploient pourtant un registre de langue familier.

3.4. Les erreurs liées à la combinatoire morphosyntaxique

Un peu plus d’une erreur sur cinq de notre corpus (21 %) concerne une propriété morphologique ou syntaxique d’une unité lexicale. Parmi les propriétés morphologiques en jeu, les plus fréquentes sont l’invariabilité et le genre du nom.

(11) ... *des gens avec plus de 6 mois d’expériences [d’expérience]* [46]

(12) ... *on peut ressentir les scènes d’actions [action] vécu du superhéro.* [54]

Les problèmes d’invariabilité touchent principalement des noms qui ne sont invariables que dans certaines de leurs acceptions, comme dans les exemples (11) et (12). Les acceptions représentées dans ces phrases renvoient en effet à des noms non comptables, qui ne peuvent par conséquent pas être employés au pluriel. Le diagnostic de telles erreurs est complexifié par le fait que les dictionnaires sont généralement peu explicites sur ce genre de propriété.

(13) *Elle aimait mieux faire moins d’argents [argent] et aimer son travail* [74]

(14) *Si l’un partait en vacance [vacances] en Floride, l’autre travaillait.* [273]

Les cas présents en (13) et (14) sont différents, puisqu’ils mettent en jeu des vocables comme VACANCES, explicitement présentés comme pluriel dans toutes ses acceptions liées à ‘période de repos’ ou ARGENT, pour lequel aucune mention relative au nombre n’est faite, mais qui n’apparaît qu’au singulier dans tous les exemples des dictionnaires consultés ⁽¹⁵⁾.

(15) À noter tout de même que *Le Multidictionnaire* mentionne que la forme plurielle *argents* au sens de ‘fonds, crédits, ressources financières’ constitue un archaïsme, mais il ne s’agit pas de l’acception en jeu dans notre exemple.

Les erreurs liées au genre nominal, comme nous pouvions nous y attendre, affectent presque exclusivement des noms commençant par une voyelle (ex. *[une] *endroit* [98], *[une] *autobus* [256], *[une] *éclair* [79, 268], *[une] *immeuble* [126], *[un] *allure* [23]) ou un *h* muet (ex. *[un] *heure* [66, 91], *[une] *habît* [120], *[une] *historique* [211]) ; par contre, nous remarquons qu'il n'est pas nécessaire que la liaison s'entende pour qu'un nom masculin soit utilisé au féminin, puisque nous avons observé des cas où un adjectif faisait écran et « bloquait » la liaison, sans pour autant empêcher l'erreur (ex. **une gigantesque éclair* [285], **un aussi grand étendue d'eau* [185]). Les deux erreurs les plus fréquentes – encore une fois à cause des thèmes de rédaction – sont ÉTÉ, utilisé au féminin par 14 élèves, et ARGENT, que sept élèves employaient systématiquement comme un nom féminin.

(15) *Il avait compris combien l'argent pouvait être **précieuse** [précieux]. [266]*

(16) *Nous pesions sur le bouton de l'ascenseur, mais **elle** [il] ne fonctionnait plus. [289]*

La plupart du temps, ces erreurs se révèlent par l'emploi d'un déterminant qui trahit la méconnaissance de l'élève relativement au genre du nom, mais dans certains cas, c'est plutôt l'accord fautif de l'adjectif dans le groupe du nom qui nous permet de constater l'erreur (ex. **l'argent amassée* [146], **l'historique professionnelle* [211]) ; plus rarement, il faut regarder dans le groupe prédicat pour constater le problème (phrase (15)), voire s'appuyer sur les pronoms de reprises (phrase (16)). On peut donc comprendre pourquoi les enseignants ont tendance à classer ces erreurs parmi les problèmes grammaticaux, puisqu'elles se manifestent dans les accords.

La structure syntaxique régie par une unité lexicale est en jeu dans 10 % des erreurs du corpus, l'utilisation d'une préposition erronée représentant le cas le plus fréquent (38,3 % des erreurs lexicosyntaxiques), principalement auprès de verbes.

(17) *Le chef s'empresse à [d'] aller chercher cette formule dans une gigantesque bibliothèque. [18]*

(18) *En plus, ils réussissent souvent **de** [à] nous convaincre. [55]*

Comme dans les exemples (17) et (18), beaucoup des problèmes observés mettent en jeu les prépositions À et DE, prépositions par excellence pour introduire les compléments indirects des verbes, mais aussi difficilement prévisibles parce que vides d'un point de vue sémantique. Un certain nombre d'erreurs de prépositions régies concernent des adjectifs et des noms :

(19) *Anxieux à [de] partir, une force le poussait quand même. [241]*

(20) *L'influence **envers** les enfants à **propas** de la publicité de gomme et de friandises [L'influence de la publicité de gomme et de friandises sur les enfants] est déjà assez forte. [134]*

Évidemment, les verbes constituent la classe de mots la plus affectée par les erreurs de régime. Nous avons observé à peu près tous les cas de figure possibles : utilisation d'un groupe prépositionnel pour compléter un verbe transitif direct (phrase (21)), d'un groupe nominal pour compléter un verbe

transitif indirect (phrase (22)), emploi transitif de verbes intransitifs (phrase (23)) et omission d'un complément d'un verbe transitif n'admettant pas l'emploi absolu (phrases (24)).

