

# Pratiques orthographiques en français d'étudiants étrangers: le cas d'étudiants hispano-américains et afghans

Jeanne GONAC'H & Clara MORTAMET

IRED, Université de Rouen & UFR Lettres et sciences humaines, Université de Rouen

In this article, we analyze and compare French spelling errors produced by two different groups of university students who have learned French in a country where French is a foreign language (FFL): students from Central and South American countries (Salvador, Peru and Argentina) and Afghanistan. Our corpus is based on errors found in papers of graduate distance-learning students, most of whom are planning on becoming French teachers in their countries. Our aim is to identify what errors are specific to FFL university students and eventually what errors are specific to the two different groups. The first results of our study confirm tendencies observed on previous research and prove that there exist specific types of errors for FFL learners. In the end, we will also discuss the limit of error categories.

## Introduction

S'interroger sur l'élaboration de méthodes d'enseignement spécifiques pour pallier les problèmes d'orthographe des apprenants en contexte FLE implique en amont d'identifier les problèmes particuliers que l'orthographe française pose aux apprenants allophones. Or, les études qui ont inventorié les erreurs spécifiques à ces apprenants d'une part sont peu nombreuses et d'autre part montrent que la majorité des erreurs des apprenants FLE, si elles sont plus nombreuses que celles de leurs pairs natifs, ne sont pas pour autant toujours différentes de celles de ces derniers (Amokrane & Brissaud, *à paraître*). Cette ressemblance entre les erreurs des natifs et des non-natifs dans des copies universitaires a été confortée dans une étude contrastive que nous avons menée récemment à partir de copies d'étudiants éthiopiens et mahorais (Mortamet & Gonac'h, 2011), étude qui s'est ajoutée à l'expérience que nous avons de l'analyse de l'orthographe des natifs monolingues et bilingues (Mortamet, 2003; Gonac'h, 2009).

Notre contribution vise ainsi à compléter l'analyse des zones orthographiques problématiques pour des apprenants en contexte FLE. Pour ce faire, nous avons choisi d'analyser les erreurs produites dans des copies d'examen rédigées par des étudiants de 3<sup>ème</sup> année de licence inscrits en sciences du langage en télé-enseignement à l'université de Rouen. En effet, par le biais de ce centre, nous avons accès à des dizaines

de copies du monde entier. Ces adultes en formation sont répartis sur tous les continents, très majoritairement dans les pays du Sud. Beaucoup de ces étudiants n'ont pas le français pour langue première, mais l'enseignent ou souhaitent l'enseigner dans leur pays. Ce corpus offre la possibilité de comparer des productions émanant de contextes d'apprentissage très divers. Dans le cadre de cette étude, nous explorons deux contextes de français langue étrangère: l'Amérique hispanophone d'une part, et l'Afghanistan d'autre part.

Dans un premier temps, nous ferons le point des travaux, relevant tant de la sociolinguistique que de la didactique, qui nous ont inspirées pour penser la question des erreurs d'orthographe chez les non-natifs. Nous présenterons ensuite la méthodologie de recueil et d'analyse des productions. Nous analyserons enfin les principales erreurs observées dans notre corpus, qu'elles relèvent de la fonction morphographique, visuographique, phonographique, ou qu'elles fassent l'objet d'interprétations mixtes.

## **1. Les erreurs d'orthographe des non-natifs comme objet de recherche**

### *1.1 Un objet au croisement de plusieurs domaines de recherche*

La question des erreurs d'orthographe des non-natifs nécessite une expertise à la fois sur l'orthographe, sur le plurilinguisme et sur l'acquisition-apprentissage des langues.

En ce qui concerne l'analyse des erreurs d'orthographe, nous nous appuyons d'abord sur les travaux de Catach (1986, 1998), mais aussi sur ceux de Brissaud (2006), de Honvault (dir., 2006), de Lucci & Millet (1994), de Manesse & Cogis (2007). Ces études, s'appuyant toutes sur l'analyse de l'orthographe française de Catach (phonographie, morphographie, idéographie, logographie), qu'elles affinent au passage, présentent l'intérêt de décrire des pratiques orthographiques dans des activités diverses (des dictées comme dans Manesse & Cogis, des écrits ordinaires comme dans Lucci & Millet, des écrits scolaires dans Brissaud), par des locuteurs divers (des élèves dans Manesse & Cogis, Brissaud, des secrétaires et des étudiants dans Lucci & Millet). Elles donnent également à voir des pratiques issues de situations plus ou moins formelles, plus ou moins publiques (des brouillons dans Lucci & Millet, des écrits scolaires dans Brissaud, Manesse & Cogis). Mais elles nous ont surtout servi ici à établir notre grille d'analyse des erreurs, et à comparer les erreurs de notre corpus avec celles rencontrées dans ces corpus de natifs. Elles ont enfin alimenté notre réflexion plus générale sur l'orthographe du français et les pratiques orthographiques des francophones (Honvault (dir.), 2006).

