



**HAL**  
open science

# L'apprentissage de l'orthographe chez les apprenants chinois du français L2 : une analyse portant sur un corpus écrit

Yilun Li

► **To cite this version:**

Yilun Li. L'apprentissage de l'orthographe chez les apprenants chinois du français L2 : une analyse portant sur un corpus écrit. Linguistique. Université Sorbonne Paris Cité, 2018. Français. NNT : 2018USPCA071 . tel-02147421

**HAL Id: tel-02147421**

**<https://theses.hal.science/tel-02147421>**

Submitted on 4 Jun 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Sorbonne Paris Cité

**Université Sorbonne Nouvelle - Paris 3**

ED 268

*Langage et langue : description, théorisation, transmission*

Thèse de Doctorat en Sciences du langage

Présentée par **Yilun Li**

***L'apprentissage de l'orthographe  
chez les apprenants chinois du français L2 :  
une analyse portant sur un corpus écrit***

**Tome 1**

Sous la direction de Mme Claire Doquet

et de Mme Zhitang Drocourt

Soutenue le 6 septembre 2018

Composition du jury :

Mme. Claire Doquet, Professeur, Université Sorbonne Nouvelle

Mme Zhitang Drocourt, Professeur émérite, Inalco

Mme Sylvie Plane, Professeur émérite, Université Paris-Sorbonne

M. Jacques David, Maître de conférences, Université Cergy

M. Fu Rong, Professeur, Université des langues étrangères de Beijing

*L'apprentissage de l'orthographe chez les apprenants chinois du français L2 : une analyse  
portant sur un corpus écrit*

**Résumé.** Dans l'apprentissage du français écrit, la norme orthographique représente un lieu de difficulté majeure. Confronté à des complexités multiples, tout scripteur aura un long trajet à parcourir avant d'arriver à la maîtrise complète des traitements orthographiques, et ceci est sans doute d'autant plus vrai lorsqu'il s'agit de s'appropriier le système orthographique du français dans une langue étrangère. Cette étude vise donc à analyser la réalisation orthographique, chez les apprenants chinois du français L2. Pour ce faire, nous examinons les données issues d'une tâche écrite semi-spontanée, en nous attachant particulièrement au profil développemental des apprenants chinois, et leurs tendances interlangagières face aux problèmes de l'orthographe lexicale, de l'accord en nombre/genre et de la morphologie verbale. En outre, une analyse des erreurs commises par les apprenants, ainsi que les explications possibles pour interpréter ces erreurs sont également présentées, en faisant référence aux caractéristiques spécifiques de l'apprentissage de la langue étrangère.

**Mots-clés :** orthographe, apprenants chinois, français L2, orthographe lexicale, morphologie écrite, erreur

*Orthographic learning by Chinese-speaking university learners of French L2: an analysis  
based on a written corpus*

**Abstract:** In the area of written French, the orthographic learning represents a major difficulty. Faced with multiple complexities, any writer will have a long way to go before arriving at the complete control of spelling, and this is probably more true when it comes to appropriating this French system in a foreign language which is typologically distinct from the French language. This study aims to analyze the orthographic realization in French L2, for Chinese learners. To do this, we examine data from a semi-spontaneous writing task, focusing in particular on the developmental profile of learners, and their interlanguage tendencies in relation to the problems of lexical orthography, number/gender agreement, and verbal morphology. In addition, an errors analysis as well as possible explanations for interpreting these errors are also presented, with reference to the specific characteristics of L2 learning.

**Keywords:** spelling, Chinese learners, French L2, lexical orthography, written morphology, error

# Remerciements

Je tiens à remercier tout particulièrement mes deux directrices de recherche, Mme Claire Doquet et Mme Zhitang Drocourt pour leur aide précieuse, leur rigueur scientifique et surtout leur gentillesse tout au long de ces années de travail.

J'adresse de chaleureux remerciements à Mme Sylvie Plane, pour ses conseils avisés, ses critiques pertinentes et son attention de tous les instants consacrée à mes travaux depuis le Master.

Je remercie également M. Jacques David pour ses conseils et ses recommandations relatives aux travaux de recherche s'agissant de l'apprentissage de l'orthographe.

Je pense aussi à M. Fu Rong qui a accepté de faire partie de mon jury et a la gentillesse de se déplacer de la Chine pour la soutenance.

Je suis très reconnaissante envers M. Serge Fleury. Son aide précieuse sur la construction du corpus et ses conseils de l'utilisation du *Trameur* m'ont permis de disposer des outils nécessaires pour bien analyser les textes écrits.

Merci à Mme Geneviève Bouyoux, Mme Marielle de Velder et M. Jean-Remy Bure. Je tiens vraiment à leur exprimer ma gratitude pour les longues heures qu'ils ont consacrées à relire et à corriger mon écrit en français.

J'ai une dette aussi envers les enseignants de l'Institut de langues étrangères N° 2 de Pékin, en particulier Hu Yu, Qiu Yinchun, Qin Qinglin, Wang Xin et Wang Mingli, envers les étudiants et anciens étudiants dont le texte a permis ma recherche.

Pendant la thèse, j'ai été vivement soutenue par mes amis : Liang Chendi, Liu Wanting, Mao Wei, Wang Yuling, Ye Kairun, Zhang Fengjiao et Zhang Fanying.

Enfin, j'adresse un ultime remerciement à mes parents.

Je remercie chacun de tout cœur.

# Abréviation

\*(astérisque) : signe marquant la graphie erronée

CE1 : Cours élémentaire première année

CE2 : Cours élémentaire deuxième année

CM1 : Cours moyen première année

CM2 : Cours moyen deuxième année

CP : Cours préparatoire

Français L1 (FLM) : Français langue maternelle

Français L2 (FLE) : Français langue étrangère

1PS : Première personne du singulier

1PP : Première personne du pluriel

2PS : Deuxième personne du singulier

2PP : Deuxième personne du pluriel

3PS : Troisième personne du singulier

3PP : Troisième personne du pluriel

# Table des matières

REMERCIEMENTS .....	3
ABRÉVIATION.....	4
TABLE DES MATIÈRES.....	5
<b>I. INTRODUCTION .....</b>	<b>14</b>
<b>0.1. INTRODUCTION GÉNÉRALE .....</b>	<b>14</b>
<b>0.2. QUESTIONS DE RECHERCHE.....</b>	<b>15</b>
<b>0.3. PLAN DE CETTE ÉTUDE .....</b>	<b>17</b>
<b>II. PARTIE THÉORIQUE .....</b>	<b>19</b>
<b>1. L'ORTHOGRAPHE DU FRANÇAIS : UN OBJET D'APPRENTISSAGE</b>	
<b>PROBLÉMATIQUE.....</b>	<b>19</b>
1.1. <i>Aspects sociolinguistiques : les Français et leur orthographe .....</i>	<i>19</i>
1.1.1. Éléments d'histoire du système de l'orthographe française .....	20
1.1.1.1. Les débuts de l'orthographe française.....	20
1.1.1.2. Le moyen français.....	23
1.1.1.3. Les deux renaissances .....	26
1.1.1.4. Vers l'orthographe d'État .....	28
1.1.1.5. L'orthographe actuelle .....	30
1.1.2. Les Français et leurs jugements sur l'orthographe .....	31
1.1.2.1. Questions de norme.....	31
1.1.2.2. L'enseignement de l'orthographe.....	33
1.1.2.3. Les réformes de l'orthographe.....	34
1.2. <i>Aspect linguistique : les spécificités de l'orthographe française .....</i>	<i>38</i>
1.2.1. Un système mixte.....	38
1.2.1.1. Présentation générale du plurisystème.....	38
1.2.1.2. La phonographie : représenter les sons de la langue.....	39
1.2.1.2.1. Les sons du français.....	39

1.2.1.2.2.	L'écriture des phonèmes : principes généraux.....	40
1.2.1.2.3.	L'écriture des phonèmes : problèmes .....	40
1.2.1.3.	La sémiographie : donner accès au sens .....	42
1.2.1.4.	La morphographie : signaler les rapports entre lexèmes.....	43
1.2.2.	Un système qui est source de difficultés pour les scripteurs.....	44
1.2.2.1.	La polyvalence du français : un compromis entre sémiographie et phonographie .....	44
1.2.2.2.	L'opacité du français.....	45
1.2.2.3.	Le degré de régularité .....	46
<b>1.3.</b>	<b><i>Aspects psycholinguistiques : orthographier le français</i></b> .....	<b>48</b>
1.3.1.	La place de la composante orthographique dans l'activité de production de texte .....	48
1.3.1.1.	L'activité de production de texte : présentation générale.....	49
1.3.1.2.	De l'orthographe des mots isolés à l'orthographe des mots en contexte .....	50
1.3.2.	Procédures et stratégies pour orthographier .....	51
1.3.2.1.	L'accès lexical .....	52
1.3.2.2.	L'accès phonologique.....	54
1.3.2.3.	D'autres stratégies pour orthographier .....	55
1.3.2.4.	La variété des procédures individuelles .....	56
1.3.3.	Les erreurs et sources d'erreur dans l'orthographe du français.....	56
1.3.3.1.	Les types d'erreurs .....	56
1.3.3.2.	Les problèmes attentionnels.....	58
<b>2.</b>	<b>ORTHOGRAPHER LE FRANÇAIS QUAND ON EST UN LOCUTEUR DE CHINOIS</b> .....	<b>62</b>
2.1.	<i>Entrer dans une langue étrangère : trois déterminants de l'apprentissage</i> .....	<i>62</i>
2.1.1.	Le développement de l'apprenant : apprendre le français comme langue étrangère.....	62
2.1.1.1.	Interlangue .....	62
2.1.1.2.	Hypothèses générales sur les processus de l'acquisition d'une langue étrangère.....	65
2.1.2.	Le système linguistique particulier : les propriétés du système français.....	66
2.1.2.1.	Versant phonographique.....	66
2.1.2.2.	Versant sémiographique.....	67
2.1.3.	L'instruction dispensée : un enseignement explicite.....	69

2.1.3.1.	Essor du français en Chine.....	69
2.1.3.2.	Motivations de l'apprentissage.....	70
2.1.3.3.	Le contexte local et le terrain.....	71
2.1.3.3.1.	Trois références pédagogiques dans les établissements d'enseignement supérieur.....	71
2.1.3.3.2.	Besoin d'évolution.....	73
2.1.3.4.	Organisation de cours.....	74
2.1.3.5.	Manuels utilisés.....	75
2.1.3.5.1.	Manuels élaborés en Chine.....	75
2.1.3.5.2.	Nouvelle génération de manuels.....	76
2.1.3.6.	Caractéristiques de l'enseignement du français.....	77
2.1.3.6.1.	Langue d'enseignement.....	77
2.1.3.6.2.	Enseignement explicatif.....	78
2.1.3.6.3.	Entraînement intensif en phonétique et en orthographe.....	79
2.1.3.6.4.	Place à l'écrit.....	79
2.2.	<i>Le capital initial du locuteur chinois : quelques spécificités de la langue chinoise et de l'écriture chinoise.....</i>	80
2.2.1.	Aspect phonographique.....	80
2.2.2.	Aspect sémiographique.....	83
2.2.3.	Apprentissage de l'écriture chinoise.....	84
2.3.	<i>L'apprentissage de l'orthographe pour un locuteur chinois.....</i>	85
2.3.1.	Les étapes majeures pour apprendre une orthographe alphabétique.....	85
2.3.2.	Apprentissage de l'orthographe lexicale.....	87
2.3.3.	Apprentissage de la morphologie.....	88
<b>III.</b>	<b>PARTIE MÉTHODOLOGIQUE.....</b>	<b>92</b>
<b>3.</b>	<b>MÉTHODOLOGIE.....</b>	<b>92</b>
3.1.	<i>Considérations méthodologiques.....</i>	92
3.1.1.	Un dispositif pour observer les performances au cours de la production écrite.....	92
3.1.1.1.	Processus de mise en texte.....	93
3.1.1.2.	Interlangue.....	94



3.1.1.3.	Deux contraintes mises en jeu dans la production de texte .....	95
3.1.1.4.	Bilan .....	95
3.1.2.	Réflexions sur l'analyse des données en interlangue .....	96
3.1.2.1.	Analyse comparative .....	97
3.1.2.2.	Analyse d'erreurs.....	99
3.1.2.2.1.	Erreurs interlinguales <i>versus</i> erreurs intralinguales.....	100
3.1.2.2.2.	Faute <i>versus</i> erreur .....	101
3.1.2.2.3.	Perspective différente de l'analyse contrastive .....	101
3.1.2.3.	Analyse de performance et études d'interlangue.....	102
3.1.2.4.	Bilan .....	103
<b>3.2.</b>	<b><i>Corpus</i></b> .....	<b>104</b>
3.2.1.	Description de la population.....	104
3.2.2.	Réalisation d'une tâche d'écriture.....	105
3.2.2.1.	Protocole expérimental.....	105
3.2.2.2.	Texte de source.....	106
3.2.2.3.	Déroulement .....	107
3.2.2.4.	Consignes à donner .....	108
3.2.2.5.	Recueil de copies.....	108
3.2.3.	Constitution du corpus.....	108
3.2.3.1.	Procédures de transcription et d'annotation.....	108
3.2.3.2.	Préparation de bases pour <i>le Trameur</i> .....	110
<b>3.3.</b>	<b><i>Analyse des données</i></b> .....	<b>113</b>
3.3.1.	Classement d'erreurs .....	116
3.3.2.	Analyse quantitative à travers <i>le Trameur</i> .....	119
3.3.3.	Démarche d'analyse.....	129
<b>IV.</b>	<b>PARTIE EMPIRIQUE</b> .....	<b>132</b>
<b>4.</b>	<b>ANALYSE DES PERFORMANCES ORTHOGRAPHIQUES</b> .....	<b>132</b>
<b>4.1.</b>	<b><i>Orthographe et phonétique</i></b> .....	<b>132</b>
4.1.1.	Différents aspects de l'orthographe lexicale .....	133

4.1.2.	Erreurs lexicales .....	137
4.1.2.1.	Classement des erreurs lexicales.....	137
4.1.2.2.	Erreurs orthographiques à dominante phonétique .....	139
4.1.2.3.	Erreurs phonographiques.....	139
4.1.2.4.	Formes approchantes.....	140
4.1.3.	Erreurs lexicales à travers quelques recherches .....	141
4.1.3.1.	En français L1.....	141
4.1.3.2.	En français L2.....	143
4.1.4.	Système phonétique du chinois.....	146
4.1.4.1.	Structure syllabique.....	146
4.1.4.2.	Consonnes .....	148
4.1.4.3.	Voyelles.....	149
4.1.4.4.	Difficultés lors de la production orale .....	151
4.1.5.	Hypothèses et méthodes d'observation.....	153
4.1.6.	Résultats.....	154
4.1.6.1.	Le profil orthographique des apprenants chinois .....	154
4.1.6.1.1.	Erreurs orthographiques à dominante phonétique.....	154
4.1.6.1.2.	Erreurs phonogrammiques .....	156
4.1.6.1.3.	Erreurs lexicales sans incidence phonique.....	157
4.1.6.2.	Analyse des erreurs .....	159
4.1.6.2.1.	Erreurs orthographiques à dominante phonétique.....	159
4.1.6.2.1.1.	Confusion de voyelles.....	159
4.1.6.2.1.2.	Confusion de consonnes .....	160
4.1.6.2.1.3.	Adjonction .....	161
4.1.6.2.1.4.	Omission .....	162
4.1.6.2.1.5.	Accent.....	163
4.1.6.2.1.6.	Mots inventés .....	165
4.1.6.2.2.	Erreurs phonogrammiques .....	167
4.1.6.2.3.	Erreurs lexicales sans incidence phonique.....	169
4.1.6.2.3.1.	Adjonction .....	169

4.1.6.2.3.2.	Omission .....	169
4.1.6.2.3.3.	Confusion.....	170
4.1.6.2.3.4.	Accent.....	171
4.1.7.	Discussions .....	172
4.1.7.1.	Surdit� phonologique.....	173
4.1.7.1.1.	Diff�rents mod�les th�oriques .....	174
4.1.7.1.2.	Difficult�s lors de la production graphique.....	177
4.1.7.2.	D�veloppement lexical des apprenants L2 .....	181
4.1.8.	Conclusions .....	187
<b>4.2.</b>	<b><i>Orthographe et syntaxe : le probl�me des accords</i></b> .....	<b>190</b>
4.2.1.	Syntaxe de l'accord : le genre et le nombre.....	190
4.2.1.1.	Orthographe et marques morphologiques .....	191
4.2.1.2.	Acquisition d'un sous-syst�me linguistique sp�cifique � l'�crit.....	192
4.2.2.	Accord du nombre .....	193
4.2.2.1.	Questions de recherche .....	194
4.2.2.2.	Expression morphologique du nombre.....	194
4.2.2.2.1.	Distinction s�mantique et syntaxique en fran�ais .....	194
4.2.2.2.2.	Le nombre en chinois et en anglais.....	196
4.2.2.3.	L'acquisition de la morphologie �crite du nombre en Fran�ais L1 .....	197
4.2.2.3.1.	L'accord en nombre : la logique de la pluralit� vs la difficult� de l'inaudible.....	198
4.2.2.3.2.	Une perspective d�veloppementale sur l'acquisition de la morphologie �crite du nombre en fran�ais L1 .....	199
4.2.2.3.3.	Vers un mod�le d'apprentissage : apprentissage d'instances et apprentissage de r�gle	201
4.2.2.4.	La morphologie �crite du nombre nominal .....	205
4.2.2.4.1.	R�sultats.....	205
4.2.2.4.2.	Discussion.....	209
4.2.2.5.	La morphologie �crite du nombre adjectival .....	218
4.2.2.5.1.	R�sultats.....	219
4.2.2.5.2.	Le placement de l'adjectif par rapport au nom � qualifier .....	220

4.2.2.6.	La morphologie écrite du nombre verbal.....	222
4.2.2.6.1.	Résultats.....	223
4.2.2.6.2.	Discussion.....	229
4.2.2.7.	Conclusion : Acquisition de l'accord en nombre des trois catégories des mots.....	231
4.2.3.	L'accord du genre.....	234
4.2.3.1.	Le genre en français.....	235
4.2.3.2.	Le traitement et l'acquisition de la morphologie du genre en français L1.....	237
4.2.3.3.	Le traitement et l'acquisition de la morphologie du genre en français L2.....	239
4.2.3.4.	Résultats.....	239
4.2.3.4.1.	Le marquage du genre pour les déterminants.....	240
4.2.3.4.2.	Le marquage du genre pour les adjectifs.....	243
4.2.3.5.	Variation intra-individuelle : ontogénèse des erreurs de genre.....	244
4.2.3.5.1.	Erreurs concernant l'attribution du genre.....	245
4.2.3.5.2.	Erreurs concernant l'accord en genre.....	247
4.2.3.6.	Discussion.....	251
4.2.3.6.1.	Taux de réussite et limites de constituants.....	251
4.2.3.6.2.	Taux de réussite et nature du modificateur.....	254
4.2.3.6.3.	Sur-généralisation et effets de L1.....	257
4.2.3.7.	Conclusion.....	258
4.3.	<i>Orthographe et morphologie : le problème des verbes</i> .....	260
4.3.1.	La morphologie verbale du français.....	260
4.3.2.	L'acquisition de la morphologie verbale en L2.....	263
4.3.2.1.	La finitude verbale.....	264
4.3.2.2.	Les itinéraires acquisitionnels.....	267
4.3.3.	Les apprenants chinois face aux problèmes de morphologie verbale.....	270
4.3.3.1.	Erreurs issues de la morphologie verbale sans altération de la valeur phonique.....	270
4.3.3.1.1.	L'homophonie en français.....	271
4.3.3.1.1.1.	Une vue historique de la question de l'homophonie.....	271
4.3.3.1.1.2.	Comment gérer les homophones en français ?.....	272
4.3.3.1.1.3.	Acquisition de l'orthographe des homophones.....	273

4.3.3.1.2.	Acquisition de la morphologie écrite des verbes.....	275
4.3.3.1.2.1.	L'acquisition de la morphologie verbale écrite en français L1 .....	277
4.3.3.1.2.2.	L'acquisition de la morphologie verbale écrite en français L2 .....	280
4.3.3.1.3.	Les formes verbales homophones en /E/ .....	281
4.3.3.1.3.1.	Questions de recherche.....	281
4.3.3.1.3.2.	Résultats .....	282
4.3.3.1.3.3.	Analyse des erreurs .....	285
4.3.3.1.3.3.1.	Erreurs non-spécifiques .....	285
4.3.3.1.3.3.2.	Erreurs d'utilisateur du dictionnaire .....	290
4.3.3.1.3.3.3.	Erreurs de sur-généralisation .....	293
4.3.3.1.3.4.	Discussions .....	295
4.3.3.1.4.	D'autres formes verbales homophones.....	296
4.3.3.1.4.1.	Lettres finales muettes.....	296
4.3.3.1.4.2.	Accord sujet-verbe.....	303
4.3.3.1.5.	Conclusions .....	307
4.3.3.2.	Erreurs issues de la morphologie verbale altérant la valeur phonique .....	311
4.3.3.2.1.	Erreurs de substitution.....	311
4.3.3.2.2.	Erreurs d'analogie .....	316
4.3.3.3.	Conclusion et discussion .....	321
4.3.3.3.1.	Erreur de substitution .....	321
4.3.3.3.2.	Erreur de sur-généralisation.....	326
4.3.3.3.3.	Le profil développemental des apprenants de trois niveaux linguistiques différents	
	328	
4.4.	<i>Pistes didactiques</i> .....	333
4.4.1.	Erreurs orthographiques altérant la valeur phonique .....	333
4.4.1.1.	Les lacunes des manuels .....	334
4.4.1.2.	Le manque de formation des enseignants.....	336
4.4.1.3.	Pistes didactiques .....	337
4.4.2.	Erreurs orthographiques relevant de la morphologie verbale .....	339
4.4.2.1.	Oral et écrit pour l'apprentissage de conjugaisons verbales .....	339

4.4.2.2.	Classification de verbes .....	340
4.4.2.3.	Utilisation des corpus dans l'apprentissage de la morphologie verbale .....	342
<b>IV.</b>	<b>CONCLUSION.....</b>	<b>344</b>
<b>5.</b>	<b>BILAN ET PERSPECTIVES .....</b>	<b>344</b>
5.1.	<i>Le développement orthographique des apprenants chinois .....</i>	<i>345</i>
5.2.	<i>Perspectives.....</i>	<i>349</i>
	<b>BIBLIOGRAPHIE.....</b>	<b>351</b>
	<b>LISTE DES FIGURES .....</b>	<b>370</b>
	<b>LISTE DE TABLEAUX .....</b>	<b>371</b>

# I. INTRODUCTION

## 0.1. Introduction générale

Depuis quelques décennies, la question de l'orthographe s'est enrichie d'une foisonnante production de contributions critiques, d'analyses et de propositions, issue de travaux de plusieurs champs scientifiques : linguistique, sociologie et sociologie du langage, sciences cognitives en général. Il nous semble nécessaire d'attirer particulièrement l'attention sur les études basées sur la coopération du linguiste et du psycholinguiste. A partir des années 80, l'essor des travaux psycholinguistiques nous permet de savoir plus, d'une part, les particularités cognitives de l'apprentissage de la langue écrite, d'autre part, les processus différents qu'utilisent les apprenants dans le développement de leurs habiletés orthographiques. Sur le versant linguistique, le progrès le plus notable concerne la découverte de la mixité des orthographe. Celles-ci ne se réduisent plus simplement aux correspondances avec les sons du langage (syllabes ou phonèmes), et elles sont désormais considérées comme un compromis entre le principe phonographique et sémiographique.

Dans le cadre spécifique du français, les travaux linguistiques mettent en évidence l'idée que l'apprentissage du français écrit et de la norme orthographique qui le régit se distingue sur plusieurs points d'autres langues (Morin, 2006). Confronté à ces complexités multiples, tout scripteur aura un long trajet à parcourir avant d'arriver à la maîtrise complète des traitements orthographiques, et ceci est sans doute d'autant plus vrai lorsqu'il s'agit de s'approprier le système orthographique du français dans une langue étrangère. La présente étude empirique vise donc à analyser la réalisation orthographique en français L2, chez les apprenants chinois, dont la langue maternelle est issue de la famille linguistique lointaine du français.

En s'insérant entre plusieurs domaines de recherches bien établis en linguistique, l'analyse menée dans cette étude s'inscrit de façon complémentaire et novatrice sur le territoire pour deux raisons. Premièrement, elle met sur un élargissement de la

recherche de l'acquisition de l'orthographe en français L1 vers l'acquisition de ce même système graphique dans le domaine du français L2. Avec le grand essor récent de l'enseignement-apprentissage du français L2 en Chine, la recherche en linguistique, ainsi qu'en didactique de langues, s'intéresse de plus en plus « aux spécificités des apprenants chinois et du contexte chinois » (Pernet-Liu, 2012). La lacune dans le domaine de l'acquisition de l'orthographe par des Sinophones en français L2 nous a poussé à créer un corpus écrit, qui contient 92 textes produits par les apprenants qui étudient le français au sein d'une université chinoise et qui ont des niveaux linguistiques différents. L'analyse de ce corpus nous permet d'examiner la réalisation orthographique en français L2 et de mieux connaître le développement linguistique spécifique à ce groupe d'apprenants, dont la langue maternelle est typologiquement distincte de la langue française. Deuxièmement, elle offre une ouverture pour de nombreuses recherches consacrées à l'acquisition de l'oral en français L2 vers l'étude du développement de l'interlangue écrite. L'écriture n'est pas seulement un moyen de communication mais elle joue également un rôle majeur dans l'acquisition d'une langue étrangère en salle de classe (voir par exemple Wolff, 2000). Dans ce travail, nous souhaitons examiner le traitement lexical et grammatical des apprenants chinois lorsqu'ils écrivent en français.

## **0.2. Questions de recherche**

L'objectif principal de cette étude repose sur deux angles spécifiques : l'un met l'accent sur la réalisation lexicale et grammaticale au niveau orthographique ; l'autre vise à décrire le développement orthographique pour les apprenants en français L2.

Les questions principales auxquelles nous souhaitons trouver les éléments de réponse en ce qui concerne la réalisation orthographique sont donc les suivantes :

1. Quels traitements orthographiques les apprenants chinois du français L2 adoptent-ils quand ils se trouvent confrontés à une tâche d'écriture ?
2. Lors de la réalisation orthographique, quels types d'erreur les apprenants chinois



commettent-ils et quelles sont les erreurs les plus fréquentes ?

3. Comment pourrait-on interpréter ces erreurs dans le cadre de notre recherche, quels éléments linguistiques et cognitifs pourraient jouer un rôle dans la réalisation orthographique des apprenants chinois ?

En comparant les performances orthographiques des apprenants de niveaux linguistiques différents, nous souhaitons enrichir nos connaissances sur le développement linguistique des apprenants de français en Chine. La présente étude est censée contribuer à une meilleure compréhension de l'acquisition du système graphique du français. Ainsi nous souhaitons mener une réflexion autour des questions suivantes :

1. Quelles sont les performances orthographiques des apprenants chinois de niveaux différents, quel scénario développemental présentent-ils dans leurs productions orthographiques ?
2. Comment les erreurs évoluent avec l'avancement dans le système de l'enseignement du français en Chine (évolution qualitative et quantitative) ?

Tout en parallèle avec ces questions de recherche principales, nous souhaitons également trouver les ressemblances et les différences entre les résultats obtenus par la présente étude et d'autres recherches similaires (menées dans les domaines du français L1 et du français L2). Cela nous permet d'identifier les tendances interlangagières spécifiques aux apprenants chinois du français L2. En outre, en s'inscrivant de façon complémentaire dans le projet de recherche *Ecriscol*, centré sur la constitution d'un corpus des écrits produits en situation scolaire français, ce travail se veut également une contribution à la mise en œuvre des protocoles de la transcription et de l'annotation des textes écrits en français L2, ainsi qu'à la mise en place de l'analyse automatique de ces textes écrits.

### 0.3. Plan de cette étude

Ce travail se divise en quatre parties. Tout d'abord se présente le **cadre théorique** dans lequel s'inscrit notre travail. Dans le premier chapitre, le système orthographique du français sera présenté dans une perspective sociolinguistique, linguistique et psycholinguistique. En rappelant comment l'orthographe française s'est construite au fil des siècles, la partie sociolinguistique cherche à expliquer « dans quelle mesure la complexité orthographique du français est en grande partie due à son histoire un peu particulière » (Jaffré, 2009). Dans la partie linguistique, nous présenterons les caractéristiques et le fonctionnement du système orthographique du français, afin de comprendre les raisons de ses difficultés. Nous décrirons de façon sommaire le processus du traitement orthographique mené par des scripteurs de français dans la partie psycholinguistique.

Le chapitre 2 mettra l'accent sur le point de départ de ce travail : comment orthographier le français quand on est locuteur de chinois. Nous présenterons en premier lieu les trois déterminants de l'apprentissage d'une langue étrangère, à savoir, le développement interlangagier de l'apprenant L2, les particularités de la langue française, ainsi que l'enseignement du français dispensé en Chine. Les caractéristiques du chinois moderne et de son écriture seront soulevées dans la partie suivante, afin de faciliter de notre analyse du comportement interlangagier des apprenants chinois de notre corpus. La partie théorique se terminera sur la mise en place des hypothèses : quels problèmes pourraient se poser aux apprenants chinois lors de leur production orthographique.

