

## **Analyse des performances orthographiques**

Dans ce chapitre, nous présenterons les principaux résultats obtenus de notre analyse du traitement orthographique des apprenants chinois en français L2. Cette partie sera divisée en trois grandes sections : orthographe et phonétique (comment les apprenants réalisent l'orthographe lexicale) ; orthographe et syntaxe (comment les apprenants réalisent l'accord du nombre et du genre) ; orthographe et morphologie (comment les apprenants réalisent les terminaisons verbales). La discussion théorique, ainsi que les scénarios psycholinguistiques possibles pour interpréter les erreurs orthographiques produites par les apprenants chinois seront intégrés dans la discussion des résultats empiriques.

### **4.1. Orthographe et phonétique**

Tout système graphique est multidimensionnel. Les deux dimensions jouant un rôle décisif sont la dimension phonographique et la dimension sémiographique. En tant qu'écriture alphabétique, le système graphique du français « consacre en son principe l'asservissement du signe graphique au signe phonique » (Martinet, 1974 : 66). Etant donné que toutes les variations d'orthographe « ne sont pas la contrepartie d'une différence au niveau phonique » (Ameur-Amokrane, 2007), la dimension sémiographique est également importante dans le système du français. Elle est présentée notamment par des morphogrammes non prononcés (morphogramme grammatical : *-s* dans le mot *pommes* ; morphogramme lexical : *-d* dans le mot *grand*), et également par des logogrammes servant à distinguer les homophones hétérographiques tels que *ver/vers*.

Néanmoins, 85% de l'orthographe du français est phonographique (Ameur-Amokrane, 2007). Autrement dit, 85% des graphèmes servent à transcrire des phonèmes. Ces phonogrammes pourraient être plus facile à maîtriser par les apprenants, en tenant compte que les phonogrammes ont un correspondant oral, et que les apprenants peuvent

s'appuyer sur leur forme orale lors de la transcription des phonogrammes. Cependant, comme Ameer-Amokrane (2007) l'indique, pour que cela soit vrai, encore faudrait-il que « les phonèmes soient maîtrisés tant sur le plan de l'articulation que de la discrimination auditive » et donc que les apprenants maîtrisent le système phonologique de la langue française. Ainsi, il est pertinent de dire que les difficultés auxquelles les apprenants sont confrontés au niveau phonique pourraient avoir des répercussions sur le maniement du système graphique. Rappelons que la situation d'apprentissage de l'écrit pour les apprenants du français L2 diffère de celle des enfants français. L'apprenant L2 apprend simultanément le français oral et écrit, tandis que l'enfant français aborde l'écrit avec une maîtrise précoce du français oral. Il nous semble donc intéressant d'examiner comment les apprenants chinois réalisent les phonogrammes lors de la production écrite.

Dans cette partie, nous nous intéressons particulièrement à deux types d'erreur, les erreurs orthographiques à dominante phonétique et les erreurs phonographiques. Notre objectif est d'étudier la répartition de ces deux types d'erreur chez les apprenants chinois, et d'examiner s'il existe les tendances interlangagières propres à ce groupe d'apprenants, qui abordent simultanément le français écrit et oral. A cet effet, nous chercherons tout d'abord à définir ces deux types d'erreur orthographique. Dans l'étape suivante, nous présenterons une revue non exhaustive de quelques études sur l'analyse de ces deux types d'erreur orthographique en français L1 et L2. Ensuite les hypothèses de recherche seront présentées à partir d'une comparaison synthétique des deux systèmes phonologiques, français et chinois. Et puis nous mènerons une analyse des données dans une perspective développementale, et essaieront de caractériser l'interlangue des apprenants en nous basant sur une analyse des graphies erronées. Les hypothèses psycholinguistiques possibles pour expliquer ces erreurs seront également discutées, en faisant référence aux caractéristiques spécifiques du traitement orthographique des apprenants de L2.

#### 4.1.1. Différents aspects de l'orthographe lexicale

L'orthographe française est traditionnellement divisée en deux branches, à savoir l'orthographe grammaticale et l'orthographe lexicale. Si l'orthographe grammaticale

repose sur un nombre fini de règles, l'orthographe lexicale, ou l'orthographe des mots du lexique, est essentiellement arbitraire et imprévisible (Manesse, 2007 : 137). Dans les systèmes orthographiques transparents, l'acquisition des correspondances phonèmes-graphèmes semble permettre aux apprenants de lire et d'écrire la totalité des mots. Le lexique mental, assurant l'identification et la production des mots, ne paraît plus nécessaire dans de tels systèmes, étant donné que la voie indirecte reposant sur les correspondances phonèmes-graphèmes suffirait. Néanmoins, les données empiriques nous suggèrent que même les utilisateurs de tels systèmes mémorisent les formes orthographiques de mots, puisque le lexique mental, ou la voie directe, leur permet de lire et d'écrire plus rapidement. Ce constat nous amène à soulever une question fondamentale : comment les apprenants acquièrent-ils la forme orthographique des mots ? Les données théoriques et empiriques nous attestent que la réalisation de l'orthographe lexicale demande les connaissances appartenant à des registres de savoir différents. Et ces connaissances sont construites à des moments différents du développement d'un apprenant.

La première d'entre ces connaissances spécifie la relation à la forme orale des mots (Manesse, 2007 : 138). Comme Fayol (2008) l'indique, la phonologie, notamment par le biais du recodage phonologique ou le déchiffrage, joue un rôle majeur dans l'apprentissage du lexique orthographique. Par exemple, le déchiffrage de  $m+ai \rightarrow mai+s \rightarrow mais+on$  permettrait de produire la forme orale /mezɔ̃/ et de mettre en correspondance cette forme orale avec la forme écrite *maison*. Ainsi, à l'aide de la production de la forme orale de mots, la pratique du déchiffrage sert soit à mettre en relation cette forme orale avec la forme écrite déjà connue par l'apprenant, soit à mémoriser la forme orthographique correspondante. Toutefois, rien n'assure qu'une seule rencontre suffirait pour la mémorisation de la forme écrite de nouveaux mots. Fayol (2008 : 185) nous suggère que, pour certains utilisateurs et certains mots, « plusieurs rencontres et déchiffrages successifs sont nécessaires pour inclure un mot dans le lexique orthographique ». Selon Pacton et *al.* (2005), la phonologie joue un rôle crucial dans le système d'écriture du français, même pour apprendre l'orthographe de mots irréguliers, lesquels « ne comportent souvent qu'une seule graphie irrégulière ». Par exemple, au lieu

de mémoriser toutes les lettres constitutives du mot *oignon*, l'apprenant pourrait apprendre seulement que le phonème /o/ correspond à *-oi* dans le mot avec une bonne connaissance des correspondances phonèmes-graphèmes.

Néanmoins, il faut indiquer que le français constitue un système orthographique opaque, où l'on constate de nombreuses irrégularités, et qu'une bonne maîtrise des correspondances phonographémiques ne permet l'écriture correcte que de la moitié des mots français (simulation informatique de Véronis, 1988). Ainsi, en français, pour une production orthographique correcte, la prise en compte des contraintes orthographiques serait nécessaire. Plus l'apprenant utilise le système orthographique du français, plus il rencontre de la difficulté à transcrire les mots français en utilisant seulement les appariements phonèmes-graphèmes. Ainsi se marque l'entrée dans la phase orthographique, où l'apprenant commence à découvrir les particularités orthographiques des mots. Foulin (2003) a décrit l'évolution des performances orthographiques des enfants : au fil des rencontres successives du même mot, la représentation de la forme du mot en mémoire se raffine au fur et à mesure, d'abord en fonction de la forme phonologique (*\*éléfan* pour *éléphant*), puis se référant aux indices orthographiques de plus en plus conformes au caractère conventionnel de l'orthographe (*\*éléfant* → *éléphant*).

Cette évolution réduisant progressivement la distance à la norme orthographique est également illustrée dans la recherche de Martinet et *al.* (1999). Les chercheuses ont mené une dictée de mots à des enfants de la première à la cinquième de l'école primaire. La dictée concerne trois séries de mots différant par leur régularité : mots réguliers (*confiture, marmite*) ; mots moyennement réguliers (*angine, bain*) ; mots très irréguliers (*habit, monsieur*). Les résultats ont montré que les performances des élèves s'améliorent en fonction de leur niveau scolaire et que les mots réguliers sont mieux orthographiés. Avec les analyses fines de formes erronées, les chercheuses ont également observé que les erreurs phonologiquement plausibles dominent de plus en plus par rapport à celles qui altèrent l'image phonologique de mots. En outre, les suites illégales de lettres ont une tendance à disparaître dans la production d'élèves en fonction de leur niveau scolaire, tandis que les indices orthographiques présents dans les mots irréguliers du français

apparaissent de plus en plus fréquemment dans la production des élèves, même si leur emplacement est quelquefois erroné (\**abhit*, \**ihver*). Ces résultats étayent donc le rôle fondamental de la phonologie pour établir les premières représentations orthographiques. Néanmoins, à part les appariements phonèmes-graphèmes, les élèves apprennent également les indices orthographiques, qui est relativement indépendant de la phonologie.