(21) *Déjà, l'amertume avait gagné **sur** [---] Amélie.* [265]

(22) *... la police avait mis une loi qui interdirait **les** [aux] paysans de s'aventurer la nuit.* [117]

(23) *... les trois garss **s'enfuirent du** [se sauvèrent du, s'enfuirent loin du] loup-garou.* [29]

(24) *Cette publicité nous fait croire que les stars **s'arrachent** pour avoir un produit comme Sunsilk [s'arrachent le produit Sunsilk].* [83]

Des unités lexicales imposant le subjonctif dans la subordonnée qu'elles introduisent ont aussi donné lieu à une vingtaine d'erreurs de régime, caractérisées par l'utilisation de l'indicatif ; il s'agit principalement des conjonctions SANS QUE [18], POUR QUE [21, 76, 92], JUSQU'À CE QUE [42, 109, 168, 287], AVANT QUE [46], MALGRÉ QUE [259] et BIEN QUE [300].

Le fait d'avoir à diagnostiquer des erreurs lexicosyntaxiques nous a confronté aux lacunes importantes des dictionnaires quant à la description du régime des unités lexicales. Certains verbes présentés comme intransitifs (ex. PARTIR) contrôlent pourtant des actants sémantiques que l'on peut tout à fait exprimer syntaxiquement sous la forme d'un complément. N'est-il pas aberrant de dire que PARTIR est un verbe intransitif au même titre qu'ÉTERNUER, qui lui ne contrôle véritablement qu'un seul actant (*X éternue*) ? On constate d'autres incohérences : alors que les verbes à un seul complément sont généralement qualifiés de transitifs directs ou indirects, les verbes ditransitifs sont simplement marqués comme transitifs, ce qui nous donne bien peu d'informations quant à la structure qu'ils contrôlent et nous oblige à nous en remettre aux quelques exemples fournis. Difficile aussi de trouver une information systématique sur la possibilité des verbes d'apparaître en emploi absolu ; on en fait parfois mention, d'autres fois pas. Et nous ne parlons ici que des verbes, mais le régime relatif à des unités lexicales appartenant à d'autres classes de mots est encore moins bien décrit. Ces constats nous poussent à douter de l'aide réelle que peuvent apporter les dictionnaires aux élèves en matière de régime, d'autant plus que ces derniers poussent rarement leur recherche jusqu'à lire attentivement les exemples et que les notions de transitivité et d'intransitivité demeurent obscures pour plusieurs.

3.5. Les erreurs liées à la combinatoire lexicale

Les erreurs liées au phénomène de collocation représentent environ le dixième du corpus. Une collocation est un cas particulier de cooccurrence de deux unités lexicales, telle que l'une des deux – la base de la collocation – est choisie librement par le locuteur d'après son sens et que l'autre – le collocatif – est sélectionné de façon contrainte pour exprimer auprès de la base un sens très général (Polguère, 2008). L'expression *faire un choix*, par exemple, est une collocation. La personne utilisant cette expression a d'abord voulu exprimer le

sens ‘choix’ en utilisant l’unité lexicale CHOIX et, en fonction de ce choix, elle a sélectionné le verbe FAIRE⁽¹⁶⁾.

La majorité des erreurs relatives à la combinatoire lexicale (78,7 %) sont liées à l’utilisation par les élèves de collocations qui ne correspondent pas aux cooccurrences qu’aurait spontanément produites un locuteur expert dans un même contexte d’énonciation. Parmi ces collocations boiteuses, plus de la moitié sont des collocations verbales, c’est-à-dire dans lesquelles le collocatif est un verbe et la base un nom. L’analyse des erreurs de collocations relevées à l’aide du formalisme des *fonctions lexicales* de la LEC⁽¹⁷⁾ nous a permis d’effectuer des regroupements qui mettent en lumière les patrons de collocations en jeu de façon récurrente dans les textes d’élèves.

Les collocations dans lesquelles le collocatif est un verbe ayant pour complément la base de la collocation sont les plus nombreuses dans les erreurs ; il peut s’agir de verbes supports sémantiquement vides (phrases (25) et (26)) ou de verbes véhiculant un sens très général (phrases (27) et (28))⁽¹⁸⁾.