La **partie méthodologique** sera consacrée à la présentation des choix méthodologiques qui ont été faits tout au long de notre étude. Nous y décrirons en détail le recueil des données, la construction du corpus, ainsi que des outils d'analyse et la typologie des erreurs dont nous nous sommes servi, dans le but de mieux éclairer les conditions dans lesquelles a été mené ce travail sur l'analyse de la réalisation orthographique des apprenants chinois en français L2.

Dans la **partie empirique**, nous présenterons les principaux résultats obtenus lors

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.

Dans la partie de la **conclusion**, nous traiterons les éléments principaux de notre analyse, ainsi que les points que nous souhaitons développer dans nos prochaines études.

Cette thèse est suivie d'un tome « **Annexe** », qui est consacré à la présentation des scans des textes produits par les apprenants chinois lors du recueil de données et de la transcription de ces textes écrits dans la version annotée.

## **II. PARTIE THÉORIQUE**

### **1. L'ORTHOGRAPHE DU FRANÇAIS : UN OBJET D'APPRENTISSAGE PROBLÉMATIQUE**

Ce chapitre a pour objectif d'exposer le système orthographique de la langue française et d'expliquer les raisons de ses difficultés. A travers un aperçu de l'évolution de l'orthographe française au fil des siècles, nous souhaitons montrer « dans quelle mesure que la complexité orthographique du français est en grande partie due à son histoire un peu particulière » (Jaffré, 2009). Ensuite, grâce à la description de la mixité du système orthographique, tissé selon le principe phonographique et le principe sémiographique, ce chapitre tente de montrer le fonctionnement de ces deux principes dans le système graphique du français. Ces deux principes sont par ailleurs à « l'origine d'une certaine opacité et donc complexité de l'orthographe française » (Willemin, 2009). Le système graphique du français se caractérise alors par « des zones de fragilité orthographique » (Pérez & Garcia-Debanc, 2016). Donc dans l'étape suivante, nous présenterons comment orthographient les apprenants de français et aborderons la question des erreurs orthographiques.

#### **1.1. Aspects sociolinguistiques : les Français et leur orthographe**

Si l'orthographe française est « motivée par des structures linguistiques originales », elle fait également « l'objet de tractations, de débats, donc l'histoire conserve des traces » (Jaffré & Brissaud, 2006). Donc il n'est pas inutile de mettre l'évolution de l'orthographe française en perspective dans la présente étude et comme Jaffré l'évoque, « la cause de la complexité orthographique du français est d'abord à rechercher dans l'histoire elle-même » (2009), une suite de décisions contradictoires et de conflits entre « Anciens » et « Modernes », entre les défenseurs « d'une orthographe au service de la

lisibilité » et ceux « d'une écriture proche de la parole » (Willemin, 2009).

### 1.1.1. Éléments d'histoire du système de l'orthographe française

#### 1.1.1.1. Les débuts de l'orthographe française

L'histoire de l'orthographe du français est étroitement associée à l'histoire de sa langue. Comme Jaffré l'indique (2009), la langue française « n'est pas un long fleuve tranquille » et on ne pourrait pas dire tout simplement que cette langue dérive du latin classique. Dès la fin du II<sup>e</sup> siècle av. J.-C., les Romains s'emparent progressivement de la Gaule, couvrant les territoires actuels de la France. A la suite de la conquête de la Gaule, la République puis l'Empire romain modifient substantiellement les peuples conquis : ils ignorent les langues celtiques de ces peuples et font en sorte que le latin devienne indispensable. Au tournant du V<sup>e</sup> siècle, l'unilinguisme latin est *de facto* atteint. Parallèlement au latin classique réservé à l'administration, à l'aristocratie et aux écoles, s'est développé un latin populaire, parlé à l'origine par les soldats romains. Ce sont de ces formes langagières populaires qu'émergeront les langues romanes.

Dès la période romaine, dans les espaces ruraux existe déjà une césure omniprésente entre le latin et les parlers locaux, situation trouvant son origine dans la diversité des substrats en présence. Cette distinction est bien plus forte et ancrée que dans d'autres pays de langue romane, contribuant de ce fait à isoler les langues gallo-romaines des autres langues romanes.

A partir du III<sup>e</sup> siècle, la situation linguistique tend à se complexifier dans le sillage des invasions germaniques et de l'effondrement progressif de l'Empire romain d'Occident. Le latin n'est dès lors plus utilisé que pour les écrits et au latin populaire est substitué de proche en proche une langue nouvelle : le roman, comprenant diverses variétés. En envahissant le nord de la Gaule, les Francs s'approprient le gallo-roman à leur façon : la germanisation des parlers romans. Par l'ajout d'éléments germaniques, de nombreux Francs adoptent la langue des vaincus. La population parle à l'époque « *rustica romana lingua* » (langue romane rustique), sujette à des variations importantes

selon les régions. Le latin classique, en revanche, « n'en reste pas moins la langue de la loi, de l'administration, de l'Église, et plus généralement de l'élite sociale » (Fayol & Jaffré, 2008 : 104). Toute la Gaule connaît ainsi une situation de diglossie pendant une longue période. En outre différencié des autres langues romanes, dérivant d'un dialecte dominant, ce gallo-roman est « une forme composite faite de divers dialectes – normand, picard, champenois, etc. – que le XIX<sup>e</sup> siècle nommera francien » (Fayol & Jaffré, 2008 : 104). Appréhendée en tant que langue véhicule de communication, le gallo-roman s'étend bientôt en pays d'oïl, puis en pays d'oc.

Selon Catach (2003 : 9), « le latin et le français ont vécu durant de nombreux siècles en état de véritable symbiose, ils étaient sentis comme une seule et même langue : on lisait le latin à la française, on écrivait le français à la latine ». Les *Serments de Strasbourg*, premiers documents de l'histoire du français, nous permettent de comprendre que les scribes « tentent d'appliquer l'orthographe latine à la langue vulgaire, à l'aide d'un alphabet déjà insuffisant pour le latin classique » (Catach, 2001 : 37). Cette symbiose d'une importance capitale caractérise ainsi l'évolution graphique du français. Catach note à ce titre les efforts d'adaptation phonétique et les hésitations du scribe dans les *Serments de Strasbourg* : *c* a deux valeurs selon la voyelle qui suit (*commun, cist*) ; *z* est en relation étymologique avec *e* dans *fazet* < *faciat*. De même, F. Brunot nous montre la sonorisation des explosives *p* et *e* (*savir, podir, auidha, cadhuna*), cependant, le *d* final étymologique dans *quid*, le *ch* dans *Christian*, le double *mm* de *commun*, les deux *l* de *nulla* sont conservés. Le mélange caractéristique d'orthographe étymologique et d'orthographe phonétique est déjà constaté. Au moins, avec quelques tâtonnements, les scribes s'efforcent d'indiquer la prononciation.

Dans cette longue histoire de la naissance de la langue française, on assiste à une réorganisation en profondeur de la structure syllabique. Avec les tendances telles que l'ouverture vocalique devant la consonne, la fermeture en syllabe ouverte, la prosodie du latin recule et s'estompe, ce qu'illustre notamment le déplacement de l'accent sur la dernière syllabe prononcée. Comme l'évolution habituelle est la conservation de la syllabe accentuée et la chute des voyelles finales autre que le *a*, on constate un affaiblissement important des consonnes ou des groupes de phonèmes internes,

aboutissant à un raccourcissement considérable du mot français. En ce qui concerne le placement ignoré ou erroné de l'accent tonique, il est particulièrement important de noter les effets produits par l'influence germanique. En raison des contacts fréquents avec les langues germaniques, le francique notamment, le gallo-romain se modifie davantage que les langues des autres espaces recomposés de l'ancien Empire romain. L'Italie ou l'Espagne connaissent ainsi des influences moindres. Avec leur accent nordique et leur système phonétique, les Francs assimilent de nouvelles tendances linguistiques au gallo-romain ; autrement dit, ils prononcent les voyelles romanes beaucoup plus fortement, ce qui a une incidence non négligeable sur l'orthographe française. De ce strict point de vue phonétique, la cohabitation linguistique du francique et du roman entraîne le déplacement de l'accent tonique, modifiant le timbre des voyelles. De ce fait, « la plupart des mots terminés par une voyelle audible dans d'autres langues romanes le sont en français par un *e* dit muet » (Fayol & Jaffré, 2008 : 105). De même, les consonnes intervocaliques ne sont pas intactes, ce qui contribue à l'usure généralisée de la structure phonétique des mots. Une multiplication des formes homophones est donc imputable à ces phénomènes. Signalons aussi que l'influence germanique est également assez importante sur le plan morphologique : les finales *-and*, *-ard*, *-aud*, *-ais*, *-er* et *-ier* sont d'origine francique, sans oublier un grand nombre de verbes en *-ir* comme *choisir*, *blanchir*.

Tous ces faits de langue constituent des illustrations éloquentes de la germanisation considérable à l'œuvre, dans le sens où le gallo-romain se différencie par bien des aspects des autres langues issues du latin, notamment au sud où les langues occitanes (ainsi que l'italien, l'espagnol, le portugais, etc.) restent plus proches de la langue latine que la langue française. Ces différenciations progressives ne seront pas sans entraîner, bien entendu, des changements majeurs sur l'orthographe de la langue française.

A partir des *Serments de Strasbourg* de 842 et jusqu'au XII<sup>e</sup> siècle, on assiste à la naissance d'une littérature en langue vulgaire et d'une transcription proprement française.

Remarquons également qu'il existe une concentration de la diffusion des écritures :

elles sont aux mains des clercs lettrés. Ces écritures, créées par cette classe peu nombreuse, sont la belle minuscule caroline. Ce qui est particulièrement frappant, c'est qu'une langue parfaitement maîtrisée et stabilisée est utilisée dans les écritures, en dépit de la multiplication des dialectes. Catach (2001 : 47) nous présente quelques caractéristiques de cette langue :

- « La formation de nombreuses diphtongues (*ai, ue, ui, eu, ou*, etc.) et triptongues (*ieu, uou*, etc.) ».
- « Des phénomènes morphologiques originaux, comme les marques de séparation systématiques du masculin et du féminin (*mort/morte*) ».
- « La formation de nouveaux modes, temps et flexions verbales : le subjonctif, le futur et le conditionnel ».
- « De nouveaux paradigmes d'articles (*li, la, les, lo*), de pronoms et de mots-outils ».

Catach signale également les défauts de cette orthographe. Par exemple, elle reste dépendante de l'étymologie et a du mal à noter les voyelles. Le système vocalique de l'ancien français, surtout les diphtongues et les triptongues, évolue rapidement. Jusqu'à la fin du XIII<sup>e</sup> siècle, l'orthographe s'efforce de suivre ces changements, puis elle se fige. Comme les transformations ne cessent jamais, on assiste à un écart de plus en plus important entre l'écrit et l'oral. En dépit de tous ces défauts, ces écritures constituent les fondations de l'orthographe d'aujourd'hui.

### **1.1.1.2. Le moyen français**

Pendant la période gothique (XIII<sup>e</sup>-XVI<sup>e</sup> siècle), on assiste à une transition entre la langue ancienne et le français moderne. Les anciennes structures disparaissent alors que les nouvelles évolutions linguistiques se développent, caractérisant ce que l'on appelle le moyen français : affaiblissement de l'ancien système nominal et verbal, transformation de la syntaxe, bouleversement de l'ordre des mots, enrichissement considérable du vocabulaire, et surtout, évolution phonétique rapide. Cette instabilité



phonique, qui est propre au langage français, éloigne plus que jamais le français du latin et des autres langues romanes. Ces multiples transformations, d'une importance capitale, ont une influence sur le français (Catach, 2003 : 14-15) caractérisée par :

- « Un raccourcissement considérable du mot français » : caractère qui est particulièrement important, non seulement par le fait qu'il s'agit généralement des mots les plus utilisés et les plus utiles à la langue, mais aussi parce qu'il crée des risques nombreux d'ambiguïté par la multiplication des homophones.
- « Un relâchement des liens entre le français et le latin ».
- « Une tendance à la coupure entre radicaux et dérivés » : les premiers subissent une usure phonétique intense, alors que les seconds sont issus plus ou moins directement des radicaux latins.

Enfin, il nous paraît essentiel de rappeler les principales tendances de la graphie de l'époque (Catach, 2001 : 78-79) :

- Tendance phonétique, d'une façon générale, dans laquelle la graphie de la quasi-totalité des orthographe, si arbitraire qu'elle apparaisse, ne s'éloigne pas de la langue parlée.
- Tendance étymologique, historique, morphologique et sémantique, qui joue un rôle particulièrement important pour la distinction des mots homophones et le problème des monosyllabes.

Après ce rappel des faits linguistiques concernant l'orthographe du XIII<sup>e</sup> au XVI<sup>e</sup> siècle, nous aborderons l'étude d'un grand tournant de l'orthographe du français.

A partir du XII<sup>e</sup> siècle, l'écriture se transforme dans une direction nouvelle, vers la gothique, dans un contexte où s'épanouit la nouvelle culture de l'écrit : avec la laïcisation des clercs, ce ne sont plus seulement des scribes de monastères qui transcrivent les manuscrits, en peu d'exemplaires, mais de nouveaux clercs, qui passent rapidement de quelques centaines à plusieurs milliers ; en France, grâce à l'utilisation

du papier de chiffon, ce qui diminue considérablement les coûts, le commerce des livres devient florissant ; avec la création des universités, la demande des textes s'accroît de jour en jour.

Du XIV<sup>e</sup> au XVI<sup>e</sup> siècle, l'orthographe « s'éloigne du phonétisme pour prendre un aspect plus idéographique » (Catach, 2003 : 21). Ce premier caractère de l'orthographe du moyen français s'explique par la contradiction de l'écriture : d'un côté, par le besoin de rapidité de l'écriture, les lettres et les mots se rejoignent ; on voit également la multiplication des ligatures, des accolements, des abréviations. De l'autre côté, par l'impératif de lisibilité, on allonge les boucles et les jambages, on emploie de nombreux signes et repères afin de mieux distinguer les traces semblables, les syllabes et les mots différents. Parmi tous ces signes et repères, il faut noter particulièrement les variantes calligraphiques de positions (Catach, 2003 : 21) : le *s* rond final et le *S* long initial ou médian ; le *i* accentué puis pointé et le *I* long ou *J*, notant, suivant leur position, *i*, *y* et *j* actuels ; le *V* initial, notant indifféremment *u* et *v*, le *u* à l'intérieur, notant également *u* et *v*, etc. De ce fait, une conséquence importante se produit : « on prend l'habitude de déchiffrer chaque lettre d'après son contexte et sa position dans le mot » (Catach, 2003 : 21). On se sert d'une série de consonnes comme repère diacritique, en puisant tout naturellement dans le latin. Ainsi, l'étymologie n'est plus une fin, mais un moyen, qui apporte non seulement l'habitude des rapprochements analogiques, mais aussi de divers procédés graphiques permettant de distinguer le sens des mots.

Aux XIV<sup>e</sup> – XV<sup>e</sup> siècles, une grande différence sur le plan littéraire se présente entre la France et d'autres pays romans : la France est « entièrement soumise au bilinguisme culturel », alors que l'Italie, par exemple, « possède déjà une littérature reconnue, qui sert très tôt de modèle » (Catach, 2003 : 23). Les copistes italiens suivent ce modèle, et les grammairiens s'appuient sur ces traditions populaires littéraires et linguistiques. Étant donné la présence d'une littérature nationale, l'évolution graphique de l'italien est moins rapide et plus régulière que le français et il ne subit pas non plus les modifications profondes que nous venons d'étudier. Dans la période suivante, avec l'apparition de l'imprimerie, et d'une véritable littérature d'ampleur nationale, de considérables transformations vont être apportées à l'orthographe française.

### 1.1.1.3. Les deux renaissances

Avec l'effondrement des structures de l'ancien français, le français du XVI<sup>e</sup> siècle connaît une évolution originale et rapide sous tous ses aspects : phonétique, morphologique, syntaxique et lexicale. N'étant pas en rupture avec le système de son ancêtre, cette langue évolue progressivement vers le français actuel.

A cette époque, centrée sur un retour direct aux textes sacrés, la religion réformée nécessite une lecture et une écriture accessibles et donc un usage des langues vulgaires. Avec le soutien du monde de l'imprimerie et en appui sur le français, les publications dans tous les domaines ont investi dans ce domaine ; on voit ainsi une première vague d'imprimés en langue française. De 1533 à 1564, la production annuelle des livres en français passe de 42 à 527 (Febvre & Martin, 1973 : 442-443). Parallèlement, une lutte pour la fixation du français écrit a lieu dans le monde de l'imprimerie. Dès le premier quart du XVI<sup>e</sup> siècle, les imprimeurs français renoncent aux caractères gothiques et adoptent des caractères romains.

Cependant, les progrès sont aussi difficiles que rapides. Comme au XVI<sup>e</sup> siècle, l'écriture de manuscrites reste fidèle aux habitudes gothiques et devient plus rapide, plus mauvaise et plus illisible ; le décalage entre les manuscrits et les imprimés ne cesse de grandir. Après quelques tâtonnements, les imprimeurs « renoncent à la copie textuelle des manuscrits, aux excès et aux signes de dégénérescence de l'écriture manuscrite » (Catach, 2003 : 26). Avec cette modernisation des impressions, émerge une orthographe moderne :

« Pour les imprimeurs d'avant-garde, l'orthographe moderne signifie avant tout la modernisation des impressions : caractères romains et italiques, bonne séparation des mots, usage modéré des majuscules, ponctuations plus riches, meilleure correction matérielle, normalisation des usages, suppression des abréviations et variantes arbitraires, mise en place des accents, etc. » (Catach, 2001 :119).

Ces réformes suivent et transforment l'aspect des textes en moins de vingt ans. Ainsi, un accord de normalisation des textes se produit progressivement entre éditeurs. Ce sont d'abord des imprimeurs d'importance secondaire publiant les textes poétiques qui suivent la nouvelle orthographe. Comme les poètes attachent de l'importance à l'aspect extérieur de leurs œuvres et à la facilité de lecture et de prononciation, ils s'adressent aux imprimeurs d'avant-garde, qui acceptent leurs exigences. Ils engagent ensuite les grands imprimeurs dans la voie des réformes. Ces derniers acceptent aussi certaines publications poétiques écrites dans l'orthographe nouvelle. Ainsi, sous l'influence des typographes et des petits ateliers, les grandes maisons s'acheminent vers « l'unité nouvelle de correction et des règles typographiques » (Catach, 2001 : 121).

Ce qui caractérise l'orthographe de la Renaissance, c'est tout d'abord « la beauté des impressions, le choix des caractères et le soin nouveau apporté à la correction » (Catach, 1968 : 250). En reprenant contact avec l'orthographe des chansons de geste, les poètes suppriment beaucoup de lettres adscrites et emploient les accents. Ils cherchent à éliminer également les lettres grecques et doubles, l'emploi de *y* pour *i*, du *z* marquant *e* fermé, du *x* muet final. Ils remplacent *en* prononçant [ã] par *an* dans de nombreux cas. Notons que ces réformes seront reprises par l'Académie au XVIII<sup>e</sup> siècle, et qu'elles sont dans la lignée de la véritable tradition du français.

A la fin du XVI<sup>e</sup> siècle, malgré les efforts des typographes d'avant-garde, les imprimeurs du roi reprennent les choses en main. De ce fait, le monde de l'édition connaît une crise. La qualité des livres subit une régression si importante que « les erreurs, les coquilles, l'ignorance des règles les plus élémentaires vont de pair avec le retour à l'orthographe la plus archaïque » (Catach, 2003 : 28). Il faudra attendre jusqu'en 1640, fondation de l'Imprimerie royale par Richelieu, pour que le monde de l'édition en France voie la fin de la crise.

A partir de la deuxième partie du XVII<sup>e</sup> siècle, on assiste à une amélioration des éditions ainsi qu'au retour des habitudes graphiques de la Renaissance chez les écrivains. Les deux traditions, « modernisme » et « conservatisme » s'opposent et s'équilibrent. Notons que les dictionnaires sont aussi partagés : en 1680, Richelet publie le premier véritable dictionnaire français, dans lequel les marques morphologiques

essentielles sont gardées, de même que les tendances analogiques et étymologiques, mais elles le sont de façon plus modérée. Cependant, Richelet simplifie les consonnes doubles, et supprime les lettres adscrites et grecques. Il emploie largement les accents, aigu, grave et circonflexe. En recueillant l'héritage des écrivains et des réformateurs, les innovations de Richelet sont en accord avec son temps et son milieu.

De nombreux auteurs et éditeurs s'intéressent aux réformes et jouent un rôle non négligeable dans ce domaine. Prenons Corneille comme un exemple : en 1663, il se déclare en faveur de l'utilisation de l'accent grave, ainsi que les lettres *j* et *v*, et préconise l'emploi du *s* au lieu de *z* en tant que signe de pluriel.

Sur le terrain de l'orthographe du XVII<sup>e</sup> siècle, la fixation devient au fur et à mesure moins aléatoire, et ainsi, le français « achève de devenir elle-même » (Catach, 2003 : 29) : les diphtongues disparaissent pratiquement et les consonnes finales ne se prononcent plus en général. Le français continue à s'ordonner et à se répandre.

#### **1.1.1.4. Vers l'orthographe d'État**

C'est finalement l'Académie française, fondée en 1635, qui est chargée de normaliser l'orthographe française par l'élaboration de son dictionnaire qui fixe l'usage du français. Les premiers académiciens, plutôt conservateurs, préconisent une orthographe plus étymologisante. La prise de position de l'Académie rend vaines toutes les options radicalement phonographiques. En 1694, la première édition de Dictionnaire paraît, près de soixante ans après la création de l'Académie française. Le Dictionnaire souffre d'une trop longue élaboration, alors que l'orthographe du français se trouve dans une époque où tous les aspects de la langue évoluent rapidement. Avec un demi-siècle de stagnation, les décisions prises en 1635, et en 1639, pour la première édition de Dictionnaire, ne sont plus conformes à l'usage de la fin du siècle. « L'Académie a un demi-siècle de retard, comme l'indique Catach, et ce retard se répercutera à chacune des éditions suivantes, jusqu'au XIX<sup>e</sup> siècle » (2001 : 169) ; le fond du problème est que la langue, qui est fixée par l'Académie, devient « classique ».

Cependant, il ne faut pas croire, comme le laisse entendre Beaulieu, qu'après un

siècle passé, l'Académie s'en tienne encore à l'orthographe de R. Estienne. Notons les changements étonnants, si l'on compare la première édition de Dictionnaire à celui de R. Estienne ou à celui de Nicot : elle fait tout d'abord la distinction entre *j* et *v*, peu systématique, mais qui permettra « la chute de toute une série de consonnes ou de signes chargés de distinguer les deux graphèmes manquants » (Catach, 2003 : 33) ; elle élimine un grand nombre de consonnes plus ou moins « étymologiques », à la finale surtout ; elle supprime aussi quantité d'anciens hiatus ; elle n'emploie pas l'accent intérieur, et conserve partout le *s* muet, mais elle écrit *baume* (pour *bausme*), *couture* (pour *cousture*), etc. ; le plus frappant, c'est qu'elle remplace *en* par *an* dans beaucoup de mots .

La deuxième édition, également sous l'autorité de Régnier-Desmarais, se situe dans la lignée de la première. De ce fait, le fossé, déjà grand, entre l'Académie et l'opinion continue à s'élargir.

Il faudra attendre la troisième édition pour que l'Académie française fasse preuve d'un esprit novateur : avec l'entrée des philosophes à l'Académie, des réformes de l'orthographe se mettent systématiquement en place (Catach, 2003 : 37) :

- En 1740, presque tous les anciens hiatus ainsi que les consonnes diacritiques internes sont supprimés ; un nouveau système compensatoire d'accentuation est mis en place ; l'usage du yod et de l'orthographe des participes est réglementé.

- En 1762, avec la consolidation des premières réformes, l'emploi de l'accent grave est mis en place ; notons également la restriction et l'ordonnancement de l'accent circonflexe ; les lettres grecques sont simplifiées et les signes du pluriel sont réglementés.

- En 1798, la simplification des lettres grecques continue et de nombreuses variantes concernant les lettres doubles sont introduites.

A l'aide de cette série de réformes réfléchies, réalisée par d'Olivet, la troisième édition de l'Académie instaure en France une orthographe moderne, qui « se libère de l'influence latine, ou plus exactement qui opte pour une position intermédiaire, en

mêlant de façon parfois hétérogène, tradition et modernité » (Fayol & Jaffré, 2008 : 111). C'est cette orthographe qui devient l'orthographe actuelle. En 1798, la cinquième édition voit le jour, marquant ainsi la fin des initiatives importantes.

#### **1.1.1.5. L'orthographe actuelle**

La seconde moitié du XIX<sup>e</sup> siècle est marquée par ces nombreuses disputes sur l'orthographe. Les deux courants s'opposent et la querelle prend de l'ampleur. C'est à cette époque que l'État commence à s'intéresser à la question de l'orthographe et publie en 1900 un arrêté du Journal officiel, suivi un an plus tard de la fameuse Circulaire qui propose une simplification de la syntaxe française. Cependant, malgré tous les efforts, aucune décision ferme n'est prise. Il faudra attendre jusqu'en 1952 pour qu'un projet significatif de réforme voie le jour : Aristide Beslais, nommé président d'une Commission ministérielle d'études orthographiques, publie le premier rapport, mais échoue très vite. En 1960, la deuxième Commission Beslais se crée. En dressant un « catalogue plutôt complet des points noirs de notre orthographe qui sont régulièrement repris, en totalité ou en partie, par les projets de simplification » (Fayol & Jaffré, 2008 : 113), cette commission essaie de concilier l'efficacité, la modération et la simplicité de l'orthographe. En 1990, afin d'introduire de la cohérence dans l'orthographe française, les Rectifications orthographiques sont publiées au Journal officiel. Cette fois, le projet est bien accueilli, à la fois par l'Académie et le Conseil supérieur de la langue française. A la fin de l'année 1990, malheureusement, une nouvelle vague de critiques se déclenche, dans le but de discréditer ces Rectifications.

Finalement, toutes ces réformes tentées depuis le XIX<sup>e</sup> siècle ont régulièrement échoué, alors que l'orthographe des autres langues européennes, comme l'espagnol, l'allemand et l'italien, est judicieusement simplifiée. Par ailleurs, la démocratisation des usages, et l'avènement de l'école obligatoire mettent en évidence le fait que l'orthographe n'est plus capable à répondre aux besoins d'une société multiforme et étendue. L'orthographe du français, comme l'indiquent Fayol et Jaffré, « aux mains d'une élite masculine, s'est avérée difficile sinon impossible à réformer » (2008 : 114).

### 1.1.2. Les Français et leurs jugements sur l'orthographe

En rappelant comment s'est construite l'orthographe du français au fil des siècles, nous souhaitons mettre en évidence le fait que ce système graphique se trouve dans « une position médiane, en mêlant de façon parfois hétérogène tradition et modernité » (Jaffré, 2009). Les travaux linguistiques et psycholinguistique d'aujourd'hui nous permettent de considérer que l'orthographe française est l'un des systèmes graphiques les plus complexes à apprendre et à utiliser, question que nous aborderons dans la partie suivante. Néanmoins, nous vivons dans un contexte social où la demande orthographique explose. Par conséquent, la question de la simplification de l'orthographe française se pose avec acuité.

Aujourd'hui, l'orthographe française et son apprentissage suscitent tant de débats et différents chercheurs traitant l'orthographe semblent soutenir l'idée de la réforme (voir par exemple Manesse & Cogis, 2007 ; Fayol & Jaffré, 2008) : il est urgent de « donner aux citoyens une orthographe plus accessible et mieux adaptée à leurs besoins réels » (Jaffré, 2009). Dans la présente étude, nous ne résoudrons pas les problèmes posés par la réforme orthographique, mais nous souhaitons profiter des réflexions que ce débat suscite également au niveau de l'enseignement-apprentissage de l'orthographe française, à savoir, comment enseigner l'orthographe du français dans les milieux scolaires et comment réagissent les apprenants de français, confrontés à ce système graphique en quelque sorte opaque. Tout ceci est dans le but de pouvoir aborder ensuite, dans le chapitre suivant, la question de l'apprentissage de l'orthographe.

#### 1.1.2.1. Questions de norme

Dans l'évolution historique de l'orthographe du français, on assiste à un débat toujours agité, sur la crise qui traverse la langue française : c'est la question de la norme du français. Les conceptions de la norme du français connaissent quatre périodes, depuis que son existence est reconnue officiellement par l'ordonnance de Villers-Cotterets en 1539.



La première période est définie, à la fois « par la conception très religieuse que l'on avait d'une langue (reflet divin de la pensée) et par le souvenir de la rude bataille qui est nécessaire à la reconnaissance du français, langue populaire contre le latin ». De ce fait, le français, langue nouvellement victorieuse, est pure, sacrée et intangible. La norme dans cette période est unique, et les différences sont des impuretés à rejeter. Cette opinion est affirmée par les grammairiens des XVI<sup>e</sup> et XVII<sup>e</sup> siècles (Malherbe, Vaugelas, etc.). La maîtrise de la langue, c'est connaître le bon usage de la langue, qui est une notion quasi sacrée, à respecter inconditionnellement.

Avec l'arrivée des travaux sur la linguistique, les conceptions de la norme du français sont un peu plus ouvertes et plus conformes à la réalité du fonctionnement social. Dès les années 50, l'idée de l'existence des variations et de la pluralité des normes apparaît. Elle est toujours liée à la notion de hiérarchie de ces normes. Autrement dit, bien que des variétés des normes existent, elles ne sont pas d'égale valeur : il existe trois niveaux de langue : bas (ou familier), courant et soutenu (littéraire, le modèle). A ce titre, l'apprentissage de la langue a pour objectif de faire connaître les niveaux de langue et faire utiliser le meilleur en toute circonstance aux locuteurs.