Ainsi nous constatons que l'orthographe « ne doit plus être vue comme une simple mémorisation de formes visuelles, mais comme une construction continue, une reconstruction du système orthographique » (Lalande, 1988 : 13). Dans son travail sur l'orthographe lexicale, Lalande nous indique que cet apprentissage est possible chez l'apprenant grâce à l'aide de deux types de connaissances : les connaissances spécifiques, qui se définissent comme « toute information en mémoire, reliée à un mot particulier du lexique personnel et susceptible d'intervenir dans la lecture et la production graphique de ce mot » (1988 : 24) et les connaissances générales, qui constituent « l'ensemble des connaissances reliées à la construction d'une graphie conforme à la langue » (1988 : 32). Pour illustrer ces deux sortes d'informations orthographiques, nous pouvons reprendre un exemple de Lalande (1988 : 26) : dans le mot *maison*, savoir que le phonème /ɛ/ correspond au graphème *-ai* concerne la connaissance d'une correspondance phonème-graphème à l'intérieur de ce mot et constitue ainsi une connaissance spécifique, et savoir que le graphème *-ai* est une des graphies possibles en français représentant le phonème /ɛ/ constitue une connaissance générale.

Étant donné le nombre de graphèmes possibles pour un phonème et le caractère polyvalent du système orthographique du français, la quantité des connaissances spécifiques à maîtriser dans l'apprentissage de l'orthographe lexicale doit être énorme. Cela explique probablement le fait que l'apprentissage de l'orthographe de mots est effectivement long. Comme Lalande (1988 : 193) l'indique, « obtenue par construction ou par combinaison de sources diverses, l'image graphique peut s'avérer conforme à l'orthographe ou s'en éloigner considérablement ». Les informations que nous venons de présenter nous permettent de mieux comprendre les procédures impliquées dans l'apprentissage de l'orthographe lexicale, ainsi que les erreurs commises par les apprenants lors de la production écrite.

### 4.1.2. Erreurs lexicales

Si les choix concernant l'orthographe grammaticale se font « en fonction des unités effectivement présentes dans la chaîne du discours », en revanche, les décisions en matière d'orthographe lexicale se prennent « dans un ensemble virtuel, un dictionnaire de formes possibles » (Manesse, 2007 : 144). Dans une tâche d'écriture telle que la composition, afin de satisfaire aux normes de l'orthographe lexicale, l'apprenant-scripteur doit choisir la forme juste, associée à la signification décidée dans le message préverbal.

Nous avons précédemment fait état des différentes compétences linguistiques impliquées dans la production de l'orthographe lexicale. Comme il s'agit d'une écriture semi-spontanée dans ce travail, les apprenants chinois devaient s'en accommoder et n'avaient pas d'autre choix que de traiter le problème de l'orthographe lexicale en s'appuyant sur toutes les compétences linguistiques dont ils disposaient. Dans leur production, nous pouvons observer les formes correctes et erronées mobilisées selon « une logique dictée par les connaissances du code orthographique emmagasinées » (Mayard, 2007). Dans une situation de dictée, comme Manesse (2007 : 145) l'indique, on n'est pas capable de dire si la forme erronée n'a pas été comprise par l'apprenant, ou si simplement elle n'était pas connue dans sa spécialité orthographique. En revanche, en produisant les graphies erronées lors d'une production écrite semi-spontanée, les apprenants ont pour la plupart du temps compris le sens de mots mais n'en connaissaient pas les caractéristiques orthographiques.

#### 4.1.2.1. Classement des erreurs lexicales

Selon Manesse (2007), les formes erronées relevant de l'orthographe lexicale pourraient être groupées en deux ensembles : erreurs de langue et erreurs lexicales au sens strict. Pour les erreurs de langue, il s'agit principalement des formes révélant que le mot n'est pas identifié pour des raisons différentes : soit les mots sont coupés, segmentés ou regroupés par erreur ; soit ils sont remplacés par un autre terme que celui du mot dicté

(dans la situation de dictée de contrôle) ; soit la représentation phonétique des mots est aberrante. Quant aux erreurs lexicales au sens strict, trois types d'erreurs sont classés dans cette catégorie : formes erronées présentant un cumul de fautes grammaticale et lexicale ; erreurs lexicales graves ; erreurs lexicales représentant les formes approchantes qui témoignent d'une bonne connaissance du système graphique français.

Dans ce travail, il nous semble nécessaire de donner une différente typologie des erreurs lexicales que celle de Manesse, puisque nos données sont issues d'une tâche d'écriture semi-spontanée. Comme nous l'avons souligné auparavant, les erreurs lexicales ont été produites par des apprenants qui ont compris le sens des mots mais n'en connaissaient pas les caractéristiques orthographiques. Ainsi, nous proposons une typologie des erreurs lexicales adaptée à notre corpus de données.

**Tableau 8.** Typologie des erreurs lexicales

Types d'erreur	Définitions	Exemples d'erreur
1	Mauvais découpage du mot	sans fonce
2	Substitutions de mots	sont font
3	Erreurs orthographiques à dominante phonétique	s'onvonce
4	Erreurs phonographiques	s'enffonssent
5	Formes approchantes	s'enfonsent

Dans ce chapitre, nous nous concentrerons sur les trois derniers types d'erreur, notamment le type 3, qui représentent une modification grave de l'image phonique de mots, et qui semble être les erreurs spécifiques des apprenants du français L2. Afin de mener les analyses fines, il nous est indispensable de comprendre la nature des erreurs commises par les apprenants. Ainsi donnerons-nous par la suite les définitions et les caractéristiques de ces trois types d'erreur.

#### **4.1.2.2. Erreurs orthographiques à dominante phonétique**

Les erreurs orthographiques à dominante phonétique, bien que classées et repérées parmi les erreurs d'orthographe, se distinguent d'autres types d'erreurs d'orthographe proprement dites par sa nature particulière. Ce type d'erreur regroupe tous les mots dont la représentation phonétique est erronée. Selon Manesse, la mauvaise transcription des sons nous amène à considérer qu'« on est bien loin de la pure et simple faute lexicale » (2007 : 73). Si dans une situation de dictée, ce type d'erreurs pourrait s'expliquer par le fait que le mot est visiblement inconnu de l'apprenant, puisque l'apprenant cherche à restituer les sons de ce mot comme s'il les percevait pour la première fois. En revanche, lors de la production écrite, ces erreurs, qui occasionnent une dégradation de la valeur phonique du mot concerné, relèvent plutôt d'une « image auditive inaccessible » (Katoozian, 2013). Cependant, il n'est pas superflu d'indiquer que dans la situation de dictée, et d'autant plus de la production écrite, les erreurs lexicales dont la représentation phonétique est aberrante relèvent d'une étape préorthographique et ce type d'erreur nécessite la correction phonétique.

Ces erreurs orthographiques à dominante phonétique concernent notamment l'omission ou l'adjonction de lettres / syllabes, la confusion des consonnes, des voyelles ou encore l'adjonction d'un accent altérant gravement la valeur phonique du mot. Autrement dit, la compréhension des bases phoniques de la part du sujet-scripteur est différente de l'image phonique du mot concerné.

#### **4.1.2.3. Erreurs phonographiques**

Cette catégorie regroupe les formes erronées dont les sons sont reproduits correctement ou presque, mais la forme graphique est loin d'être maîtrisée. Ainsi se trouvent dans cette catégorie toutes appariements phonème-graphème erronés avec incidence phonique pour les graphies polysémiques ; mauvaise notation du yod et d'autres semi-voyelles ; déplacement de lettres voisines (inversion-confusion) ; erreur de double consonne altérant l'image phonique du mot ; omission de lettres finales muettes altérant la valeur phonique



du mot ; omission du *e* caduc, possédant une valeur auxiliaire qui modifie ainsi la valeur phonique du mot.

Il faut remarquer que la distinction entre les erreurs orthographiques à dominante phonétique et les erreurs phonographiques est essentielle. Le premier type d'erreur relève d'une étape préorthographique, alors que le second concerne les erreurs purement lexicales. Malgré la modification de l'image phonique du mot occasionnée par l'un ou l'autre type, le premier concerne une mauvaise compréhension des bases phoniques de la part du sujet-scripteur, ou une image auditive inaccessible par le scripteur lors de la production écrite, tandis que le second résulte d'un non-respect des règles d'associations phonèmes-graphèmes.

#### 4.1.2.4. Formes approchantes

A l'intérieur de la catégorie des erreurs lexicales, Manesse (2007 : 75) établit une limite relativement nette entre fautes graves et formes approchantes. Ces dernières sont considérées comme des « quasi-réussites », puisque les formes approchantes témoignent d'« une bonne connaissance du système graphique français, ainsi que des moules graphiques les plus fréquents de la langue » (Manesse, 2007 : 76). Autrement dit, l'apprenant choisit une transcription plausible pour représenter les sons du mot, mais qui n'est pas la bonne.

Le modèle ci-dessous, proposé par Simon (1924 : 183), montre le chemin de transition entre les erreurs orthographiques à dominante phonétique, les erreurs orthographiques lexicales avec incidence phonique, les erreurs lexicales sans incidence phonique et la forme correcte :

**Tableau 9.** Modèle de transition de Simon

Mot : nécessaire	Phrase de début	*nassassaire
	Transition	*nésessaires
	Forme approchante	*nessésaire
	Perfection	nécessaire

### 4.1.3. Erreurs lexicales à travers quelques recherches

A travers une revue de quelques recherches sur l'étude et l'analyse des erreurs orthographiques en français L1 et en français L2, nous pouvons tracer une ligne de démarcation concernant la répartition de ce type d'erreur entre les deux peuples différents.

#### 4.1.3.1. En français L1

En français L1, Girolami-Boulinier (1984) a mené une dictée auprès des élèves du CE1 à la 3<sup>e</sup> dans la région parisienne. Le tableau ci-dessous montre que les erreurs phonétiques, assez fréquentes au début de la scolarité, deviennent le type le moins fréquent à partir du CM2. En revanche, les fautes d'usage (erreurs lexicales sans incidence phonique) persistent pendant toute la scolarité.