(25) *Ils (= les publicitaires) **portent** [ont, exercent] une très mauvaise influence pour toutes les sociétés.* [82]

(26) *J’ai regardé sur internet pour voir quel emploi je pourrait **faire** [exercer, occuper ; obtenir].* [272]

(27) *... une fantastique idée lui **remît** [redonna] le sourire.* [213]

(28) *Ces créatures sanguinaires s’attaquaient aux habitants de petits villages pour **réduire** [assouvir, satisfaire, étancher] leur soif de sang.* [161]

Un certain nombre d’erreurs de combinatoire lexicale correspondent à des collocations dans lesquelles la base constitue le sujet syntaxique du collocatif verbal erroné, comme dans les phrases suivantes.

(29) *Un grand fracas **se créer** [retentit, résonna] puis un silence **se fit entendre** [régna ; s’installa, tomba].* [268]

(30) *Depuis ce jour d’autres rumeurs parlants d’attaques de deux Vampires **se dirent** [se repandirent, circulèrent, coururent] à travers le pays.* [119]

Les collocations dans lesquelles le collocatif (souvent adjectival) constitue un intensificateur de la base sont aussi fréquentes dans les erreurs des élèves.

(31) *... en moins de 30 secondes, elle avait tombée dans un **grand** [profond] sommeil.* [76]

(32) *Ça lui a permis de pouvoir aller au coeur du volcan, sans ressentir la chaleur **immense** [étouffante, accablante, intense] qu’il y avait.* [110]

(16) Pour se convaincre du caractère contraint des collocations, nous n’avons qu’à penser à l’unité lexicale DÉCISION, synonyme du nom CHOIX, mais qui contrôle un collocatif différent : on *prend une décision* de la même façon qu’on *fait un choix* (et non **prendre un choix* ou **faire une décision*).

(17) Pour une introduction aux fonctions lexicales, voir Mel’čuk *et coll.* (1995).

(18) Dans les exemples, la base de la collocation est soulignée et le collocatif erroné apparaît en gras.

De façon générale, nous constatons donc que les problèmes liés aux collocations mettent presque exclusivement en jeu des bases nominales et correspondent à des collocatifs sémantiquement très « légers » traduisant des sens plutôt vagues ('exister', 'avoir lieu', 'commencer', 'finir', 'causer', 'intensification', etc.). Comme la façon d'exprimer ces sens vagues est contrainte par le choix de la base, on peut comprendre que les élèves produisent des cooccurrences qui arrivent à traduire le sens visé, mais marquées par rapport à l'usage : **faire son plan* [178] plutôt que *mettre son plan à exécution*, **avoir des changements* [203] plutôt que *connaître, subir des changements*, **monter les ventes* [55] plutôt que *mousser, augmenter les ventes*, **mettre une loi* [117] plutôt que *passer, adopter une loi*, **tomber en éruption* [187] plutôt qu'*entrer en éruption*, **force puissante* [199] plutôt que *force incroyable, surhumaine* ⁽¹⁹⁾. Nous croyons que la hantise que peuvent développer certains élèves à force de se faire répéter par des enseignants d'éviter certains verbes (AVOIR, FAIRE, PRENDRE, etc.) n'est peut-être pas étrangère à quelques-unes des erreurs observées, du type **s'offrir de grosses dépenses* plutôt que *faire de grosses dépenses* [280] ou **créer un curriculum vitae* [260] plutôt que *faire, rédiger un curriculum vitae*.

Par ailleurs, soulignons aussi que les dictionnaires dont disposent les élèves en classe sont plus ou moins aidants en ce qui concerne les collocations. Les exemples fournis dans l'article de la base comprennent souvent un petit nombre de collocatifs parmi les plus fréquents, mais pour la majorité des erreurs observées, la consultation du dictionnaire n'aurait pas été d'une grande utilité pour trouver une collocation plus heureuse... d'autant plus que beaucoup d'élèves auraient sûrement comme réflexe de chercher plutôt une justification sémantique dans l'article du collocatif, ce qui s'avère peu utile pour plusieurs types de collocations, notamment les collocations verbales.

4. Les erreurs lexicales et la correction des enseignants

Comme nous l'avons mentionné plus haut, les trois quarts des 300 rédactions recueillies avaient déjà été corrigées par les enseignants au moment de notre collecte de données. Nous avons observé que tous les textes étaient corrigés de la même façon : quatre colonnes à la droite de chaque page permettent aux enseignants d'inscrire le nombre d'erreurs de chaque type présentes à chaque ligne (1. *accord* ou *grammaire* ; 2. *usage* ou *orthographe* ; 3. *syntaxe* ; 4. *punctuation*). C'est en convertissant ces nombres d'erreurs en cotes que les enseignants attribuent à la fin une note pour les critères de langue de la rédaction. Le vocabulaire, quant à lui, est toujours évalué qualitativement, par une appréciation globale du vocabulaire utilisé dans la rédaction, ce qui explique qu'il n'y ait pas de colonne précise pour cet aspect de l'évaluation.