En ce qui concerne les questions de plurilinguisme et de didactique des langues, nous nous appuyons essentiellement sur les travaux développés par Porquier & Py (2005), Selinker (1994), Tarone (2005), et ceux menés dans le cadre du CECR. Les réflexions nées autour de la notion de *compétence plurilingue* (Coste, Moore & Zarate, 1997: 105-106) ont ouvert la voie à une nouvelle conceptualisation des liens entre compétence et pratique des langues présentes dans le répertoire sans faire intervenir l'idée de l'influence négative d'une langue sur l'autre. Coste (2002: 120) souligne ainsi que les compétences de toutes les ressources langagières présentes dans le répertoire sont mises en relation. Il fait l'hypothèse que les *éventuelles* différences entre les compétences des non-natifs et des natifs sont liées à la "circulation interlinguistique" des répertoires. Avec Coste, nous faisons donc l'hypothèse que les langues présentes dans le répertoire des non-natifs peuvent affecter leur orthographe, et donner lieu à des erreurs "différentes" de celles des natifs – la "différence" pouvant porter sur les types d'erreurs, ou sur la proportion de chaque type d'erreurs dans l'ensemble des erreurs. Cette hypothèse implique aussi que les pratiques orthographiques des apprenants peuvent être affectées différemment lorsque les répertoires langagiers sont distincts. C'est ce que nous nous attacherons à regarder ici en comparant les pratiques d'étudiants afghans et hispano-américains en français.

## 1.2 *Quelques particularités des non-natifs*

Dans une précédente recherche (Mortamet & Gonac'h, 2011), nous avons comparé des copies rédigées dans un contexte FLE (l'Ethiopie) et dans un contexte FLS (Mayotte) du point de vue de l'orthographe. Notre hypothèse de départ était alors que les étudiants en contexte FLE produiraient des erreurs différentes de celles des étudiants en contexte FLS. Nos résultats ont plutôt infirmé cette hypothèse: il apparaît des erreurs spécifiques dans ces deux contextes, qui les différencient des natifs mais pas entre eux. En particulier, nous avons montré que les non-natifs, quel que soit leur contexte d'apprentissage (FLE et FLS), produisaient *en proportion* davantage d'erreurs d'accord en genre que ce que l'on rencontre dans des corpus de natifs.

Dans une étude portant sur des natifs, Brissaud *et alii* (2006) indiquent par exemple que les erreurs d'accord en genre figurent parmi les plus fréquentes, mais qu'elles ne sont pas plus fréquentes que celles portant sur les accords en nombre ou sur les cas d'homophonie verbale. Dans l'étude de Manesse & Cogis (2007: 111ss.), qui porte également sur un corpus de natifs, les accords en genre apparaissent même quelque peu plus réussis que les accords en nombre.

Dans notre corpus d'erreurs produits par les étudiants éthiopiens et mahorais, les erreurs liées au genre sont non seulement plus nombreuses mais aussi différentes de celles des natifs. De plus, elles ne relèvent pas uniquement de l'orthographe (*un problème majeure, la méthodologie audio-visuel*) mais aussi, bien souvent de la morphologie (*la problème, un public instruite*, etc.). A partir de ces résultats, nous avons fait l'hypothèse que l'apprentissage du genre grammatical en français, que ce soit par l'orthographe ou par la morphologie, posait davantage problème aux étudiants non-natifs qu'à leurs pairs natifs notamment parce que le marquage du genre grammatical dans leurs langues premières n'existe pas ou ne fonctionne pas de la même façon.

Nous appuyant sur cette hypothèse de circulation interlinguistique, nous avons voulu ici prolonger ce travail dans cette direction en comparant les erreurs produites par des étudiants dont la langue première ou principale est typologiquement proche du français (l'espagnol) avec celles des étudiants dont la langue première ou principale est très éloignée du français (le dari).

## 2. Méthodologie

### 2.1 *Choix du corpus*

Nous avons constitué deux groupes d'étudiants: huit étudiants afghans d'une part et huit étudiants hispano-américains (6 Salvadoriens, 1 Péruvienne et 1 Argentine) d'autre part. Parmi les étudiants afghans, cinq sont des femmes; ils ont autour de 30 ans à l'exception d'une personne qui en a 60. Parmi les étudiants hispano-américains, sept sont des femmes. Ils sont âgés entre 25 et 30 ans et une femme a plus de 60 ans.

Pour chaque étudiant, nous avons analysé cinq de leurs copies d'examens de troisième année de licence de Sciences du langage parcours FLE. Nous avons toujours choisi des copies rédigées de façon libre (dissertations, commentaires de texte, etc.).