**Tableau 10.** Répartition des erreurs orthographiques par type (Girolami- Boulinier, 1984)

Types d'erreur	Niveaux (%)					
	Fin CE1	CE2	CM1	CM2	6 <sup>e</sup>	4 <sup>e</sup> -3 <sup>e</sup>
Fautes linguistiques	26	48	68.5	37	52	29.5
Fautes d'usage	45.5	25	13	13	13	31
Fautes de genre et de nombre	9	5	3.5	44	27.5	29.5
Fautes phonétiques	19.5	22	15	6	7.5	10
Total	100	100	100	100	100	100

La même observation a été faite aussi par Chervel & Manesse (1989), qui ont établi un corpus de 3048 copies de dictées des élèves français du CM2 à la 3<sup>e</sup>. De plus, les chercheurs ont comparé les résultats avec ceux d'un corpus de 3114 dictées au 19<sup>e</sup> siècle (du même texte de source et auprès des élèves du même niveau).

**Tableau 11.** Répartition des erreurs orthographiques par type (Chervel & Manesse, 1989)

Types d'erreur	Définitions	Siècles	
		XIX	XX
1	Mauvais découpage du mot ; mot sauté ou tronqué	7.1%	2.8%
2	Aberration dans la représentation des sons	3.3%	2.3%
3	Substitution de mot	10%	6.5%
4	Cumul de fautes grammaticale et lexicale	3.7%	5.5%
5	Fautes d'orthographe grammaticale : la catégorie grammaticale n'est pas présentée	20.1%	21.8%
6	Fautes d'orthographe grammaticale : la catégorie grammaticale est mal présentée	11.3%	13.1%
7	Faute grave d'orthographe lexicale	6.6%	10.5%
8	Faute d'orthographe lexicale : forme approchante	18.2%	18.6%
9	Faute sur les signes orthographiques et les majuscules	19.6%	18.9%
<b>Total</b>		100	

Le tableau ci-dessus nous présente le comportement orthographique des élèves pris dans leur ensemble, en le comparant à celui des élèves du 19<sup>e</sup> siècle. Entre 19<sup>e</sup> siècle et 20<sup>e</sup> siècle, nous constatons que les comportements orthographiques des élèves n'ont pas substantiellement changé. Autrement dit, la répartition des erreurs en différents types reste grosso modo la même. En ce qui concerne les erreurs orthographiques à dominante phonétique (type 2 dans le tableau ci-dessous), il s'agit dans les deux corpus des erreurs les moins fréquentes. En revanche, les chercheurs ont indiqué que l'erreur lexicale la moins grave (forme approchante) semble résister plus que les autres.

Une autre enquête identique menée par Manesse et *al.* (2007) auprès de 2767 élèves français du même niveau a conduit à un résultat analogue.

A travers l'ensemble de ces recherches en français L1, nous voyons que les erreurs orthographiques à dominante phonétique sont de faible fréquence chez les apprenants français et si elles existent, elles « s'observent aux niveaux bas de la scolarité et

disparaissent avec l'âge des élèves » (Katoozian, 2013). Néanmoins, les erreurs lexicales sans incidence phonique semblent persister pendant toute la scolarité.

#### 4.1.3.2. En français L2

Les recherches des performances orthographiques des apprenants allophones nous indiquent cependant une tendance différente. Rey-von Allmen a mené une étude du comportement orthographique auprès dix adolescents hispanophones et italoophones de 13 à 15 ans de travailleurs migrants en Suisse. En analysant les résultats obtenus en deux années scolaires successives le chercheur a constaté que des erreurs à dominante phonétique sont fréquentes chez les apprenants non-francophones. En outre, le chercheur a comparé les résultats de cet échantillon avec ceux d'un groupe d'élèves francophones de 12 à 13 ans. Cette comparaison montre que des erreurs à dominante phonétique sont exceptionnelles chez les élèves francophones. En revanche, on rencontre plutôt des erreurs phonogrammiques.

Après une analyse des erreurs orthographiques de 27 étudiants libyens de langue française (en dictée et en rédaction), Ghellai (1997) a dressé le tableau comparatif ci-dessous, qui nous montre que les erreurs phonétiques occupent le 2<sup>e</sup> rang en dictée, et qu'elles sont aussi fréquentes en rédaction.

**Tableau 12.** Répartition des erreurs orthographiques chez les étudiants libyens du français L2 (Ghellai, 1997)

Types d'erreur	Instruments d'observation	
	Dictée (%)	Rédaction (%)
Phonétiques	27.3	12.4
Phonogrammiques	35.5	13
Morphogrammiques	25.5	24
Syntaxiques	0	31.7
Logogrammiques	10.2	8
Idéographiques	0	6.6

Lettes non justifiables	1.7	0
Grammaticales	0	4
Total	100	100

A la suite d'une étude des erreurs orthographiques, basée sur 531 copies de rédaction des élèves tunisiens de la 1<sup>ère</sup> année secondaire, Lassad (2011) a fait la même observation : les erreurs phonétiques (type 2) sont les plus fréquentes juste après les erreurs d'orthographe grammaticale (type 5). Notons également que les erreurs phonétiques sont plus nombreuses que les erreurs lexicales au sens strict.

**Tableau 13.** Répartition des erreurs orthographiques chez les élèves tunisiens du français L2 (Lassaad, 2011)

Types d'erreur	Définitions	(%)
1	Mauvais découpage du mot ; mot sauté ou tronqué	4.8
2	Aberration dans la représentation des sons	24.9
3	Substitution de mot	2.7
4	Cumul de fautes grammaticales et lexicales	1.7
5	Orthographe grammaticale : la catégorie grammaticale n'est pas présentée ou mal présentée	35.8
6	Orthographe lexicale	18.1
7	Signes orthographiques et majuscules	11.5
Total		100

La recherche de Amokrane (2011) porte sur les performances orthographiques des étudiants algériens de quatre années de licence du français L2 à travers les tâches d'écriture différentes, à savoir la dictée, la rédaction et la copie. Le chercheur a séparé les erreurs en deux grands ensembles : les erreurs non spécifiques aux étudiants algériens et les erreurs spécifiques aux étudiants algériens. Quant à ce dernier groupe, le chercheur a remarqué l'omniprésence des erreurs phonétiques aussi bien en dictée et en copie qu'en rédaction.

A la suite de l'étude des copies d'examen rédigées par 8 étudiants afghans et 8 étudiants hispano-américains de 3<sup>e</sup> année de licence en sciences du langage, Gonac'h et al. (2011) ont noté que les erreurs orthographiques à dominante phonétique sont fréquentes et particulières chez ce groupe d'étudiants non-francophones<sup>27</sup>.

Katoozian (2013) a mené une dictée auprès 313 étudiants iraniens de licence en langue française. Le tableau suivant de la répartition des erreurs orthographiques nous permet de constater que les erreurs orthographiques à dominante phonétique restent parmi les trois types les plus fréquents (avec les erreurs de non-reconnaissance du mot et les erreurs d'orthographe grammaticale). En revanche, les erreurs lexicales avec ou sans incidence phonique sont beaucoup moins fréquentes.

**Tableau 14.** Répartition des erreurs orthographiques chez les étudiants iraniens en français L2

(Katoozian, 2013)

Types d'erreur	Définitions	(%)
1	Non-reconnaissance du mot en tant qu'unité lexicale	25.6
2	Erreur phonétique	22.6
3	Substitution de mot	0.7
4a	Erreurs de grammaire et de lexique avec incidence phonique	1.4
4b	Erreurs de grammaire et de lexique sans incidence phonique	4
5	Orthographe grammaticale	26.3
6	Orthographe lexicale avec incidence phonique	3.3
7	Orthographe lexicale sans incidence phonique	8.9
8a	Accents, tréma, cédille	6.5
8b	Apostrophe, trait d'union, majuscules	0.8
Total		100

<sup>27</sup> Spécifiquement vocaliques pour les étudiants afghans et consonantiques pour les étudiants hispano-américains

L'ensemble des observations nous montre que les erreurs orthographiques à dominante phonétiques sont fréquentes et particulières aux apprenants du français L2. Selon ces chercheurs (Amokrane, 2011 ; Kattozian, 2013), ce type d'erreur résulte des dissemblances entre le système phonologique du français et celui de la langue maternelle. Pour mieux comprendre les difficultés auxquelles les apprenants chinois sont confrontés lors de la perception et de la production du français, il nous semble important de décrire par la suite le système phonétique du chinois.

#### 4.1.4. Système phonétique du chinois

Le chinois standard (mandarin standard) désigne une norme de langue. Il est utilisé par les médias, l'Etat et pour l'enseignement (S. Landron, J. Gao, Y.-C. Chang & Y. Tian, 2017) en République populaire de Chine (RPC). Au niveau de la phonologie, il a largement été inspiré par le mandarin de Pékin. Quant à la transcription des sinogrammes, elle s'effectue généralement avec le Pinyin, mais aussi avec le bopomofo<sup>28</sup>.

##### 4.1.4.1. Structure syllabique

Avant d'entrer dans la description des formes phonologiques du chinois, il nous semble nécessaire d'introduire en premier lieu la composition syllabique du chinois, car la description des sons du chinois standard s'organise autour de la syllabe, unité phonique de base (Drocourt, 2007 : 155).