A partir des années 70, une pédagogie pluri-normative est adaptée dans le milieu scolaire. Selon elle, les variétés sont d'égale valeur et sont classées dans des registres, qui dépendent des situations de communication. Les locuteurs doivent connaître divers registres de langage et choisir entre eux en fonction de la situation de communication. En réalité, cette pédagogie n'a que peu de sens, puisque les choix langagiers dépendent non seulement des données objectives (statut social du destinataire, lieu de la communication, etc.), mais aussi du projet de communication, de ses enjeux et de l'effet que le locuteur veut produire.

Actuellement, on parle de pédagogie pluri-normaliste : c'est-à-dire, l'apprentissage de la langue consiste à prendre conscience des enjeux véritables d'une situation de communication, et à associer les choix langagiers au projet de communication, à l'effet à produire et au résultat attendu.

### **1.1.2.2. L'enseignement de l'orthographe**

Avec une complexité indiscutable, la maîtrise de l'orthographe en France est un signe de distinction : une bonne orthographe « incarne le respect du bien commun, à l'égal de la politesse, du bon goût ou d'autres qualités du vivre-ensemble » (Manesse & Cogis, 2007 : 23). En revanche, son ignorance est liée aux jugements bien au-delà de l'orthographe à proprement parler. Une personne qui a une mauvaise orthographe « court le risque de n'être pas estimée fiable » (Manesse & Cogis, 2007 : 23).

En plus de l'intérêt qu'elle suscite dans l'opinion publique, l'orthographe est aussi l'affaire du milieu scolaire. Sans doute est-elle le domaine le plus emblématique de l'école et après un siècle et demi d'école obligatoire, l'enseignement de l'orthographe évolue progressivement. A la fin du XXI<sup>e</sup> siècle, les inspirateurs de la rénovation pédagogique lancent publiquement une campagne de réforme de l'enseignement du français. De ce fait, les priorités pédagogiques sont désormais inversées. L'orthographe et la grammaire sont invitées à passer au second rang de l'enseignement, et l'accent est mis sur la compréhension des textes, car il s'agit de « fournir le jugement, le raisonnement, le cœur et même le goût de l'élève » (Chervel & Manesse, 1989 : 144).

Cependant, l'enseignement de l'orthographe et de la grammaire « ne disparaît pas des préoccupations officielles et encore moins des pratiques des maîtres » (Chervel & Manesse, 1989 : 147). Il subit toutefois une évolution décisive dans son esprit : une conception nouvelle s'impose sur l'enseignement de l'orthographe et de la grammaire. Intégré à l'ensemble de l'enseignement du français, l'apprentissage de l'orthographe et des règles de grammaire commence par les textes que l'on lit. Et la mise en orthographe change aussi de nature. Les automatismes, les mécanismes et la mémorisation, sans disparaître totalement, s'effacent progressivement derrière des pratiques plus souples (Chervel & Manesse, 1989 : 150-151). Jusqu'à la fin du XX<sup>e</sup> siècle, l'enseignement du français n'a pas à subir de profondes modifications. Avec le prolongement de la scolarité, l'acquisition de l'orthographe s'étale sur 8 ans. Autrement dit, on retire l'exclusivité de l'enseignement de l'orthographe à l'école primaire et le collège contribue aussi à la consolidation de la connaissance de l'orthographe lexicale et

grammaticale. On assiste ainsi à une modification importante de l'enseignement du français : l'enseignement de l'orthographe et de la grammaire occupe « une place centrale, mais non plus la place unique » (Chervel & Manesse, 1989 : 153).

### **1.1.2.3. Les réformes de l'orthographe**

L'orthographe du français demeure longtemps « Janus à deux visages » (Fayol & Jaffré, 2008 : 18), l'un tourné vers le passé, enraciné dans l'étymologie latine, l'autre qui assure la communication écrite du plus grand nombre. Et les deux courants, l'un conservateur et l'autre réformateur, continuent à s'affronter.

L'un résiste aux changements de l'orthographe, en s'appuyant sur l'idée de l'immobilisme orthographique : certains partisans sont immodérément attachés à leur écriture nationale, ce qui peut être interprété comme une sorte de « chauvinisme ». Un certain nombre d'entre eux confond l'orthographe et la langue : « la question orthographique, dit Burney, toute secondaire qu'elle semble, présente des aspects philosophiques, politiques et sociaux dont les deux parties en présence ne sont pas toujours bien conscientes : l'Académie française et la bourgeoisie traditionnaliste sont opposées dans leur ensemble à la réforme et à l'orthographe de la cuisinière... ». Parmi les arguments des conservateurs, on voit également « la crainte d'occulter la lecture de notre littérature nationale », comme l'explique Catach (2003 : 90), c'est l'attachement « aux règles poétiques de la prosodie et de la métrique, à la physionomie des mots, pour tout dire, à la tradition » qui compte. Autrement dit, les conservateurs vouent un attachement très fort, perçu d'un point de vue esthétique, à l'orthographe. Selon eux, elle est « sentie comme ayant sa beauté, son charme propre dans sa complexité même » (Cibois, 1991). Notons aussi la domination personnelle. Il n'est pas facile de pénétrer dans notre système de l'orthographe ; cependant, lorsqu'on le possède, on découvre un monde riche de sens et d'analogies. Ces partisans s'interrogent récemment sur l'ordre juridique : qui peut réformer l'orthographe ? A cette question, l'écrivain Jean d'Ormesson répond fièrement en 1989 : « Ce n'est ni le gouvernement, ni les syndicats, ni un comité, ni même l'Académie. C'est le peuple ». Il faut cependant bien noter que

la France est un pays où « le moindre écart d'orthographe populaire est durement sanctionné » et où « l'État a réglé au cours des siècles les problèmes de langue, surtout écrite, de façon strictement centralisée » (Catach, 2003 : 91).

Cependant, l'autre courant veut changer l'orthographe du français. De nombreux phonétistes, de Meigret à A. Martinet, en passant par Brunot et Cl. Blanche-Benveniste insistent sur le fait que les modifications orthographiques ne doivent pas s'arrêter en chemin. Les partisans de ce courant indiquent également les positions intermédiaires des conservateurs, en montrant qu'en pratique, même les plus conservateurs reconnaissent « la nécessité d'effectuer quelques concessions aux plaintes exprimées par l'opinion publique » (Catach, 2003 : 91). Ils oublient cependant que, en théorie, tout le monde est d'accord pour faire une réforme, mais dans la pratique, les gens hésitent : quelle réforme faudrait-il faire et par où commencerait-elle ?

Une centaine d'interviews ont été réalisées, entre 1987 et 1989, par les auteurs d'*Orthographe mon amour* (Millet, Lucci & Billiez, 1990). Ils ont réalisé une enquête auprès d'enseignants, de secrétaires, d'apprentis, de gens du livre, et d'élèves. Afin d'avoir une vision globale des jugements portés par le public, leurs questions concernaient « la crise de l'orthographe », « les souvenirs d'apprentissage », les « difficultés » et les « fautes », ainsi que le thème de la réforme. Les résultats ont été intéressants : de nombreuses personnes ont du mal à définir l'orthographe (« l'orthographe pour moi, dit l'un... c'est comme ça ») ; elles ne semblent pas hostiles à l'orthographe, au contraire, elles portent plutôt une attitude réservée, contradictoire, avec cette consolation que « personne n'est infaillible ». De nombreux enquêtés trouvent l'orthographe française belle et avouent qu'ils ont du plaisir à écrire. En ce qui concerne la réforme, certains déclarent avec lucidité qu'il faudrait simplifier le système de l'orthographe, surtout les accents, les accords des verbes pronominaux et des consonnes doubles. Cependant, ils reculent devant une application et apparemment ne semblent pas y songer. Parmi tous les interrogés, il faut bien noter l'avis des parents d'élèves. Dans l'intérêt de leurs enfants, ils sont à la fois pour et contre la réforme : une scolarité mieux réussie, oui ; une baisse de niveau, non. Ils sentent que ces exigences orthographiques démesurées et sans réels profits constituent un obstacle aux progrès de

leurs enfants. Inversement, ils refusent de renoncer à cette orthographe, considérée comme une valeur historique et culturelle, même au prix de difficultés.

Quant aux enfants, les trois-quarts des interrogés disent qu'ils vivent dans les affres au moment de la dictée : « ça m'agace », « ça fait mauvaise impression », « c'est pas que j'ai honte, mais... ça me gêne ». Les adolescents, surtout les apprentis, selon les auteurs, sont plus contestataires. Certains d'entre eux déclarent ouvertement : « c'est idiot », « c'est nul », « on se casse la tête pour rien ». Les autres, plus nombreux, trouvent l'orthographe « normale », mais celle du français est trop compliquée. Notons que c'est surtout l'apprentissage qui leur pose problème : « c'est un dédale », « c'est fait pour induire en erreur ». Cependant, les jeunes sont perplexes quand on leur parle de réforme. Selon eux, c'est trop tard pour la réapprendre. Et ils rejettent l'écriture phonétique : « c'est plus compliqué ». Le plus étonnant, c'est que ce sont les élèves en échec scolaire qui se montrent parfois les plus conservateurs : ils disent qu'ils « en ont bavé », mais plus important, ils ne savent rien de la langue, seulement quelques bribes. C'est justement à cause de cette ignorance qu'ils sont en situation d'échec. Après réflexion, les trois-quarts des jeunes interrogés acceptent les modifications de l'orthographe comme la suppression des consonnes doubles, du *h* des mots grecs, mais n'acceptent pas, par exemple, les atteintes grammaticales, comme suppression du *s* du pluriel. Comme les grands, les jeunes pensent que c'est inutile d'y songer. L'orthographe, pour eux, c'est trop surveillé, trop important, et donc impossible. Comme le soulignent les auteurs, en conclusion : « La confusion entre orthographe et langue apparaît comme le nœud du processus de fétichisation de l'orthographe ».

En 1989, la revue *Lire* lance également une enquête publique sur un échantillon national concernant le problème de la réforme orthographique. On trouve la présence d'un socle commun entre les favorables et les hostiles à l'orthographe (Cobois, 1991) : l'orthographe fait partie de notre patrimoine (86% le pense dans l'échantillon national), elle est difficile (73%) mais sans être un casse-tête inutile (83%), en matière de réforme on pourrait envisager d'en retoucher quelques bizarreries (76%). On peut s'apercevoir que, en respectant l'orthographe, les Français prennent en même temps conscience de ses difficultés et des améliorations marginales.

Cobois note également que l'on est hostile à la « réforme » mais pas à des « modifications ». Selon lui, c'est parce que l'idée de réforme est associée à une réforme radicale, phonétique, utopique, impossible, même quand on la souhaite. De ce fait, on comprend mieux que le Conseil supérieur de la langue française publie au JO en 1990 un rapport, qui abandonne le terme de « réforme » et porte dans son titre le terme de « rectifications ».

Afin de juger de ces modifications proposées par le Conseil, une enquête a été faite en 1990 auprès d'une centaine d'étudiants. Malgré l'existence d'une opposition entre l'orthographe d'enracinement et l'orthographe fonctionnelle, pour la majorité des répondants, la clé pour résoudre le problème de l'orthographe est d'articuler ensemble l'enracinement et la fonctionnalité, c'est-à-dire qu'il faut respecter les racines culturelles de l'orthographe et à la fois, sans étouffer la nécessaire adaptation des outils culturels au présent (Cobois, 1991).

En concluant, l'orthographe du français est « le produit d'une histoire séculaire...qui en codifie les principes, au terme d'une bataille remportée par le camp de la tradition et du conservatisme » (Jaffré, 2009). Au fil des siècles, les événements linguistiques, que nous avons évoqués auparavant, ont une forte influence sur la phonétique et la morphologie du français, puis dans le même mouvement son orthographe. La langue française est par conséquent isolée des autres langues romanes et son orthographe se caractérise par « une grammaire écrite qui n'est pas relayée par la phonographie » (Jaffré, 2009), dont la présence rend ce système graphique compliqué à apprendre et à utiliser. En outre, avec la révolution électronique, les besoins de la communication écrite augmentent énormément et la place de l'écrit est en train de se transformer à l'époque actuelle. Parallèlement avec les besoins de scripteurs, la question de la simplification de l'orthographe française se pose avec acuité et l'orthographe, ainsi que son enseignement-apprentissage, suscitent tant de débats aujourd'hui. Afin de comprendre les problèmes posés par l'apprentissage de l'orthographe française, il nous semble indispensable de mettre en évidence, dans la partie suivante, les caractéristiques de ce système graphique.

## **1.2. Aspect linguistique : les spécificités de l'orthographe française**

Toute écriture entretient une relation étroite avec la langue et se caractérise généralement par deux principes majeurs, qualifiés respectivement de phonographique et de sémiographique. Relevant d'une sphère alphabétique, l'orthographe française dispose d'une solide base phonographique. Cependant, pour des raisons historiques, comme nous l'avons évoqué auparavant, l'orthographe du français est moins régulière que celle des autres langues romanes venues du latin. Mais elle n'en reste pas moins rentable, parce qu'elle est également dotée de l'une des plus complexes sémiographies qui soient. Nous essaierons donc de décrire ce système linguistique mixte en présentant respectivement ces deux dimensions : phonographique et sémiographique.

### **1.2.1. Un système mixte**

#### **1.2.1.1. Présentation générale du plurisystème**

Il est bien connu que l'orthographe française se présente sous la forme d'une collection de règles et d'exceptions. Les études conçues par Catach et son équipe de HESO (Catach, 1980) font, cependant, une description du système linguistique du français en mettant l'accent sur la cohérence, c'est-à-dire en montrant que, malgré sa complexité, le système graphique français est bien structuré et fait l'objet d'une description raisonnée.

Cette approche scientifique met en évidence trois éléments très importants : le phonème, le morphème et le lexème (mot), et ils entrent eux-mêmes en combinaison syntaxique (la phrase). Le système orthographique français, constitué essentiellement de ces trois catégories, est donc un plurisystème ou un ensemble de sous-systèmes (Catach, 2003 : 53-66) :

Le centre de ce plurisystème est constitué des phonogrammes, c'est-à-dire des graphèmes qui correspondent directement aux phonèmes. Cette zone du système comprend elle-même un noyau essentiel, les archigraphèmes ainsi qu'une partie annexe,

les graphèmes positionnels, dont nous parlerons en détail plus loin. Ensuite sera abordée la zone des morphogrammes : désinences, flexions verbales, préfixes, suffixes, etc. Prononcés ou non, ces morphogrammes sont considérés comme des marques de série ou de sens. La troisième zone est constituée par des logogrammes, qui jouent un rôle annexe. Les logogrammes servent dans le système à distinguer les homonymes, qui sont en petit nombre par rapport à tous les lexèmes qui existent dans le lexique et le discours, mais qui troublent considérablement les mécanismes d'interprétation. Enfin, la dernière zone est celle des lettres étymologiques ou historiques. Si le critère fonctionnel ne permet pas de justifier de toutes les lettres utilisées dans le système graphique du français, c'est à cause de l'existence de ce reliquat. Ainsi, il est à noter le maintien des lettres grecques dans les mots comme *théâtre*, *rhume*, ou des lettres latines comme le *p* de *prompt* et de *sculpteur*, ainsi que les consonnes doubles qui font appel soit à la connaissance du latin (*appeler*, *effarer*, *offenser*), soit à celle de l'histoire de la langue (dans les mots *sonner*, *donner*, *honneur*, etc., les doubles consonnes notent l'ancienne nasalisation).

Le premier sous-système (phonogrammique) est au centre de tout le système et est relativement stable ; le deuxième (morphogrammique) couvre d'assez larges zones de la chaîne écrite ; le troisième (logogrammique) ainsi que les lettres étymologiques et historiques constituent une zone partielle et marginale.

### **1.2.1.2. La phonographie : représenter les sons de la langue**

#### **1.2.1.2.1. Les sons du français**

Les sons produits par la voix sont en nombre infini, comme ceux de l'enfant dans son babil. Cependant, au fur et à mesure qu'il apprend à parler, l'homme devient incapable d'articuler les sons qui n'appartiennent pas à son système linguistique et il perd ainsi une « aptitude naturelle » au profit d'une « aptitude culturelle » (Jakobson, 1963). Les sons produits par la voix n'ont aucune valeur linguistique, mais dès lors qu'ils sont sélectionnés par une langue, ils acquièrent une valeur phonétique, et deviennent ce que l'on appelle le phonème.



Le phonème est la plus petite unité de signifiant phonique dont la variation modifie la valeur du signe à l'oral. La langue française standard dispose de 36 phonèmes : 17 consonnes, 16 voyelles et 3 semi-consonnes. Ce chiffre peut varier selon les régions et l'âge des locuteurs.

#### 1.2.1.2.2. L'écriture des phonèmes : principes généraux

Si le phonème se manifeste à l'orale, le graphème quant à lui se manifeste à l'écrit. Le graphème est donc la plus petite unité de signifiant graphique destiné à transcrire le phonème. Avec des correspondances entre phonèmes et graphèmes, l'orthographe pourrait se structurer et permettre aussi à la compétence orthographique de se construire. Quant aux phonogrammes, ils sont les graphèmes qui correspondent directement aux phonèmes. Considérés comme le centre du plurisystème du graphique français, les phonogrammes comprennent un noyau d'archigraphèmes. Comme les graphèmes de base du français, les archigraphèmes sont les « formes maximales stables » (Catach, 2003 : 62), c'est-à-dire que parmi tous les phonogrammes, leur occurrence est la plus élevée, et on les trouve dans toutes positions, en toute place du mot. Par exemple, les phonogrammes *e*, *é*, *è* et *ê* ont comme archigraphème le *e* ; et le *e*, lui, a les *é*, *è* et *ê* comme graphèmes positionnels. Ceci indique les valeurs diverses que l'archigraphème *e* peut prendre en telle ou telle position. Il y a 33 archigraphèmes en français, qui sont A – E – I – O – U – EU – OU – AN – IN – ON – UN – ILL – Y – OI – OIN – P – B – T – D – C – G – F – V – S – Z – X – CH – J – L – R – M – N – GN.

#### 1.2.1.2.3. L'écriture des phonèmes : problèmes

En français, la complexité de l'écriture des phonèmes s'explique par le fait que la correspondance entre les phonèmes et les graphèmes n'est pas biunivoque, c'est-à-dire que comme le nombre de phonèmes et celui des graphèmes n'est pas identique, leurs relations manifestent certaines distorsions. Dans le cas particulier des phonogrammes, la difficulté vient du fait qu'un phonème dispose de phonogrammes concurrents entre lesquels le scripteur doit choisir. Ce phénomène linguistique tient essentiellement, comme l'expliquent Fayol et Jaffré, à la genèse des orthographes : « avec le temps,

certaines distinctions phoniques peuvent disparaître, en créant un surplus graphique » (2008 : 86). C'est ainsi qu'est née la polyvalence graphique, notion sur laquelle nous reviendrons plus loin. En français par exemple, *o*, *au* et *eau* sont des phonogrammes concurrents, qui jouent des rôles différents. Et pour sélectionner le bon phonogramme, le scripteur doit avoir recours à divers procédés de transcription.

Montonen (2001) montre que le français dispose de trois types de notations pour transcrire les sons français : le signe simple, comme le *p* et le *a* dans le mot *papa* ; le signe complexe, par exemple, dans les mots *feuille*, *feuillage* et *feuille*, le digramme *eu*, le trigramme *ill* et le quadrigramme *ille*, étant tous trois des signes complexes ; le signe accompagné d'un élément auxiliaire, c'est-à-dire accent, cédille ou lettre diacritique, comme dans les mots *à*, *ça* et *cette*. Ces éléments auxiliaires servent à unir ou disjoindre les graphèmes complexes. Par exemple, dans le mot *algue*, le *gu* est un graphème complexe, mais dans le mot *aiguë*, *gu* correspond à deux graphèmes. Pour bien transcrire ces graphèmes, des lois de position doivent être prises en compte. Prenons l'exemple des mots *égaler* et *effort* : on écrit *égaler* avec accent pour préciser la prononciation du *e*, alors que dans le mot *effort*, les lettres qui suivent le *e* indiquent la prononciation et rendent l'accent inutile. Parmi les lois de position, les lois d'accentuation sont les plus régulières, alors que les autres lois sont, de natures diverses, soit inhérentes au système, soit aléatoires, produits du temps et de l'usage.

La sélection des phonogrammes ne répond pas seulement au niveau linguistique, mais aussi au lexique. Si l'on reprend l'exemple de *o*, *au* et *eau*, on voit que *o*, comme l'archigraphème, est la forme la plus neutre : le timbre est soit ouvert, comme dans le mot *vol*, soit fermé, comme dans le mot *voleur* ; et il est également présent dans toutes les positions du mot, *oreille*, *doré*, *métro*, etc. Quant aux phonogrammes *au* et *eau*, ils sont considérés comme des formes marquées, car ils apportent un surcroit d'information. La lettre *a* dans *au* et *eau* a généralement une cause étymologique : « elle rappelle un ancien *a* qui se retrouve dans des mots appartenant à une même famille et qui partagent une même racine » (Fayol & Jaffré, 2008 : 116). Ainsi, plus la polyvalence graphique est élevée, plus la sélection d'un bon phonogramme implique un traitement non phonographique, qui répond à une compétence orthographique supérieure chez le

scripteur.

### **1.2.1.3. La sémiographie : donner accès au sens**

Comme nous l'avons indiqué auparavant, la phonographie fournit les mécanismes de base au bon fonctionnement de l'orthographe. Cependant, l'orthographe ne se résume pas à une simple notation des sons du langage et elle représente également le sens linguistique. Cette mécanique sert à construire un autre principe de l'écriture, le principe sémiographique, qui est tout aussi essentiel que le principe phonographique. Pour se doter d'un sens linguistique, l'orthographe doit reconfigurer les mécanismes phonographiques afin de leur donner une forme fonctionnelle. Il nous semble nécessaire d'introduire ici une notion importante, le potentiel sémiographique des langues : « La raison d'être de l'orthographe est de donner à voir le sens linguistique, la base phonographique s'y prête plus au moins bien selon qu'elle rend visible ou non les unités lexicales et grammaticales d'une langue » (Fayol & Jaffré, 2008 : 94). De ce fait, les orthographe, comme celle du turc ou du coréen, qui transparaissent essentiellement dans la phonographie, disposent d'un potentiel sémiographique très élevé. Cependant, pour certaines langues qui présentent une homophonie importante, source majeure d'une ambiguïté graphique, la base phonographique seule ne permet plus d'assurer une bonne représentation sémiographique. En l'occurrence, l'orthographe doit avoir recours aux procédés complémentaires, par exemple l'ajout d'éléments spécifiques : lettres silencieuses des orthographe alphabétiques ou éléments sémantiques de l'orthographe du chinois.

La sémiographie n'est donc pas homogène et présente différentes facettes : elle est située sur un plan qui « englobe l'écriture d'unités lexicales, ou logographie, et l'écriture d'unités grammaticales, ou morphologie » (Jaffré, 2002). Les formes du principe sémiographique repose sur deux tendances majeures :

« La première, la moins complexe, se limite à la réanalyse des phonogrammes, qui sont associées pour constituer des entités plus larges,

sans que des adjonctions soient nécessaires (sémiographies mineures). La deuxième tendance s'efforce en revanche de compenser les insuffisances sémiographiques de la base phonographique, et de ses formes homophones, en ajoutant les éléments spécifiquement graphiques (sémiographies majeures) » (Fayol & Jaffré, 2008 : 97).

Dans le cas du français, l'orthographe n'échappe, comme toutes les autres orthographe, n'échappe pas à une réanalyse sémiographique de la phonographie. Si l'on prend l'exemple du mot *nouveau*, on voit qu'un trigramme *eau* correspond, d'un point de vue phonographique, au phonème /o/. Mais d'un point de vue morphologique, « une relation de type *nouveau/nouvelle* révèle une nouvelle fonction de ce trigramme, qui motive le *e* de *eau* » (Jaffré, 2003). On constate ainsi l'existence, dans l'orthographe française, de phonogrammes à fonction morphologique, que nous appelons morphogrammes. Ces unités très accessibles constituent les représentations lexicales mais aussi l'interface entre phonographie et sémiographie. A ce titre, elles bénéficient « d'un double traitement potentiel puisque leur présence peut être détectée par la mise en relation avec un phonème ou par une analyse morphologique » (Jaffré, 2003). Mais la sémiographie peut se manifester d'une autre façon plus directe, en utilisant des marques qui échappent à la phonologie. C'est sur ce versant des représentations grammaticales que les choses se compliquent, concernant principalement les morphogrammes, notion que nous développerons plus loin.

#### **1.2.1.4. La morphographie : signaler les rapports entre lexèmes**

La morphographie du français est l'aspect orthographique qui correspond à la morphologie écrite. Théoriquement, toute écriture comprend « des aspects qui s'écartent d'une stricte correspondance avec des unités phoniques pour prendre en compte les signes linguistiques » (Jaffré, 2003). Et le français offre à cette dimension sémiographique un exemple tout à fait différent.

Comme nous l'avons abordé auparavant, aucune écriture n'échappe à une

réanalyse sémiographique de la phonographie. Et la discordance entre le nombre de phonèmes et de graphèmes du système de la langue française conduit à une polyvalence graphique motivée par la morphologie. Cependant Fayol et Jaffré montrent que « la complexité sémiographique va bien au-delà de la polyvalence phonographique pour recourir à des graphies propres qui peuvent être des désinences verbales (le *t* de *il sait*), des marques de nombre (le *s* de *deux maisons*), ou de genre (le *e* de *amie*) ; des marques d'accord (*les petits chats jouent*), etc. » (2008 : 116). On voit ainsi que la morphologie écrite du français contient des unités qui échappent à toute phonographie et qui jouent très souvent un rôle grammatical essentiel. Bien que ces morphogrammes ne jouissent d'aucune autonomie sémiographique, ils deviennent des unités à part entière dans le système orthographique. Si l'on prend l'exemple de l'ensemble des désinences verbales, elles concernent, le plus souvent, les vestiges de l'ancien système graphique qui sont disparus aujourd'hui : le *t* de *il veut*, le *s* de *tu parles*, etc. sont des éléments qui font double emploi avec les pronoms *il*, *tu*, etc. On peut citer également l'exemple du *s* du pluriel dans *les petits chats* ; ce morphogramme, dépourvu de correspondances phoniques, signale les rapports entre les trois lexèmes.

## 1.2.2. Un système qui est source de difficultés pour les scripteurs

### 1.2.2.1. La polyvalence du français : un compromis entre sémiographie et phonographie

La phonographie a pour finalité de « fournir aux orthographes un matériau de base systématique, économique et récurrent, à partir duquel elles façonnent des signes linguistiques conformes aux exigences sémiologiques de l'écrit » (Fayol & Jaffré, 2008 :100). Pour bien comprendre le compromis entre sémiographie et phonographie de l'orthographe, il nous semble nécessaire d'expliquer une notion importante, qui est la polyvalence graphique. Étant donné que le nombre de graphèmes est supérieur à celui des phonèmes, la quasi-totalité des orthographes alphabétiques présente plus ou moins une polyvalence graphique. Les psycholinguistes se sont particulièrement intéressés à

ce phénomène, en distinguant des orthographe transparentes (finnois, grec, italien, espagnol, allemand) et des orthographe opaques (portugais, français, norvégien, suédois, néerlandais, danois et anglais).

Dans le cas particulier du français, la polyvalence graphique contribue, d'un côté, à complexifier la phonographie, en obligeant le scripteur à faire une sélection parmi les phonogrammes concurrents et d'un autre côté, elle se trouve investie de fonctions morphologiques ou distinctives, c'est-à-dire qu'elle « constitue un avantage sémiographique en fournissant un matériau visuellement fonctionnel » (Fayol & Jaffré, 2008 : 97). Cependant, dans le système du français, la structure phonologique n'arrive pas à rendre compte de la totalité des variations lexicales et morphologiques. L'orthographe française doit donc inventer des procédés spécifiquement graphiques pour tenter de compenser les insuffisances phonologiques. Le plus souvent, ces procédés se présentent sous des formes diverses qui sont tributaires de l'histoire de l'orthographe, ce qui constitue une source majeure des difficultés du scripteur. C'est ce qui caractérise l'orthographe française, un des systèmes sémiologiques les plus spécifiques et les plus opaques.

#### **1.2.2.2. L'opacité du français**

Pour les scripteurs, la difficulté d'acquisition de l'orthographe sera fonction du degré d'opacité ou de transparence de la langue, qui représente le degré de correspondance entre l'orthographe et la phonologie de cette langue, c'est-à-dire entre la manière dont elle s'écrit et la manière dont elle se prononce. Les orthographe qui ont plus de chances d'être transparentes sont celles qui sont proches du modèle idéal où la correspondance entre graphèmes et phonèmes est univoque, ou qui sont récentes. Mentionnons ici le cas du turc, dont l'orthographe est transparente grâce à la notation systématique des variantes morphologiques.

Le français, en revanche, se situe parmi les systèmes les plus opaques en orthographe. Si les correspondances graphèmes-phonèmes sont régulières et fiables à environ 95% (dans le sens de la lecture), elles ne le sont qu'à moins de 70% dans le

sens phonèmes-graphèmes (dans le sens de l'écriture) (Peereman & Content, 1999 ; Ziegler, Jacobs & Stone, 1996). Véronis (1988) a estimé que, à partir de simulations basées sur les correspondances phonèmes-graphèmes, seulement 52.7% des formes graphiques du français pouvaient être prédites à partir de leur forme phonologique.