Les 224 copies corrigées nous ont permis d'observer de quelle façon les enseignants traitent les erreurs lexicales de leurs élèves lors de la correction. Ce

(19) L'intensification avec l'adjectif SURHUMAIN semble tout de même plus forte que celle apparaissant dans l'erreur commise par l'élève.

Figure 2. Prise en compte des erreurs par les enseignants dans leur correction

sous-corpus de textes comportait 1954 erreurs selon notre analyse, en incluant cette fois les erreurs d'orthographe et de flexion (absentes de notre discussion jusqu'à présent). La figure 2 indique, en pourcentages, dans quelles proportions les erreurs de chaque grande classe ont été pénalisées, relevées seulement ou non relevées par les enseignants ayant corrigé les rédactions.

Ce qui frappe à première vue est le nombre important d'erreurs non relevées par les enseignants (37,6 %). Il ne faut pas s'alarmer devant cette statistique, puisque l'on se rappelle que le terme *erreur* désigne pour nous tant les écarts unanimement condamnables en vertu d'une norme claire (par exemple une erreur de régime verbal) que certaines maladroites ; il se peut donc fort bien que les enseignants aient choisi de « laisser passer » certaines erreurs moins « graves », qui ne méritaient pas d'être sanctionnées, ni même soulignées, considérant la quantité importante d'erreurs de tous types présentes dans les textes (syntaxe, accords grammaticaux, etc.). On retrouve aussi un certain nombre d'erreurs lexicales relevées mais non pénalisées (6,9 %), ce qui s'explique sans doute par le fait que le vocabulaire est généralement évalué de façon qualitative dans les grilles de correction utilisées ; ainsi, ces erreurs relevées, même si elles n'ont pas été comptabilisées au même titre que les erreurs relevant des autres catégories mentionnées plus haut, ont donc sans doute eu un impact sur la note attribuée aux élèves pour le critère « vocabulaire ». Il faut aussi souligner que beaucoup des rédactions recueillies constituaient des examens de fin d'année scolaire et n'étaient par conséquent pas remises aux élèves ; on peut donc supposer que les enseignants ont limité leurs annotations sur ces copies, d'où le faible pourcentage d'erreurs relevées mais non pénalisées.

On constate que les classes d'erreurs lexicales sur lesquelles semblent se concentrer les enseignants sont les erreurs formelles (76,2 % d'erreurs relevées ou pénalisées) et celles liées aux propriétés morphosyntaxiques des unités lexicales (57,3 % d'erreurs relevées ou pénalisées) ; il s'agit des erreurs qui correspondent aux grandes divisions de leurs grilles de correction et qui relèvent

de dimensions linguistiques soumises à une norme généralement claire. Comme nous nous y attendions, les erreurs liées aux propriétés morphologiques des unités lexicales (principalement *genre nominal* et *invariabilité*) sont comptabilisées par les enseignants parmi les erreurs de grammaire neuf fois sur dix. Les erreurs de régime sont quant à elles considérées comme des problèmes de syntaxe dans 95 % des cas. On peut aisément comprendre que les enseignants classent ainsi ces types d'erreurs, puisqu'elles se manifestent dans les accords et la syntaxe, mais il y a lieu de s'interroger sur la pertinence d'un tel classement pour l'élève, qui doit alors départager par lui-même ce qui relève de propriétés lexicales et ce qui appartient plutôt au système grammatical de la langue.

Des 270 erreurs sémantiques présentes dans la partie du corpus corrigée par les enseignants, seulement 91 (33,7 %) ont été relevées, dont moins de la moitié ont été pénalisées (41). Étonnamment, les trois quarts de ces erreurs ont été classées parmi les problèmes syntaxiques, peut-être parce les problèmes sémantiques se manifestent souvent à travers la cooccurrence de deux unités lexicales (ex. **les vacances d'été firent enfin leur apparition* [285], **le soleil [...] éblouissait ses yeux* [224]). Cette hypothèse vaut aussi pour les erreurs de collocations pénalisées dans ce sous-corpus, puisqu'elles ont aussi été classées en syntaxe dans presque tous les cas (ex. **l'excursion venait à sa fin* [259], **Marie-Roxanne se retira de l'eau* [280]). Soulignons finalement que les quelques erreurs de registre pénalisées ont aussi été considérées comme des problèmes syntaxiques (ex. ****Tout d'un coup**, il commence à pleuvoir et à venté* [48]).