Pour illustrer la structure syllabique du chinois standard, Drocourt a fait une comparaison entre le français et le chinois. Dans chacun des deux tableaux ci-dessous, la première ligne présente la phrase découpée en syllabe, la deuxième ligne, les morphèmes correspondant aux syllabes de la première ligne, et la troisième ligne, les mots composés de la phrase. Le tableau nous permet de constater que pour la phrase en français, il n'existe « aucune coïncidence unitaire entre les trois niveaux » (Drocourt, 2007 : 155). Même entre la syllabe [nu] et le mot *nous*, il n'y pas de coïncidence étant donné que la liaison

---

<sup>28</sup> Système de transcription utilisé à Taiwan, qui transcrit les sons avec des signes non latins.

fait basculer le –s final du mot *nous* à la syllabe suivante. Comme Drocourt nous l’indique, la coïncidence pourrait avoir lieu avec des mots monosyllabiques, tels que *la*, *quatre*, qui sont cependant peu nombreux en français. D’une façon générale, « la frontière morphémique ne correspond ni à celle de la syllabe ni à celle du mot » (Drocourt, 2007 : 155). Néanmoins, la coïncidence est beaucoup plus régulière en chinois. Avec le tableau ci-dessous, nous constatons que « la syllabe au niveau de la phonation correspond au morphème au niveau du sens » (Drocourt, 2007 : 156). Autrement dit, le chinois fonctionne essentiellement avec des monosyllabes<sup>29</sup>.

**Tableau 15.** Rapport entre syllabe, morphème et mot en français (Drocourt, 2007 : 155)

(i) Nous aimons lire les magazines.

Niveau	Recoupement								
Syllabation	[nu]	[zɛ]	[mɔ̃]	[lir]	[le]	[ma]	ga	zin]	
Morphème	1 <sup>e</sup> PL	(PL)	aim-	ons	li-	re	le	PL	magazine(-s muet)
Mot	nous	aimons			lire	les	magazines		

**Tableau 16.** Rapport entre syllabe, morphème et mot en chinois (Drocourt, 2007 : 156)

(ii) Women ai kan huabao.

Nous aimons lire les magazines

Niveaux	Recoupement					
Syllabation	[wo]	[mən]	[ai]	[k <sup>h</sup> an]	[xua]	[pau]
Morphème	1 <sup>e</sup> SG	PL	aimer	lire	image	journal
Mots	women	ai	kan	huabao		

Ainsi, en chinois, l’organisation syllabique est un élément capital de toute description phonologique. A la différence des nombreuses langues européenne, où les phonèmes se combinent assez librement entre eux, la structure syllabique du chinois est

<sup>29</sup> Les mots chinois peuvent atteindre plusieurs syllabes, tels que *women* (*nous*), *huabao* (*magazine*), mais les mots monosyllabiques sont majoritaires.



très rigide : les phonèmes s’organisent en une séquence définie (Alleton, 2010 : 12), et on ne peut obtenir qu’environ quatre cents syllabes différentes (sans l’intervention des tons). Généralement, la syllabe est formée de (S. Landron, J. Gao, Y.-C. Chang & Y. Tian, 2017) :

- Une initiale ou attaque optionnelle, qui est le phonème se trouvant au début de la syllabe. Notons que l’initiale est toujours « une unité phonématique simple, assumé une seule consonne » (Drocourt, 2007 : 157) ([lao], vieux ; [ai], aimer<sup>30</sup>) ;
- Un noyau obligatoire, qui peut être une monophthongue ([la], tirer), une diphtongue ([lai], venir) ou une triphthongue ([xiao], petit) ;
- Une coda optionnelle limitée aux nasales [n] ou [ŋ], qui est placée après une monophthongue ou une diphtongue ascendante ;
- Un ton que porte la rime.

#### 4.1.4.2. Consonnes

Le système consonantique du chinois compte 21 consonnes, qui se distinguent selon leur mode d’articulation et leur lieu d’articulation :

Consonnes occlusives :

	Sourdes Non-aspirées		Sourdes Aspirées	
Bilabiale	[p]	爸 [pa]	[p <sup>h</sup> ]	怕 [p <sup>h</sup> a]
Dentale	[t]	大 [ta]	[t <sup>h</sup> ]	它 [t <sup>h</sup> a]
Vélaire	[k]	卡 [ka]	[k <sup>h</sup> ]	卡 [k <sup>h</sup> a]

Consonnes fricatives :

---

<sup>30</sup> Sans attaque, la syllabe commence directement par une voyelle.

	Sourdes		Sonores	
Labio-dentale	[f]	发 [fa]		
Alvéolaire	[s]	仨 [sa]		
Rétroflexe	[ʃ]	沙 [ʃa]	[z]	热 [ʒ]
Alvéo-palatale	[ç]	西 [çi]		
Vélaire	[x]	哈 [xa]		

Consonnes affriquées :

	Sourdes		Sourdes	
	Non-aspirées		Aspirées	
Alvéo-palatale	[tʃ]	几 [tʃi]	[tʃʰ]	七 [tʃʰi]
Rétroflexe	[tʃ]	扎 [tʃa]	[tʃʰ]	茶 [tʃʰa]
Alvéolaire	[ts]	砸 [tsa]	[tsʰ]	擦 [tsʰa]

Consonnes nasales :

	Sonores	
Bilabiale	[m]	妈 [ma]
Dentale	[n]	那 [na]
Vélaire	[ŋ]	昂 [aŋ]

Consonne latérale :

	Sonore	
Alvéolaire	[l]	拉 [la]

#### 4.1.4.3. Voyelles

Le chinois standard compte 5 voyelles phonémiques :

	Non-arrondies		Arrondies	
Haute	[i]	里 [li]	[y]	绿 [ly]
			[u]	路 [lu]
Mi-basse			[ə]	乐 [lə]
Basse	[a]	拉 [la]		

Il existe de nombreuses diphtongues et triphthongues en chinois standard, où les voyelles subissent des règles d'assimilation et leur réalisation varie en fonction de la voyelle précédente et/ou suivante (S. Landron, J. Gao, Y.-C. Chang & Y. Tian, 2017).

Diphtongues descendantes :

[ai]	来 [lai]
[au]	老 [lau]
[əu]	楼 [ləu]
[əi]	累 [ləi]

Diphtongues ascendantes :

[iə]	列 [liə]
[ia]	俩 [lia]
[yə]	略 [lyə]
[uə]	罗 [luə]
[ua]	花 [xua]

En se combinant avec les nasales [n] et [ŋ], elles ont des réalisations différentes :

[an]	蓝 [lan]	[iɛn]	连 [liɛn]	[yɛn]	元 [yɛn]	[uan]	乱 [luan]
[aŋ]	狼 [laŋ]	[iaŋ]	两 [liaŋ]	[uaŋ]	广 [kuaŋ]		

[ən]	本 [pən]		
[ɤŋ]	冷 [lɤŋ]	[uɤŋ]	翁 [uɤŋ]
[in]	林 [lin]		
[iŋ]	零 [liŋ]		
[un]	论 [lun]		
[yn]	寻 [ɥyn]		
[uŋ]	工 [kuŋ]	[iuŋ]	凶 [ɥiuŋ]

#### 4.1.4.4. Difficultés lors de la production orale

Étant donné les différences existant dans les systèmes consonantiques et vocaliques des deux langues, les apprenants chinois sont confrontés aux difficultés spécifiques lors de la production orale (S. Landron, J. Gao, Y.-C. Chang & Y. Tian, 2017). Ces difficultés dépendent également de la structure syllabique plus complexe en français qu'en chinois.

##### 1) Difficultés liées aux voyelles absentes du système en L1

- Les voyelles nasales du français ne sont pas les mêmes que les rimes chinoises [in], [iŋ], [an], etc. Ainsi, les apprenants chinois ont tendance à réaliser les voyelles nasales du français comme « des voyelles orales nasalisées avec un appendice nasal [n] ou [ŋ] » (Landron & al., 2017) : [ã] prononcé [aŋ] ; [õ] prononcé [oŋ] ; [ɛ̃] prononcé [an] ou [ən].
- Les apprenants chinois peinent à réaliser les voyelles arrondies antérieures [ø] et [œ], qui sont prononcés comme [ɤ], [y] ou [u]. Par exemple : *deux enfants* [duzãfã].

##### 2) Difficultés liées aux voyelles phonétiquement différentes

- Les voyelles moyennes [e] [ɛ] [o] [ɔ] peuvent être diphtonguées par les apprenants chinois : [e] prononcé [ei] ; [ɛ] prononcé [ai] ; [o] prononcé [ou] ; [ɔ] prononcé [au].

##### 3) Difficultés liées aux consonnes absentes du système en L1

- Généralement, l'opposition de voisement pose des difficultés aux apprenants chinois, puisque 17 consonnes du chinois standard sont sourdes ou sourdes aspirées parmi les 21 consonnes et qu'il n'y a que trois consonnes sonores communes aux deux systèmes consonantiques ([m], [n] et [l]). Ainsi, les consonnes sonores en français [b] [d] [g] [v] [z] [ʒ] sont produites souvent comme des sourdes [p] [t] [k] [f] [s] [ʃ] : gâteau [kato] ; et les consonnes sourdes parfois comme des sonores : *manquer* [mãʒe].
- [ʒ] est prononcé comme les consonnes affriquées [tʃ] ou [dʒ] : *bonjour, je ...* [pɔ̃tʃux tʃ].
- La consonne uvulaire ou roulé [R] constitue un phonème difficile à acquérir quelle que soit sa position. Lors de la production orale, elle est fréquemment non prononcée. Quand l'apprenant y fait une attention particulière et la prononce, elle est souvent réalisée non voisée [x] et avec une intensité de friction excessive : rouge [xuʒx] ; en gris [ãkxi].

#### 4) Difficultés liées aux consonnes phonétiquement différentes

- Étant donné qu'en chinois, les consonnes [p/t/k] et [p<sup>h</sup>/t<sup>h</sup>/k<sup>h</sup>] s'opposent par l'aspiration, les consonnes [p/t/k] en français peuvent respectivement être produites comme des aspirées [p<sup>h</sup>/t<sup>h</sup>/k<sup>h</sup>] par les apprenants chinois : *a peint* [ap<sup>h</sup>ɛ̃].