En général, l'opacité du français est liée à plusieurs difficultés de transcription : tout d'abord, c'est l'existence des allographes, c'est-à-dire des graphèmes différents pour transcrire un même phonème. Nous reviendrons sur ce point plus loin. Deuxièmement, le français est difficile à transcrire parce qu'il existe de nombreuses consonnes doubles et le plus souvent, ces consonnes sont absentes au niveau phonologique, comme dans le mot *flamme*. Signalons également l'existence non négligeable de nombreux homophones, hétérographes, comme *sain* et *saint*, ou homographes, comme *rame* de bateau, de métro et de papier. Une autre source d'ambiguïté du système orthographique français se situe au niveau morphosyntaxique, comme les marques du nombre, qui sont le plus souvent, sans contrepartie phonologique.

### **1.2.2.3. Le degré de régularité**

Il n'existe aucune communauté linguistique dans laquelle l'apprentissage de l'orthographe repose sur la mémorisation pure et simple des signes graphiques. Pour être facile à apprendre, toute orthographe doit disposer d'une base phonographique avec des correspondances systématiques entre phonèmes et graphèmes. Généralement, le principe phonographique se manifeste sous différentes formes selon les langues. La différence la plus notable concerne le degré de régularité. Et comme il existe une correspondance systématique entre les unités phoniques et les unités graphiques, la régularité sera « d'autant plus grande que le nombre respectif de ces unités sera proche » (Fayol & Jaffré, 2008 : 86). Une écriture serait absolument régulière si chacun des phonèmes dont elle disposait correspondait à une graphie propre, lettre ou caractère, et si cette graphie était différente de toutes les autres.

Cependant, pour des raisons historiques, ce cas idéal ne se présente quasiment

jamais et à ce titre, la relation entre ces deux unités présente donc certaines distorsions. Lorsque le nombre de phonèmes est supérieur à celui des graphèmes, cette écriture serait, théoriquement, plus facile à encoder qu'à décoder, c'est-à-dire qu'elle serait plus facile à écrire qu'à lire. Il en va autrement pour l'écriture ayant à sa disposition un nombre de graphèmes supérieur à celui des phonèmes. Le décodage serait alors plus aisé que l'encodage, et la lecture serait plus facile que l'écriture. C'est en général ce qui se passe avec les langues alphabétiques. Pour des raisons qui tiennent à la genèse et à l'évolution historique des orthographe, certaines distinctions phoniques peuvent disparaître et un surplus graphique est ainsi créé.

La langue française standard dispose d'une trentaine de phonèmes et de plus de cent graphèmes. Cette distorsion entre le nombre de phonèmes et de graphèmes est révélatrice de la complexité de la phonographie du français et elle se fait au détriment du scripteur, en l'obligeant à sélectionner le bon phonogramme entre plusieurs phonogrammes concurrents. Comme ces phonogrammes concurrents jouent des rôles distincts, la sélection demande au scripteur des traitements supérieurs non phonographiques.

L'équipe HESO a réalisé un sondage statistique en 1972 pour avoir une évaluation chiffrée des phonogrammes du français. Tous les phonogrammes ont été relevés et classés à partir des cinq textes. On a alors observé la prédominance quasi absolue d'un certain nombre d'unités. Pour les voyelles, on parvenait aux chiffres suivants (Catach, 2003 : 63) :

- A : 92% des transcriptions des phonèmes [a] ;
- E : 78% des transcriptions du [e] ; 61% des transcriptions du [ɛ] ;
- I : 98% des transcriptions du [i] ;
- O : 75% des transcriptions du [o] ou du [ɔ] ;
- U : 100% des transcriptions du [y] ;
- EU : 93% des transcriptions du [ø] ou du [œ] ;
- OU : 98% des transcriptions du [u] ;
- ON : 92.8% des transcriptions du [ɔ̃] ;



- UN : 97% des transcriptions du [œ̃], etc.

Dans l'ensemble des textes, le rendement des phonogrammes atteint 82,51%. En revanche, il existe indubitablement des difficultés concernant quelques phonèmes. Si l'on prend l'exemple du phonème [ã], il est noté indifféremment par les graphèmes *an* et *en*, et [ɛ̃], noté par *in* (44%), mais aussi par *ain* (21%) et *en* (23%). Mais avec les lois de position, on pourrait essayer de dissiper l'ambiguïté de ce dernier graphème, lorsqu'il transcrit [ɛ̃], il est le plus souvent précédé de *i*, *e*, *y*, comme dans *chien*, *lycéen*, *moyen*.

Cependant, la correspondance est rompue ici dans les deux sens dès lors que plusieurs phonèmes sont transcrits de façon semblable, et que le même graphème correspond à plusieurs phonèmes, ou au phonème zéro (Catach, 2003 : 64).

Notons que le réseau des graphèmes, même pour cette partie centrale de l'orthographe qui transcrit les sons, est complexe. Pour les scripteurs, il faudrait connaître non seulement les principales représentations graphiques des phonèmes, mais aussi les lois qui régissent leur emploi. De plus, l'habitude de tenir compte des lettres qui précèdent et suivent le graphème, de sa position dans la syllabe, dans le mot, dans le groupe de mots, etc. est également nécessaire.

### **1.3. Aspects psycholinguistiques : orthographier le français**

Après avoir décrit le système orthographique français et son évolution, il nous semble crucial de poser une question sur la gestion des connaissances orthographiques : comment orthographions-nous ?

#### **1.3.1. La place de la composante orthographique dans l'activité de production de texte**

Comme « toute connaissance ne vaut que par les réponses qu'elle produit » (Ducard, Honvault & Jaffré, 1995 : 97), l'orthographe a pour objectif majeur de donner forme

graphique à un message. En l'occurrence, la production écrite implique le contrôle de cette forme. La compétence orthographique fonctionne donc dans la production écrite. Comme le soulignent Allal et ses collaborateurs, la situation de production écrite sert « de point de départ et de cadre à la construction et à l'exploitation de compétences orthographiques en cours d'écriture et lors de la révision différée du texte » (Allal, Rouiller, Saada-Robert & Wegmuller, 1999).

### **1.3.1.1. L'activité de production de texte : présentation générale**

Selon les modèles classiques de la production textuelle, il existe un découpage linéaire du processus d'écriture en phases distinctes : planification, rédaction, révision (voir par exemple, Murray, 1978). Cependant, pour les modèles hiérarchiques, proposés par Hayes et Flower en 1980, la révision est plutôt une opération ou un ensemble d'opérations susceptibles d'intervenir à n'importe quelle étape de la production d'un texte. Witte souligne également que la révision n'intervient pas uniquement dans les transformations apportées lors de la relecture d'un brouillon (révision différée), mais aussi pendant la planification du texte (révision prétextuelle), ainsi que tout au long de la rédaction du brouillon (révision on-line intégrée à l'écriture) (1985). Sur la base des modèles hiérarchiques, les chercheurs étudient les opérations intervenant dans la révision (voir par exemple, Bereiter & Scardamalia, 1987 ; Hayes, Flower, Schriver, Stratman & Carey, 1987) :

1. La détection d'éléments textuels qui « posent problème » ;
2. L'identification (ou le diagnostic) de la nature du problème ;
3. Les démarches résultant des opérations précédentes, à savoir :
  - la décision de ne rien modifier
  - la modification du texte existant (corrections d'erreurs et changements visant une légère amélioration)
  - la réécriture (impliquant des transformations/changements plus profondes/profonds du texte)

Généralement, les travaux de révision peuvent porter sur toutes les composantes d'un texte : super- et macrostructure, contenu sémantique, éléments lexicaux, structures syntaxiques, organisateurs textuels, aspects de mise en texte (linéarisation), y compris transcription orthographique (Allal, Rouiller, Saada-Robert & Wegmuller, 1999). Selon Allal *et al.*, la révision prétextuelle met le plus souvent l'accent sur les composantes notamment sémantique et macrostructurale. En l'occurrence, une révision prétextuelle au niveau orthographique est une réflexion du métalangage, telle qu'une modification du champ sémantique, ou une décision de changer l'entrée dans le texte afin d'éviter certaines difficultés orthographiques. Dans la révision on-line, l'orthographe est un élément crucial dont il faut tenir compte : le scripteur relit constamment ce qu'il est en train d'écrire et y apporte des corrections. Et la révision différée a pour objectif d'assurer la conformité de l'orthographe aux conventions de la langue écrite.

### **1.3.1.2. De l'orthographe des mots isolés à l'orthographe des mots en contexte**

Pour mieux comprendre le fonctionnement de la compétence orthographique, il nous semble utile de disposer au préalable de certaines informations sur l'architecture mentale. S'interrogeant sur le mode d'organisation des informations, certaines approches sont localisatrices. Elles regroupent les connaissances selon leur degré de spécialisation et font de la structure mentale une juxtaposition de domaines autonomes et la connaissance humaine est ainsi, organisée en modules ; d'autres optent pour l'interactivité et privilégient les relations entre savoirs, car toute connaissance implique des processus multiples et les associations de ces processus sont organisées d'une façon structurale, ou cognitive, ou encore métacognitive, et elles sont centrées à la fois sur le contenu des savoirs spécialisés et la gestion mentale de ces savoirs. L'activité mentale de l'adulte peut se décrire comme « la combinaison de modules spécialisés et de réseaux qui relient ces modules » (Ducard, Honvault & Jaffré, 1995 : 100).

Il s'ensuit que le cerveau n'est pas seulement une machine à emmagasiner des

connaissances et qu'il permet aussi à l'homme de réaliser la gestion des connaissances. Pendant l'apprentissage, si l'on veut acquérir une certaine compétence, il est indispensable d'associer l'étude de la structure du savoir et les activités qui mettent en œuvre cette structure. Au niveau orthographique, il faut donc bien distinguer les savoirs orthographiques et leur utilisation. Étant donné que l'automatisation d'un savoir implique que l'on le mette au service d'un but qui le dépasse, disposer des connaissances orthographiques ne vise pas à orthographier un mot isolé, mais à orthographier des mots dans leur contexte. C'est-à-dire que les connaissances orthographiques sont au service d'une production écrite. A ce titre, les déficits dans la production graphique ne correspondent pas nécessairement aux déficits des connaissances orthographiques et ils correspondent aussi aux déficits des compétences orthographiques. Et la compétence orthographique à son tour repose à la fois sur l'automatisation et sur l'efficacité des savoirs linguistiques et la capacité à gérer ces savoirs.

Comme Jaffré (1995) le décrit, la gestion des savoirs combine deux processus : l'activation et le contrôle. Dès qu'un besoin se fait sentir, l'activation, processus complexe, permet d'aller chercher l'information, à l'oral comme à l'écrit. Et l'évocation d'un mot donné entraîne « celle de mots ayant avec le précédent un rapport analogique que les similitudes soient phoniques, graphiques, sémantiques, etc. » (Ducard, Honvault & Jaffré, 1995 : 100). Après, une phase de contrôle sera donc nécessaire pour faire le choix le plus approprié parmi toutes les réponses concurrentes que l'activation fait émerger. Dans la dimension orthographique, le processus de contrôle fait le bon choix par l'analyse d'un ensemble d'informations linguistiques. C'est pourquoi l'accord graphique d'un mot dépend de « son appartenance à une classe morphologique et de sa distribution syntaxique » (Ducard, Honvault & Jaffré, 1995 : 101).

### 1.3.2. Procédures et stratégies pour orthographier

Comment procédons-nous lorsque nous devons transcrire un mot, en passant de sa

forme orale à sa production écrite correcte ? Il existe deux grandes catégories de circonstances où les situations de transcription sont analysées intuitivement : premièrement, s'il s'agit d'orthographier un mot connu et déjà rencontré à plusieurs reprises (en lecture ou en écriture), le scripteur est capable de récupérer la forme orthographique de ce mot directement en mémoire. D'une manière générale, une telle transcription se fait sans erreur<sup>1</sup>. Notons que « la connaissance ainsi mobilisée est spécifique au mot et ne se généralise donc pas » (Fayol & Jaffré, 2008 : 138). Deuxièmement, le scripteur entend une forme phonologique nouvelle. En l'occurrence, la solution la plus usuelle pour le scripteur est de conserver temporairement cette forme phonologique en mémoire. Ensuite, il s'agit de décomposer la forme phonologique en segments et de faire la transcription segment par segment, en donnant à chacun une lettre ou un graphème, sans perdre de vue la forme phonologique d'ensemble ; autrement dit le scripteur transcrit le mot par association entre phonèmes et graphèmes.

Au cours des trois dernières décennies, les neuropsychologues ont analysé de nombreux cas de patients cérébrolésés et ont proposé deux modèles concernant le traitement des orthographes alphabétiques. D'après le premier modèle, dit voie d'adressage ou voie lexicale, les items à lire ou à écrire seraient « directement récupérés dans un lexique orthographique, sorte de dictionnaire mental qui stockerait les mots connus, réguliers ou irréguliers » (Fayol & Jaffré, 2008 : 143). Pour ce qui est de l'autre voie, dite phonologique, voie indirecte ou encore voie d'assemblage, les items feraient l'objet de conversion entre phonèmes et graphèmes.

### **1.3.2.1. L'accès lexical**

Dans la mesure où le lien entre la forme du mot et sa signification est arbitraire, la nécessité d'un dictionnaire mental ou lexique mental, qui correspond à « l'ensemble des mots déjà connus et susceptibles d'être reconnus ou produits » (Fayol & Jaffré, 2008 : 143) semble être évidente ; mais comment est-il organisé et construit ?

---

<sup>1</sup> Lorsque le scripteur commet une erreur, il est en mesure de la repérer rapidement et la corriger lui-même.

D'énormes recherches ont été conduites ces dernières décennies afin de déterminer les caractéristiques de ce lexique mental. Segui Juan (1992) montre que l'existence de certaines capacités élémentaires chez les locuteurs d'une langue permettent d'affirmer que l'organisation du lexique mental est multidimensionnelle. Par exemple, un locuteur français peut accéder à son lexique et le consulter à partir de différentes informations linguistiques.

A partir de l'observation des patients cérébrolésés qui souffrent d'atteintes spécifiques de l'orthographe, on émet l'hypothèse de l'existence de quatre lexiques (Fayol & Jaffré, 2008 : 144) : le lexique phonologique d'entrée (en audition), le lexique phonologique de sortie (en production de la parole), le lexique orthographe d'entrée (en lecture), le lexique orthographe de sortie (en production orthographique). Maintenant, dès lors que les troubles de lecture sont généralement associés à des troubles d'écriture qui présentent de fortes similitudes, il est pertinent de supposer qu'il existe deux lexiques, phonologique et orthographique, ou un seul lexique qui fusionne les différents types d'informations relatives aux mots (Fayol & Jaffré, 2008 : 144). Néanmoins, certains modèles, dits connexionnistes, défendent l'idée qu'il n'existe pas de lexique mental : le traitement lexical repose sur les « associations plus ou moins fréquentes entre des séquences sous-lexicales, suites de sons ou de lettres de taille inférieure au mot » (Fayol & Jaffré, 2008 : 145).

Comment départager ces différentes idées concernant le lexique mental ? À travers l'observation des patients de dysgraphie phonologique, qui sont capables de lire ou d'écrire des mots connus mais pas des mots qu'ils n'ont jamais rencontrés, il est facile de considérer l'existence d'au moins un lexique mental dans lequel les items seraient stockés afin que ces patients puissent les récupérer directement. L'hypothèse qu'il existe un seul lexique permet difficilement d'expliquer le fait qu'un grand nombre d'individus sont capables de lire de nombreux mots qu'ils ne parviennent pas à orthographier correctement. En revanche, cette hypothèse de lexique unique a l'avantage de rendre compte de ce que la lecture de mots orthographiés de façon erronée, conduit à la survenue d'erreurs, même chez les lecteurs qui sont bons en orthographe. L'idée qu'il n'existe pas de lexique mental (ni unique, ni multiple) est également

plausible si l'on prend en compte de la fragilité des représentations orthographiques stockées en mémoire. Plusieurs recherches montrent que des adultes voient « leurs performances orthographiques influencées positivement ou négativement par la présentation de formes orthographiques respectivement correctes ou erronées » (Fayol & Jaffré, 2008 : 146). Cependant, dans l'état actuel de nos connaissances, il reste bien difficile de trancher entre ces différentes théories du lexique mental et des recherches supplémentaires devraient être entreprises.

### **1.3.2.2. L'accès phonologique**

Lorsque les items à orthographier ne peuvent être récupérés directement en mémoire, il faut recourir aux conversions entre phonèmes et graphèmes. C'est ce que l'on appelle un déchiffrage pour la lecture et une transcription phonétique pour l'écriture. Bien que cette voie phonologique soit toujours possible, elle présente des difficultés spécifiques au niveau de son fonctionnement. C'est une procédure coûteuse en attention puisque les formes phonologiques à transcrire doivent être temporairement stockées dans la mémoire tout au long de la transcription. Soulignons aussi que ces mémoires pourraient « se trouver spécifiquement affectées par des troubles mais aussi présenter des difficultés spécifiques de fonctionnement chez certains individus » (Fayol & Jaffré, 2008 : 147). Par exemple, certains individus éprouvent des difficultés à maintenir en mémoire temporaire des séquences phonologiques longues.

Il existe trois types d'opérations que subissent les entités stockées temporairement en mémoire (Fayol & Jaffré, 2008 : 148-149) : premièrement, l'item d'entrée devrait être segmenté en unités, et cette segmentation pourrait être syllabique, phonémique, voire en unités moins familières comme l'attaque (début d'une syllabe) ou la rime (fin d'une syllabe). Deuxièmement, dès que la segmentation d'entrée est effectuée, la correspondance des segments d'entrée avec des unités de la modalité de sortie doit être mise en œuvre. C'est ce que l'on appelle l'opération d'appariement et elle dépend des règles d'associations entre phonèmes et graphèmes mais aussi des probabilités de leurs mises en relation. Enfin, à la suite de l'opération d'appariement, les segments cibles de

la modalité de sortie seraient assemblés pour former l'item de sortie. Notons que chacune de ces opérations (segmentation, appariement, assemblage) pourrait être affectée par des troubles.

En l'état actuel, les chercheurs pensent que l'évolution de la compétence en lecture et en écriture est de moins en moins influencée par la phonologie, au fur et à mesure que le niveau de maîtrise de l'écrit augmente (Fayol & Jaffré, 2008 : 149). La voie phonographique ne serait qu'exceptionnelle et serait écartée au profit d'autres stratégies, dont nous parlerons dans la partie suivante.

### **1.3.2.3. D'autres stratégies pour orthographier**

Entre la voie lexicale et phonographique, plusieurs stratégies sont envisageables. Tout d'abord, nous pourrions recourir à une analogie orthographique lorsque la forme sonore d'un mot à transcrire ressemble à une autre que nous avons déjà rencontrée et transcrite. Si nous devons transcrire un nom nouveau dont la forme sonore est [kobar], nous pourrions l'écrire de la manière phonologique la plus simple *cobar*, ou nous pourrions aussi l'écrire *cobard*, par analogie avec *foulard*, *dard*, *lard*... Dans d'autres situations, nous pourrions saisir les informations qui impliquent le sens du mot à transcrire. Si le nom [kobar] est présenté dans un contexte qui se réfère à un métier et que nous apprenons que la pratique de cette activité est le verbe [kobarte], nous pourrions bien évidemment transcrire *cobart*. L'ajout du *t* permet d'établir une relation de sens avec le verbe, ou d'autres mots de la même famille morphologique comme *corbartage*, *cobarteur*... Il s'agit toujours d'une analogie, mais cette fois, la transcription repose sur « des dérivations associées au sens des mots ou sur la morphologie dérivationnelle » (Fayol & Jaffré, 2008 : 139).

Cependant, la transcription des seules formes phonologiques, particulièrement en français, ne suffit pas à produire une orthographe correcte. Cela s'explique par le fait que, en français, le marquage du nombre et dans une moindre mesure du genre, oblige à mobiliser des marques, *-s* pour les noms et les adjectifs, *-nt* pour les verbes, qui n'ont pas de correspondance phonologique. Et ces marques de morphologie dite flexionnelle



sont très fréquentes en français.

#### **1.3.2.4. La variété des procédures individuelles**

En résumé, une production orthographique pourrait s'effectuer en mobilisant différentes stratégies. Comme le remarquent Rittle-Johnson et Siegler (1999), l'orthographe n'est pas un domaine de connaissances dans lequel une quelconque stratégie garantit la réussite puisque chaque stratégie présente non seulement intérêt, mais aussi risques et limites. Notons que « ni le recours aux correspondances phonèmes-graphèmes, du fait de l'existence de nombreux mots irréguliers, ni la récupération directe en mémoire, en raison de la fragilité des représentations orthographiques, n'assure la réussite » (Fayol & Jaffré, 2008 : 152).

Pour bien produire l'orthographe, un éventail de procédures serait nécessaire : la remémoration de la forme orthographique des mots, l'analogie orthographique, le recours aux règles, l'emploi des correspondances entre phonèmes et graphèmes, l'emploi du dictionnaire, etc. La réussite en orthographe demande au scripteur non seulement la connaissance de ces stratégies, mais aussi la capacité à les mobiliser et à les mettre en œuvre avec discernement. Ainsi, nous pourrions comprendre les variations de performance des mêmes individus, puisque, dans une situation donnée, la performance en orthographe dépend « à la fois de la tâche, des contraintes incluant l'état de fatigue et la pression sociale, et des connaissances préalables, y compris relativement à ses propres capacités » (Fayol & Jaffré, 2008 : 141).

### **1.3.3. Les erreurs et sources d'erreur dans l'orthographe du français**

#### **1.3.3.1. Les types d'erreurs**

En ce qui concerne l'orthographe, lire et écrire ne posent pas les mêmes difficultés, puisque ces deux activités présentent d'importantes différences : d'abord, par rapport aux configurations écrites (lettres, mots et signes de ponctuation), qui sont, en lecture,

concrètes, permanentes et donc faciles à définir, les configurations sonores à transcrire sont « plus variables d'une région, d'un individu et d'un moment à l'autre, moins concrètes, plus évanescences, et donc plus difficiles à isoler, discriminer et identifier » (Fayol & Jaffré, 2008 : 123). De ce fait, la performance orthographique en écriture varie selon l'accent ou le dialecte qu'on a.

Ensuite, la lecture, c'est aller d'un nombre donné de graphèmes ou de lettres à un nombre plus limité de phonèmes. En revanche, en écriture, il s'agit d'aller d'un nombre limité de phonèmes à un nombre plus élevé de graphèmes. Par exemple, en français, on utilise environ 130 unités de lettres pour transcrire une trentaine de phonèmes. Ainsi se créent les homophones, mots qui se prononcent de manière équivalente, mais qui se transcrivent différemment. En production orthographique, ces homophones induisent des difficultés spécifiques, étant donné que le scripteur ne peut pas s'appuyer tout simplement sur la forme phonologique pour transcrire la forme orthographique.

Enfin, selon l'individu ou le dialecte, la prononciation présente une importante variabilité. En général, nous acceptons ces variations, souvent inaperçues, sans problèmes. Mais la production orthographique présente une tendance inverse, qui est rigide et supporte très peu de variations.

Ces différences entre lire et écrire rendent la production orthographique particulièrement complexe et la dernière nécessite donc « une vigilance continue de part de ceux qui rédigent et l'existence à l'école d'un curriculum, c'est-à-dire un programme spécifique pour l'instruction » (Fayol & Jaffré, 2008 : 125). C'est pourquoi même les plus experts en écriture, en dépit de leurs nombreuses compétences, ne sont pas à l'abri de faire des erreurs. Notons que ces erreurs nous informent de façon utile sur « le fonctionnement de l'expertise en orthographe, en illustrant la distinction entre les savoirs orthographiques et leur gestion » (Ducard, Honvault & Jaffré, 1995 : 100). Pour bien analyser ces erreurs, il nous semble nécessaire de disposer préalablement des connaissances sur la classification des erreurs. Nous avons eu recours ici à celle de Bexis Köhler (1993), une des versions issues des recherches de Catach (1986) (cité par Ducard, Honvault & Jaffré, 1995). Cette typologie nous permet « un affinement sensible, fondé linguistiquement, de la description des erreurs orthographiques »

(Masseron & Lusste-Chaa, 2008). Ce classement d'erreurs présente une exhaustivité : tous les types d'erreurs orthographiques sont représentés. Au total, il regroupe sept catégories d'erreurs, déterminées à partir des fonctions des graphèmes (Weck & Fayol, 2009) :

- Les erreurs phonétiques : omissions (*\*mercrdi* pour *mercredi*), confusions (*\*ballan* pour *ballon*), inversions (*\*quarte/quatre*) et adjonctions (*\*quartre/quatre*) de graphèmes ;
- Les erreurs phonogrammiques qui résultent d'une non connaissance du code graphique du français : certaines erreurs altèrent la valeur phonique du mot (*\*assiette* pour *assiette* ; *\*gider* pour *guider*), d'autres non (*\*aciette* pour *assiette* ; *\*anfant* pour *enfant*) ;
- Les erreurs morphogrammiques grammaticales qui résultent du non-respect des morphogrammes grammaticaux : marques du genre et du nombre (*petit* pour *petite* ; *\*chevaus* pour *chevaux*), et terminaisons des verbes (*\*j'é* pour *j'ai* ; *\*ils chan- tais* pour *ils chantaient*) ;
- Les erreurs morphogrammiques lexicales qui résultent du non-respect des morphogrammes lexicaux : en particulier, préfixes et suffixes (*\*lentemant* pour *lentement*), lettres justifiables par dérivation (*\*enfan* pour *enfant* ; *\*petid* pour *petit*) ;
- Les erreurs logogrammiques ont trait essentiellement à la figure des mots et permettent surtout de distinguer les homophones (*sait* pour *c'est* ; *a* pour *à* ; *ver* pour *verre*) ;
- Les erreurs concernant les graphèmes non fonctionnels : essentiellement consonnes doubles (*\*dificile* pour *difficile*) et lettres finales (*\*paraplui* pour *parapluie* ; *\*longtemp* pour *longtemps*) ;
- Les erreurs idéogrammiques concernant surtout les traits d'union (*\*vingt quatre* pour *vingt-quatre*) et les majuscules (hormis les majuscules liées à la ponctuation).

### **1.3.3.2. Les problèmes attentionnels**

Les recherches sur les performances orthographiques des adultes montrent que les

adultes n'échappent pas à la difficulté orthographique. Des chercheurs ont analysé les erreurs orthographiques recensées dans 420 pages d'écrits ordinaires en essayant de montrer qu'il existe dans un système linguistique donné, indépendamment des scripteurs et des situations, des zones de fragilité spécialement favorables à l'occurrence de fautes orthographiques (Fayol & Jaffré, 2008 : 127-130). Les données recueillies par les chercheurs mettent en évidence l'existence de zones à risque particulièrement propices à l'apparition d'erreurs : la marque du pluriel, les variations de doubles consonnes, la présence d'homophones, la morphologie du verbe. Néanmoins, les données montrent que certaines situations sont plus favorables que d'autres à la survenue de fautes : les situations formelles sont associées le plus souvent à des erreurs sur la morphologie verbale, alors que les situations informelles présentent plus d'erreurs sur l'accord en nombre des noms. Signalons également que certaines populations produisent plus d'erreurs que d'autres : plus le niveau de formation scolaire est élevé, plus la fréquence d'erreurs diminue, sans toutefois disparaître. Les chercheurs indiquent aussi qu'il existe des interactions entre ces différents facteurs : par exemple, les différentes situations n'ont pas les mêmes incidences sur les différents individus, instruits ou non.

Pour approfondir les études sur la performance orthographique des adultes, les chercheurs se concentrent sur la question des accords (Fayol & Jaffré, 2008 : 157-166). La gestion des accords pose particulièrement problème puisque la plupart d'entre eux ne se prononcent pas à l'oral. Notons que trois catégories d'erreurs retiennent l'attention des chercheurs, ce qui leur permet de mieux comprendre le processus qui gère l'accord chez les adultes.

D'abord, ce sont les erreurs qui consistent en l'omission de la flexion : *les jeune(s) ne font plus attention ; les étudiants entre (nt) dans la salle*. Ceci s'observe souvent chez ceux qui utilisent les claviers. Selon les chercheurs, il s'agit d'un problème d'attention. L'orthographe du français est un système dont les marques d'accord sont le plus souvent silencieuses, la gestion d'accord requiert donc toujours beaucoup d'attention. Si une activité (par exemple, taper sur le clavier d'ordinateur, en nous obligeant à tourner le regard régulièrement du clavier vers l'écran, diminue ainsi notre

attention portée au texte au cours de la réalisation) capte l'essentiel voire la totalité de l'attention, il n'en reste plus ou plus assez pour la gestion d'accords. Lorsque la surcharge d'attention est maximale, les erreurs surviennent.

Ensuite, ce sont des erreurs qui consistent à accorder le verbe non pas avec le sujet, mais avec le mot, nom ou pronom, qui le précède immédiatement : *les chiens du voisin arrive (nt)* ; *la mère des enfants s'inquiètent*. Les spécialistes les appellent une « erreur d'attraction ». Lorsque l'on focalise l'attention sur le verbe, les adultes sont capables de repérer et corriger ce type d'erreur. Cependant, elle passe le plus souvent inaperçue. Il faut noter qu'il est facile d'induire ces erreurs chez les adultes qui sont habitués à rédiger sans faute. Les chercheurs leur demandent de transcrire une phrase, telle que *le chien des voisins arrive*, sous la dictée et de réaliser une seconde tâche en transcrivant. En l'occurrence, le nombre d'erreurs s'élève fortement, alors qu'aucune erreur ne survient avec *le chien du voisin arrive* ou *les chiens des voisins arrivent* si les adultes les transcrivent sans une seconde tâche.