### 2.2 *Méthodologie d'analyse*

Pour chaque étudiant, nous avons relevé l'ensemble des erreurs produites et les avons ensuite classées selon différentes catégories<sup>1</sup>. Dans un premier temps, nous analyserons ici plus particulièrement les catégories suivantes: les problèmes liés à l'accord en genre qui relèvent de l'orthographe (*un nouvelle emploi*) et de la morphologie (*son rigueur*)<sup>2</sup> (§3);

<sup>1</sup> Pour une présentation complète des catégories, voir Mortamet & Gonac'h, 2011.

<sup>2</sup> Ces deux types d'erreurs étant très représentés (tous les étudiants en font) et répétés dans les mêmes copies, il nous a semblé que ces formes devaient être analysées ensemble. Pour

les problèmes d'accord en nombre qui relèvent de l'orthographe (*ces communication*) (§4); les erreurs liées aux lettres hors système (*auteur, manière*) (§5); les problèmes qui relèvent de la phonographie (*contrasens, distiné*) (§6). Nous analyserons aussi les erreurs qui peuvent avoir des origines différentes et qui peuvent donc être interprétées de deux façons différentes (§7).

Nous n'analyserons pas les écarts portant sur les diacritiques et les auxiliaires d'écriture (accents, traits d'union, apostrophes), les césures, les "ratés graphiques", ni sur la ponctuation, ces zones d'écarts soulevant chacune des questions particulières, difficiles à articuler avec celles présentées ici<sup>3</sup>. Par ailleurs, nous écarterons aussi certaines catégories très faiblement représentées dans nos copies. Il s'agit notamment des accords sur les verbes conjugués, des anglicismes, des erreurs de morphographie lexicale (*aspet*), des erreurs sur les morphogrammes (*judicière*).

Une fois toutes ces formes relevées et classées, nous avons compté les occurrences dans chaque catégorie pour chaque étudiant (toutes copies confondues). Cela nous a permis d'établir des diagnostics de chaque étudiant, de regarder les écarts les plus fréquemment rencontrés chez chacun d'entre eux, ceux qu'ils avaient en commun, ceux qui les distinguaient. Par contre, nous n'avons pas fait la somme par groupe des erreurs de chaque type. Nous avons en effet voulu adopter une analyse qualitative, en établissant un diagnostic par étudiant, et en comparant ensuite les deux ensembles de diagnostics. La comparaison de nos deux populations, que nous présentons maintenant, ne s'appuie donc pas sur un diagnostic global de chaque groupe d'étudiants, mais sur un ensemble d'évaluations individuelles qui composent chaque groupe.

---

la même raison, nous avons analysé ensemble tous les problèmes de genre, qu'ils portent sur les participes passés, les adjectifs, les attributs ou les noms. Nous avons fait la même chose avec les erreurs concernant le nombre.

<sup>3</sup> Les diacritiques et auxiliaires d'écriture ne font pas partie, ou à des degrés divers, du plurisystème graphique: si les accents aigus et graves sur le <e> et le tréma présentent pour certains encore une dimension phonographique, elle est dans certains cas discutée, et ne vaut plus pour les autres diacritiques et auxiliaires que sont les traits d'union, les apostrophes, l'accent circonflexe. La question de la césure, absente de notre corpus du fait du niveau avancé de nos scripteurs, pose des questions diverses, entre morphographie et visuographie. Les "ratés graphiques" (répétitions de syllabes comme dans *institutution*, omissions de syllabes dans *instution*), relativement nombreux, sont pour l'essentiel dus à la contrainte physique de l'écriture manuscrite et difficilement interprétables, *a priori* du moins, comme des erreurs d'orthographe. La ponctuation enfin relève d'un autre niveau, celui de la phrase et du texte.

### 3. Les erreurs portant sur le genre

Dans l'ensemble du corpus, les erreurs liées à l'accord en genre qui relèvent de l'orthographe et de la morphologie sont un peu plus nombreuses que les erreurs liées à l'accord en nombre. Il faut tout de suite souligner que les étudiants afghans contribuent largement à ce volume et que parmi eux, certains produisent largement plus d'erreurs que d'autres (l'un d'entre eux produit plus de 30 erreurs alors qu'on en relève 140 dans l'ensemble des copies de ce groupe). Autrement dit, sur ce point, l'hétérogénéité interne au groupe est plus forte chez les étudiants afghans que chez les étudiants hispano-américains.

Il n'en reste pas moins que ces résultats rejoignent ce que nous avons déjà constaté, à savoir que les variations liées au genre sont les plus nombreuses. Elles sont bel et bien surreprésentées dans les productions des étudiants dont le français n'est pas la langue première par rapport aux productions des natifs (cf. *supra*, §1.2).