Le taux d'erreurs relevées ou pénalisées pour les erreurs sémantiques (33,7 %) ou liées à la combinatoire lexicale (38,2 %) ou pragmatique (32 %) sont sensiblement les mêmes et laissent une plus grande place aux erreurs relevées mais non pénalisées, sans doute parce que ces types d'erreurs font partie de l'évaluation qualitative du vocabulaire faite par les enseignants. On peut tout de même regretter l'importante proportion de problèmes sémantiques non relevés, qui dépasse les 65 % ; en ne soulignant pas les à-peu-près lexicaux de leurs élèves, les enseignants contribuent en quelque sorte à l'*approximatisme* lexical qu'ils leur reprochent souvent. Le même commentaire pourrait s'appliquer aux problèmes de collocation, à la nuance près que l'on comprend mieux que les enseignants ne relèvent pas des problèmes concernant un phénomène linguistique dont ils ne sont peut-être pas eux-mêmes conscients (la collocation) et qu'ils auraient de la difficulté à justifier auprès de leurs élèves (Zimmerman, 1993 ; Polguère, 2004 ; Anctil, 2011).

Pour ce qui est des annotations utilisées par les enseignants relativement au vocabulaire, on constate qu'elles sont vagues et peu systématiques. Elles se concentrent principalement sur la répétition, dimension non considérée dans notre analyse des erreurs. Les « mots » mal utilisés – c'est-à-dire les erreurs sémantiques, mais aussi certains problèmes de collocation et de registre –, lorsqu'ils donnent lieu à des annotations, sont quant à eux soulignés, barrés ou entourés d'un *petit nuage* et généralement accompagnés d'une annotation (« V », « voc », « vocabulaire » ou « ? ») dans la marge ou au-dessus du terme problématique. Très rares sont les cas où l'annotation est plus précise (par exemple « *familier* » ou « *anglicisme* ») ou ceux pour lesquels on propose une correction. Globalement, nous pouvons affirmer que les annotations fournissent

peu d'indices aux élèves sur la nature du problème lexical et les façons d'y remédier et que par conséquent, elles contribuent peu à l'amélioration de leur vocabulaire.

5. Des pistes didactiques

Notre description détaillée des erreurs lexicales des élèves en production écrite nous conduit à proposer certaines pistes didactiques pour l'enseignement du lexique. Soulignons d'entrée de jeu la pertinence de se servir des erreurs des élèves en rédaction comme point de départ pour des activités lexicales ; elles permettent de confronter les élèves à certaines lacunes dans leur vocabulaire et donnent du sens aux enseignements lexicaux en les liant à des contextes réels de production langagière.

La diversité des types d'erreurs observés et le nombre important d'unités lexicales en jeu militent en faveur d'approches didactiques visant la compréhension de phénomènes lexicaux transversaux (collocation, régime verbal, registres de langue, polysémie, synonymie, etc.), l'identification de « zones sensibles » dans l'utilisation du vocabulaire et l'utilisation d'ouvrages de référence. Cette conception plus globale de l'enseignement lexical s'inscrit d'ailleurs tout à fait dans l'esprit d'une véritable didactique du lexique, dont l'objectif principal est une meilleure compréhension du système lexical plutôt qu'un accroissement quantitatif du vocabulaire des élèves (Simard, 1994 ; Polguère, 2004 ; Vancomelbeke, 2004).

Le nombre élevé d'erreurs sémantiques dans notre corpus devrait nous convaincre de l'importance de travailler davantage le sens en classe, dimension à laquelle les enseignants avouent ne consacrer que peu de temps (Dreyfus, 2004 ; Anctil, 2011). La prépondérance des erreurs sémantiques mettant en jeu une unité lexicale erronée véhiculant un sens proche du sens visé nous porte à croire qu'il serait pertinent de travailler en classe sur des ensembles de termes synonymes. Un tel travail permettrait aux élèves d'affiner leur conception de la synonymie, notamment en prenant conscience que la synonymie exacte est à peu près inexistante et qu'il est important de bien saisir les nuances de sens entre des termes présentés comme des synonymes dans les ouvrages de référence. Des activités sur la synonymie permettraient non seulement aux élèves d'élargir leur vocabulaire et de développer leur compétence métasémantique, mais leur démontrerait aussi la pertinence d'utiliser conjointement le dictionnaire des synonymes et le dictionnaire « régulier ».