Enfin, ce sont les erreurs constituées par la substitution d'un mot par un autre, qui a la même prononciation mais une orthographe différente : *Le vagabond a des poches. Il les fouille(s)*. Notons que ces erreurs ne surviennent qu'avec certains verbes qui ont un homophone nominal : *il juge/le juge* ; *il fouille/la fouille*, etc. Les analyses montrent que le pourcentage d'erreurs est le plus élevé lorsque l'homophone nominal est plus fréquent que le verbe : *Le postier prend des paquets. Il les timbre(s)*. Concernant le mécanisme des erreurs de ce type, l'interprétation la plus plausible revient à considérer qu'au lieu de recourir à la règle ou au calcul de l'accord avec le sujet, les adultes récupèrent directement dans leur mémoire certains mots déjà tout fléchis. Il s'agit des mots qui portent souvent « la flexion nominale pour les verbes ayant un homophone nominal, surtout si cet homophone nominal est plus fréquent, et donc dominant » (Fayol & Jaffré, 2008 : 161). Ce mécanisme est adaptatif : il permet de maximiser les réussites en orthographe en minimisant le coût en temps et en attention. Néanmoins, il n'évite pas toutes les erreurs.

En résumé, toutes ces analyses nous permettent de mieux comprendre les mécanismes de la gestion de l'accord. D'abord, il faut noter que les adultes disposent

des règles d'accord de l'adjectif, du nom et du verbe. Cependant, la plupart du temps, ils ne s'y réfèrent pas, à cause du coût élevé en temps et en attention. Ensuite, pour effectuer les accords, ils mobilisent soit « une procédure simplifiée, vraisemblablement automatisée, l'accord avec le mot qui précède », soit « une procédure de récupération en mémoire d'une forme toute fléchie » (Fayol & Jaffré, 2008 : 165). Ces procédures, rapides et peu coûteuses en attention, réussissent le plus souvent. Néanmoins, comme nous l'avons dit plus haut, elles n'évitent pas toutes les erreurs. Mais si les adultes disposent de suffisamment de temps et d'attention pour mobiliser la règle explicite et procéder au calcul d'accord, ils sont en mesure de corriger l'erreur, même avant de l'avoir transcrite. Évidemment, le coût en attention serait très élevé, voire aux dépens d'une autre activité.

Il s'agit donc de distinguer entre les savoirs orthographiques et leur gestion. Les adultes sont généralement capables d'accorder les mots mais, dans un texte donné, à un moment particulier de la production écrite, ils ne sont pas en mesure de contrôler les effets du contexte syntaxique. En l'occurrence, ils doivent mobiliser d'autres compétences qui sont coûteuses en attention. Par ailleurs, la mise en œuvre de ces compétences exige des conditions particulières : « l'exercice de la compétence orthographique a besoin d'un certain espace mental et plus la part de calcul est grande, plus cet espace doit être important » (Ducard, Honvault & Jaffré, 1995 : 100). Ces problèmes, propres aux adultes, nous donnent un avant-goût de ceux que nous allons rencontrer avec des jeunes adultes chinois d'acquisition de l'orthographe française.

## **2. ORTHOGRAPHER LE FRANÇAIS QUAND ON EST UN LOCUTEUR DE CHINOIS**

### **2.1. Entrer dans une langue étrangère : trois déterminants de l'apprentissage**

Dans les orthographe alphabétiques, l'apprentissage dépend de la conjonction de trois éléments : le développement de l'apprenant, le système linguistique particulier et l'instruction dispensée. Cependant, l'acquisition de l'orthographe française présente des particularités quand on est un locuteur chinois, dont la langue maternelle est une langue syllabique.

#### **2.1.1. Le développement de l'apprenant : apprendre le français comme langue étrangère**

L'apprentissage d'une langue dépend tout d'abord du développement de l'apprenant. Pour apprendre une langue, l'apprenant doit acquérir des habitudes verbales, mais ces données linguistiques ne sont pas le facteur déterminant de l'apprentissage. Ce qui importe, c'est « le matériau linguistique mis à la disposition de l'apprenant et contrôlé par celui-ci sous l'effet d'un programme interne » (Gaonac'h, 2004 : 124). Ce programme interne conduit à une progression construite par l'apprenant lui-même.

##### **2.1.1.1. Interlangue**

Remarquons que le programme interne pose des problèmes particuliers s'il s'agit de l'apprenant d'une langue étrangère. Signalons ici la notion des dialectes idiosyncrasiques<sup>2</sup> des apprenants de langue étrangère, évoquée par Corder (1980). Ces

---

<sup>2</sup> Corder définit la notion de dialectes idiosyncrasiques en les distinguant avec des dialectes sociaux. Les dialectes sociaux sont les dialectes qui constituent la langue d'une communauté, alors que les dialectes idiosyncrasiques comprennent des règles qui sont particulières à un individu et qui ne sont pas interprétables de manière évidente par la connaissance des dialectes sociaux. Selon Corder, les dialectes idiosyncrasiques se répartissent en quatre catégories : le langage poétique, le discours des aphasiques, le langage de l'enfant qui apprend sa langue maternelle,

dialectes sont par nature instables, puisqu'ils sont liés à « des étapes transitoires dont la destinée est d'être remise en cause par la confrontation à la réalité linguistique » (Gaonac'h, 2004 : 125). Les dialectes idiosyncrasiques partagent certains traits de deux dialectes sociaux, ou de deux langues, et au niveau de l'apprentissage d'une langue étrangère, ce sont les traits du système de la langue maternelle des apprenants, et ceux du système de la langue-cible. Aussi Corder n'exclut-il pas le rôle possible des interférences dans le développement d'un système transitoire. Celui-ci se compose de trois sous-systèmes (Gaonac'h, 2004 : 125) :

- Une partie du système de la langue maternelle ;
- Une partie du système de la langue-cible ;
- Un système de règles n'appartenant ni à l'un ni à l'autre de ces deux systèmes, donc spécifique du dialecte idiosyncrasique constitué à un moment donné.

Les notions introduites par Corder sont développées par d'autres chercheurs sous différents vocables, notamment l'« interlangue » (Selinker, 1972). L'interlangue est définie, selon Selinker, comme « un système linguistique productif par référence aux deux autres systèmes linguistiques que constituent la langue maternelle de l'apprenant et la langue étrangère » (Rosen & Porquier, 2003). Ce système linguistique séparé n'est jamais entièrement assimilable au système de la langue étrangère<sup>3</sup> et il est susceptible de se fossiliser<sup>4</sup>. Signalons que la formation de ce système résulte de la mise en œuvre

---

et celui des apprenants de langue étrangère.

<sup>3</sup> Selinker considère que très peu d'adultes réussissent véritablement l'apprentissage d'une langue étrangère. Ceux-ci (5%) réactivent une structure linguistique latente qui est identique à celle qui préside à l'acquisition de la langue maternelle. Le reste (95%) active une structure psychologique latente lorsqu'ils tentent de traiter une langue pour l'apprentissage. Dans ce cas, un véritable apprentissage est en fait impossible, car l'objectif ici est de forger une compétence formelle, explicite et complète de la langue : d'une part parce qu'une compétence ne s'enseigne pas, d'autre part parce qu'un tel apprentissage ne peut en fait porter que sur une norme particulière de la langue et non sur ses aspects diversifiés et mouvants (Gaonac'h, 2004 : 126).

<sup>4</sup> Selon Selinker, l'interlangue est figée si ses stratégies de communication induisent l'apprenant à penser qu'il connaît suffisamment le système de la langue-cible pour communiquer. Ce phénomène peut se produire au niveau individuel ou au niveau d'un groupe social (Gaonac'h, 2004 : 127).



d'un certain nombre de processus, qui constitue la structure psychologique latente : transfert d'éléments de la langue maternelle, transfert d'apprentissage, stratégies d'apprentissage, stratégies de communication et surgénéralisation des règles de la langue-cible (Gaonac'h, 2004 : 127).

Notons également une des hypothèses liée fortement à la notion d'interlangue. C'est l'existence possible d'une théorie commune qui peut rendre compte à la fois des acquisitions en langue maternelle et en langue seconde. On pourrait même dire qu'il existe un dispositif inné d'acquisition. Les chercheurs proposent une interprétation génétique des acquisitions en langue étrangère en montrant l'existence d'une cohérence entre l'apprentissage de l'enfant ou de l'adulte en langue seconde et celui de l'enfant en langue maternelle (voir par exemple, Bailey, Madden & Krashen, 1974 ; Dulay & Burt, 1974 ; Krashen, Sferlazza, Feldman & Fathman, 1976). Cependant, l'explication génétique de l'acquisition d'une langue étrangère est remise en cause si l'on tient compte du rôle des situations d'acquisition. Selon les défenseurs de ce point de vue (voir par exemple, Selinker, Swain & Dumas, 1975 ; Taylor, 1974 ; Hakuta & Cancino, 1977), la nature du développement linguistique dépend du contexte d'acquisition, et ce contexte n'est pas figé : « Il dépend lui-même des acquisitions linguistiques réalisées, puisque celles-ci modifient l'usage de chacune des langues par l'apprenant, le statut qu'il leur attribue, bref le type de contact entretenu entre les deux langues du point de vue de l'apprenant lui-même » (Gaonac'h, 2004 : 137).

De surcroît, le rôle de l'âge des apprenants fait aussi partie des débats entre psychologues : certaines recherches (voir par exemple, Dulay & Burt, 1974 ; Krashen, Sferlazza, Feldman & Fathman, 1976) laissent penser que les interférences dues à la langue maternelle seraient plus fortes chez les adultes que chez les enfants, et que généralement, l'apprentissage d'une langue étrangère serait d'autant plus difficile qu'on est plus âgé. Cependant, d'autres chercheurs conduisent à l'effet facilitateur de l'âge (voir par exemple, Fathman, 1975 ; Olson & Samuels, 1973 ; Ervin-Tripp, 1974).

### **2.1.1.2. Hypothèses générales sur les processus de l'acquisition d'une langue étrangère**

A partir de l'analyse des recherches qui précède, Gaonac'h propose certaines hypothèses générales sur les processus qui seraient mis en œuvre dans les situations d'acquisition d'une langue étrangère : « la nature des tâches liées à l'acquisition serait un déterminant essentiel des fonctionnements cognitifs sous-jacents. Elles portent sur la multiplicité des processus d'acquisition, et sur les caractéristiques des stratégies qui en contrôlent la gestion » (2004 : 144).

Il s'ensuit que l'acquisition des langues étrangères dépend à la fois des connaissances linguistiques que possèdent les apprenants et de la façon dont ils utilisent effectivement ces connaissances. Basées sur le traitement de l'information fournie par les données de la perception, les représentations des connaissances, qui sont stockées en mémoire, s'accroissent, se modifient et évoluent. Dans l'apprentissage de la langue étrangère, les connaissances linguistiques nécessaires à chaque compétence en langue étrangère sont « accumulées, codées et mises en mémoire de façon à être utilisées dans les traitements langagiers différents correspondant à chacune des compétences » (Barbé, 2005). Pour la mise en œuvre, cette analyse des connaissances linguistiques est indissolublement liée avec une autre composante procédurale, qui est le contrôle des traitements langagiers, qui remplit trois fonctions principales (Barbé, 2005) : sélection, intégration et prise en compte des contraintes en temps réel.

Ces deux composantes sont interdépendantes, mais chacune d'elles joue un rôle spécifique dans des aspects différents des processus d'apprentissage et de traitement du langage. Autrement dit, « chaque utilisation spécifique du langage (conversation, lecture, etc.) nécessite un niveau de compétence différent dans chacune des composantes procédurales. Chaque apprenant possède un niveau de maîtrise spécifique pour chacune de ces composantes. Une tâche ne peut être réalisée que dans des conditions où les contraintes de traitement ne dépassent aucun de ces deux niveaux » (Barbé, 2005).

Pour réaliser des progrès dans l'apprentissage du langage, l'apprenant doit développer ces compétences de manière délibérée et précise. Notons également que la performance des apprenants sur ces compétences est fragile, et que leur acquisition ne se maintient que si elles sont régulièrement exercées. Et une maîtrise suffisante de chacune de ces deux compétences procédurales permet de renforcer leur automatisation, qui pourra conduire à des réalisations langagières satisfaisantes ayant des différents objectifs.

Pour une production écrite, qui impose un « recours constant à des stratégies attentionnelles d'utilisation de connaissances lexicales et de règles syntaxiques » (Barbé, 2005), les connaissances linguistiques, la compétence d'analyse de ces connaissances ainsi que la compétence de contrôle des traitement langagiers et l'exercice de cette compétence seraient dans ce cas absolument nécessaires.

### **2.1.2. Le système linguistique particulier : les propriétés du système français**

Pour fonctionner, toute orthographe fait appel à « des lois qui en allègent la maîtrise et l'utilisation » (Jaffré, 2005). Pour être facile à apprendre, l'orthographe doit être aussi proche que possible de la notation des sons d'une langue ; pour être utile, elle doit donner accès au sens en tenant compte de sa structure lexicale et grammaticale. En l'occurrence, comme nous l'avons dit plus haut, toute orthographe est un compromis entre phonographie et sémiographie.

#### **2.1.2.1. Versant phonographique**

Malgré sa réelle complexité, l'orthographe française présente des similitudes avec d'autres orthographe alphabétiques au niveau phonographique. Si la phonographie réside dans la correspondance entre unité phonique et unité graphique, la biunivocité (chaque phonème correspond à un graphème et un seul, et réciproquement) est un critère de référence. Bien que l'orthographe française soit évidemment éloignée de cet idéal, elle est plus régulière que celle de l'anglais et occupe donc une position

intermédiaire, entre l'anglais et d'autres orthographe plus régulières, comme celle de l'italien et de l'espagnol. Et cet écart à la biunivocité entraîne une complexification des relations phonographiques à l'orthographe française, c'est la polyvalence graphique, ce qui veut dire qu'un phonème peut correspondre à plusieurs graphèmes et inversement. De ce fait, quantité de dysfonctionnements se présentent, et ils brouillent fortement la relation entre compétences orthographiques et compétences phonographiques<sup>5</sup>. Cependant, il est imprudent de conclure que l'orthographe du français se caractérise par une exception phonographique. Bien qu'elle soit fort éloignée de la biunivocité, elle n'est pas la seule dans ce cas. L'orthographe de l'anglais présente en effet une polyvalence graphique plus importante que celle du français. Si l'on quitte la famille alphabétique et compare l'orthographe du français à celle du japonais ou du chinois, le phénomène apparaît plus nettement.

#### **2.1.2.2. Versant sémiographique**

En revanche, le versant sémiographique de l'orthographe du français présente la plus grande spécificité, surtout en ce qui concerne les relations entre l'écrit et la morphologie. En Europe, l'orthographe du français est sans doute celle où « le décalage entre la morphologie de l'oral et celle de l'écrit (morphographie) est le plus important » (Jaffré, 2005).

La notion de variation morphographique est considérée comme un excellent indicateur du degré de difficulté d'une orthographe. Quand cette variation est minimale, les mots ou leurs équivalents se manifestent sous des formes graphiques stables, et la maîtrise de l'orthographe implique que l'on gère ces formes stables en les agencant selon les lois de la syntaxe. C'est en général ce qui se passe pour le chinois.

Quand cette variation est importante, la maîtrise de l'orthographe réside dans la

---

<sup>5</sup> Il est évident que la maîtrise de l'orthographe française doit passer par une phase alphabétique. Néanmoins, différenciée des autres orthographe assez régulières (finnois, allemand, etc.), la compétence phonographique n'aboutit pas d'emblée à des formes graphiques qui sont conformes à la norme orthographique. Pour bien orthographier, un traitement non phonographique serait nécessaire.

relation entre phonographie et morphographie. Et « plus cette relation est marquée et plus la gestion de l'information graphique est supportée par la coprésence d'une information phonologique » (Jaffré, 2005). C'est dans une large mesure ce qui se passe pour la plupart des langues romanes, exception faite du français.

Comme langue romane, le français présente une variation morphographique importante. Cependant, pour des raisons essentiellement historiques, il est sans aucun doute le moins roman des langues romanes. Sous l'effet de l'invasion des Francs, dont la langue était germanique, le français, langue romane, a subi un assemblage linguistique, notamment une usure importante des secteurs grammaticaux.

Exposée en permanence aux conflits entre les défenseurs d'une orthographe phonique et les zéloteurs de surcharges graphiques, la langue française se caractérise ainsi par les suppléments graphiques qui font largement appel à l'étymologie. A ce titre, l'orthographe du français n'est plus « un calque de la parole », mais « donne aux mots une forme stable qui libère en partie la prononciation » (Jaffré, 2005). C'est pourquoi la langue française présente un déficit phonographique bien plus important que d'autres langues romanes, comme le montre clairement le tableau ci-dessous.

**Tableau 1.** La conjugaison du verbe « prendre »

	Prendre		
	Chinois	Français	Espagnol
1PS	Wo (je) na	Je prends	Tomo
2PS	Ni (tu) na	Tu prends	Tomas
3PS	Ta (il) na	Il prend	Toma
1PP	Wo men (nous) na	Nous prenons	Tamamos
2PP	Ni men (vous) na	Vous prenez	Tomais
3PP	Ta men (ils) na	Ils prennent	Toman

A travers le tableau de la conjugaison du verbe *prendre* de trois différentes langues, on voit que, contrairement à ce que se passe pour le chinois, les mots du français et de

l'espagnol sont soumis à de constantes variations. Cependant, à la différence de l'espagnol, dont l'information grammaticale relève de la phonographie, le français subit un déficit, voire une absence phonographique au niveau morphographique, ce qui rend la gestion de la morphologie à l'écrit problématique. Il oblige les scripteurs du français à apprendre à contrôler les variations morphographiques en analysant le contexte morphosyntaxique dans lequel ils se trouvent. Notons que « plus l'écrit est spécifique – moins il est phonographique – et plus la part de calcul à fournir est importante » (Jaffré, 2005). Ce qui explique que les scripteurs, même très bons en orthographe, commettent des fautes lorsqu'ils affrontent plusieurs problèmes en même temps<sup>6</sup>.

Il s'ensuit que le décalage important entre la morphologie de l'oral et celle de l'écrit, qui existe dans l'orthographe du français, rend nécessaire « la construction d'une compétence métalinguistique qui excède largement l'explication d'une grammaire de la langue parlée » (Jaffré, 2005). Cela explique également que l'enseignement du français langue étrangère en Chine met l'accent sur l'apprentissage explicite. Chez les apprenants chinois, dont la langue maternelle présente une variation morphographique minimale, la connaissance métalinguistique est déficitaire, voire absente, problème dont nous parlerons plus loin.

### 2.1.3. L'instruction dispensée : un enseignement explicite

#### 2.1.3.1. Essor du français en Chine

Avant d'aborder la didactique du FLE dans la situation chinoise, il nous semble indispensable de présenter la place du français dans la société et dans l'enseignement-apprentissage de langues étrangères en Chine. En plein essor, le français est, depuis la fin des années 90, la 4<sup>ème</sup> langue étrangère apprise par les Chinois<sup>7</sup> (Pernet-Liu, 2012).

---

<sup>6</sup> Par exemple, les scripteurs doivent gérer des idées, écrire un texte, et mettre l'orthographe en même temps. Dans ce cas-là, la surcharge cognitive rend le contrôle de l'orthographe impossible.

<sup>7</sup> Notons que le français se trouve loin derrière l'anglais, mais aussi derrière le russe et le japonais, qui restent des

D'après Hardy-Zhang (2010), il y a, en 2010, au moins 100 000 apprenants de français. Selon Saillard (2006), les effectifs ont doublé en cinq ans (50 000 en 2005). Ces apprenants de français se répartissent dans quatre milieux scolaires : plus de 6 000 dans l'enseignement secondaire, environ 30 000 dans le milieu universitaire, plus de 25 000 dans les Alliances françaises, et plus 25 000 dans des établissements privés ou semi-privés (Pernet-Liu, 2012). Selon Hardy-Zhang (2010), il existe également 15 000 chinois qui apprennent le français par eux-mêmes.

Si l'on se concentre sur les établissements d'enseignement supérieur, le français était déjà présent dans 175 universités en 2006 et 60 universités sont classées d'excellence (Saillard, 2006). En ce qui concerne les universités ayant un département de français (français comme discipline de spécialité), le nombre est passé d'une trentaine en 2002 à plus de 90 en 2012 (Pernet-Liu, 2012). Pernet-Liu nous indique que la répartition de ces universités sur l'ensemble du territoire chinois n'est pas égale (suivant le développement économique du pays, plus fort à l'Est qu'à l'Ouest). Selon une enquête publiée par le consulat de France à Shanghai<sup>8</sup>, dans la circonscription consulaire (Shanghai et les provinces environnantes), 14 221 étudiants chinois apprennent le français dans 47 établissements universitaires en 2014. D'après Pernet-Liu (2012), la région consulaire de Shanghai compte 358 établissements d'enseignement supérieur, parmi lesquels 12 ont un département de français. Autrement dit, cette région concentre 13% des départements de français. En dépit de ces disparités régionales, la Chine est le pays où le nombre de départements de français dans les universités progresse le plus vite au monde<sup>9</sup>.

### **2.1.3.2. Motivations de l'apprentissage**

D'après Pernet-Liu (2012), les raisons principales du choix du français par les apprenants chinois sont les suivantes :

---

choix linguistiques privilégiés par les Chinois (Pernet-Liu, 2012).

<sup>8</sup> <https://cn.ambafrance.org/Etat-des-lieux-de-l-enseignement-du-francais-langue>

<sup>9</sup> <http://www.consulfrance-pek.in.org/Inauguration-de-l-Annee-de-la-langue-francaise-en-Chine.html> [archive]

- Recherche d'une alternative à l'anglais ;
- Attrait de l'image de la France (découverte d'un univers culturel ; attraction des établissements d'enseignement supérieur français, etc.)
- Atout pour la carrière professionnelle (perspectives d'emplois en Chine dans des entreprises étrangères, mais aussi dans des entreprises chinoises en Afrique francophones)
- Projet d'émigration vers le Canada, par le Québec.

Avec les motivations susmentionnées, nous constatons que l'objectif prioritaire des apprenants chinois de français porte sur la capacité de communiquer et d'agir professionnellement au contact des étrangers avec leur savoir-faire langagier et multiculturel. Ainsi se crée une situation contradictoire entre les besoins sociaux des apprenants et la méthodologie de l'enseignement de langues étrangères pratiquée traditionnellement en Chine. En décrivant le contexte et le terrain didactique chinois, nous analyserons cette situation contradictoire en détail dans la partie suivante.

### **2.1.3.3. Le contexte local et le terrain**

Ce qui se passe en général pour l'enseignement du français en milieu institutionnel chinois, c'est comme le décrit Beacco (2000), l'effort essentiel qui porte souvent sur la langue, sauf au niveau avancé, et la civilisation française qui constitue une annexe. En Chine, actuellement, les apprenants de français se répartissent principalement entre les Alliances françaises, les institutions privées, ainsi que les départements universitaires de français. Ce sont sur ces derniers que porte notre travail.

#### **2.1.3.3.1. Trois références pédagogiques dans les établissements d'enseignement supérieur**

En ce qui concerne l'enseignement du français en milieu universitaire chinois, trois références pédagogiques ont été établies vers la fin du 20<sup>e</sup> siècle, à savoir le programme national promulgué par le Ministre de l'Éducation, le manuel très influent intitulé *Le*



*Français*, ainsi que le test national TFS, qui forment un tout d'une grande cohérence interne, en reposant essentiellement sur les aspects linguistiques de la langue.

En Chine, les départements de français doivent établir leur cursus conformément au programme national. Notons que le programme a été construit à partir des caractéristiques des étudiants chinois, et leur fonction est de maintenir l'enseignement universitaire dans son ensemble à un niveau qualitatif très élevé pour correspondre aux besoins de l'État (Wang, 2015).

Depuis 2004, les étudiants de français passent en fin de deuxième année de Licence le TSF4, et depuis 2009, passent en fin de la dernière année le TFS8<sup>10</sup>. Pour entrer en deuxième cycle des études (Master), les étudiants doivent passer un concours national. Tous ces tests nationaux ont pour objectif de certifier le niveau d'un étudiant dans la langue française, essentiellement sa connaissance des structures langagières et du vocabulaire. Ainsi, la compétence linguistique est un excellent indicateur du niveau de langue des étudiants, et également un élément déterminant dans la réussite de leurs futures études et de leurs débouchés professionnels.

Quant au manuel *Le Français*, il a été élaboré afin de faciliter l'enseignement et les tests à l'échelon national (Ma, 2010). Après sa publication pour la première fois en 1987, *Le Français* devient le manuel le plus répandu dans les universités chinoises.

Toutefois, ces références pédagogiques présentent plusieurs défauts. Si elles mettent l'accent sur les connaissances linguistiques, littéraires et culturelles, la formation des compétences communicatives est négligée. D'ailleurs, l'élaboration des activités pédagogique et d'évaluation représente notamment une passivité ou une unilatéralité (Wang, 2015). Comme Wang l'indique, le TFS4 n'inclut que des sujets « passifs » ou « unilatéraux », et les candidats feront tous les sujets par écrit et n'auront pas d'occasion de faire face à un jury, de répondre à ses questions ou de réagir à ses remarques.

---

<sup>10</sup> TFS4 : Test national du français enseigné à titre de spécialité, niveau IV; TSF8: Test national du français enseigné à titre de spécialité, niveau VIII.

#### 2.1.3.3.2. Besoin d'évolution

Si les trois références pédagogiques représentent une grande cohérence interne pour faciliter l'enseignement-apprentissage de français dans les universités, sont-elles adaptées au contexte actuel de la société chinoise ? Notons qu'avec les changements sociaux intervenus en Chine, l'objectif principal des jeunes étudiants chinois est d'être capables de « communiquer et d'agir professionnellement au contact des étrangers » (Humeau, 2012). Il semble que la langue d'apprentissage en cours de français ne permet pas aux apprenants chinois une utilisation pragmatique de la langue. La situation contradictoire ainsi créée entre la méthodologie pédagogique et les besoins sociaux nécessite donc des réformes dans l'enseignement supérieur chinois (Fu, 2006).

Actuellement, le besoin d'évolution des pratiques pédagogiques est partagé par le Ministre chinois de l'Éducation, ainsi que par de nombreux acteurs dans le domaine de l'enseignement-apprentissage du français dans le milieu universitaire chinois. A la lecture des thèmes choisis par la revue *Synergies Chine* depuis une décennie (par exemple, « Pour une recherche-action en didactique du FLE dans la Chine d'aujourd'hui » pour l'édition 2011 ; « Tradition et innovation dans le domaine du FLE » pour l'édition 2016), nous constatons que la réflexion sur l'évolution s'engage en Chine et que nous commençons à entrer dans une période de réformes.

Notons que les débats méthodologiques sont beaucoup influencés par des problématiques émergeant en contexte européen, notamment par la perspective actionnelle du CECR, qui vise à réduire la distance entre la langue d'apprentissage et celle d'usage. Tout cela pousse les chercheurs et les enseignants à réfléchir sur la tradition didactique chinoise pour l'enseignement-apprentissage de langues étrangères (Wei, 2011). Dans son travail, Pu (2011) nous indique que les caractéristiques principales de l'approche actionnelle et celles de la culture d'apprentissage chinoise sont très dissemblables. Une des différences porte sur les représentations du rôle de l'enseignant et des apprenants. Les cours sont souvent donnés passivement : l'enseignant parle la plupart du temps et les apprenants l'écoutent. On constate que les apprenants prennent rarement la parole spontanément, et que l'interaction sollicitée génère souvent une réponse brève ou un silence (Sagaz, 2011). De plus, l'apprentissage

se limite à l'espace scolaire. Dans les universités chinoises, comme l'évoque Pu, il manque une formation à l'autonomie qui prépare l'apprenant à la vie sociale et professionnelle.

Néanmoins, il n'est pas superflu de remarquer que dans ces débats de méthodologies, si l'accent est mis sur la compétence à communiquer, c'est l'aisance à communiquer à l'oral qui retient l'attention essentielle des chercheurs. En revanche, l'écriture et son apprentissage sont « peu envisagés dans sa capacité à relier l'univers d'apprentissage et la vie sociale » (Pernet-Liu, 2012).

#### **2.1.3.4. Organisation de cours**

Les cours proposés dans les départements universitaires sont majoritairement des cours de langue française, et quelques cours de littérature, de civilisation et plus rarement, de cours de français spécialisé (traduction, ou français des affaires, etc.). L'équipe pédagogique se compose majoritairement de professeurs chinois (un à trois professeurs français par département), et la plupart des cours sont assumés par les professeurs chinois, en utilisant majoritairement des manuels chinois<sup>11</sup>, problème dont nous parlerons en détail plus loin (Xu, 2005).

Selon l'observation de Cuet et Marguerie (2007), les pratiques de classe dans les universités chinoises correspondent à ce que l'on appelle méthode traditionnelle ou encore méthode grammaire-traduction : présentée sous une forme d'un ensemble de règles grammaticales et d'exceptions, la langue est enseignée soit par une séquence de grammaire, soit par une séquence de vocabulaire. La grammaire est « explicite, déductive le plus souvent (de la règle aux exemples)<sup>12</sup> et bien qu'introduite dans le texte de la leçon, elle est en grande partie décontextualisée » (Cuet & Marguerie, 2007).