#### 3.1 Particularités des étudiants afghans

Les étudiants afghans se distinguent particulièrement par leurs erreurs portant sur le genre morphologique, catégorie à laquelle ils contribuent beaucoup plus que les étudiants hispano-américains. Nous avons repéré dans cette catégorie plusieurs cas de figure: soit les étudiants utilisent les marques du masculin alors que les marques du féminin sont attendues soit l'inverse; les marques du masculin peuvent être portées par le déterminant: *son rigueur*; par l'adjectif: *la compétence communicatif* et peuvent aussi être multiples comme dans: *le revu* pour *la revue*. Les marques du masculin portées par le déterminant sont les erreurs les plus fréquentes et les plus partagées chez les étudiants afghans (7 étudiants sur 8). Le nombre d'erreurs de ce type varie de façon importante à l'intérieur du groupe puisque l'un d'entre eux en fait dix, et que cinq autres n'en font qu'une. Les marques du masculin (ou l'absence de marque du féminin) sur l'adjectif et les marques "multiples" concernent quatre étudiants.

Parmi ces erreurs, certaines indiquent que des terminaisons nominales régulières, qui sont plutôt typiques d'un genre, ne sont pas reconnues par les étudiants. C'est le cas par exemple des terminaisons du féminin *-ion* (mission), *-ologie* (méthodologie), *-ure* (candidature).

Les marques du féminin sont aussi souvent utilisées quand celles du masculin sont attendues. Alors que les marques du masculin portaient plus souvent sur le déterminant seulement (*le façon*), les marques du féminin les plus partagées sont celles qui portent uniquement sur l'adjectif (*diverses domaines*). Ensuite, on compte les erreurs qui impliquent l'ajout

des marques du féminin à la fois sur les déterminants et les adjectifs (*une phénomène sociale*) que cinq scripteurs différents produisent. Viennent ensuite les erreurs sur le déterminant seul (*la couloir*), et enfin les marques sur les participes passés, qu'un seul étudiant produit (*l'apprentissage s'est développée*).

Nous avons distingué les marques du féminin et du masculin afin de faire émerger des tendances: soit à "masculiniser" ou à l'inverse à "féminiser". Toutefois, cette distinction ne s'est avérée pertinente que pour un étudiant afghan qui a tendance à féminiser. Pour tous les autres, on ne repère aucune tendance particulière.

Comme signalé plus haut, les confusions sur le genre au niveau morphologique sont beaucoup moins nombreuses pour les étudiants hispano-américains. Seulement la moitié des étudiants produit des erreurs de ce type et parmi eux, une étudiante – péruvienne – se détache en en produisant beaucoup plus que les autres. Les erreurs ne sont pas liées à un calque de l'espagnol à l'exception de "*La petite chaperon rouge*" (*Caperucita Roja, féminin en espagnol*). Ainsi, par exemple, un étudiant écrit "*la texte*" alors qu'en espagnol "texte" est aussi masculin (*el texto*).

La comparaison des deux contextes d'apprentissage révèle que la question du genre ne pose pas les mêmes problèmes aux étudiants afghans et aux étudiants hispano-américains. Alors que les étudiants afghans utilisent tous, souvent et à différents niveaux, des marques du féminin ou du masculin quand l'inverse est attendu, seules quelques erreurs apparaissent dans le corpus de copies des Hispano-américains. Le marquage du genre grammatical dans les langues premières peut être un des motifs d'explication de cette différence entre les deux groupes. Ainsi, en dari, langue que tous nos étudiants pratiquent, le genre grammatical n'est pas marqué, alors qu'en espagnol il l'est. Les quelques erreurs produites par les Hispano-américains peuvent être liées au fait que les genres grammaticaux du français et de l'espagnol ne correspondent pas toujours. Pour les Afghans, on peut penser que les erreurs sont liées au surcoût cognitif que représente l'apprentissage du genre grammatical. Ainsi, on peut même penser que c'est la logique de la "loterie" qui domine dans leurs copies lorsqu'ils ne connaissent pas le genre. Si cette interprétation est juste, il y a fort à parier que chez les Afghans étudiés, les méconnaissances du genre des noms utilisés sont plus nombreuses que celles qui apparaissent ici: si les cas de loterie malchanceuse sont remarquables, ceux de loterie chanceuse se confondent avec les formes issues d'une bonne maîtrise du genre.

### 3.2 *Erreurs communes aux deux groupes*

Pour les étudiants afghans comme pour les étudiants hispano-américains, les problèmes d'accord orthographique liés au genre portent essentiellement sur l'absence d'accord au féminin de l'adjectif (*une notion flou; des études supérieurs*) et aussi et non moins souvent sur l'ajout de lettres qui correspondent grammaticalement au féminin sur l'adjectif (*un nouvelle emploi; les mots inconnues*). Là encore, on peut trouver dans les langues premières une explication à cette convergence de pratiques. Si les hispanophones rejoignent ici les Afghans, c'est que malgré leur usage d'un genre grammatical, celui-ci ne se manifeste jamais uniquement orthographiquement: quand le genre n'est pas porté par l'adjectif à l'oral (comme c'est le cas pour *negro/negra*), la forme écrite est épiciène (*suave/suave*) mais on ne connaît pas de cas comme *noir/noire*. Les erreurs sur les participes passés et les attributs qui ont été intégrées à cette catégorie ne sont pas nombreuses (*elle est basé*) et restent donc marginales. Un étudiant hispano-américain se distingue néanmoins ici car, contrairement aux autres, ses erreurs portent essentiellement sur les participes passés. Pour autant, cette particularité ne signale pas pour nous une incompétence particulière de cet étudiant sur la question des participes passés. En effet, autant il est difficile de ne pas employer de noms, et donc de masquer la non-reconnaissance de leur genre, autant l'emploi des participes passés peut être évité ou du moins limité dans des textes tels que ceux que nous étudions. Cet étudiant fait davantage d'erreurs car il utilise davantage de participes passés, et sa particularité tient surtout à sa prise de risques. Enfin, sauf exception, on ne relève pas non plus de tendance à masculiniser ou à féminiser tant au niveau des groupes qu'au niveau des étudiants.