Le travail lexical sur des groupes de termes synonymes nous paraît d'ailleurs fructueux à plusieurs égards, puisqu'il permet de contraster d'autres aspects de la combinatoire restreinte. Par exemple, travailler sur les différences sémantiques entre RÉAGIR [à] et ses synonymes S'OPPOSER [à], SE DRESSER [contre] et S'INDIGNER [de] permet de faire ressortir les différences de régime de ces verbes. Demander à un élève de consulter le dictionnaire pour trouver un synonyme de JOB peut se révéler une bonne occasion d'aborder la question des registres de langue en constatant que BOULOT et GAGNE-PAIN, tout comme JOB, appartiennent plutôt au registre familier, alors que leurs synonymes

EMPLOI, TRAVAIL et PROFESSION font partie du registre standard. Attirer l'attention des élèves sur les collocatifs des synonymes étudiés peut permettre de prendre conscience du caractère arbitraire des collocations : pourquoi dit-on *pratiquer un métier*, mais *occuper un emploi* (et non **occuper un métier* ou **pratiquer un emploi*) alors que MÉTIER et EMPLOI sont synonymes ? Évidemment, il ne s'agit pas de se limiter au travail sur des synonymes. Travailler sur des groupes d'unités lexicales liées sémantiquement à travers les notions de *famille de mots* et de *champ sémantique* pour aborder différents aspects de la maîtrise lexicale (régime, collocation, registre, etc.) renforce la conception du lexique comme un réseau structuré et organisé par différents types de liens.

Un travail beaucoup plus systématique sur le dictionnaire devrait aussi être effectué en classe. Une proportion importante des erreurs de notre corpus auraient pu être évitées si les élèves l'avaient consulté, à condition qu'ils en aient une connaissance plus profonde puisque, aux dires des enseignants rencontrés, la recherche dans le dictionnaire demeure pour eux une tâche fastidieuse presque seulement orientée vers l'orthographe lexicale (Anctil, 2011). Sensibiliser les élèves à la richesse des informations présentées dans les dictionnaires, leur enseigner la structure des articles, leur expliquer les différentes marques d'usage utilisées, leur montrer les informations apparaissant de façon implicite dans les exemples (régime, collocations) et comparer différents ouvrages devraient constituer des contenus essentiels d'un enseignement lexical.

En ce qui concerne la correction des productions écrites par les enseignants, une plus grande systématisme dans le traitement des erreurs lexicales profiterait sans doute aux élèves. Toutes les erreurs induisant un problème de compréhension (en majorité des erreurs sémantiques) devraient être relevées et une proposition de correction devrait être fournie chaque fois qu'il est difficile pour l'élève d'y accéder par lui-même (ex. barbarisme, collocation douteuse) ; en indiquant à l'élève l'unité lexicale correcte à utiliser dans un contexte précis, l'enseignant l'expose à une occurrence de cette unité lexicale, contribuant ainsi, ne serait-ce que minimalement, à l'accroissement de son vocabulaire... si tant est bien sûr que l'élève lise la correction proposée ! L'utilisation de codes de correction plus précis concernant l'emploi du vocabulaire (ex. *anglicisme*, *terme familier*, *combinaison de mots*, etc.) pourrait aussi orienter davantage les élèves sur la façon de corriger leurs erreurs ; il serait intéressant que l'annotation des erreurs liées aux propriétés morphologiques ou syntaxiques fasse ressortir le caractère lexical de ces problèmes en apparence grammaticaux.

Conclusion

La classification d'erreurs lexicales, comme d'autres l'ont constaté avant nous (Patris et Vansnick, 1992 ; Luste-Chaa, 2009), demeure une entreprise difficile. Néanmoins, la typologie que nous avons développée s'est révélée très opérationnelle, notamment parce qu'elle constitue un outil purement descriptif. Elle nous a permis de constater la prépondérance des erreurs sémantiques et des problèmes liés à la combinatoire restreinte, notamment aux propriétés

morphologiques et syntaxiques des unités lexicales. Il est important de rappeler que les erreurs lexicales ne peuvent constituer à elles seules un baromètre de la compétence lexicale des élèves ; il faudrait compléter ce portrait en considérant aussi la richesse et la variété du vocabulaire utilisé ⁽²⁰⁾.

Bien entendu, les pistes didactiques esquissées à grands traits à partir de la description de notre corpus et des corrections fournies par les enseignants demanderaient à être étoffées en tenant compte des sources des erreurs observées, dimension que nous n'avons pu traiter ici. Tout un travail reste aussi à faire pour proposer aux enseignants des activités lexicales concrètes à mener en classe avec leurs élèves, dont certaines, dans un souci d'intégration des activités de vocabulaire à celles d'écriture, prendraient comme point de départ les erreurs de ces derniers. Il ne faut pas oublier non plus que l'animation d'activités de vocabulaire et l'analyse de problèmes lexicaux demandent une compréhension profonde du fonctionnement du lexique ; la formation des maîtres est donc aussi à reconsidérer (Polguère, 2004 ; Tremblay, 2009) si nous voulons les outiller convenablement afin qu'ils puissent se servir des erreurs de vocabulaire des élèves comme tremplins vers des enseignements lexicaux.