---

<sup>11</sup> Certains manuels édités en France (par exemple, *Reflets, Taxi*) sont utilisés en complément, essentiellement pour la pratique de l'oral. Cependant, il faut noter que la plupart de ces manuels sont traduits par les professeurs chinois, en ajoutant, par exemple, les explications grammaticales, la liste des nouveaux mots.

<sup>12</sup> A partir de l'observation des manuels utilisés aux cours, on voit que la progression grammaticale est le plus souvent linéaire, du simple au complexe. Au lieu d'être expliqué dans plusieurs leçons, un point de grammaire est étudié dans sa totalité en une seule fois.

Quant au vocabulaire nouveau, il est présenté « sous forme de listes que l'étudiant devra apprendre par cœur » (Cuet & Marguerie, 2007). Indiquons également que la traduction joue un rôle capital dans l'enseignement de la langue : la compréhension du texte de la leçon s'appuie sur le travail de traduction, et les mots nouveaux sont systématiquement traduits en chinois.

Dans sa recherche, Perche (2011) met en lumière une distinction entre les choix méthodologiques du groupe d'enseignants chinois et ceux du groupe d'enseignants francophones natifs. Si le premier groupe se caractérise par « une juxtaposition d'emprunts aux différentes méthodologies avec une tendance importante à la méthodologie grammaire-traduction », le second groupe, en revanche, a essentiellement recours à un « éclectisme empirique pratique à tendance communicative » (Pernet-Liu, 2012). D'après une enquête menée par l'Ambassade de France en 2006, dont Perche reproduit les résultats, 60 des enseignants francophones en Chine n'ont aucune formation en FLE.

Selon Pernet-Liu (2012), la distinction entre deux groupes d'enseignants joue un rôle important dans l'enseignement du français en Chine, notamment en ce qui concerne la répartition des cours : les pratiques à l'oral sont souvent confiées à des enseignants français, ainsi qu'en ce qui concerne la répartition des rôles : les enseignants chinois assument les contraintes institutionnelles<sup>13</sup>, alors que leurs collègues français ont une certaine liberté d'initiatives dans leurs classes.

### **2.1.3.5. Manuels utilisés**

#### **2.1.3.5.1. Manuels élaborés en Chine**

En l'état actuel, les universités chinoises utilisent largement les manuels chinois, c'est –à-dire les manuels de français édités en Chine par des auteurs chinois. Parmi ces manuels, l'exemple le plus célèbre est le manuel *Le Français* de Ma (1987). Relevant essentiellement de la méthodologie grammaire-traduction, le manuel *Le Français*

---

<sup>13</sup> Par exemple, préparer les étudiants à l'examen national de niveau, utiliser les manuels imposés d'en haut.

propose une structure stricte à suivre (Cuet & Marguerie, 2007). Chaque leçon se divise de la façon suivante : texte/dialogue introduisant les notions de grammaire et de vocabulaire traitées ensuite dans la leçon, liste de tous les mots nouveaux de la leçon avec traduction en chinois, point de grammaire, puis exercices divers : exercices lacunaires, exercices structuraux, résumés, etc.

Notons que les textes choisis dans ce manuel, notamment dans le premier tome s'adressant aux débutants, sont principalement des documents fabriqués. Le dialogue ci-dessus concerne l'extrait d'un texte utilisé dans le manuel.

- *Qu'est-ce que c'est, Annie ?*
- *C'est une valise.*
- *Une valise ? Est-ce ta valise ?*
- *Oui, c'est ma valise.*
- *C'est ta valise ? Mais où vas-tu ?*
- *Moi ? Devine !*

Ce petit texte vise à démontrer à l'apprenant comment les phrases interrogatives sont construites en français. Cependant, il est fabriqué d'une façon artificielle et aussi très éloigné des situations communicatives auxquelles l'apprenant pourrait être confronté dans la vie réelle. A partir d'une analyse des textes choisis par ce manuel, nous constatons que les indices permettant à l'apprenant de connaître la situation de communication, tels que l'identité des interlocuteurs, le but de l'interaction, les règles socioculturelles contraignant l'interaction, sont souvent absents. Les dialogues suivent davantage les schémas d'interaction chinois que les schémas d'interaction français (Humeau, 2010), et les échanges entre les interlocuteurs ne semblent pas avoir d'objectifs communicatifs précis, autres que de faire apparaître essentiellement du vocabulaire et de la grammaire.

#### 2.1.3.5.2. Nouvelle génération de manuels

Néanmoins, la situation évolue depuis plus d'une décennie. L'exemple le plus

significatif est la méthode *Reflets*, recommandée par l'Ambassade de France au début des années 2000. Cette méthode a ensuite été publiée en 2005 dans une version bilingue. Avec l'ajout des contenus que les traducteurs chinois jugent manquants dans l'édition en français, telles que la phonétique, la systématisation grammaticale, les listes de vocabulaire traduit, les notes explicatives sur des faits de culture (Pernet-Liu, 2012), cette méthode en version chinoise est très appréciée par apprenants et enseignants.

Le succès de la méthode *Reflets* en Chine témoigne par la suite d'une ouverture sur des manuels dits « communicatifs » et « français » (Pernet-Liu, 2012). Signalons ici la méthode *En Route*, utilisée dans l'université où nous avons recueilli les données de recherche. Les auteurs de cette méthode se sont beaucoup inspirés du manuel français *Studio 100*, qui présente constamment une idée très attirante : faire apprendre le français par la résolution de problèmes. On retrouve dans la méthode *En Route* une structure relativement classique des manuels chinois de langues : phonétique-documents-lexique-grammaire-exercices. Mais cette structure classique, à dominante traditionnelle, est équilibrée d'une part en contenu par la présence de la communication, de lexiques thématiques, des réflexions (inter)culturelles au niveau du contenu, et d'exercices à penchant actionnelle, et d'autre part en méthodologie d'apprentissage fondée sur l'esprit de l'apprenant responsable et autonome (l'auto-apprentissage et la métacognition).

Ainsi, dans un « contexte de débats méthodologiques au sujet de l'adaptation du matériel didactique chinois ou français à la culture d'apprentissage des apprenants chinois et aux contraintes institutionnelles » (Pernet-Liu, 2012), nous constatons une résistance de la méthodologie traditionnelle sur le terrain (Wang, 2005 ; Hu, 2011).

### **2.1.3.6. Caractéristiques de l'enseignement du français**

#### **2.1.3.6.1. Langue d'enseignement**

Signalons d'abord que la langue d'enseignement est principalement le chinois, surtout

pour les débutants<sup>14</sup>. Cela s'explique partiellement par le fait que les étudiants inscrits en licence de français sont pour la majorité de vrais débutants en langue française à leur entrée en université. Pour eux, la langue française est un domaine tout nouveau parce que ces étudiants viennent principalement des écoles secondaires où l'anglais est enseigné comme la seule langue étrangère obligatoire<sup>15</sup>. Si nous tenons compte du fait que ces étudiants sont face à un système linguistique entièrement nouveau, il ne semble pas anti-pédagogique de leur donner au début un enseignement basé sur leur langue maternelle.

#### 2.1.3.6.2. Enseignement explicatif

Les cours de français présentent souvent une systématisation grammaticale avec un enseignement très explicatif. Ce qui se passe souvent dans un cours de français, c'est que l'enseignant « explique la règle, montre comment l'appliquer et le travail des étudiants consiste à se familiariser avec ces données nouvelles, les apprendre et s'entraîner à les appliquer » (Robert, 2002). L'adoption de ce type d'enseignement s'explique partiellement par la distance entre les deux systèmes linguistiques. Une morphologie développée de la langue française constitue la première difficulté de l'apprentissage des étudiants chinois, dont la langue maternelle présente une variation morphologique minimale. Autrement dit, l'acquisition du chinois ne prépare pas les apprenants au métalangage grammatical. Les informations grammaticales leur fournissent donc les bases métalinguistiques qui leur font défaut. Et l'enseignement explicite est donc, dans une certaine mesure, une des façons les plus efficaces.

Si l'enseignement du français en Chine privilégie une approche structuraliste de la langue, c'est aussi parce que les apprenants comprennent ainsi l'enseignement ou l'apprentissage des langues. On observe qu'au lieu de découvrir des notions, de les organiser et de conceptualiser, les apprenants chinois se sont habitués à attendre de

---

<sup>14</sup> En général, les professeurs français assument deux heures de cours par semaine pour une classe. Ils s'attachent au début à l'expression orale, au fur et à mesure que le niveau de français des étudiants s'élève, ils s'attachent davantage à l'écriture et à la civilisation française.

<sup>15</sup> Exception faite des écoles des langues étrangères où l'anglais est remplacé par le français. Notons que ces écoles des langues étrangères occupent une place beaucoup moins importante en Chine que les écoles habituelles.

l'enseignant des règles simples et logiques qu'ils pourront appliquer ensuite. D'après Robert (2002), les étudiants chinois sont à la recherche d'une compétence linguistique et d'un savoir culturel, et pour remplir ces besoins langagiers, ils privilégient, surtout au niveau débutant, l'acquisition d'une grammaire et d'un vocabulaire de base. Il s'ensuit qu'ils sont plus sensibles à une approche progressive et structuraliste, concentrée sur la compétence linguistique, qu'à une approche fonctionnelle, qui a pour objectif de développer la compétence communicative.

#### **2.1.3.6.3. Entraînement intensif en phonétique et en orthographe**

Une autre caractéristique de l'enseignement du français en Chine concerne un entraînement intensif de la phonétique et des règles orthographiques. Du fait de l'éloignement linguistique entre le français et le chinois, l'apprenant adulte chinois en français montre souvent des difficultés de perception phonologique, qui réduisent considérablement, notamment au début de l'apprentissage, ses capacités de discrimination et de catégorisation phonologiques. De plus, le français présente une forte polyvalence graphique et une variation morphographique importante. Avec toutes les particularités de l'orthographe du français, sa maîtrise se situe au centre de l'apprentissage du français écrit. Ainsi, tous ces éléments obligent les apprenants chinois à accorder une grande importance à la correction phonéto-phonologique et à l'entraînement en orthographe. D'ailleurs, les étudiants chinois trouvent souvent que l'organisation des premières leçons présentée dans les manuels français semble toujours « trop dense en contenu et trop rapide en rythme » (Hu, 2010). Donc, un entraînement intensif de la phonétique et de l'orthographe permet de dédoubler le contenu des premières leçons, et d'offrir par conséquent un rythme pédagogique souple et adapté aux débutants chinois.

#### **2.1.3.6.4. Place à l'écrit**

En respectant une tradition chinoise de privilégier l'écrit, les enseignants de français proposent régulièrement les différentes tâches d'écriture aux étudiants chinois, telles dictée, composition, réécriture. Au niveau de l'évaluation, un examen de français



commence souvent par une dictée, et se termine par une composition.

Un autre exercice souvent pratiqué dans les cours de français est la traduction. À part l'exercice de traduction traditionnel, les enseignants de français en proposent également un autre, qui est lié fortement au contexte. Ce dernier consiste à déverbaliser, après avoir compris le texte à traduire, puis à reformuler ou à ré-exprimer le message que le texte véhicule. Ce type d'exercice sert à enlever la mauvaise habitude de la traduction mot-à-mot chez les apprenants chinois. De plus, cette activité « traduction-déverbalisation » (Hu, 2010) favorise le développement de la compétence métalinguistique. Par exemple, dans le manuel *En Route*, nous trouvons un exercice portant sur la traduction des questions fréquemment posées en classe de langue, qui invite les apprenants à réfléchir sur les différences entre français et chinois pour construire les formes interrogatives. Autrement dit, la traduction les aide à prendre conscience des similitudes et des différences entre les deux systèmes de langue.

## **2.2. Le capital initial du locuteur chinois : quelques spécificités de la langue chinoise et de l'écriture chinoise**

En dépit de toutes les qualités spécifiques de l'écriture chinoise, elle fonctionne comme texte et se lit comme toute autre écriture. Autrement dit, elle « renvoie au signe linguistique tout entier, forme phonique et signifié, dans leur nécessaire cohésion » (Alleton, 2002 : 187). L'orthographe du chinois, comme toute autre orthographe, est considérée comme le résultat d'un compromis entre la phonographie et la sémiographie. Bien entendu, l'orthographe du chinois actualise ces deux principes à sa façon particulière.

### **2.2.1. Aspect phonographique**

L'écriture chinoise se présente sous forme d'une suite d'éléments séparés les uns des autres par des espaces égaux. Ces éléments sont le noyau de l'écriture chinoise, et nous les appelons « caractères ». Selon la définition d'Alleton (2002 : 11), les caractères chinois sont « des formes graphiques indépendantes, isolées matériellement les unes

des autres par un espace, et invariable en ce sens que leur tracé ne change pas, quelles que soient les formes environnantes ». Ces caractères sont invariables, c'est-à-dire leur tracé ne change pas selon les formes qui les précèdent ou les suivent. Dans l'écriture chinoise, cette unité graphique minimale « correspond à une syllabe et a pour signifié un morphème » (Alleton, 2002 : 188). Cependant, pour une langue alphabétique, une unité minimale, c'est la lettre, qui a pour signifié un son, ou idéalement un phonème. A ce titre, un caractère chinois est équivalent à un morphème ou un mot.

Au niveau phonologique, l'unité essentielle, en chinois, c'est la syllabe, qui est également invariable<sup>16</sup>. De ce fait, le chinois est une langue syllabique. Bien qu'on puisse analyser les mots chinois en phonème, l'organisation de la syllabe est un élément capital de toute description. Contrairement à ce qui se passe dans une langue comme le français, où les phonèmes se combinent assez librement entre eux, les phonèmes en chinois « s'organisent en une séquence définie dans la syllabe, celle-ci étant, de plus, affectée d'un ton » (Alleton, 2002 : 12).

Les caractères chinois comportant un élément phonétique sont appelés *xíngshēngzì* (phonogrammes). Sagart (2006) explique la formation de ces phonogrammes de la façon suivante :

« Les *xíngshēngzì* trouvent leur origine dans la pratique plus ancienne des *jiǎjiè* 假借, caractères employés pour noter des mots de même prononciation que leur référent original. Ainsi le caractère archaïque pour "femme", *nǚ* 女 dont la prononciation était \*<sup>b</sup>nra? en chinois archaïque, était aussi utilisé pour noter par *jiǎjiè* d'autres mots de prononciation voisine, tels le pronom "tu, toi" *rú* (\*<sup>b</sup>na? ), le mot *nú* "esclave" (\*<sup>a</sup>na), le mot "comme" *ru<sub>2</sub>* (<sup>b</sup>na), et d'autres encore. Dans l'écriture moderne, les caractères correspondant à ces mots contiennent tous le caractère *nǚ* 女 comme phonétique. On voit bien ici la continuité entre l'ensemble des mots

---

<sup>16</sup> Comme le décrit Alleton (2002), la syllabe ne connaît pas de modifications fonctionnelles, mais la séquence phonique est insensible à ce qui l'entoure.

notés habituellement au moyen d'un même *jiǎjiè* dans l'écriture chinoise ancienne et l'ensemble des caractères comprenant le même élément phonétique dans l'écriture moderne ».

Contrairement aux *jiǎjiè*, les *xíngshēngzì* comportent un élément sémantique, dont la fonction est de différencier les graphies homophones de sens différent. Le tableau suivant permet de montrer une série de *xíngshēngzì* :

**Tableau 2:** phonogrammes avec la phonétique de base : zhǔ 主 (Drocourt, 2015)

Prononciation	zhù	zhù	zhù	zhù	zhù	zhù	zhǔ
Graphie	住	注	柱	驻	炷	蛀	拄
Sens du sémantique	homme	eau	bois	cheval	feu	ver	main
Sens du morphème	habiter	verser	colonne	camper	encens	ronger	s'appuyer

Dans l'écriture chinoise moderne, les *xíngshēngzì* sont considérablement majoritaires, représentant jusqu'à 90% de la totalité des caractères attestés. Cependant, si les *xíngshēngzì* sont fondés sur « une sorte de syllabaire tout en cherchant à différencier les graphies des homophones », le principe de ce syllabaire est « imparfait » dès son origine (Drocourt, 2015). Sagart (2006) explique les imperfections de la manière suivante :

« Les scripteurs de l'antiquité disposaient d'un jeu de phonétiques pour représenter la prononciation désirée des caractères. L'ensemble de ces phonétiques formait une sorte de syllabaire imparfait, dans lequel certaines syllabes avaient plusieurs phonétiques possibles, et d'autres aucun. Lorsque plusieurs phonétiques étaient possibles pour la même syllabe, il y avait une tendance à spécialiser chacun dans certains détails de prononciation tels que

le type de syllabe, l'aspiration de l'initiale, ou encore la présence ou l'absence de certains affixes. Les scripteurs antiques recherchaient avant tout la meilleure adéquation phonétique possible entre la prononciation d'un caractère et celle de son phonétique. Il leur arrivait certes de créer des caractères contenant des phonétiques décrivant mal la prononciation du mot qu'ils transcrivaient...mais c'était toujours faute de mieux, faute d'un meilleur phonétique ».

Notons également qu'en chinoise moderne, la plupart d'*xíngshēngzì* sont créés à l'époque des Royaumes Combattants et Qin. Et la prononciation de référence, issue du chinois archaïque, est bien différente du chinois standard moderne. Ainsi, avec l'existence de tons en chinois, les éléments phonétiques n'indiquent la prononciation des *xíngshēngzì* que d'une façon approximative. Tous ces éléments empêchent « le système de s'orienter vers une phonographie plus économique et plus transparente » (Drocourt, 2015).

En un mot, la phonographie de l'orthographe du chinois est bien différente de celle des langues alphabétiques. Bien que dans les langues comme le français, un graphème puisse correspondre à plusieurs phonèmes et vice versa, l'indication donnée par un graphème est immédiate, c'est-à-dire que pour connaître la prononciation d'un graphème, on n'a pas besoin de considérer d'autres signes graphiques. En tant que système graphique, l'écriture chinoise n'est pas idéographique et son phonétisme « n'en est pas pour autant absent ni totalement opaque » (Drocourt, 2015)

### 2.2.2. Aspect sémiographique

La sémiographie de l'orthographe du chinois est basée pour l'essentiel sur les indicateurs sémantiques, qui joue surtout un « rôle lexicographique en permettant le classement de caractères souvent homophones » (Fayol & Jaffré, 2008 : 100). Cependant, ces indices sémantiques ne sont pas directs non plus. Comme l'explique Alleton, certains de ces éléments sémantiques peuvent être utilisés isolément avec un

sens propre. Néanmoins, comme nous l'avons montré dans le tableau 2, quand les indicateurs sémantiques forment seulement une partie du caractère, le sens qu'ils ont comme caractère indépendant n'est pas toujours pertinent. Autrement dit, le sens d'*xíngshēngzì* est « celui du morphème auquel il renvoie – qui n'est pas la résultante d'une composition d'éléments » (Alleton, 2002 : 189).

Par conséquent, nous ne pouvons accéder au sens d'un caractère que par l'identification de ce caractère et connaître le sens des éléments qui constituent un caractère ne nous permet pas de savoir ce que ce caractère désigne. Par exemple, un caractère chinois est constitué de la coexistence d'un élément référé généralement à la notion de femme et d'un élément associé à la notion de toit. A partir du sens de ces deux éléments, on aurait peu de chance de deviner que ce caractère désigne la paix.

Il s'ensuit que le signe graphique du chinois, le caractère, renvoie à la totalité du signe linguistique, c'est-à-dire à la fois à une forme phonique et à un sens.

### 2.2.3. Apprentissage de l'écriture chinoise

Observons bien la différence entre l'orthographe du chinois et les orthographes alphabétiques par les processus d'apprentissage de l'une et de l'autre. Si l'on prend le français comme exemple, l'apprentissage du système alphabétique commence par la découverte de vingt-six formes graphiques et l'apprentissage des correspondances lettre-son. Un enfant français de moins de quatre ans est rarement capable de lire parce qu'il ne pourra pas lire « avant d'avoir saisi ce mécanisme relativement abstrait et avant de connaître toutes les lettres ou presque » (Alleton, 2002 : 15). En revanche, dès qu'il a acquis cette technique, il sera en mesure de prononcer tous les mots en français qu'il rencontrera. La connaissance des nouveaux mots par la suite dépend de ses conversations et du biais de textes écrits. Nous appelons ces procédés « lecture globale ».

Par contre, un enfant chinois, dès l'âge de deux ans, est capable de lire quelques caractères. Néanmoins, comme en chinois, une syllabe donnée s'écrit d'autant de façons différentes qu'elle a de sens distincts : savoir lire un caractère pour chaque

syllabe de son parlé ne sert pas à grand-chose pour l'enfant. Notons que la connaissance de plus de mille caractères permet à l'enfant chinois de lire facilement. Cependant, si un chinois, sachant même plus de cinq mille caractères, rencontre un caractère qu'il ne connaît pas, il n'arrive pas à déduire ni la prononciation, ni le sens de ce caractère. Par analogie avec des caractères qu'il connaît déjà, l'enfant chinois peut tout au plus faire une hypothèse sur quelques lectures plausibles et le genre de chose signifiée.

De surcroît, comme langue morphosyllabique, dont la variation morphographique est minimale, le chinois est appris d'une façon assez particulière (Robert, 2002) :

- L'apprentissage de l'écriture prend beaucoup de temps ;
- L'ordre des mots constitue la base quasi unique de la syntaxe ;
- Les procédés fondamentaux d'apprentissage s'appuient essentiellement sur l'imitation d'un modèle.

Ce modèle d'apprentissage, qui est basé sur la mémorisation par l'écriture, et qui manque de préparation élémentaire de métalangage grammatical, n'est pas sans incidence sur l'apprentissage d'une langue alphabétique chez les locuteurs chinois.

### **2.3. L'apprentissage de l'orthographe pour un locuteur chinois**

#### **2.3.1. Les étapes majeures pour apprendre une orthographe alphabétique**

Dans l'apprentissage d'une orthographe donnée, la mise en évidence de phases fiables est difficile à définir, puisque l'apprentissage dépend largement de la conjonction de trois éléments (évoqués au-dessus) : développement de l'apprenant, système linguistique particulier et instruction dispensée. Néanmoins, la plupart des chercheurs s'accordent sur un découpage en trois périodes dans l'apprentissage d'une orthographe alphabétique, tel que celui du français (Fayol & Jaffré, 2008 : 168) : Très tôt, vers trois ou quatre ans, les enfants se trouvent dans une phase préalphabétique, et ils commencent à distinguer l'écriture du dessin et se font une idée, même vague de ce

qu'est la lecture. Ensuite, les enfants sont à la recherche du principe alphabétique et des correspondances élémentaires entre graphèmes et phonèmes et entre phonèmes et graphèmes. Enfin, dans la troisième phase, les enfants prennent en compte au fur et à mesure les contraintes orthographiques, comme des régularités d'associations de séries de lettres ou de la forme spécifique de certains mots.

Il s'avère cependant que la situation d'apprentissage de l'orthographe des apprenants chinois se différencie de celle des enfants francophones sur plusieurs points (Hedbor & Ågren, 2006). Une différence fondamentale concerne le cadre linguistique des deux formes d'apprentissage : les apprenants chinois (pour la plupart au moins) entreprennent l'apprentissage du français oral et écrit simultanément, alors que les enfants français apprennent à parler beaucoup plus précocement qu'ils n'apprennent à écrire. A ce titre, la procédure phonocentrée, une première façon d'orthographier presque uniquement centrée sur l'oral, semble absente chez les apprenants chinois. Par ailleurs, comme les apprenants chinois sont plus âgés au moment d'apprendre le français, ils font preuve d'un niveau plus élevé de capacité cognitive. Ces divergences entraînent évidemment des différences en ce qui concerne l'acquisition du système orthographique chez les apprenants chinois.

Selon la psychologie cognitive, la production écrite en langue étrangère manifeste également des différences. Comme le décrit Barbier (1998), le même individu, qui possède une capacité cognitive limitée, ne peut utiliser cette capacité de la même façon en langue étrangère et en langue maternelle. En langue maternelle, les processus de bas niveaux, liés à la mobilisation des connaissances linguistiques (lexique, grammaire, etc.) sont automatisés, alors qu'en langue étrangère, ces connaissances ne sont que partiellement automatisées en fonction du degré de maîtrise de la langue. Ainsi, « le degré de mobilisation et d'efficacité de ces processus serait restreint en situation de rédaction en langue étrangère » (Barbier, 1997 : 363). Par exemple, la production écrite en langue étrangère est plus coûteuse en temps, parce que l'apprenant passe plus de temps à mobiliser ses compétences de bas niveau, tel que la récupération des mots ou la mise en œuvre de la grammaire. Il s'ensuit que « la production écrite en langue étrangère ne peut pas toujours être prise en charge de la même façon qu'en langue

maternelle ».

### 2.3.2. Apprentissage de l'orthographe lexicale

Dans les systèmes orthographiques transparents, l'acquisition du principe alphabétique et des correspondances entre phonèmes et graphèmes permet aux locuteurs de lire et d'écrire la totalité des mots. En l'occurrence, la disposition d'un lexique orthographique, ou dictionnaire mental, notion dont nous avons parlé précédemment, ne serait pas utile. Cependant, il s'avère que même dans les orthographe régulières tel que l'espagnol ou l'italien, les locuteurs disposent de ce lexique orthographique pour lire ou écrire plus rapidement. Certes, les locuteurs s'appuient moins sur l'utilisation de ce lexique que ceux ayant les systèmes alphabétiques opaques.

A partir des travaux qui ont été menés sur l'apprentissage de l'orthographe en langue maternelle, quelques éléments pourraient être donnés sur l'acquisition de l'orthographe lexicale (Fayol & Jaffré, 2008 : 183-195) : il faut noter tout d'abord que les effets de fréquence et d'analogie (voir par exemple, Bosse et *al.*, 2003 ; Martinet et *al.*, 2004) sont mis en évidence très précocement. Autrement dit, les enfants mémorisent très tôt les formes orthographiques. Ainsi, un lexique orthographique se constitue chez les enfants, dont la taille augmente progressivement avec la pratique de lecture des enfants. Simultanément, les enfants retirent du corpus qu'ils lisent des régularités statistiques consistant en « unités sous-lexicales reposant essentiellement sur les associations entre lettres contigües »<sup>17</sup>. Signalons également que la connaissance de ces régularités reste implicite et inconsciente. On voit ainsi que les processus d'orthographier les mots, les représentations et leurs fonctionnements sont identiques chez les enfants et les adultes.

---

<sup>17</sup> Les enfants acquièrent les régularités dites graphotactiques de manière implicite avant même d'avoir compris leur fonction. Par exemple, en français, au moins trois graphèmes : *o*, *au* et *eau* peuvent transcrire le phonème [o]. Cette polyvalence graphique rend la production écrite difficile. Mais la recherche des régularités statistiques de l'orthographe indique que le graphème *eau* n'apparaît jamais au début de mot tandis qu'il se trouve souvent en fin de mot. Cette régularité est déjà notée par les enfants en première primaire et ces derniers sont capables de la mettre en œuvre dans leur production, afin de produire les mots inconnus.



Deux questions doivent être posées alors. D'une part, s'il existe deux types de représentation lexicale : un lexique mental des vrais mots et un lexique des unités sous-lexicales, de quelle manière fonctionnent-ils ensemble ? S'ils fonctionnent en synergie, comment résoudre des conflits alors que la forme lexicale et les unités sous-lexicales ne concordent pas ? D'autre part, avons-nous besoin de postuler l'existence d'un lexique mental ? Notons que dans certaines occurrences, la forme lexicale mémorisée subit des instabilités. Les expériences montrent qu'il est facile d'induire des erreurs chez des adultes qui ne les commettent pas. De ce fait, dans la représentation mentale, il pourrait n'être qu'une combinaison de fragments plus ou moins souvent associés par l'usage et plus ou moins rapides à mobiliser.

La réflexion sur ces questions pourrait ouvrir de nouvelles perspectives en ce qui concerne l'acquisition de l'orthographe lexicale dans le domaine du FLE, qui a reçu peu d'attention de la part des chercheurs. Une conception dominante dans l'enseignement du français en Chine tend à réduire l'apprentissage de l'orthographe lexicale à un processus d'enrichissement de la connaissance spécifiques des mots. Pour bien orthographier, les apprenants doivent transcrire les nouveaux mots plusieurs fois pour que leur représentation orthographique soit stockée dans la mémoire. Ce travail monotone et décontextualisé ignore le fait que le processus d'enrichissement des connaissances orthographiques est progressif et qu'il résulte de plusieurs facteurs : le développement de l'usage de ces connaissances, des évolutions stratégiques et enfin la formation de représentations orthographiques dans le lexique.

### 2.3.3. Apprentissage de la morphologie

L'acquisition de la morphologie joue un rôle crucial pour la compréhension et pour la production de nombreuses langues écrites, surtout dans le cas du français, où le système de l'orthographe se caractérise par de nombreuses marques morphologiques qui n'ont pas de contrepartie phonologique. Linguistiquement, la morphologie se répartit en deux catégories : la morphologie flexionnelle, qui concerne des accords en nombre et en genre ainsi que des temps du verbe ; la morphologie dérivationnelle, qui permet de

fabriquer des mots nouveaux à partir d'une base et d'affixes, préfixes et suffixes.