### 3.3 *Bilan*

Pour les deux groupes, les erreurs de genre au niveau morphologique et au niveau orthographique portent en majorité sur le syntagme nominal. Le genre est une des difficultés majeures pour les étudiants afghans, au niveau morphologique et au niveau orthographique. Les étudiants hispano-américains rejoignent les étudiants afghans uniquement sur les erreurs au niveau orthographique. Par rapport aux hypothèses qui pourraient éclairer ce résultat, la question du genre grammatical dans les langues premières nous semble être une des plus pertinentes.

## 4. Le nombre

La question des accords en nombre relève essentiellement de l'orthographe (*les langue*) et seulement quelques erreurs marginales sont



d'ordre morphologique (*le document authentique oraux*); nous laisserons donc ces dernières de côté. Les erreurs liées à l'accord en nombre sont la 2<sup>ème</sup> catégorie la plus nombreuse pour l'ensemble des étudiants des deux groupes.

#### 4.1 *Des accords en nombre désordonnés chez les étudiants afghans*

Parmi les différents types d'erreurs portant sur le nombre, nous n'avons pas été étonnées de l'absence de <s> sur le nom quand celui-ci est précédé d'un déterminant au pluriel (*les texte*) ni de l'absence de <s> sur l'adjectif (*des documents authentique*) pas plus que de l'absence de <s> sur les deux (*des raison typique*). En revanche, l'absence de <s> sur le nom alors que le déterminant et l'adjectif sont marqués au pluriel (*les texte littéraires*) nous a plus surprises. Ce type d'erreurs est en effet le deuxième le plus partagé par les étudiants afghans (cinq d'entre eux en produisent) et il n'est donc pas marginal. Ce non-marquage du nom pris "en sandwich" entre deux formes au pluriel suggère deux interprétations possibles. Ces erreurs peuvent être mises en relation avec le fait que les compétences des étudiants ne sont pas systématisées sur ce point: l'apprentissage est en cours, et l'attention est portée sur la marque du pluriel du déterminant et de l'adjectif en priorité. Ces erreurs peuvent aussi être éclairées par le marquage du pluriel en dari. En effet, en dari, à l'oral et à l'écrit, le pluriel est porté par le déterminant uniquement (des langues étrangères: *zaban hay kharji*; une langue étrangère: *yak zaban kharji*). Au contraire, en français, le pluriel est marqué par l'ensemble des composants du syntagme nominal à l'écrit, mais seulement par le déterminant à l'oral, à l'exception des pluriels dits "irréguliers" (*des chevaux*) et de la liaison entre l'adjectif antéposé et le nom (*des petits amis*). Ces différences de marquage peuvent être sources de confusion pour l'apprenant qui n'est pas habitué à marquer le nombre par un morphogramme, mais toujours par un phonogramme.

Du point de vue des étudiants (et non plus des types d'erreurs), on remarque que les quatre étudiants qui produisent le plus d'erreurs d'accord en nombre font aussi partie de ceux qui font le plus d'erreurs sur le genre. Généralement, la tendance au niveau de la fréquence des erreurs de genre rejoint celle des erreurs de nombre.

#### 4.2 *Les erreurs attendues des étudiants hispano-américains*

Les Hispano-américains produisent beaucoup moins d'erreurs et moins d'erreurs de types différents que les étudiants afghans: le type d'erreur le plus fréquent est, de même que pour les Afghans, celui de "*ces communication*". On ne rencontre pas non plus beaucoup d'erreurs sur les

participes passés et les attributs; elles se concentrent essentiellement sur le syntagme nominal.

Dans l'ensemble ces erreurs sont marginales dans ce corpus. Un étudiant se distingue en produisant davantage d'erreurs de ce type. Il faisait aussi partie de ceux qui produisaient le plus d'erreurs sur le genre au niveau orthographique. De la même façon que dans le corpus afghan, la tendance au niveau de la fréquence des erreurs sur le nombre rejoint donc celle des erreurs sur le genre.