#### 5) Difficultés liées aux structures syllabiques du français

- Les groupes consonantiques sont très courants en français, qui peuvent apparaître en position initiale, médiane ou finale de mot. Cependant, la phonologie du chinois ne permet pas une suite de deux ou trois consonnes en position initiale ou finale (Chang, 1987). Ainsi, les apprenants chinois ont une difficulté considérable à prononcer les groupes de consonnes, notamment lorsqu'ils contiennent le son [R]. La stratégie de réparation adoptée par les apprenants est alors d'insérer une voyelle entre deux consonnes (par exemple : [ə]) ou simplifier les groupes consonantiques par l'omission de la consonne finale : *phrase* [faz].

- Les apprenants chinois ont également des difficultés à réaliser les syllabes fermées. Notons que les syllabes du chinois standard finissent généralement par une voyelle, et que seulement deux consonnes nasales [n] et [ŋ] peuvent se produire à la fin des syllabes comme coda optionnel de rimes. Les apprenants chinois peuvent transmettre cette habitude dans leur prononciation du français, de sorte que la consonne en coda peut ne pas être réalisée lors de leur production orale, notamment s'il s'agit de la consonne [R] : matière [matie] ; ou une voyelle supplémentaire est ajoutée après une consonne en coda pour former une nouvelle syllabe : *pomme* [pɔmɔ].

#### 4.1.5. Hypothèses et méthodes d'observation

Le centre d'intérêt de cette partie est de savoir comment les apprenants chinois du français L2 mettent en œuvre l'orthographe des mots du vocabulaire dans la production écrite et si la phonologie joue un rôle majeur même si les apprenants ne sont pas obligés de transmettre le matériau phonique en matériau graphique dans l'activité d'écriture. En comparant les performances orthographiques des élèves francophones et des apprenants du français L2, nous constatons que les comportements orthographiques différents se dégagent chez les deux groupes d'apprenants. Comme nous l'avons indiqué auparavant, les erreurs orthographiques à dominante phonétiques sont fréquentes chez les apprenants du français L2, tandis que ce type d'erreurs ont tendance à disparaître chez les élèves français en fonction de leur niveau scolaire. En outre, les élèves français commettent les erreurs lexicales sans incidence phonique pendant toute leur scolarité.

Ainsi, nos deux questions de recherche principales dans cette partie – c'est-à-dire les questions de savoir (i) quel scénario développemental les apprenants chinois présentent dans leur production orthographique des mots du vocabulaire et (ii) quelles sont les difficultés ou tendances particulières aux apprenants chinois – seront abordées de manière pluridimensionnelle. Afin d'apporter des éléments de réponse, nous vérifierons d'abord le profil orthographique des apprenants chinois qui se trouvent aux trois niveaux scolaires différents. Nous analyserons ensuite les graphies erronées et creuserons les causes

possibles pour les expliquer et enfin, nous décrirons les caractéristiques interlangagières des apprenants chinois dans le domaine de l'orthographe lexicale.

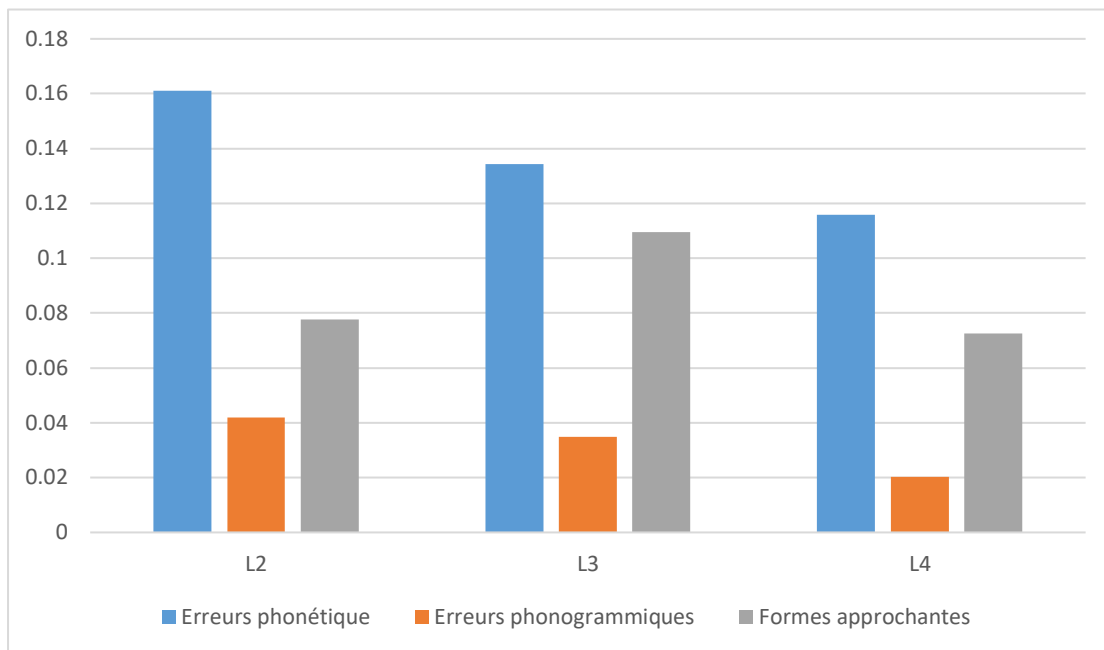
#### 4.1.6. Résultats

##### 4.1.6.1. Le profil orthographique des apprenants chinois

L'observation des graphies produites par les apprenants chinois nous permet d'une part de conclure à l'amélioration continue des performances au cours de la scolarité à l'université, et d'autre part, de repérer les tendances propres aux apprenants chinois dans les formes erronées.

###### 4.1.6.1.1. Erreurs orthographiques à dominante phonétique

La figure 15 permet de visualiser la répartition des différents types d'erreur lexicale chez les trois groupes d'apprenants chinois. Chez les apprenants de tous les trois niveaux, les erreurs orthographiques à dominante phonétique sont beaucoup plus nombreuses que les erreurs phonogrammiques et les formes approchantes. En effet, les erreurs orthographiques à dominante phonétique restent parmi les types les plus fréquents (avec les erreurs portant sur un élément morphogrammique et ayant une incidence sur la prononciation ; et sans incidence sur la prononciation). Ces remarques confirment les autres recherches portant sur l'orthographe lexicale des apprenants du français L2. A la différence des élèves français, le poids relatif important des erreurs orthographiques à dominante phonétique dessine une tendance particulière chez les apprenants allophones.



**Fig. 15 :** Comparaison de la répartition des erreurs

En termes de la répartition, ce type d'erreur régresse, mais légèrement, au cours de la scolarité. Nous constatons que le poids relatif diminue de 16.1% à 11.6%. Néanmoins, si l'on observe le nombre moyen d'erreurs commises par apprenant présenté dans le tableau suivant, les erreurs orthographiques à dominante phonétique semblent résister au cours des trois ans d'études de français : un apprenant de L2 faisait 1.35 erreur de ce type, là où un apprenant de L4 en faisait 1.48. Bien évidemment, ces chiffres ne nous permettent pas de rendre compte de la diversité des performances entre les apprenants, car il s'agit des moyennes. Mais ils nous permettent de préciser la tendance lourde : le profil développemental ne se dessine pas nettement chez les apprenants chinois au cours des trois années d'études de français.

**Tableau 17.** Nombre des erreurs orthographiques phonétiques par niveau

	L2	L3	L4	Total
Nombre des erreurs phonétiques	54	27	40	121
Nombre total des erreurs orthographiques	335	201	345	891
Pourcentage	16.1%	13.4%	11.6%	13.6%
Nombre d'apprenant par groupe	40	24	28	92



Nombre moyen d'erreur par groupe	1.35	1.125	1.48	1.32
----------------------------------	------	-------	------	------

A la différence des apprenants chinois, les élèves français se dégagent un profil orthographique complètement différent. Le tableau ci-dessous nous permet de constater que les erreurs orthographiques à dominante phonétique régressent en effectif moyen au cours de la scolarité obligatoire entre le CM2 et la troisième (Manesse & Cogis, 2007). Soulignons que ce type d'erreur a un poids relatif moindre en 1987 qu'en 2005. Néanmoins, le nombre d'erreurs diminue nettement pour les élèves en 1987 autant que pour les élèves en 2005.

**Tableau 18.** Nombre moyen d'erreurs phonétiques par classe (Manesse & Cogis, 2007)

	CM2	6 <sup>e</sup>	5 <sup>e</sup>	4 <sup>e</sup>	3 <sup>e</sup>
1987	0.42	0.30	0.18	0.08	0.05
2005	1.14	0.80	0.59	0.41	0.33

#### 4.1.6.1.2. Erreurs phonogrammiques

En ce qui concerne les erreurs phonogrammiques, elles ne sont pas nombreuses chez les apprenants chinois : nous n'avons trouvé que 28 occurrences dans notre corpus (14 occurrences pour L2 ; 7 occurrences pour L3 et L4). En termes de répartition, le tableau suivant nous permet de lire le poids relatif assez modeste de ce type d'erreur chez les trois groupes d'apprenants. Quant au nombre moyen d'erreur commise par apprenant, un apprenant faisait 0.3 erreur de ce type. Si nous observons le profil développemental, ce type d'erreur n'évolue que faiblement entre L2 et L4, tant en termes de répartition qu'en terme du nombre moyen d'erreurs.