(20) Certaines méthodes de calcul de la richesse lexicale ont été proposées, mais principalement dans le cadre de travaux en langue seconde (voir par exemple Arnaud, 1984; Engber, 1995; Read, 2000).

Références

- AGUSTÍN LLACH, M. P. (2004). « Lexical inconsistencies and deficiencies in English second language learners : proposal of classification ». *Interlinguistica*, vol. 15, n° 1, 83-92.
- (2005). « A critical review of the terminology and taxonomies used in the literature on lexical errors ». *Miscelánea : a journal of english and american studies*, vol. 31, 11-24.
- ANCTIL, D. (2005). *Maîtrise du lexique chez les étudiants universitaires : typologie des problèmes lexicaux et analyse des stratégies de résolution de problèmes lexicaux*. Mémoire de maîtrise non publié, Université de Montréal, Montréal.
- (2011). *L'erreur lexicale au secondaire : analyse d'erreurs lexicales d'élèves de 3^e secondaire et description du rapport à l'erreur lexicale d'enseignants de français*. Thèse de doctorat non publiée, Université de Montréal, Montréal.
- ANTIDOTE HD (Version 3 [logiciel]). (2009). Montréal : Druide informatique.
- ARNAUD, P. J. L. (1984). « The Lexical Richness of L2 Written Productions and the Validity of Vocabulary Tests ». Dans T. Culhane, C. Klein-Braley & D.K. Stevenson (Eds.), *Practice and Problems in Language Testing*. Colchester : University of Essex, 14-28.
- (1992). « Objective lexical and grammatical characteristics of L2 written compositions and the validity of separate components tests ». Dans P. J. L. Arnaud & H. Béjoint (Eds.), *Vocabulary and applied linguistics*. London : Macmillan.
- BAILLY, D. (1998). *Les mots de la didactique des langues : Le cas de l'anglais (lexique)*. Paris : Ophrys.
- BASTUJI, J. (1978). « Les théories sur le vocabulaire, éléments pour une synthèse ». *Pratiques*, n° 20, 75-89.
- BRENT, E. (1999). « Vers l'élaboration de normes pédagogiques du français ». *Terminogramme. La norme du français au Québec. Perspectives pédagogiques*, n° 91-92, 117-130.
- CORDER, S. P. (1981). *Error Analysis and Interlanguage*. Oxford : Oxford University Press.
- DE VILLERS, M.-É. (2009). *Multidictionnaire de la langue française* (5^e éd.). Montréal : Québec Amérique.
- DEBYSER, F., HOUIS, M. et ROJAS, C. (1967). *Grille de classement typologique des fautes*. Paris : Bureau pour l'enseignement de la langue et de la civilisation françaises à l'étranger (BELC).
- DENANCE, M. (1999). « L'évaluation des fautes de vocabulaire à l'évaluation de l'épreuve uniforme ». *Correspondance (périodique en ligne)*, vol. 4, n° 4.
- DREYFUS, M. (2004). *Pratiques et représentations de l'enseignement du lexique à l'école primaire*. Communication présentée au 9^e colloque de l'AIRDF, Québec, 26 au 28 août 2004.
- ENGBER, C. A. (1995). « The Relationship of Lexical Proficiency to the Quality of ESL Compositions ». *Journal of Second Language Writing*, vol. 4, n° 2, 139-155.