## 5. Les erreurs visuographiques

Parmi les erreurs visuographiques, nous analyserons ici uniquement celles qui concernent les lettres hors système. Il s'agit toujours d'ajouts et d'omissions de lettres étymologiques et de consonnes doubles chez les Afghans et les Hispano-américains.

C'est la première catégorie d'erreurs qui n'oppose pas les deux groupes d'étudiants. On relève une trentaine d'erreurs pour chaque groupe: ajout ou omission de <h>: *autheur, authentique*; ajout ou omission de <n>: *mentioner, manière*; etc. Nous avons même relevé une erreur identique dans les deux groupes: *synonime*. De manière générale, il n'existe pas pour chaque étudiant de tendance à ajouter ou à omettre des lettres étymologiques ou des consonnes doubles.

Sur ce point, nous n'avons pas relevé dans notre corpus plus d'erreurs ou des erreurs différentes de celles produites par des natifs (telles qu'elles apparaissent dans les études déjà citées de Brissaud *et alii*, 2006, Manesse & Cogis, 2007, Jaffré & Bollengier, 1995). Autrement dit, les lettres hors système, qui constituent une des spécificités de l'orthographe française, ont été dans l'ensemble rapidement intégrées au système orthographique des apprenants FLE. Notons au passage que ce résultat nuance sinon contredit l'opinion de certains spécialistes de l'orthographe qui appellent à une réforme de l'orthographe sur ces lettres hors système en priorité. Finalement, malgré leur caractère non motivé, elles posent plutôt moins de difficultés, y compris chez les apprenants de FLE, que la question des accords ou de la transcription des phonèmes.

## 6. Les erreurs phonographiques

Dans cette catégorie, nous avons classé les erreurs du type: *répétation* pour *répétition*.

### 6.1 *Des phonogrammes vocaliques problématiques pour les étudiants afghans*

Les étudiants afghans produisent en majorité des erreurs au niveau des graphèmes chargés de transcrire les phonèmes vocaliques. Par exemple, on compte pour le seul phonème /E/<sup>4</sup> quatre correspondants graphiques différents ("*des interjonction*", "*portinent*", "*logiciale*", "*distiné*"), pour le phonème /O/, deux graphèmes différents ("*communité*"; "*emboucher*") ainsi que pour le phonème /i/ ("*déchotomie*", "*répétation*"), etc. Sept des huit étudiants afghans produisent au moins une erreur de type phonographique. La répartition de ces erreurs est très inégale au sein du groupe. Ainsi, l'un d'entre eux produit un quart du total des erreurs de ce type.

Ce type d'erreur est le plus attendu dans les copies de locuteurs non-natifs, et on les trouve d'ailleurs aussi dans les copies des Ethiopiens (Mortamet & Gonac'h, 2011), mais surtout dans celles des Maghrébins analysées par Amokrane et Brissaud (2011). Il s'agit d'une conséquence naturelle de la complexité du système vocalique du français.

### 6.2 *Quelques phonogrammes consonantiques chez les étudiants hispano-américains*

La zone phonographique ne pose pas de difficultés comparables aux étudiants hispano-américains. Et, contrairement au groupe des Afghans, ce sont davantage quelques graphèmes consonantiques qui leur posent problème, en particulier le phonème /s/ (*s'agissait*, *accessible*) et les <ss> pour *suffissant* par exemple. On peut ajouter que la voyelle nasale /ɛ̃/ est souvent traduite à l'écrit par le graphème <an> (*salvadorian*). Enfin, il apparaît nettement que la plupart des erreurs de type phonogrammique sont des calques de l'espagnol: *acent* (< *acento*); *contra* (< *contra*); *divide* (< *divide*); *peuple* (< *pueblo*). Ainsi, la parenté linguistique de l'espagnol avec le français, si elle a pu favoriser certaines compétences orthographiques, peut aussi, comme ici, freiner l'apprentissage de formes pourtant assez courantes.

## 7. Les erreurs donnant lieu à plusieurs interprétations

Il convient pour terminer de revenir sur les erreurs pouvant relever de plusieurs catégories, sans que l'on puisse à aucun moment déterminer si l'erreur relève de l'une ou de l'autre, c'est-à-dire d'un problème visuographique, phonographique ou encore morphographique. Ces erreurs

<sup>4</sup> Nous prenons ici pour référence le système phonologique restreint, qui neutralise l'opposition /e/ /ɛ/ (/E/), et /o/ /ɔ/ (/O/).

ne constituent pas un bon indicateur de différences ou spécificités entre les groupes: elles ne permettent en effet ni de caractériser nos corpus – on en trouve probablement dans d'autres corpus, y compris chez des natifs – et elles ne permettent pas non plus de distinguer nos deux groupes, ni même de distinguer des étudiants à l'intérieur de chaque groupe. Pour autant, il nous semble important de les présenter et de les analyser ici, parce qu'elles interrogent les catégories établies, et qu'elles posent la question de l'interprétation des erreurs, préalable nécessaire mais pas toujours possible à leur remédiation.