**Tableau 19.** Nombre des erreurs orthographiques phonogrammiques par niveau

	L2	L3	L4	Total
Nombre des erreurs phonogrammiques	14	7	7	28
Nombre total des erreurs orthographiques	335	201	345	891
Pourcentage	4.2%	3.5%	2%	3.1%

Nombre d'apprenants par groupe	40	24	28	92
Nombre moyen d'erreurs par groupe	0.35	0.29	0.25	0.3

Pour les élèves français, la recherche de Manesse & al. (2007) nous montre que les erreurs phonogrammiques sont plus lourdes en poids relatif. Néanmoins, un profil développemental assez évident se dessine chez les élèves français, étant donné que le nombre moyen d'erreurs par élève diminue nettement entre CM2 et la troisième, tant en 1987 qu'en 2005.

**Tableau 20.** Nombre moyen d'erreurs phonétiques par classe (Manesse & Cogis, 2007)

	CM2	6 <sup>e</sup>	5 <sup>e</sup>	4 <sup>e</sup>	3 <sup>e</sup>
1987	1.50	1.25	0.92	0.62	0.41
2005	1.26	1.02	0.81	0.72	0.59

#### 4.1.6.1.3. Erreurs lexicales sans incidence phonique

Le troisième type d'erreur dont nous discutons dans cette partie concerne les erreurs lexicales sans incidence phonique, autrement dit les formes approchantes. Le nombre d'erreurs commises par les apprenants chinois ne change pas substantiellement entre les trois niveaux différents. En outre, ce type d'erreur représente un profil développemental différent que les deux types d'erreur que nous avons discuté précédemment. En termes de répartition, les erreurs lexicales sans incidence phonique sont plus lourdes en poids relatif au niveau avancé qu'au niveau initial. Cette remarque n'est pas infirmée si nous observons le nombre moyen d'erreurs par apprenant : un apprenant de L2 faisait 0.65 erreur, tandis qu'un apprenant de L3 en faisait 0.97 et qu'un apprenant de L4 en faisait 0.89. Ainsi l'observation du tableau ci-dessous nous permet de faire deux remarques : d'une part, les erreurs lexicales sans incidence phonique semblent résister au cours des trois années d'études de français ; d'autre part, les apprenants chinois au niveau avancé ont tendance à faire plus d'erreurs de ce type que les apprenants au niveau initial.

**Tableau 21.** Nombre des erreurs lexicales sans incidence phonique par niveau

	L2	L3	L4	Total
Nombre des formes approchantes	26	22	25	73
Nombre total des erreurs orthographiques	335	201	345	891
Pourcentage	7.7%	10.9%	7.2%	8.2%
Nombre d'apprenants par groupe	40	24	28	92
Nombre moyen d'erreurs par groupe	0.65	0.97	0.89	0.79

Selon le tableau suivant, les erreurs lexicales sans incidence phonique ont un poids relatif plus important chez les élèves français que chez les apprenants chinois. Néanmoins, chez les élèves français, ce type d'erreur régresse au cours de la scolarité entre le CM2 et la troisième. Indiquons également ces erreurs n'évoluent que faiblement et qu'elles persistent au cours de toute la scolarité obligatoire.

**Tableau 22.** Nombre moyen d'erreurs phonétiques par classe (Manesse & Cogis, 2007)

	CM2	6 <sup>e</sup>	5 <sup>e</sup>	4 <sup>e</sup>	3 <sup>e</sup>
1987	1.99	1.87	1.66	1.45	1.32
2005	2.35	2.22	2.09	2.14	1.93

En résumé, à partir d'une analyse comparative, nous constatons que les comportements orthographiques des apprenants chinois sont substantiellement différents de ceux des élèves français. Les erreurs orthographiques à dominante phonétique ont un poids relatif assez lourd chez les apprenants chinois. En outre, au cours de toute la scolarité, les erreurs de ce type persistent. Cependant, en termes de répartition, le poids relatif des erreurs phonétiques est moindre au niveau avancé qu'au niveau initial. Les erreurs phonogrammique occupent en revanche un poids relatif modeste, et elles n'évoluent que faiblement en fonction du niveau scolaire. Quant aux erreurs lexicales sans incidence phonique, les erreurs de ce type sont beaucoup moins lourdes en poids relatif que les erreurs orthographiques à dominante phonétique. De plus, ce type d'erreurs résistent durant toute la scolarité, et il nous semble que les apprenants au niveau avancé

ont tendance à faire plus d'erreurs de ce type que les apprenants du niveau initial.

A partir de ce scénario développemental global des apprenants chinois, nous proposerons dans la partie suivante les analyses fines des graphies erronées commises par les apprenants.

#### **4.1.6.2. Analyse des erreurs**

Pour dessiner le profil orthographique complet des apprenants chinois, il nous est indispensable de comprendre la nature des erreurs qu'ils ont commises.

##### **4.1.6.2.1. Erreurs orthographiques à dominante phonétique**

Les erreurs de cette catégorie concernent l'omission ou l'adjonction de phonèmes, la confusion de consonnes, de voyelles, ainsi que l'omission ou l'adjonction d'un accent (aigu ou grave), qui altère la valeur phonique du mot.

##### **4.1.6.2.1.1. Confusion de voyelles**

Les erreurs de cette sous-catégorie pourraient concerner la nasalisation d'une voyelle orale :

- (10) \*monquer (moquer) (L2-01-003) ; \*prenx (preux) (L2-01-020) ; \*ingnorait (ignorait) (L3-011) ; \*ignanrance (ignorance) (L4-027) ; \*chevancha (chevaucha) (L2-01-008 ; L3-006) ; \*rancontait (racontait) (L4-022) ; \*audance (audacieux) (L4-026).

La dénasalisation d'une voyelle nasale est également trouvée dans notre corpus :

- (11) \*racontre (rencontre) (L3-002) ; \*plaisaitant (plaisantant) (L3-013) ; \*loitaine (lointaine) (L3-009) ; \*recontra (rencontra) (L4-016 ; L4-019 ; L4-025 ; L4-027).

Les erreurs pourraient concerner simplement la confusion des deux voyelles orales :

- (12) [a] et [ə] : \*maledroitentant (maladroitement) (L2-01-007) ;  
 [a] et [u] : \*jouleux (jaloux) (L2-01-019) ;  
 [a] et [ɛ] : \*craignait (craignait) (L2-01-008) ; \*réclamant (réclamant)(L2-01-023) ; \*chaisa (chassa) (L3-003) ; \* réal (réel) (L4-017)  
 [i] et [e] : \*offécier (officier) (L2-01-022)  
 [o] et [ø] : \*châteux (château) (L2-02-014)  
 [ɔ] et [œ] : \*meures (morts) (L2-02-015)  
 [y] et [e] : \*bienvené (bienvenu) (L2-02-016)  
 [ɛ] et [i] : \*crignait (craignait) (L3-015 ; L3-016)  
 [ɔ] et [u] : \*trope (troupe) (L4-022)  
 [i] et [ɛ] : \*volontir (volontaire) (L4-019)  
 [ɔ] et [ə] : \* honerez (honorez) (L3-020)

#### 4.1.6.2.1.2. Confusion de consonnes

Dans cette sous-catégorie, les erreurs relèvent d'une confusion entre deux consonnes :

- (13) [r] et [s] : \*admisait (admirait) (L2-01-001) ;  
 [r] et [t] : \* rites (rires) (L3-018)

Pour les erreurs suivantes, il s'agit de la confusion de consonne sonore avec la consonne sourde correspondante.

- (14) [k] et [g] : \*blaque (blague) (L2-01-003)

Dans les formes erronées ci-dessous, le groupe consonantique a été réduit à la consonne simple :

- (15) \*aigne (aigre) (L2-02-005)  
 \*monte (montrés) (L4-008)

\* rencontré (rencontré) (L3-016 ; L4-008 ; L4-009)

L'erreur suivante concerne la confusion de la consonne nasale [ɲ] avec une autre consonne nasale [n] :

(16) \*manifiques (magnifiques) (L4-009)

L'exemple (17) ci-dessous cumule plusieurs erreurs. D'abord, le scripteur a remplacé la consonne *c* [k] par un groupe consonantique *cr* [kr]. Deuxièmement, la voyelle simple *u* [y] a été aussi remplacé par *-iet*. Ainsi, la forme produite par le scripteur est difficile à identifier sans le texte de source.

(17) \*Mais si on avait appris au chevalier l'emploi de la lance, de  
{1\_QUOTE\_écriet}\_{1\_QUOTE\_écu}. (L2-01-008)

#### 4.1.6.2.1.3. Adjonction

Cette erreur concerne l'adjonction de la lettre muette *e*, qui possède une valeur auxiliaire, en altérant la valeur phonique du mot :

(18) \*forête pour forêt (L2-01-007)

Les erreurs suivantes concernent l'adjonction d'un phonogramme supplémentaire :

(19) \* adventures (aventures) (L2-01-013)

\* deveniaire (devenir) (L2-02-012)

\* admiera (admira) (L3-009)

\* isolié (isolé) (L3-016)

\* chevalerie (chevalerie) (L4-011)

\* gentilmment (gentiment) (L4-023)

Les erreurs pourraient également concerner l'adjonction de plusieurs phonogrammes :

- (20) \* parlentait (parlait) (L2-01-023)
- \* dominant (dormant) (L2-02-015)

#### **4.1.6.2.1.4. Omission**

Pour les erreurs suivantes, il s'agit de l'oubli d'un seul phonogramme.

- (21) \* orginaire (originaire) (L2-01-017)
- \* consils (conseils) (L2-02-006)
- \* assie (assise) (L2-02-007 ; L4-011)
- \* pas (pays) (L2-02-003)
- \* émereillé (émerveillé) (L3-012)
- \* ressembler (ressemble<sup>31</sup>) (L4-026)
- \* irrié (irrité) (L4-026)

Les erreurs dans l'exemple (22) pourraient également concerner l'omission d'une syllabe, ou de plusieurs phonogrammes.