L'enseignement du français donné en Chine consacre un temps considérable à la morphologie, en tenant compte que le chinois possède une variation morphographique minimale. Signalons que la dimension morphologique fait l'objet d'un enseignement explicite en Chine. L'attention des apprenants est attirée par l'enseignant sur les marques du nombre, du genre, les désinences des verbes conjugués, et leurs associations avec les catégories syntaxiques. L'instruction est directe : elle fait appel à un métalangage permettant de catégoriser les mots en noms, adjectifs, verbes, et à un effort de mémorisation des apprenants<sup>18</sup>. Autrement dit, « une fois que ces relations entre classes syntaxiques et marques sont constatées et énoncées sous forme de règles explicites, les apprenants ont à réaliser des exercices (compléter, transformer, transcrire) pour les consolider en mémoire » (Fayol & Jaffré, 2008 : 200).

Un enseignement ainsi dispensé pose une série de problèmes. Fayol et Jaffré mettent en question l'efficacité relative de cet enseignement orthographique. Dans les cours de langue, les apprenants effectuent les exercices d'application de règles donc l'objectif est la réalisation d'une orthographe correcte, tandis que la production écrite réelle implique le plus souvent la rédaction d'un texte. Dans une telle activité, l'attention de l'apprenant porte essentiellement sur le contenu, l'organisation du texte, le choix lexical et par conséquent, la forme orthographique n'est plus sa préoccupation principale. Ainsi surgissent les erreurs orthographiques et cela donne aux enseignants le sentiment que l'instruction dispensée n'est pas efficace.

Néanmoins, il faut noter que la survenue d'erreurs est probablement inévitable, puisque le scripteur ne peut pas gérer simultanément toutes les dimensions de la production écrite en raison de la capacité limitée d'attention ou de mémoire. De ce point

---

<sup>18</sup> Si l'on prend l'exemple de conjugaison de verbe, les verbes se répartissent en trois catégories : le premier groupe (verbes terminés par *-er*), le deuxième groupe (verbes terminés par *-ir*) et le troisième groupe (verbes conjugués irrégulièrement). Comme les verbes des deux premiers groupes se conjuguent régulièrement, le travail des apprenants consiste à prendre les règles de conjugaison par cœur. Quant aux verbes irréguliers, les apprenants cherchent à retenir les conjugaisons de toutes les personnes grammaticales de chaque verbe rencontré par cœur. Dans les manuels de français édités en Chine, il existe des tableaux de conjugaison des verbes bien organisés.

de vue résulte une double nécessité : apprendre à relire afin de détecter et de corriger les erreurs ; être en mesure de mobiliser moins d'attention pour la réalisation orthographique.

Sur l'acquisition de la morphologie en langue maternelle, aussi bien flexionnelle que dérivationnelle, les chercheurs inclinent à l'idée que deux processus fonctionnent de façon indissociable (Fayol & Jaffré, 2008 : 198-210) : d'un côté, la mémorisation de formes orthographiques fréquentes, éventuellement sous leur forme fléchie ; de l'autre côté, l'extraction de régularités permettant la résolution partielle de la transcription des mots.

Néanmoins, en langue étrangère, le débat sur l'acquisition de la morphologie est dominé par une discussion sur « la représentation mentale et le traitement des mots composés de plusieurs morphèmes » (Hedbor & Ågren, 2006). Le premier cas est défendu par le courant théorique connexionniste. Comme la mémoire est un réseau associatif puissant, dynamique et hautement structurée, tous les mots y sont gardés individuellement et « interconnectés entre eux dans des schémas associatifs basés sur des traits sémantiques, phonétiques et morphologiques » (Hedbor & Ågren, 2006). A ce titre, le traitement morphologique des apprenants est particulièrement influencé par la fréquence d'un mot. Le deuxième cas est défendu par l'idée dualiste de l'existence d'un double système morphologique (voir par exemple, Pinker, 1998 ; Clahsen, 1999) : le processus régulier s'appuie sur l'application de règles morphologiques, alors que le processus irrégulier repose sur la récupération directe en mémoire d'unités « toutes faites ».

Les travaux de Hedbor et Ågren (2006) semblent fournir des éléments de réponse. Les deux chercheurs ont abordé le travail d'analyser la façon de s'appropriier la morphologie flexionnelle du nombre à l'écrit chez les apprenants suédois de FLE. A partir de leurs observations, deux types de traitement morphographique existent chez les apprenants de FLE : l'application de la règle et la récupération d'instances pour les formes les plus fréquentes. Précisément, les chercheurs ont observé que les apprenants produisent plus facilement l'accord nominal (-s) que l'accord verbal (-nt) et que l'accord nominal du pluriel est appris assez rapidement, sans que les apprenants en

soient conscients.

Selon Hedbor et Ågren, ce phénomène pourrait s'expliquer principalement par deux facteurs : d'une part, un caractère sémantique est motivé par l'accord nominal, c'est-à-dire pour les apprenants, le -s est porteur d'une différenciation de sens qui joue un rôle important pour la compréhension du message, tandis que le morphème -nt ne change en rien l'action verbale. La difficulté de respecter l'accord en nombre de l'adjectif, qui est observée en FLE, contribue à renforcer cette explication. En absence de contenu sémantique, les apprenants éprouvent des difficultés avec l'accord adjectif, bien que la marque du pluriel soit la même dans le cas du nom et de l'adjectif. D'autre part, les possibilités de transfert pourraient aussi expliquer les performances des apprenants. La langue suédoise possède un accord nominal en nombre à l'oral et à l'écrit, alors qu'il ne distingue morphologiquement aucune personne grammaticale sur le verbe.

Évidemment, ces éléments ne sont pas sans incidence sur l'apprentissage de la morphologie en FLE et ils contribuent à expliquer l'acquisition précoce de l'accord nominal. A partir de l'analyse de transcription des mots irréguliers, les chercheurs ont vu une tendance à généraliser le modèle flexionnel régulier dans le domaine du verbe (racine + nt, *\*ils veulent*) et du nom (racine + s, *\*les cadeaux*). Ce phénomène est interprétable à la fois par une approche connexionniste et dualiste. Pour le premier, les morphèmes -s et -nt seraient si fréquents dans l'input et dans la production des apprenants qu'ils « formeraient des schémas associatifs très puissants qui contamineraient même des items qui, normalement, ne font pas partie du groupe régulier ». Pour la seconde approche, ce phénomène serait un suremploi de la règle par défaut. Autrement dit, la règle de l'accord en nombre s'applique à tout item pour les apprenants sans connaître l'existence des exceptions.

# III. PARTIE MÉTHODOLOGIQUE

## 3. Méthodologie

Avant d'aborder les écrits des apprenants chinois, il nous semble pertinent de nous expliquer d'abord sur les choix méthodologiques. En vue d'atteindre les objectifs décrits dans l'introduction de cette étude, il a été indispensable de créer un corpus écrit permettant d'observer les pratiques orthographiques des apprenants chinois du français L2. Étant donné que la création d'un corpus relève de maints paramètres pratiques et méthodologiques, nous entamons, dans ce chapitre, une discussion sur les considérations méthodologiques, ainsi que sur les choix des méthodes pour la constitution du corpus et les analyses de données. Cette discussion nous semble tout à fait importante pour la compréhension de ce travail et pour l'interprétation des données.

### 3.1. Considérations méthodologiques

Les données sont fondamentales dans les recherches en science du langage. Au cours de notre travail sur les données nous avons dû prendre position par rapport aux différents aspects méthodologiques que nous discutons brièvement ici. Premièrement, le recueil des données (3.1.1.) : il s'agit d'adopter une méthode adéquate aux objectifs de notre recherche tout en tenant compte des contraintes et problématiques propres aux apprenants du français L2. Nous évoquerons ensuite les problèmes méthodologiques qui s'imposent à l'analyse des données (3.1.2.) et les choix théoriques ayant permis leur description et la mise au jour des résultats.

#### 3.1.1. Un dispositif pour observer les performances au cours de la production écrite

La production langagière est « une habileté spécifiquement humaine » (Segui & Ferrand, 2000). Si la production orale est extrêmement rapide et efficace (Fayol, 2002),

l'écriture demande plus d'effort à des scripteurs que celle à l'oral, et elle « s'avère laborieuse et parfois avec un succès qui n'est pas garanti » (Bonin, 2002). Notons que la production verbale sous la modalité écrite n'a fait, comme l'écrit Fayol (2007), que très récemment l'objet de recherches scientifiques, plus récemment encore quand il s'agit de son acquisition, contrairement à l'intérêt porté à la production verbale orale ou à la lecture. Malgré ce relatif retard, les études consacrées à la production verbale écrite ont augmenté de façon significative au cours des trois dernières décennies. Dans le cadre spécifique du français, le groupe de recherche GDR CNRS 2567<sup>19</sup> se situe dans ce contexte, en fédérant les différents chercheurs nationaux impliqués dans l'étude de la Production Verbale Écrite. Toutefois, dans le cadre spécifique du français L2, les recherches conduites sur la production écrite sont en effet relativement rares quand le système graphique de la langue source est très éloigné de celui du français, tels que le chinois et le français, que rien ne semble rapprocher.

Comment les locuteurs de chinois font-ils pour apprendre à orthographier en français ? Comment gèrent-ils les différences entre deux systèmes graphiques si opposés ? Comment s'approprient-ils le français dans « un milieu scolaire (universitaire) complètement coupé du contexte social de la langue cible » (Trevisiol, 2003) ? Ce fut le point de départ qui a motivé notre travail. Nous souhaitons ainsi créer un dispositif pour observer les performances orthographiques des apprenants chinois étudiant le français en Chine.

### **3.1.1.1. Processus de mise en texte**

Depuis le modèle fondateur de Hayes et Flower (1980), il est généralement admis que « la production verbale écrite met en œuvre des processus récursifs qui sollicitent une activité langagière complexe et de haut niveau » (Plane, Rondelli & Venerin, 2013). Parmi d'abondants travaux sur la production écrite, on voit un intérêt particulier porté sur l'un des processus d'écriture, la révision, surtout dans le champ de la psychologie

---

<sup>19</sup> <http://www.gdr-pve.fr/>

cognitive, qui a pour objectif d' « affiner les formalisations de l'activité scripturale en s'appuyant sur des expérimentations menées dans un cadre très contrôlé » (Plane, 2015). Pourtant, l'autre grand processus, la textualisation est rarement abordé ou du moins partiellement étudiés par les chercheurs. On voit notamment les recherches menées par les linguistes à travers les analyses génétiques des brouillons d'écrivains (Lumbroso, 2004 ; Hamon, 2009) ou des textes d'élèves (Fenoglio & Boucheron-Pétillon, 2002 ; Doquet-Lacoste, 2003). En essayant de reconstituer la genèse d'un texte, ces études se sont notamment attachées à caractériser les singularités des scripteurs, et il nous semble donc particulièrement intéressant d'enrichir les connaissances portant sur les variations dans les processus de mise en texte par des apprenants-scripteurs, pour identifier plus précisément les différentes manières dont ils procèdent dans ce processus et des types de difficultés rencontrées, afin, dans une perspective didactique, de construire des dispositifs d'enseignement différenciés et appropriés (Plane, Rondelli & Venerin, 2013).

### **3.1.1.2. Interlangue**

Si l'étude porte sur les variations dans les processus de mise en texte par des apprenants-scripteurs d'une langue étrangère, le travail serait enrichi par un autre élément très important, l'interlangue, qui est, selon Selinker (1972), « un système linguistique productif par référence aux deux autres systèmes linguistiques que constituent la langue maternelle de l'apprenant et la langue étrangère » (Rosen & Porquier, 2003). De ce fait, pour mettre en lumière les processus impliqués dans la production langagière chez un apprenant allophone, il faut prendre en compte les types et la quantité de l'input, ainsi que la perception de l'apprenant de l'input à travers ses connaissances linguistiques préalables (Perdue, 1984 :26), qui illustre la complexité des facteurs à considérer quand il s'agit d'expliquer l'acquisition d'une langue étrangère. Comme l'indique Benazzo (2005), « la prise en compte d'un seul type de facteur, qu'ils soient d'ordre purement linguistique ou d'ordre purement cognitif », ne permet pas de mettre en lumière l'acquisition d'un élément de la langue étrangère. C'est l'interaction de ces facteurs, variable suivant le stade atteint par l'apprenant, qui apporte des éclaircissements sur le

processus acquisitionnel.

### **3.1.1.3. Deux contraintes mises en jeu dans la production de texte**

Dans la production écrite, il existe deux niveaux de difficultés auxquelles les apprenants-scripteurs d'une langue étrangère seraient confrontés : d'une part celles que pose le système linguistique, venant de l'input, ainsi que de leurs connaissances linguistiques préalables ; d'une part celles de la psycholinguistique, venant de la tâche d'écriture, d'autre part. Ces difficultés sont antagonistes, non hiérarchisables entre elles, et offrent des degrés de liberté différents (Plane, Olive & Alamargot, 2010).

### **3.1.1.4. Bilan**

Étant donné les deux niveaux de contraintes mises en jeux dans la production écrite, il nous semble important de proposer un dispositif empruntant aux deux champs, pour observer les variations orthographiques produite par les apprenants chinois du français L2 dans une production écrite réelle. En vue d'éliminer, dans la mesure que possible des variables interférentes, un certain nombre de facteurs doivent être pris en compte lors de la création du dispositif pour recueillir les données :

- La langue maternelle de tous les apprenants est la même et ils ont approximativement le même âge.
- Les apprenants ont un profil linguistique similaire en ce qui concerne les langues étrangères qu'ils ont maîtrisées.
- Les apprenants choisis dans la recherche étudient le français dans un même établissement.
- A partir d'une perspective développementale, il nous semble important de recueillir les données auprès les apprenants se trouvant à des niveaux langagiers différents.
- Les textes écrits recueillis doivent être issus de la même tâche d'écriture.
- Pour faciliter une analyse interindividuelle, il nous semble intéressant que les



apprenants fassent une tâche d'écriture demi-spontanée à partir d'un même matériau écrit.

- Tous les textes doivent être écrits dans une salle de classe, avec la présence du professeur des apprenants.
- L'expérimentateur doit être le même.
- Pour diminuer le stress des apprenants devant une tâche difficile, l'accent doit être mis sur l'importance du contenu et non de la forme doit être accentuée.
- Il faut préciser que la tâche à accomplir n'est pas évaluée. Le résultat ne comptant pas pour la note du français.
- Le temps pour accomplir la tâche doit être le même pour les différents groupes d'apprenants.

En examinant leurs textes écrits, ainsi que leurs performances orthographiques, nous souhaitons identifier les traitements orthographiques mis en œuvre par les apprenants chinois lors d'une production écrite en français, et caractériser l'interlangue de ces apprenants dans la production écrite par sa variabilité et sa systématisme.

### 3.1.2. Réflexions sur l'analyse des données en interlangue

Dans le domaine de l'acquisition et de l'apprentissage des langues secondes, les différents types d'analyse des données sont utilisés par les chercheurs afin de mieux comprendre le processus d'acquisition. Si les théories adoptées par les études qualitatives constituent souvent un biais lié à la subjectivité du chercheur, les méthodes, quantitatives et le recours aux statistiques tentent de réduire le biais du traitement des données. Toutefois, ces méthodes ont elles-mêmes leurs limites pour interpréter les données. De fait, l'articulation entre les données et la méthodologie ne s'avère pas toujours aller de soi. Cela nous oblige à questionner sans cesse la méthodologie à travers les objectifs de la recherche. Telle est l'orientation que nous entendons donner à la partie suivante, en nous attachant particulièrement à une présentation des méthodes différentes utilisées pour le traitement des données dans le domaine de l'acquisition de

langues secondes.

### **3.1.2.1. Analyse comparative**

La linguistique contrastive est née dans les années 1950. En considérant l'interférence causée par les différences structurelles entre la langue maternelle de l'apprenant et une langue étrangère comme un des principaux obstacles à l'apprentissage de cette langue étrangère, les premières études contrastives sont suscitées dans une perspective d'application. Fries (1945: 9) indique que « the most effecient materials are those that are based upon a scientific description of the language to be learned, carefully compared with a parallel description of the native language of the learner ».

Notons que la linguistique contrastive est étroitement liée à son utilité. Comme l'indique Debyser (1970), les études contrastives « peuvent rencontrer en chemin des problèmes théoriques intéressants, permettre la vérification d'hypothèses, et susciter des descriptions qui manquaient mais ne se justifient véritablement que par les services qu'elles peuvent rendre ». Malgré que l'accent des études contrastives soit mis sur la diversité des langues, la théorie est une et permet, en outre, non seulement la description mais aussi la comparaison. En d'autres termes, ces études disposent d' « un instrument théorique et métalinguistique assez général et assez unitaire » pour rendre comparable des langues différentes (Debyser, 1970).

La linguistique contrastive suppose au départ que l'apprentissage d'une langue étrangère ne pose pas les mêmes problèmes que l'apprentissage de la langue maternelle. C'est dans le cadre de cette hypothèse psycholinguistique fondamentale que les recherches contrastives sont entreprises. Afin de répondre aux besoins d'une pédagogie spécifique des langues étrangères, la linguistique contrastive s'appuie principalement sur les études de la nature et du rôle des erreurs dans l'apprentissage d'une langue étrangère. Autrement dit, elle a pour l'objectif « de prévoir, de décrire et d'expliquer les erreurs et les difficultés » rencontrées par les apprenants d'une langue étrangère dues à l'influence de leur langue maternelle. En partant de la conviction que les différences linguistiques pourraient être utilisées pour prévoir les difficultés

d'apprenants, Weinreich (1953 : 1) affirme que plus la distance entre langue cible et langue maternelle est grande, plus les formes et les modèles s'excluant mutuellement dans les deux systèmes sont nombreux, plus les problèmes d'apprentissage de la langue cible s'élèvent. Autrement dit, dans le domaine où les deux langues sont similaires, un transfert positif se produirait, alors qu'un transfert négatif ou une interférence aurait lieu dans le domaine où les deux systèmes se distinguent l'un de l'autre.

Cependant, il faut remarquer que les analyses contrastives ne sont pas que des listes de prédictions binaires sous la forme suivante : similitude/différence = facilité/difficulté (Larsen-Freeman & H.Long, 1990 : 53). Le tableau suivant présente une version simplifiée de la hiérarchie de difficulté (Stockwell, Bowen & Martin, 1965a), illustrant une analyse contrastive sophistiquée :

**Tableau 3.** Hiérarchie de difficulté (Larsen-Freeman & H.Long, 1990 : 54)

Type de difficulté	L1 : anglais ; L2 : espagnol	Exemple
Séparé		
Nouveau		Faire l'accord du genre grammatical
Absent		<i>Do</i> comme une marque de temps
Combiné		His/her est réalisé comme une seule forme <i>su</i>
Correspondance		<i>-ing = -ndo</i>

Les chercheurs nous montrent, à partir de ce tableau, que le point linguistique le plus facile à maîtriser serait celui où la langue cible correspond à la langue maternelle au niveau de structure et de sens. Progressivement plus difficiles sont les points

linguistiques conflués, où plusieurs formes de L1 convergent en une en L2, la forme qui se présente en L1 mais est absente en L2, ainsi que le point linguistique tout nouveau en L2. Le domaine le plus difficile à maîtriser relève de la division, où une forme en L1 se traduit par deux ou plusieurs formes en L2.

Ajoutons également que dans le cadre de la linguistique contrastive, le travail de recherche comporte systématiquement trois étapes : la première réside à analyser d'une manière précise des structures de chaque langue. La seconde étape à réaliser est d'observer le comportement des deux langues, dans l'exploitation de leur matériel. Précisons ici que « l'investigation doit porter en priorité sur les traductions, c'est-à-dire sur les équivalences d'énoncés » (Delen Karragaç, 2012). La troisième étape est portée sur les prévisions des erreurs éventuelles. A la suite d'une interprétation des erreurs étudiées, certaines conclusions seront proposées.

Néanmoins, lorsque les prédictions provenant des analyses contrastives ont été finalement soumises à des tests empiriques, de graves lacunes ont été constatées (voir, par exemple, Chamot, 1978 ; Arabski, 1979). Comme l'indiquent Long et Sato (1984), « the most fatal flaw of the CAH was the dubious assumption that one could depend solely upon an analysis of a linguistic product to yield meaningful insight into a psycholinguistic process, i.e. second language learning » (cité par Larsen-Freeman & H.Long, 1990 : 56). Malgré les critiques, les analyses contrastives continuent à être menées dans le domaine de l'acquisition de langues étrangère, afin d'identifier les transferts positifs ou négatifs provenant de la langue maternelle.

### **3.1.2.2. Analyse d'erreurs**

En partant des résultats décevants des enquêtes empiriques qui cherchent à vérifier les hypothèses provenant de l'analyse contrastive, Wardhaugh (1970) propose une distinction entre une version forte et une version faible pour examiner ces hypothèses. Comme l'indiquent Larsen-Freeman et H.Long (1990 : 57), pour la version forte, la prédiction d'erreurs dans l'apprentissage d'une langue étrangère est basée sur une analyse contrastive a priori de la langue maternelle et de la langue cible. Et comme nous

l'avons mentionné plus haut, ce type de prédiction n'est pas toujours vérifiée. En revanche, dans la version faible, les chercheurs commencent par les erreurs de l'apprenant et expliquent un sous-ensemble de ces erreurs en faisant appel aux similitudes et aux différences existant entre langue maternelle et langue cible. Selon Larsen-Freeman et H.Long, les analyses contrastives pourraient ne pas être utiles a priori, mais elles posséderaient encore un pouvoir explicatif a posteriori. De ce fait, les chercheurs pourraient s'en servir dans une approche plus large d'interpréter les erreurs commises par l'apprenant, à savoir l'analyse des erreurs.

#### 3.1.2.2.1. Erreurs interlinguales *versus* erreurs intralinguales

Notons que l'apprenant L2 commet toujours les erreurs dues à l'impact de sa langue maternelle, et Richards (1971) les appelle erreur interlinguale. Néanmoins, il faut remarquer qu'un grand nombre d'erreurs similaires sont commises par les apprenants L2, indépendamment de leur langue maternelle. Il s'agit donc des erreurs intralinguales, provenant des caractéristiques intrinsèques d'une langue étrangère.

Corder (1967) soutient que les erreurs des apprenants sont de grand intérêt pour l'étude du processus de l'apprentissage de langues. En classant les erreurs commises par les apprenants, les chercheurs pourraient en déduire les stratégies d'apprentissage adoptées par eux. Ainsi sont motivées de maintes études sur les typologies d'erreurs<sup>20</sup> :

- Erreurs relevant de la surgénéralisation (Richards, 1971), où l'apprenant omet de respecter les limites d'une règle. Par exemple : \* *I wonder where are you going*. Ici, l'apprenant a probablement surgénéralisé la règle de l'inversion sujet-auxiliaire, et l'a appliquée incorrectement dans une relative.
- Erreurs relevant de la simplification (George, 1972), telles que l'omission de la marque du pluriel pour un nom se trouvant dans un contexte pluriel. Par exemple : \* *I study English for two year*.
- Erreurs relevant de la communication (Selinker, 1972). Ce type d'erreur se produit

---

<sup>20</sup> Nous avons repris les exemples utilisés dans le travail de Larsen-Freeman et H.Long (1990 : 59) pour illustrer les différents types d'erreur.

lorsque l'apprenant a recours à des stratégies communicatives. Par exemple, l'apprenant utilise la forme \**airball* pour le mot *balloon*. Dans ce cas, l'apprenant étiquette incorrectement l'objet « ballon », mais il communique avec succès le concept souhaité.

- Erreurs relevant d'une induction (Stenson, 1974). Souvent se produit ce type d'erreur lorsque deux éléments linguistiques sont présentés par l'enseignant d'une manière qui crée une confusion. Par exemple, \**She cries as if the baby* (pour *She cries like a baby*). Cette erreur s'explique probablement par le fait que l'expression *as if* est défini comme *like* par l'enseignant, et que l'enseignant n'a pas expliqué la modification structurelle nécessaire que demande l'expression *as if*.

#### 3.1.2.2.2. Faute *versus* erreur

Les tentatives d'identifier et de classer les erreurs commises par les apprenants sont d'une importance significative pour les recherches dans le domaine de l'acquisition langagière, puisqu'elles attribuent un nouveau statut aux erreurs (Larsen-Freeman et H.Long, 1990 : 59).

Corder (1967) fait une distinction entre la notion d'erreur et celle de faute. Les fautes sont liées à la performance, et causées par la fatigue, l'excitation, etc. On a par conséquent la possibilité de faire soi-même la correction parce qu'on connaît les règles. En revanche, les erreurs sont liées à la compétence, et relèvent d'une méconnaissance de règles. Autrement dit, l'erreur est souvent un produit involontaire, et l'apprenant ne peut pas la corriger lui-même. Selon Larsen-Freeman et H.Long, les erreurs sont un reflet du stade actuel du développement langagier de l'apprenant (1990 : 59). Ainsi, il est de mise de penser que l'appellation « erreur » renvoie à un intérêt pour le processus d'apprentissage langagier propre à chaque apprenant, et que « faute » est, en revanche, utilisé pour décrire une déviation à la norme.

#### 3.1.2.2.3. Perspective différente de l'analyse contrastive

A la différence de l'analyse contrastive, l'analyse des erreurs permet de réinterpréter le rôle de l'apprenant qui est confronté aux difficultés à apprendre une langue étrangère.

Les erreurs ne résultent davantage d'une intrusion des habitudes L1 sur lesquelles l'apprenant n'a aucun contrôle. Au lieu d'être receveur passif des inputs de la langue cible, l'apprenant est activement engagé dans le processus d'apprentissage : il traite les données, génère les hypothèses, puis les teste et les affine (Larsen-Freeman et H.Long, 1990 : 61). Et l'analyse des erreurs se base donc sur la description puis l'exploitation des traits observés concernant des erreurs produites par l'apprenant, afin d'expliquer le système intermédiaire propre à cet apprenant.

Néanmoins, il faut indiquer qu'en décrivant le processus d'apprentissage de langues étrangère, l'analyse des erreurs présente également les insuffisances. En se concentrant uniquement sur les erreurs, les chercheurs n'ont plus une vue d'ensemble : ils étudient les erreurs commises par l'apprenants, mais pas ses réussites. En outre, il est toujours difficile, voire impossible d'identifier la source unitaire d'une erreur. Etant donné les insuffisances que présente la perspective de l'analyse des erreurs, on constate son incorporation dans l'analyse de performances. Cette dernière vise à décrire les performances interlangagières de l'apprenant et elle ne se limite pas à analyser seulement les erreurs.

### **3.1.2.3. Analyse de performance et études d'interlangue**

Pour un apprenant L2, la performance concerne la mise en œuvre de ses connaissances linguistiques dans les activités communicatives. Il s'agit à la fois de l'émission, où l'apprenant fait des phrases, et de la réception, où l'apprenant comprend des phrases. Dans l'apprentissage de langues étrangère, la performance d'un apprenant est souvent considérée comme une indication indirecte de sa compétence (Richards, 2002). Si l'analyse des erreurs se concentre uniquement sur les produits langagiers de l'apprenant et constitue souvent un biais lié à la subjectivité du chercheur, l'analyse des performances, en tant qu'approche globale, vise à mettre au jour le processus d'apprentissage d'une langue étrangère.

Depuis les années 1970, les tentatives ont été faites pour effectuer une analyse linguistique de la langue de l'apprenant. Nous constatons notamment les études (voir,

par exemple, Dulay & Burt, 1973 ; 1974) sur l'ordre d'acquisition de morphèmes de l'anglais, qui mettent un accent fort sur le produit linguistique. A partir d'une approche plus éclairante, les chercheurs commencent à découvrir comment l'apprenant traite les données langagières afin de les acquérir. Ainsi le glissement général de l'intérêt vers le relevé des processus sous-jacents de l'apprentissage langagier fait-il une transition naturelle vers les études d'interlangue (Ringbom, 1994).

Dans les années 1990 les études d'interlangue connaissent un fort développement. Les sujets de recherche sont : le développement du langage de l'apprenant, les stratégies utilisées par l'apprenant, les variations systématiques de la langue de l'apprenant, et l'analyse du transfert (Ringbom, 1994).

#### **3.1.2.4. Bilan**

En retraçant le développement des différents modes de recherche en acquisition d'une langue seconde, nous constatons que chaque nouveau type d'analyse élargit notre perspective et apporte sa propre contribution. Ainsi, comme l'indiquent Larsen-Freeman et H.Long (1990 : 73), il serait injuste de dire que chaque type d'analyse remplace son prédécesseur. Au contraire, on estime que chaque type d'analyse subsume ce qui est venu avant.

Dans notre travail, nous souhaitons exploiter les données orthographiques des apprenants chinois à l'aide du logiciel *Le Trameur*. Les analyses seront concentrées sur les variétés diachroniques et synchroniques produites par les apprenants de niveaux scolaires différents. A partir d'une analyse quantitative, nous mènerons une enquête sur les performances des apprenants en fonction des différentes catégories d'orthographe, à savoir l'orthographe lexicale et l'orthographe grammaticale. En outre, une analyse qualitative des données nous permet d'identifier les contraintes linguistiques et cognitives possibles qui influencent les apprenants chinois. En faisant appel à des méthodes d'analyse différentes, nous souhaitons mettre en lumière le processus de la construction de connaissances orthographiques chez les apprenants chinois du français L2.



## 3.2. Corpus

En décrivant le corpus écrit analysé dans ce travail, il convient en premier lieu de présenter les apprenants chinois du français L2 participant dans l'étude ainsi que la tâche d'écriture utilisée pour recueillir des textes en français. Ensuite, nous aborderons les opérations à effectuer pour constituer le corpus, telles que la transcription et l'annotation des textes, la création de base pour le logiciel *Le Trameur*. Nous noterons également dans cette partie les difficultés que nous avons rencontrées lors de la préparation de base à l'aide du Trameur.