### 7.1 *Visuographie et morphographie grammaticale*

Ces erreurs impliquent essentiellement deux lettres de notre système: le <e> et le <s> en position finale, qui assurent de nombreuses fonctions, comme pour ce qui nous intéresse ici une fonction morphologique – la marque du féminin (une amie), celle du pluriel (des copains) – et une fonction visuographique (le mois; la vie). Nous avons ainsi relevé une vingtaine d'erreurs sur le <e> et une dizaine d'erreurs sur le <s> assurant ces fonctions.

Prenons l'exemple de l'omission du <e> dans *mond*, *norm* et *etap*. Dans ce cas on peut dire que cette omission est liée à une méconnaissance des règles de combinaison et de position des lettres pour constituer les graphèmes. La règle en l'occurrence est que si le phonème consonantique est en position finale de mot, le graphème consonantique est accompagné d'un <e> diacritique qui le marque en tant que phonogramme; en l'absence de ce <e>, il est fréquent d'avoir affaire à un morphogramme. C'est le cas par exemple de *porte* / *port*. Cette règle vaut surtout pour les phonogrammes *t*, *d*, *s*, avec quelques exceptions, mais est beaucoup moins systématique pour *r* ou *l* (*jovial* mais *fidèle*). Ce type d'erreurs est souvent relevé chez les jeunes apprenants natifs qui ne maîtrisent pas encore ces règles. Elles sont plus rares chez les apprenants qui connaissent les règles et sont habitués à l'image de ces terminaisons de mots.

Pour les exemples d'omission tels que *écrire* et *facil*, qui, nous l'avons vu, concernent des phonogrammes qui fonctionnent moins systématiquement avec un <e> diacritique final, on peut penser que les étudiants font une analogie dans le premier cas avec des infinitifs comme *finir*, *partir* et dans le second cas avec des adjectifs terminés par <l> comme *amical*.

A l'inverse, lorsque l'étudiant ajoute des <e> quand ils ne sont pas nécessaires, on peut penser qu'il "surapplique" la règle de position précédente. C'est le cas en particulier des <e> diacritiques ajoutés après des consonnes liquides comme dans *erreure*, par analogie à *heure*. On peut aussi penser que les étudiants les ajoutent par analogie à ce qui se fait

pour certains noms d'animés, comme *professeure*. Dans ce cas ils fonctionnent comme un genre référentiel et non grammatical.

Nous retrouvons également des analogies de ce type quand les scripteurs omettent un <s> comme dans: *le moi* (quand *mois* est attendu) ou quand ils en ajoutent comme dans: *le soucis*. Il semble évident ici que l'existence de formes comme *souris* les rend familières à l'œil.

### 7.2 *Phonographie et morphographie grammaticale*

Une erreur telle que *la compréhension écrit* peut relever à la fois du cas précédent – méconnaissance des règles de combinaison graphique – et d'un cas d'omission de la marque du féminin sur un adjectif qualificatif non épïcène. Dans le second cas, il s'agit d'un problème de morphologie orthographique (§3).

### 7.3 *Phonographie et morphographie lexicale*

Nous avons relevé deux erreurs dans le corpus hispano-américain, *linguistique* pour linguistique et *briève* pour brève, qui nous font penser d'abord que l'image acoustique des mots n'est pas stable dans le répertoire de ces apprenants et que leur graphie est donc non maîtrisée. En même temps, on peut considérer que le raisonnement qu'ils ont adopté pour orthographier ces mots relève de la morphographie lexicale: *brièvement* – *briève* (en espagnol: *brevemente* – *breve*).

On le voit, ces derniers cas d'erreurs sont impossibles à interpréter de façon univoque. On ne peut se livrer à une interprétation, et donc tenter une remédiation ciblée, en l'absence d'un échange avec les scripteurs, ou du moins de tests spécifiques complémentaires. Ces erreurs étant par ailleurs assez rares, mais aussi combinées avec des erreurs relevant de l'une des interprétations possibles, on ne peut guère s'appuyer non plus sur des régularités de scripteurs.

## 8. Conclusion

Ce travail a permis de montrer qu'il existe bien des zones de l'orthographe qui sont moins bien maîtrisées par les étudiants FLE que par les étudiants natifs et qu'il convient, donc, de construire des méthodologies d'apprentissage spécifiques pour ces publics. L'analyse contrastive des copies des étudiants afghans et hispano-américains a d'abord permis de dégager des points communs à chacun des groupes: accords en genre et en nombre, ajout / omission de lettres hors système, choix de phonogrammes ne correspondant pas au phonème. Les problèmes d'accord en nombre ainsi que les erreurs liées aux lettres hors système ne sont pas spécifiques des non-natifs, on en relève également dans les copies des étudiants

natifs. Il nous semble que ces erreurs sont davantage à mettre en liaison avec la spécificité et la complexité de l'orthographe française qu'avec la question de la maîtrise du français par les scripteurs. Les erreurs liées à la transcription des phonèmes sont quant à elles spécifiques aux étudiants non-natifs. Contrairement aux autres, ces erreurs sont à mettre en lien avec la circulation interlinguistique des répertoires des apprenants.