- (22) \* désperé (désespérée) (L2-01-023)
- \* honez (honorez) (L2-02-003)
- \* honnablement (honorablement<sup>32</sup>) (L2-01-016)
- \* saissait (saisissait) (L3-022)

---

<sup>31</sup> Le scripteur a omis de transcrire le groupe consonantique [bl], et n'a mis que la consonne simple [b].

<sup>32</sup> La forme erronée cumule les erreurs de deux différents types. D'abord, le scripteur a doublé la consonne *n*, qui n'altère pas la valeur phonique du mot, mais qui n'est pas retenu par la norme. Deuxièmement, le scripteur a omis de transcrire les phonèmes [ɔ] et [r] de la forme normée, qui l'a rendue ainsi aberrante.

Les formes erronées suivantes cumulent deux types d'opération en même temps. L'exemple (23) concerne la confusion des voyelles orales [a] et [o], ainsi que l'omission du phonogramme *e* [ə] :

(23) \* chvaché (chevauché) (L2-01-005)

L'erreur dans l'exemple (24) concerne la confusion des voyelles orales [ə] et [œ], ainsi que l'omission du phonogramme *i* [i] :

(24) \* chevaleur (chevalerie) (L2-02-015)

Il s'agit de la confusion des voyelles orales [a] et [ɛ], ainsi que de l'omission des phonogrammes *ier* [je] :

(25) \* chevail (chevalier) (L2-02-015)

La forme erronée dans l'exemple (26) concerne la confusion des voyelles orales [ə] et [ɛ], ainsi que l'omission du phonogramme *i* [i] :

(26) \* chevalaire (chevalerie) (L2-02-015 ; L2-02-016)

Dans l'exemple (27), le scripteur suivant a omis de mettre la lettre finale muette *e* et également remplacé la consonne finale *r* par *t* :

(27) \* contrait (contraire) (L2-01-023)

#### **4.1.6.2.1.5. Accent**

Pour les erreurs que nous allons citer ci-dessous, il s'agit de l'utilisation erronée des accents (ou leur non-utilisation) qui aboutit à des fautes phonétiques (A.Girolami-Boulinier, 1984 : 132).



Les erreurs dans l'exemple (28) concernent un accent aigu sur la lettre *e* qui, dans la forme normée, est « dépourvue d'accent et impossible d'être accentuée sous peine d'être altérée phonétiquement » (Katoozian, 2003) :

- (28) \* facé (faire face à) (L2-01-003)
- \* bonné (bonne) (L2-01-011)
- \* révétu (revêtu) (L2-01-007)
- \* imprésionné (impressionné<sup>33</sup>) (L2-02-015)
- \* régardé (regardé) (L3-004)
- \* élevé (élevé) (L4-006)
- \* intélligent (intelligent) (L4-018)

Les formes erronées dans l'exemple (29) concernent tous la non-utilisation de l'accent aigu.

- (29) \* isole (isolé) (L2-02-010)
- \* ecu (écu) (L2-01-006)
- \* reussi (réussi) (L2-02-007)
- \* realiser (réaliser) (L3-003)
- \* etait (était) (L3-010)
- \* reveilla (réveilla) (L3-013)
- \* ecoutez (écoutez) (L4-005)
- \* preparatiions (préparations) (L4-009)
- \* elevé (élevé) (L4-012)
- \* étonne (étonné) (L4-013)
- \* deménager (déménager) (L4-018)
- \* determination (détermination) (L4-026)

---

<sup>33</sup> Le scripteur a également omis de doubler la consonne *s* à l'intervocalique, et pour conséquence a introduit une modification de prononciation : de [s] à [z]. Il s'agit d'une erreur issue de la méconnaissance des lois de position.

\* élevé (élevé) (L4-028)

Les erreurs suivantes relèvent de l'omission de l'accent grave :

(30) \* tres (très) (L2-02-012)

\* mere (mère) (L2-01-018)

\* jete (jète) (L4-026)

\* leve (lève) (L4-026)

Dans l'exemple (31), le scripteur a omis de transcrire le phonème  $[\varepsilon]$  de la dernière syllabe du mot. Dû à la mémorisation insuffisante de la forme graphique du mot, le scripteur a placé un accent circonflexe sur la voyelle *o* de la première syllabe.

(31) \* Ils ont passé une nuit dans {un}\_ {une} {fôrt}\_ {forêt}. (L2-01-018)

#### **4.1.6.2.1.6. Mots inventés**

En examinant les erreurs orthographiques à dominante phonétique, nous avons trouvé un groupe de graphies erronées intéressantes, qui concernent les formes graphiques inexistantes en français et construites par les étudiants chinois. En utilisant certaines règles de dérivation du français, ils ont inventé les « mots faux » à partir d'une base existante.

Dans l'exemple (32), le scripteur a produit une forme adjectivale inexistante \* *couragé* à partir du nom *courage* :

(32) \* couragé pour courageux (L2-01-011).

Les scripteurs dans l'exemple (33) ont produit des formes nominales \* *raconte*, \* *venge*, et \* *moques*, inexistantes en français, à partir de leur base verbale, *raconter*, *venger*, et *moquer* :

- (33) \* la raconte de la cour du roi. (L4-01) ;  
\* venge (vengeance) (L4-019) ;  
\* moques (moqueries) (L4-027).

Dans l'exemple (34), le scripteur a construit une forme adjectivale \* *chevalière*, qui n'existe pas en français, à partir du nom *chevalier* :

- (34) \* l'aventure chevalière (L4-025).

Il est probable que le scripteur dans l'exemple (35) a confondu la catégorie du mot *talent*. Il l'a considéré comme adjectif, et pour conséquence créé une forme nominale inexistante à partir de cette base adjectivale :

- (35) \* talence pour talent (L2-01-003).

Nous avons trouvé la même opération dans la forme erronée suivante. Il nous semble que le scripteur a omis d'identifier la nature nominale du mot *audace* et qu'il a construit un nom inexistant \* *audaceté* à partir du mot *audace* :

- (36) \* audaceté pour audace (L4-027)

Dans l'exemple (37), le scripteur a construit une forme nominale inexistante \* *démandation* à partir du verbe *demander*. De plus, il a ajouté un accent aigu sur la lettre *e* de la première syllabe qui modifie l'image phonique du mot :

- (37) \* démandation pour demande (L3-003).

Les scripteurs dans l'exemple (38) ont inventé les formes verbales du 3<sup>ème</sup> groupe, en suivant cependant les règles de construction des verbes du 1<sup>er</sup> groupe.

- (38) \* meurer pour mourir (L2-02-013)
- \* prenner pour prendre (L2-02-014)
- \* devener pour devenir (L3-016)

Dans l'exemple (39), les scripteurs ont pris directement les formes de l'anglais :

- (39) \* expérencé pour expérimenté (L2-02-008)
- \* offer pour offrir (L2-01-003)
- \* palace pour palais (L4-022)
- \* explanation pour explication (L4-027)

#### 4.1.6.2.2. Erreurs phonogrammiques

Les erreurs de cette catégorie relèvent du domaine de l'orthographe lexicale mais le maniement des graphèmes fait par les scripteurs altère la valeur phonique des mots concernés. Le fait que les scripteurs n'ont pas été en mesure de restituer correctement l'orthographe des mots s'explique par une mauvaise maîtrise du système graphique. Plus précisément, les erreurs sont issues de la méconnaissance ou de la mémorisation insuffisante des phonogrammes, des lois de position, etc. (Katoozian, 2013).

Les formes erronées suivantes concernent le maniement des graphèmes qui altèrent la valeur phonique des mots. Il s'agit soit d'un oubli ou d'un ajout, soit d'une confusion, ou d'un déplacement de lettres voisines. Comme les mots concernés sont tous simples et s'emploient couramment par les apprenants chinois<sup>34</sup>, l'identification des mots par eux ne fait aucun doute, et les erreurs relèvent ainsi d'un lapsus lors de la production du texte.

- (40) \* das (dans) (L2-01-012)

---

<sup>34</sup> Le mot *chevalier* n'est pas couramment utilisé par les apprenants de FLE. Mais comme il est dérivé du nom cheval, et que nous avons trouvé les formes bien transcrites dans le texte du même scripteur, nous avons aussi classé la forme erronée \* *chatelier* dans cette catégorie de lapsus.

- \* ave (avec) (L2-02-014 ; L4-019)
- \* Qquand (quand) (L2-02-002)
- \* mebrassé (embrassé) (L2-02-014)
- \* saulé (salué) (L2-02-016)
- \* choes (chose) (L2-02-017)
- \* pus (puis) (L3-022)
- \* devat (devant) (L4-011)
- \* joile (jolie) (L4-018)

Les erreurs ci-dessous concernent l'omission du *e* caduc où il possède une valeur auxiliaire, introduisant ainsi une modification de la valeur phonique des mots. Plus précisément, son absence ôte la prononciation aux consonnes qui le précèdent.

- (41) \* troup (troupe) (L2-02-010 ; L3-005 ; L3-013 ; L4-009)
- \* compt (compte) (L2-02-013)
  - \* rapid (rapide) (L2-02-016)

Pour les formes erronées suivantes, il s'agit d'une erreur de doubles consonnes altérant la valeur phonique du mot :

- (42) \* vera (verra) (L3-006)
- \* plaissantant (plaisantant) (L3-014)
  - \* chaisa (chassa) (L3-003)

Dans l'exemple (42), il s'agit de la méconnaissance des lois de position pour les erreurs suivantes :

- (43) \* ca pour ça (L1-01-003)

\*innocant pou innocent<sup>35</sup> (L4-004)

\*commenca pour commença<sup>36</sup> (L4-027)

#### 4.1.6.2.3. Erreurs lexicales sans incidence phonique

Cette catégorie ressemble également les erreurs relevant de l'orthographe lexicale mais il s'agit d'erreurs qui ne modifient pas la prononciation des mots. Autrement dit, la forme produite par le scripteur est une des variantes possibles linguistiquement mais non retenue par la norme orthographique (Ducard, Honvault & Jaffré, 1995 : 61).