- GRANGER, S. (2003). « Error-tagged Learner Corpora and CALL : A Promising Synergy ». *CALICO*, vol. 20, n° 3, 465-480.
- GRANGER, S. et MONFORT, G. (1994). « La description de la compétence lexicale en langue étrangère : perspectives méthodologiques ». *AILE*, vol. 3, 55-75.
- GUILBERT, L. (1972). « Peut-on définir un concept de norme lexicale ? ». *Langue française*, n° 16, 29-48.
- HAMEL, M.-J. et MILIĆEVIĆ, J. (2007). « Analyse d'erreurs lexicales d'apprenants du FLS : démarche empirique pour l'élaboration d'un dictionnaire d'apprentissage. ». *Revue de l'Association canadienne de linguistique appliquée (ACLA)*, vol. 10, n° 1.
- JAMES, C. (1998). *Errors in language learning and use : exploring error analysis*. London : Longman, Coll. Applied Linguistics and Language Studies.
- LEGENDRE, R. (2005). *Dictionnaire actuel de l'éducation* (3^e éd.). Montréal : Guérin.
- LENNON, P. (1991). « Error : Some Problems of Definition, Identification and Distinction ». *Applied Linguistics*, vol. 12, n° 2, 180-196.
- LUSTE-CHAA, O. (2009). *Les acquisitions lexicales en français langue seconde : conceptions et applications*. Thèse de doctorat non publiée, Université Paul-Verlaine, Metz.
- MARQUILLÓ LARRUY, M. (2003). *L'interprétation de l'erreur*. Paris : CLE International, Coll. Didactique des langues étrangères.
- MEL'ČUK, I. (1997). *Vers une linguistique sens-texte : leçon inaugurale faite le vendredi 10 janvier 1997*. Paris, Collège de France : Chaire internationale.
- MEL'ČUK, I., CLAS, A. et POLGUÈRE, A. (1995). *Introduction à la lexicologie explicative et combinatoire*. Louvain-la-Neuve : Duculot, Coll. Universités francophones.
- MILIĆEVIĆ, J. (2006). « A Short Guide to the Meaning-Text Linguistic Theory ». *Journal of Koralex*, vol. 8, 187-233.
- MILIĆEVIĆ, J. et HAMEL, M.-J. (2007). « Un dictionnaire de reformulation pour apprenants du français langue seconde ». *Revue de l'Université de Moncton, Numéro hors série « Actes des colloques annuels de l'Association de linguistes des provinces atlantiques (APLA) 2005 (Moncton) »*, 145-167.
- NATION, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge : Cambridge University Press, Coll. Cambridge applied linguistics series.
- PATRIS, J. et VANSNICK, N. (1992). « Comment aborder les fautes lexicales en classe ? ». *Enjeux*, n° 26, 46-55.
- POLGUÈRE, A. (2004). « Savoir consulter un dictionnaire, c'est bien ; savoir comment le construire, c'est mieux ». *Québec français*, n° 134, 68-70.
- POLGUÈRE, A. (2008). *Lexicologie et sémantique lexicale. Notions fondamentales*. (2^e éd.). Montréal : Les Presses de l'Université de Montréal, Coll. Paramètres.
- PORQUIER, R. (1977). « L'analyse des erreurs : problèmes et perspectives ». *Études de linguistique appliquée*, vol. 25, 23-43.
- PRÉFONTAINE, C. (1998). *Écrire et enseigner à écrire*. Montréal : Les Éditions Logiques.

- READ, J. (2000). *Assessing vocabulary*. Cambridge : Cambridge University Press, Coll. Cambridge language assessment series.
- REBOUL-TOURÉ, S. (2002). « L'enseignement du lexique au collège ». *Le Français Aujourd'hui*, vol. 141, 123-128.
- REUTER, Y., COHEN-AZRIA, C., DAUNAY, B., DELCAMBRE, I. et LAHANIER-REUTER, D. (2007). *Dictionnaire des concepts fondamentaux des didactiques*. Bruxelles : de Boeck.
- REY, A. (1972). « Usages, jugements et prescriptions linguistiques ». *Langue Française*, vol. 16 (décembre), 4-28.
- RICHARDS, J. (1976). « The role of vocabulary teaching ». *TESOL Quarterly*, vol. 10, 77-89.
- ROCQUET, J.-P. (1998). *Ils manquent de vocabulaire ! : lexique et vocabulaire à l'école maternelle et élémentaire*. Marne : CDDP [Centre départemental de documentation pédagogique] Marne.
- SIMARD, C. (1994). « Pour un enseignement plus systématique du lexique ». *Québec français*, n° 92, 28-33.
- SIOUFFI, G. et STEUCKARDT, A. (2001). *La norme lexicale*. Montpellier : Université Paul-Valéry Montpellier III.
- TREMBLAY, O. (2009). *Création d'une ontologie des connaissances métalexicales pour l'élaboration d'un module de cours en didactique du lexique destiné aux futurs maîtres au primaire en français langue maternelle*. Thèse de doctorat non publiée, Université de Montréal, Montréal.
- VANCOMELBEKE, P. (2004). *Enseigner le vocabulaire*. Paris : Nathan, Coll. Les repères pédagogiques. Série Formation.
- VERREAULT, C. (1999). « L'enseignement du français en contexte québécois : de quelle langue est-il question ? ». *Terminogramme. La norme du français au Québec. Perspectives pédagogiques*, n° 91-92, 21-40.
- ZIMMERMAN, C. B. (1993). *Teachers Perceptions of and Strategies with Lexical Anomalies*. Communication présentée au Annual Meeting of the Teachers of English to Speakers of Other Languages, Atlanta, April 13-17.
- ZIMMERMANN, R. (1986). « Classification and distribution of lexical errors in the written work of German learners of English ». *Papers and Studies in Contrastive Linguistics*, vol. 21, 31-40.