L'analyse contrastive a aussi permis de dégager des spécificités de chacun des groupes. C'est à partir de ces spécificités que nous pouvons maintenant apporter des éléments de discussion à notre hypothèse initiale – celle selon laquelle la distance typologique des langues présentes dans le répertoire des apprenants a un effet sur les formes qu'ils produisent en français. A propos du genre morphologique, il apparaît bien qu'il fait l'objet de plus d'erreurs chez les étudiants afghans que chez les étudiants hispano-américains. On peut penser que l'apprentissage du genre grammatical pose davantage de problèmes aux étudiants afghans parce que le dari n'a pas de genre grammatical, contrairement au français et à l'espagnol. Et si les étudiants hispano-américains font parfois des erreurs sur certains mots qui ont le même genre grammatical en espagnol et en français, c'est parce que le genre grammatical des deux langues n'est pas systématiquement symétrique. Néanmoins, il nous semble bien que la proximité du français et de l'espagnol joue un rôle dans la "meilleure" maîtrise de l'orthographe des étudiants hispanophones. Il conviendra de confirmer ce résultat dans l'analyse de nouveaux corpus. Enfin, cette analyse a montré que certaines interprétations des écarts d'orthographe étaient nécessairement multiples, et qu'une remédiation en orthographe devait aussi tenir compte de la diversité des voies qui mènent aux mêmes erreurs.

## Bibliographie

- Amokrane, S. & Brissaud, C. (à paraître): L'apprentissage de l'orthographe en FLM et FLE/FLS: entre convergences et spécificités. In: *Langues modernes*.
- Brissaud, C. (2006): Le cas des homophones. In: *Cahiers pédagogiques*, 440, 47-48.
- Brissaud, C., Chevrot, J-P. & Lefrançois, P. (2006): Les formes verbales homophones en /E/ entre 8 et 15 ans: contraintes et conflits dans la construction des savoirs sur une difficulté orthographique majeure du français. In: *Langue Française*, 151. Paris (Larousse), 74-93.
- Catach, N. (1986): *L'orthographe française. Traité théorique et pratique*. Paris (Nathan).
- Catach, N. (1998): *L'orthographe, Que sais-je?* Paris (PUF) 8e édition.
- Coste, D. (2002): Compétence à communiquer et compétence plurilingue. In: Castellotti, V. & Py, B. (éds): *La notion de compétence en langue, Notions en questions*, 6. Lyon (ENS éditions), 115-123.

- Coste, D., Moore, D. & Zarate, G. (1997): Compétence plurilingue et pluriculturelle. Vers un Cadre Européen Commun de référence pour l'enseignement et l'apprentissage des langues vivantes. Strasbourg (édition du Conseil de l'Europe).
- Gonac'h, J. (2009): Des variations orthographiques dans les productions de jeunes bilingues turc-français et turc-anglais. In: *Langage et société*, 130, 105-122.
- Honvault-Ducrocq, R. (dir.) (2006): *L'orthographe en questions*. Rouen (PUR).
- Jaffré, J-P. & Bollengier, F. (1995): Le lycée professionnel et l'orthographe. In: Anis, J. & Cusin-Berche, F. (éds.): *Difficultés linguistiques des jeunes en formation professionnelle courte*, numéro spécial *Linx*, 279-289.
- Lefrançois, P. (2001): Le point sur les transferts dans l'écriture en langue seconde. In: *Canadian Modern Language Review*, 58. Toronto (University of Toronto Press).
- Lucci, V. & Millet, A. (dirs.) (1994): *L'Orthographe de tous les jours. Enquête sur les pratiques orthographiques des Français*. Paris (Champion).
- Manesse, D. & Cogis, D. (2007): *Orthographe, à qui la faute?* Paris (ESF).
- Mortamet, C. & Gonac'h, J., (2011): Variation orthographique en français: le cas des non-natifs. In: Schnaffer, I. & Bertrand, O. (dirs.): *Variétés, variations et formes du français*. Paris (Les éditions de l'école Polytechnique).
- Mortamet, C. (2003): *La diversité à l'université, Analyse sociolinguistique de copies et de discours d'étudiants entrant à la faculté de Lettres et Sciences humaines de Rouen*. Thèse de doctorat, Université de Rouen.
- Porquier, R. & Py, B. (2005): *Apprentissage d'une langue étrangère: contexte et discours*. Didier (Paris).
- Selinker, L. (1994): *Reconsidering interlanguage*. London/New York (Longman).
- Tarone, E. (2005): Fossilization, social language and language play. In: Han, Z. & Odlin, T. (eds.): *Studies of fossilization in second language acquisition*, Clevedon, *Multilingual Matters*, 157-172.