##### 4.1.6.2.3.1. Adjonction

Les erreurs suivantes concernent l'adjonction d'une lettre finale, qui n'est pas prononcée et n'influe pas non plus sur la prononciation du phonogramme voisin.

(44) \* rois (roi) (L2-01-001 ; L2-02-010)

\* bienvenue (bienvenu) (L2-01-008)

\* désire (désir) (L2-01-017)

\* claires (clairs) (L3-001 ; L3-009)

\* rêvetire (revêtir) (L4-002)

##### 4.1.6.2.3.2. Omission

Il s'agit d'une omission de la lettre finale du mot, qui n'introduit pas de modification de la valeur phonique du mot concerné.

(45) \* armés (armes) (L2-01-003)

\* toujours (toujours) (L2-01-003 ; L2-02-005 ; L4-017 ; L4-018 ; L4-019)

\* héro (héros) (L3-020)

---

<sup>35</sup> Le scripteur a choisi le graphème *-ant*, au lieu de *-ent*, pour transcrire la voyelle nasale [ã]. Notons que la graphème *-c* sert à transcrire la consonne [k] devant *-a*, *-o* et *-u*, tandis qu'il transcrit la consonne [s] devant *-e*, *-i*, et *-y*. Ainsi cette modification altère la prononciation du graphème voisin *-c* : de [s] à [k].

<sup>36</sup> Le scripteur a omis de placer une cédille sous la lettre *c*, altérant ainsi l'image phonique du mot *ça*. Il s'agit soit d'un simple lapsus, soit d'une méconnaissance des lois de position : le graphème *-c* se prononce [k] devant *a*, *o* et *u*.

- \* depui (depuis) (L4-019)
- \* pay (pays) (L4-028)
- \* longtemp (longtemps) (L4-028)

L'erreur pourraient également relève de l'omission du *h* muet :

(46) \* *bonneur* (bonheur<sup>37</sup>) (L2-02-005).

#### 4.1.6.2.3.3. Confusion

Les erreurs suivantes concernent le doublement des consonnes, qui n'altère pas la valeur phonique des mots concernés.

- (47) \* interressé (intéressé) (L2-01-011)
- \* équipe (équipe) (L2-01-023 ; L4-023)
  - \* marrier (marier) (L2-01-023)
  - \* guère (guerre<sup>38</sup>) (L2-01-016)
  - \* deverrait (devrait) (L2-02-011)
  - \* jeunne (jeune) (L3-005)
  - \* honeur (honneur) (L3-014 ; L4-006)
  - \* enemie (ennemi) (L4-019)
  - \* home (homme) (L4-026)
  - \* gentilmment (gentiment) (L4-023)

Les erreurs suivantes relèvent d'une confusion des graphèmes concurrents correspondant au même phonème. En connaissant déjà le moule graphique du mot, les

---

<sup>37</sup> Le scribeur a omis de transcrire la lettre muette *h*, qui est faussement étymologique, et issue de la confusion avec *heure* venue du *hora*. De plus, le scribeur a doublé la consonne *n*, probablement par analogie avec le mot *honneur*.

<sup>38</sup> Le scribeur ci-dessous a omis de mettre les doubles consonnes *-rr* du mot *guerre*, qui ne modifie pas cependant la valeur phonique du mot. De plus, le scribeur a ajouté un accent grave sur la lettre *e*, qui est, nous semble-t-il, une hypercorrection. Pour bien transcrire la voyelle [ɛ], le scribeur a utilisé le phonogramme *è* et pallié ainsi l'insécurité linguistique qu'il a ressenti pour le phonogramme *e*.

scripteurs ont proposé un autre phonogramme possible. Ainsi, les formes produites sont linguistiquement correctes mais non retenues par la norme.

- (48) \* surlendemin (surlendemain) (L2-02-002)
  - \* conseillé (conseillé<sup>39</sup>) (L2-01-023)
  - \* recommandations (recommandations) (L3-002 ; L3-009 ; L3-014 ; L3-015)
  - \* déscidé (décidé) (L4-007)
  - \* convaincre (convaincre) (L4-011)
  - \* seurs (sœurs) (L4-009)

Les erreurs suivantes concernent une confusion de lettres finales muettes :

- (49) \* jalous (jaloux<sup>40</sup>) (L2-02-012)
  - \* puit (puis) (L3-014)

#### **4.1.6.2.3.4. Accent**

Les erreurs de cette sous-catégorie concernent la plupart du temps l'omission ou l'ajout de l'accent circonflexe :

- (50) \* faché (fâché) (L2-01-006 ; L2-02-005 ; L4-002 ; L4-012 ; L4-019)
  - \* dît (dit) (L4-002)

Les erreurs pourraient également relèvent de l'omission de l'accent cédille :

- (51) \* apercut (aperçut) (L3-004)
  - \* décu (déçu) (L3-021)

---

<sup>39</sup> Il existe deux graphies possibles pour transcrire le yod issu d'un ancien *l* mouillé : *-il* (nom masculin *conseil*) et *-ill* (verbe *conseiller*). La forme erronée produite par le scripteur est contaminée par la forme nominale : *conseil*.

<sup>40</sup> Le scripteur a remplacé la lettre finale muette *x* par la lettre *s*, qui subit probablement l'effet d'analogie par les formes dérivées du mot *jaloux*, comme *jalouser*, *jalousie*...



En résumé, à la différence des erreurs phonogrammiques et des erreurs lexicales sans incidence phonique, les erreurs orthographiques à dominante phonétique sont beaucoup plus nombreuses et présente une plus grande variété. Si les erreurs phonogrammiques commises par les apprenants chinois concernent essentiellement les doubles consonnes, les lettres finales muettes, ou la méconnaissance de lois de position, problèmes qu'on trouve également dans la production des élèves français, les erreurs orthographiques à dominante phonétique représentent une grande hétérogénéité. Même si les apprenants connaissent les mots en question, ils ne pourraient pas toujours accéder à l'image phonique exacte des mots. L'errance sur la manière de transcrire les mots dont l'image phonique était inaccessible pour les apprenants lors de leur production écrite est attestée par le nombre de formes variées qu'ils suscitent dans notre corpus. Nous trouvons dans cette catégorie les mots d'une syllabe dont un son a été mal rendu, soit qu'il ait été omis, \* *assie* pour *assise*, soit qu'il ait été remplacé par un autre, \**châteux* pour *château*. Y sont également classés les mots de deux ou trois syllabes profondément altérées dans leur représentation phonétique : \* *chevalaire* pour *chevalerie*. L'importance et la persistance des erreurs orthographiques à dominante phonétique constituent une tendance orthographique interlangagière des apprenants chinois. En demandant l'ontogénèse des erreurs de ce type, nous présenterons dans la partie suivante les scénarios psycholinguistiques possibles pour interpréter la tendance orthographique des apprenants chinois. Et les facteurs divers impliquant dans l'acquisition de l'orthographe lexicale seront également discutés.

#### 4.1.7. Discussions

Pour expliquer la prédominance des erreurs orthographiques à dominante phonétique qui sont spécifiques aux apprenants du français L2, il faut remarquer d'abord que la situation d'apprentissage de l'écrit des apprenants du français L2 diffère de celle des enfants français. Si l'enfant français aborde l'écrit avec une maîtrise précoce du français oral, l'apprenant L2 apprend simultanément le français oral et écrit. Autrement dit, l'enfant

français connaît déjà la forme parlée de mots et peut simplement ignorer leur forme écrite, tandis que l'apprenant du français L2 « se trouvent devant deux inconnues : l'oral et l'écrit » (Katoozian, 2014). Catach (1982 : 35) démontre ces situations comme suit :

**Tableau 23.** Situations comparées d'un enfant francophone et d'un enfant étranger face à

l'apprentissage de l'oral et de l'écrit (*Légende : Langue 1, langue maternelle ; langue 2, langue cible*)

Niveau	Francophones	Non francophone	
Oral	Langue orale (supposée connue)	Langue orale 1 (supposée connue)	Langue orale 2 (inconnue)
Écrit	Écrit (inconnu)	Écrit 1 (connue ou inconnu)	Écrit 2 (inconnu)

#### 4.1.7.1. Surdit  phonologique

Pour les  l ves franais, ils sont en contact permanent avec la forme orale de leur langue maternelle. Ainsi, le nombre d'erreurs orthographiques   dominante phon tique diminue avec l' ge des  l ves franais. Autrement dit, en contact permanent avec l'image auditive correcte de mots, les apprenants franais ne produisent plus les erreurs phon tiques. Si elles persistent dans leur production langagiere, elles seraient dues   des facteurs divers : « fatigue, probl mes d'ordre auditif, particularit s dialectales ou socioculturelles, s quelles de retard de parole, etc. » (Girolami-Boulinier, 1984 : 131).

Neanmoins, en  tudiant le franais dans l'universit  chinoise, les apprenants chinois de notre corpus abordent simultan ment le franais oral et  crit. De plus, ils ne sont pas en contact permanent avec la forme orale de la langue franaise. Selon les recherches d'Amokrane (2011) et de Katoozian (2013, 2014), cette inaccessibilit  de l'image auditive du mot aboutit   une fausse appr ciation des phon mes de la langue cible, qui explique l'apparition des erreurs orthographiques   dominante phon tique. Pour expliquer la pr sence de cette mauvaise perception phonologique chez les apprenants L2, il existe plusieurs mod les th oriques fond s sur l'analyse du fonctionnement cognitif.