### 3.2.1. Description de la population

Au total, 99 étudiants chinois ont participé à l'expérimentation : 45 d'entre eux sont au début de leur deuxième année d'étude du français, 24 sont à la 3<sup>e</sup> année, et 30 sont à la 4<sup>e</sup> année.

Ces étudiants chinois sont tous dans le département de langue et littérature françaises de l'Université des Langues étrangères N° 2 de Pékin. La plupart d'entre eux ont commencé à apprendre l'anglais dès la première année de leur scolarité. Rappelons que le recrutement des élèves en Chine dans un lycée se fait après neuf années de scolarité obligatoire. Les élèves ayant terminé leur cursus au lycée peuvent accéder à l'université après un concours d'entrée très sélectif organisé par le Conseil Suprême de l'Enseignement Supérieur. Après avoir passé le concours, les élèves arrivent en général à l'université à l'âge de 18 ou 19 ans sans connaître un mot français. Ils commencent donc à apprendre de façon intensive le français.

Notons que ces étudiants ayant passé un concours national d'entrée à l'université viennent de différentes régions chinoises. Jusque vers les années 1990, la grande partie de la Chine connaît une situation diglossique : le chinois standard (*putonghua*), « joue le rôle de variété haute » ; et « la langue locale est variété basse » (Saillard & Boutet, 2008). Après une politique de véhicularisation de *putonghua* de près de cinquante ans, on constate actuellement en Chine un phénomène de « vernacularisation ». Saillard et

Boutet (2008) explique ce phénomène de façon suivante :

« On constate cependant depuis peu en Chine (Saillard, 2002 ; 2004) des phénomènes d'appropriation du *putonghua* par les locuteurs d'autres variétés du chinois. Cela se traduit par des mutations tant fonctionnelles (utilisation du *putonghua* pour des fonctions relevant de la vie privée) que formelles (emprunts morphosyntaxiques et lexicaux aux autres langues chinoises qui donnent lieu à des variétés locales de chinois standard). La diffusion et la vernacularisation du *putonghua*, langue véhiculaire à l'échelle de la Chine, sont tributaires non seulement d'un facteur spontané, la mobilité des populations ... mais aussi d'une planification gouvernementale visant à imposer le standard dans toutes les situations de la vie publique et, ce, notamment dans le domaine de l'éducation et des médias ».

Avec la promotion par intermédiaire des médias et de l'enseignement, le *putonghua* est considéré comme « la seule langue officielle de la scolarisation », ainsi que « le seul vecteur d'apprentissage des compétences écrites » (Saillard & Boutet, 2008). Par conséquent, on constate que les enfants chinois maîtrisent de plus en plus tôt le *putonghua*.

### 3.2.2. Réalisation d'une tâche d'écriture

#### 3.2.2.1. Protocole expérimental

L'expérimentation dont nous allons présenter les premiers éléments a été mise en place dans une perspective didactique, qui a pour visée de caractériser les modes d'écriture usités par les apprenants chinois au milieu scolaire. L'organisation de cette expérimentation a donc obéi à deux contraintes :

L'expérimentation s'est située dans une perspective écologique. Autrement dit,

elle s'est déroulée dans le cadre ordinaire des classes des apprenants, et avec la collaboration de leurs enseignants. Ce protocole interdit ainsi le recours à des matériels sophistiqués tels que les dispositifs de suivi oculaire, afin d'apporter le moins de perturbation possible, et conduit à proposer des activités proches de celles auxquelles les apprenants sont habitués (Plane, 2011).

L'expérimentation avait pour objectif d'obtenir les informations sur les différentes manières dont les scripteurs perçoivent et analysent les inputs, en utilisant leurs connaissances linguistiques préalables afin de produire leur propre texte. Cette visée interdit donc que les opérations des scripteurs interviennent sur des substrats hétérogènes, ce qui aurait été le cas « si la consigne d'écriture avait lassé aux scripteurs la liberté d'inventer de toutes pièces le texte qu'ils allaient écrire » (Plane, 2011). De ce fait, nous avons proposé aux apprenants participant à l'expérimentation un même substrat initial comme point de départ, à savoir un conte, dont nous avons décidé les caractéristiques.

### **3.2.2.2. Texte de source**

Le texte source utilisé pour cette expérimentation est issu de la légende, *Les chevaliers de la table ronde* (2010). L'extrait que nous avons choisi, *Perceval et le chevalier Vermeil*, a l'intérêt d'offrir aux apprenants chinois un univers exotique, qui ne fait pas partie de leur répertoire culture classique : Perceval le Gallois, le type du chevalier « aventureux », autour du roi Arthur et de la Table Ronde, est toujours prêt à se battre pour défendre les dames en difficulté. A travers tous les événements que ce conte narre, les apprenants chinois pourraient se plonger dans un monde étrange et remonter à des temps anciens. En revanche, il s'agit d'une version adaptée pour les jeunes locuteurs d'aujourd'hui, où la plupart des difficultés lexicales et syntaxiques sont déjà lissées.

Ce texte montre une grande diversité au niveau de temps verbaux. Dans la narration, les temps dominants sont le passé simple et l'imparfait, alors que dans le discours direct des personnages, nous trouvons le présent, le passé composé, ainsi que le future simple. Outre le mode de l'indicatif, on trouve des formes verbales au

subjonctif et au conditionnel. Il faut indiquer que nous avons élaboré deux versions de texte : l'une, destinée au groupe de débutants, dans laquelle nous avons remplacé systématiquement le passé simple utilisé dans la narration par le passé composé, en tenant compte que ce groupe des apprenants, ayant une année d'études de français, n'a pas encore appris le passé simple ; et l'autre, destinée à deux autres groupes de niveau scolaire plus haut, dans laquelle tous les temps verbaux sont gardés.

En outre, le texte est riche de propos des personnages, qui sont tous rapportés au discours direct, ce qui devrait nous permettre d'observer comment les apprenants font parler les personnages dans leur propre texte, qui est une tâche rédactionnelle à la fois cruciale et problématique.

Le dernier élément à noter concerne la variété lexicale. Dans le texte existent les mots se référant aux objets du moyen âge, tels qu'*épée* et *armure*, qui pourraient constituer un problème pour les apprenants chinois. Cependant, il nous semble intéressant de voir comment les apprenants récupèrent le sens de ces mots inconnus à partir du contexte, comment ils gèrent ces difficultés dans leur production de texte, qui pourraient transformer probablement les mots anciens aux objets quotidiens.

### **3.2.2.3. Déroulement**

Le dispositif comporte deux phases :

1. Lecture par les étudiants du texte du conte intitulé *Perceval et le chevalier Vermeil*

Le texte a été diffusé aux étudiants, avec une feuille blanche, dans laquelle les étudiants pourraient prendre les notes pendant la lecture.

Durée de la lecture : 20 minutes

2. Restitution par écrit du texte lu, immédiatement après la lecture

Les étudiants avaient reçu la consigne de lire attentivement le conte, de noter tous les éléments qui leur sont parus importants dans la feuille blanche, et de s'efforcer de mémoriser le conte car il leur serait demandé de le mettre par écrit sitôt après la lecture.

Durée de l'écriture : 30 minutes

#### **3.2.2.4. Consignes à donner**

Les consignes suivantes sont données en français aux apprenants chinois lors de la tâche d'écriture :

Vous disposez de 20 minutes à lire un texte. Il s'agit d'une histoire qui se passe aux temps anciens. Puis, vous aurez à produire un texte en vous appuyant sur cette histoire. Vous pouvez prendre les notes sur cette feuille blanche pendant votre lecture. On va reprendre le texte au bout de 20 minutes, mais vous pouvez garder vos notes. Puis, vous disposez de 30 minutes à produire votre propre texte.

#### **3.2.2.5. Recueil de copies**

Après la tâche d'écriture, nous avons collecté au total 92 textes, 40 textes venant des étudiants de la 2<sup>e</sup> année de licence (23 du premier groupe, 17 du deuxième groupe), 24 textes de la 3<sup>e</sup> année, et 28 textes de la 4<sup>e</sup> année. En raison du non-aboutissement de la tâche d'écriture, les textes produits par sept étudiants ne sont pas compris dans notre corpus final<sup>21</sup>.

### **3.2.3. Constitution du corpus**

#### **3.2.3.1. Procédures de transcription et d'annotation**

Ces textes écrits sont scannés et dénommés pour faciliter la suite de la recherche. Le nom doit indiquer :

---

<sup>21</sup> Parmi ces sept étudiants, 4 étudiants n'ont pas rendu leur texte produit après 30 minutes d'écriture ; un étudiant n'a produit qu'une phrase de dizaine de mots, sans commettre les erreurs orthographiques ; deux étudiants ont produit un texte en expliquant pourquoi ils ne voulaient pas faire cette tâche d'écriture.

- Le niveau de groupe : L2/L3/L4
- L'identification du groupe : 01/02
- Le numéro de l'étudiant dans le groupe : 001, 002...

Ce système de dénomination nous permet de spécifier l'origine de chaque écrit. Et les données ainsi numérisées constituent la première partie de notre corpus (V1).

Les données primaires sont par la suite transcrites via un éditeur de texte brut (Sublime Texte pour Mac) sous une forme au plus près de la copie scannée. Comme nous voudrions établir des transcriptions qui puissent être exploitées ensuite par *le Trameur*, en privilégiant la linéarisation du texte, certaines opérations faites par les étudiants, telles que suppression, ajout, remplacement, nécessitent un codage particulier. Ici, nous prenons la convention de transcription établie par le projet d'*Ecriscol*. Et les données ainsi informatisées constituent la deuxième partie de notre corpus (V2).

**Tableau 4:** Convention de transcription

Élément à transcrire	Proposition de codage
Suppression	[caractères supprimés]
Ajout	~~caractères ajoutés~~
Remplacement de texte	+R[met]+R+//se_mis//+R
Segment illisible	#xxx#
Retour à la ligne	§

Les données primaires donnent lieu également à l'établissement de la troisième partie de notre corpus (V3) : les annotations. Il s'agit ici d'annoter les éléments qui comportent les erreurs orthographiques. Les données primaires sont ainsi converties en données linguistiquement « normées », qui nous permettent à construire des bases textométriques pour *le Trameur*.

Pour annoter un mot ou une suite de mots dont on souhaite rectifier l'orthographe, on procède de la manière suivante :

- Isoler le segment à annoter entre deux signes { }
- Placer ensuite le tiret bas \_
- Écrire le segment normé, entre les deux signes { }

La proposition de codage pour toutes les ponctuations dans les annotations (virgule, point, apostrophe etc.) :

- {s'a}\_{s'est} : {s\_QUOTE\_a}\_{s\_QUOTE\_est}

### 3.2.3.2. Préparation de bases pour *le Trameur*

Avant d'importer les données secondaires dans *le Trameur*, il faut tout d'abord concaténer les 92 fichiers en un seul. Le fichier résultant est ensuite soumis à différentes procédures de nettoyage, formatage etc. Il nous faudrait également insérer la balise de partie sous la forme suivante :

<partie="L2">

<p="L2-01-001-B">

La base ainsi obtenue a un format compatible avec celui d'une base textométrique importable dans *le Trameur*. Comme la figure 1 l'illustre, chaque item dans le corpus est composé de 4 couches d'annotation pour le moment :

- Annotation n°1 : forme initiale. Dans notre corpus, la première couche d'annotation concerne la forme produite par les apprenants.
- Annotation n°2 : soit le lemme construit via *Treetagger*<sup>22</sup>, soit l'opération de

---

<sup>22</sup> Le programme Treetagger : « système d'étiquetage automatique des catégories grammaticales des mots avec lemmatisation. Il permet aussi de générer et de gérer des annotations multiples sur les unités du texte (et de traiter les niveaux d'annotations visés) » (<http://www.tal.univ-paris3.fr/trameur/>).

transformation réalisée sur la forme initiale. Par exemple, si l'apprenant a mis dans le texte *chevaliers*, ce sera donc le lemme *chevalier* mis dans la deuxième couche d'annotation.

- Annotation n°3 : catégorie grammaticale construite par *Treetagger*. Si nous reprenons l'exemple *chevaliers*, il s'agit donc de la catégorie du nom dans la troisième couche d'annotation.

- Annotation n°4 : forme annotée. Si la forme initiale produite par l'apprenant relève d'une erreur orthographique, cette forme erronée sera corrigée. Ainsi, ce sera la forme corrigée qui se trouve dans la quatrième couche d'annotation. En revanche, si la forme produite par l'apprenant est correcte, la forme de la première couche coïncide avec celle de la quatrième couche.

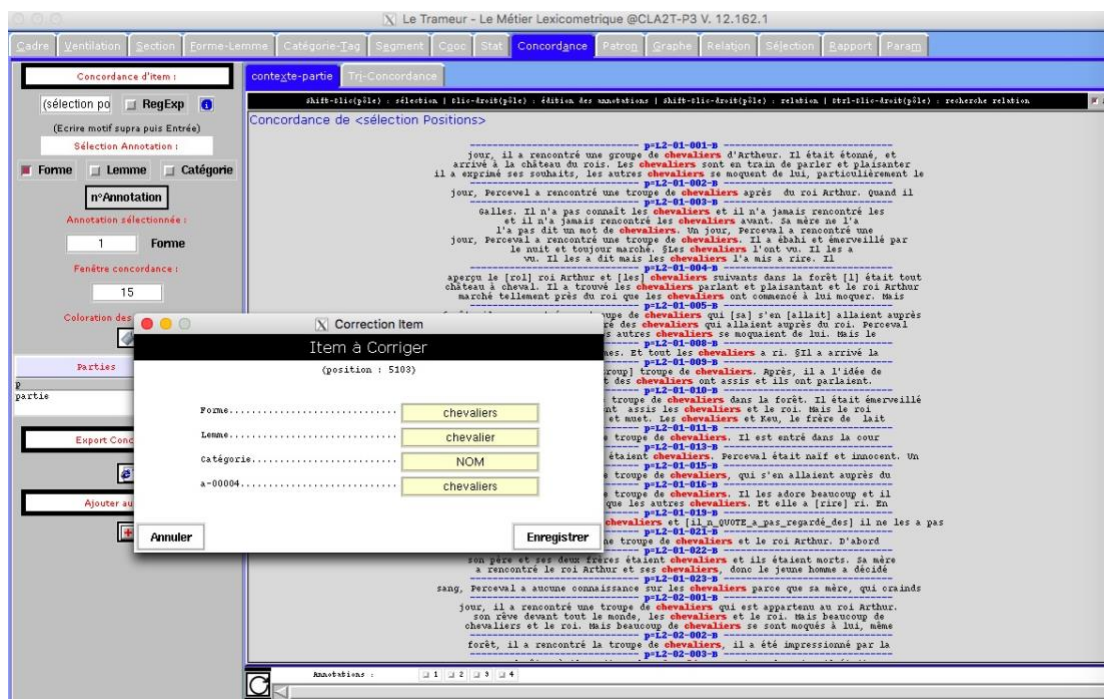


Fig. 1 : Quatre couches d'annotation pour l'item

Notons que la troisième couche d'annotation de notre base concerne les catégories grammaticales de mots, qui joue un rôle important dans l'analyse suivante. Cependant, pour un certain nombre de mots, leur catégorie ne peut pas être reconnue automatiquement par *le Trameur* pour les raisons suivantes :



- Majuscule : *le Trameur* groupe tous les mots commençant par une lettre majuscule dans la catégorie de nom propre. Cependant, la plupart des cas concerne la lettre initiale d'une phrase.

- Élision : en français, on observe la suppression de la voyelle finale d'un mot devant un mot commençant par une voyelle ou un *h* muet. Cette élision se marque par l'apostrophe. En raison de la suppression de la voyelle finale et l'ajout de l'apostrophe, *le Trameur* ne peut plus reconnaître la catégorie de mot. Par exemple :

(1) *On dit que la fille n'allait pas rire...*

Avec la suppression de la voyelle finale *e* et l'ajout de l'apostrophe, *le Trameur* ne peut plus reconnaître la catégorie du mot *ne*.

- Auxiliaire : pour le classement des catégories, nous observons une tendance à confondre l'auxiliaire qui sert à former les temps composés et le verbe principal de la phrase, surtout quand les deux éléments sont séparés. Par exemple :

(2) *\*Keu n'a pas voulu accepté le fait.*

Au lieu d'être classé comme auxiliaire, *a* est considéré comme verbe à l'indicatif présent, en raison de l'insertion de l'adverbe *pas*.

- Autres : si la catégorie d'un certain nombre de mots est mal reconnue, c'est parce que la phrase comporte une structure complexe. Par exemple :

(3) *\*Mais Perceval sauvé les autres avec les armés que le roi lui a donné.*

Dans cette phrase, le mot *que* introduit une relative subordonnée, et il s'agit donc d'un pronom relatif. Cependant, ce mot *que* est classé dans la catégorie de conjonction.

Le nettoyage de ces catégories mal reconnues est nécessaire quand on prépare les bases pour une analyse textuelle. Pour certaines erreurs qui se présentent d'une manière systématique, nous pouvons la corriger automatiquement dans *le Trameur*. Par exemple : le mot *chevalier* est systématiquement classé dans la catégorie d'adjectif. Avec *le Trameur*, on peut afficher une liste de toutes les occurrences du mot *chevalier*, et la correction de cette erreur sera appliquée automatiquement à toutes les occurrences.

Cependant, pour le problème posé par, par exemple, la majuscule de la lettre initiale de phrase, on est obligé d'intervenir à chaque occurrence. Au lieu de mettre la

lettre initiale en majuscule pour marquer le commencement d'une phrase, il nous semble pertinent de proposer un autre codage dans la suite de recherche, afin d'éviter un travail de nettoyage qui est assez coûteux en temps.

Le travail de nettoyage concerne également les mots qui pourraient appartenir à deux catégories grammaticales, notamment en ce qui concerne les mots outils, tels que *le, la, les*. Se situant avant le nom, ces mots monosyllabiques sont déterminants, tandis qu'ils sont classés dans la catégorie du pronom, en collant au verbe. Ainsi, chaque occurrence demande une analyse particulière du contexte syntaxique. Et si la catégorie grammaticale est mal reconnue, une intervention manuelle est obligatoire. Par exemple, dans la phrase *il s'est décidé de les joindre et d'être un chevalier*, le mot *les* doit être classé dans la catégorie de pronom :

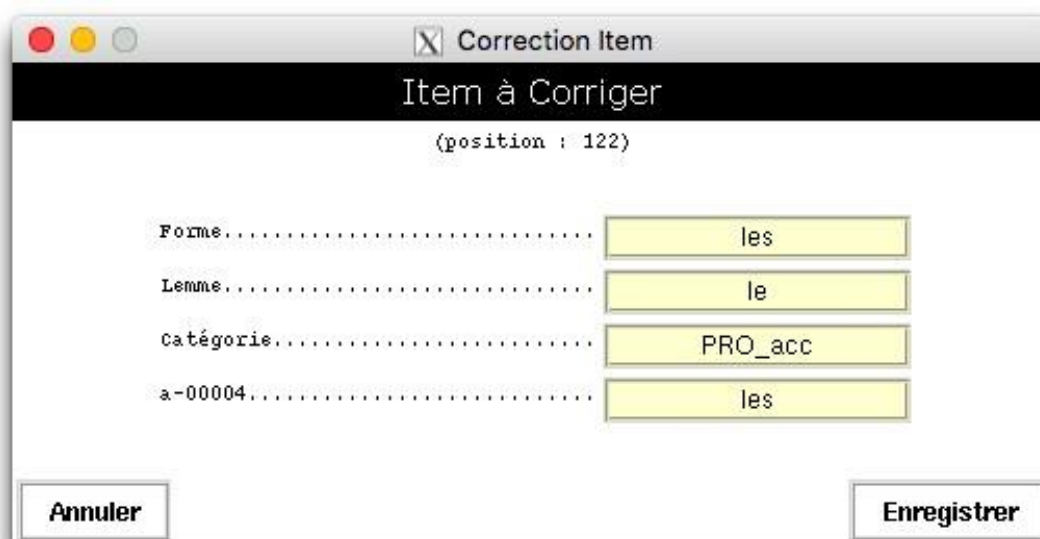


Fig. 2 : Correction des annotations.

### 3.3. Analyse des données

Dans la démarche d'analyse, nous lisons chacune des copies successivement et soulignons les mots qui comportent une erreur. Les formes erronées sont ensuite caractérisées. Pour dessiner le profil orthographique des apprenants de notre corpus, il nous paraît nécessaire de comprendre la nature des erreurs qu'ils ont commises et de

porter un regard aussi objectif que possible sur eux. Autrement dit, il faut donner un critère descriptif qui qualifie la nature de l'erreur. Nous établissons donc notre typologie des erreurs, en référence à la recherche de Manesse et *al.* (2007), qui s'appuie elle-même sur la typologie établie par Catach (2003). Le classement des erreurs de Manesse et *al.* est le suivant :

**Tableau 5.** Typologie des erreurs d'orthographe (Manesse & Cogis, 2007 : 88)

Types d'erreur	Définition	Exemples d'erreur
1	Mauvais découpage du mot ; mot sauté ou tronqué	Sans fonce
2	Aberration dans la représentation des sons	S'onvonce
3	Substitution de mot	Sont font
4	Cumul de fautes grammaticale et lexicale	S'enffons
5	Faute d'orthographe grammaticale ; la catégorie grammaticale n'est pas représentée ou est mal représentée	S'enfonce S'enfonces
6	Orthographe lexicale	S'enfonssent
7	Orthographe lexicale : forme approchante	S'enfonsent S'anfonsent
8	Signes orthographiques et majuscules	S'enfonçent

Notre corpus et celui de Manesse et *al.* présentent certaines similitudes : il s'agit des corpus à peu près homogènes. En d'autres termes, les deux corpus concernent des ensembles de textes qui ont répondu à la même consigne et dont les scripteurs sont proches d'un point de vue scolaire (Masseron & Luste-Chaa, 2008). Ce type de corpus présente donc un vocabulaire prédictible, y compris du point de vue de l'environnement morphosyntaxique. Néanmoins, nous apportons certaines modifications au classement d'erreurs conçu par Manesse et *al.* (2007), en nous basant sur les caractéristiques de notre corpus :

- Les textes recueillis dans notre corpus sont issus de la même tâche d'écriture semi-spontanée : après une lecture d'un conte intitulé « Perceval et le chevalier Vermeil » (20 minutes), les étudiants ont été invités à reconstituer par écrit le texte lu (30 minutes). A la différence du corpus de Manesse et *al.*, dont les textes ont été collectés après une dictée, les erreurs de type « mots sautés »<sup>23</sup> ne sont pas trouvées dans notre corpus.
- Selon le classement ci-dessus, les erreurs de type 2 regroupent les graphies dont la représentation phonétique est aberrante. Selon Manesse et al. (2007 : 73), ce type d'erreur concerne le mot qui est visiblement inconnu de l'élève et l'élève tente de reconstituer les sons de ce mot qu'il semble percevoir pour la première fois. Les erreurs de ce type sont également observées dans notre corpus. Cependant, étant donné la nature de la tâche d'écriture adopté par notre étude, ces graphies erronées ne pourraient pas être interprété par la méconnaissance de mots.
- Rappelons que l'apprenant chinois apprend simultanément le français oral et écrit, tandis que l'enfant français aborde l'écrit avec une maîtrise précoce du français oral. Ainsi, les erreurs orthographiques altérant la valeur phonétique de mots sont inévitables pour les apprenants chinois qui n'ont pas encore maîtrisé le système phonologique du français. Par conséquent, nous souhaitons apporter un affinement de la description des erreurs orthographiques en mettant l'accent sur cette spécificité des apprenants en français L2.
- Les erreurs de grammaire relèvent de deux domaines : les accords (en nombre et en genre) et la conjugaison du verbe. Le premier présente une grande homogénéité, car il s'agit de l'application de règles régulières, enseignées dès le début de l'enseignement du français. Les erreurs relevant de la conjugaison verbale sont, en revanche, plus hétérogènes et notre corpus présente une diversité importante de temps et de modes. Ainsi nous traiterons séparément les deux domaines de l'orthographe grammaticale.

---

<sup>23</sup> Il s'agit d'omission des mots dictés, qui touche généralement l'identification de ces mots dans la chaîne parlée.

- Selon la typologie ci-dessus, le dernier type d'erreur concerne les signes orthographiques (l'apostrophe, le trait d'union) et majuscules. Ces erreurs ne seront pas traitées dans notre corpus et les erreurs de l'accent sont classées dans la catégorie de l'orthographe lexicale.

Nous retenons au total quatre catégories, sept types d'erreurs. Nous proposerons un tableau récapitulatif dans la partie suivante.

### 3.3.1. Classement d'erreurs

**Tableau 6.** Classement d'erreurs d'orthographe

Types d'erreur	Définitions
1	Mauvais découpage du mot
2	Substitution de mot
3a	Erreur orthographique à dominante phonétique
3b	Erreur phonogrammique
3c	Erreur d'orthographe lexicale sans altération de la valeur phonique
4a	Erreur d'orthographe grammaticale : la marque du nombre ou du genre n'est pas présentée ou est mal présentée
4b	Erreurs d'orthographe grammaticale : les désinences verbales

Chaque forme erronée doit être rattachée à l'un des types d'erreur. Cependant, il arrive qu'une forme erronée comporte deux erreurs en même temps. Par exemple, la forme erronée \* *les consil* (pour *les conseils*) comporte une erreur d'orthographe lexicale ainsi qu'une erreur d'accord, et cette forme erronée sera donc recensée deux fois dans le relevé.

- Mauvais découpage du mot, qui touche généralement à la reconnaissance du mot comme entité dans la chaîne parlée. Les étudiants ne se rendent pas compte de l'existence indépendante d'une unité lexicale, soit qu'ils ne sachent pas la

séparer par des blancs de ses voisins de gauche ou de droite, soit qu'ils la tronçonnent en deux ou trois unités (Manesse & Cogis, 2007 : 73). Les erreurs de ce type relèvent surtout, selon Ros Dupont (2000), d'une difficulté de délimitation, et elle est « le fait des apprentis qui découvrent l'écrit ». Cela explique les formes erronées observées chez les apprenants chinois, qui commencent juste de découvrir l'écrit de la langue française. Le plus souvent, les erreurs surviennent au moment où les étudiants sont confrontés à une production d'écrit réelle : « la transcription d'un énoncé totalement pensé et proposé par lui, dont les éléments constitutifs ne leur sont pas forcément tous connus » (Ros Dupont 2000). Par exemple, L'erreur suivante relève de l'hypersegmentation : le verbe *embrassa* a été coupé en deux segments.

(4) \* *Sa mère {l\_QUOTE\_en\_brassa}\_{l\_QUOTE\_embrassa} et lui dit que {le}\_{la} mort honorable. (L3-004)*

- Substitution de mot, qui concerne la forme qui présente, correctement orthographiée, un autre mot que le mot attendu. Notons que ce n'est pas à proprement parler une erreur d'orthographe (Manesse & Cogis, 2007:74). S'agit-il toujours d'une mauvaise compréhension du sens du mot, ou une faute d'inattention assimilable à un lapsus. D'une manière générale, il est impossible de le dire. Par exemple, dans notre corpus, il existe des erreurs concernant les homophones *cours* et *cour*.

(5) \* *Finalelement Perceval a quitté {le}\_{la} {cours}\_{cour} de roi. (L2-02-013)*

Les deux mots ont la même prononciation mais les sens différents. Au niveau morphosyntaxique, ils sont tous des noms, mais ont les genres différents. L'orthographe de ce couple homophone lexical peut être déterminée par l'utilisation de mots de la même famille : la consonne finale *s* de *cours* peut être

déduite des mots comme *course*, *coursier*, *courser*, etc. Cependant, le stockage lexical des scripteurs de notre corpus vient d'être établi, soit ils ne connaissent pas encore les mots dérivés, soit ils ne sont pas capables de faire un lien entre eux. Dans ces cas, seul le recours à la mémoire leur permet de choisir la forme adéquate, mais chez le scripteur ci-dessus, cette mémoire est déficitaire.

Dans la catégorie des erreurs d'orthographe lexicale, il existe trois types d'erreur :

- Erreur orthographique à dominante phonétique, qui concerne les formes graphiques dont la représentation phonétique est aberrante. La mauvaise transcription des sons par les scripteurs nous amène à considérer que ces formes erronées sont bien loin de la pure et simple faute lexicale. Selon Cogis (2005 : 17-18), les erreurs de ce type, où la valeur phonique est dégradée, correspondent « à une erreur dans le langage “oral” » et relève d'« un travail sur l'audition et l'articulation ». Autrement dit, elles sont dues à une mauvaise production orale. Et pour y remédier, il faut assurer l'oral, afin d'asseoir, chez les scripteurs, une connaissance précise des différents phonèmes. Par exemple, la forme erronée ci-dessous cumule deux erreurs. Pour la première, il s'agit de la confusion de voyelles orales [a] et [u] ; quant à la deuxième, il s'agit de la confusion entre [u] et [ø].

(6) \* *Le frère d'Arthur est {jouleux}\_ {jaloux}*. (L2-01- 019)

- Erreur d'orthographe lexicale altérant la valeur phonique, qui relève du domaine de l'orthographe lexicale mais le choix des graphèmes 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 maîtrise imparfaite 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). Par exemple, la forme erronée ci-dessous concerne l'omission du *e* caduc où il possède une valeur auxiliaire, introduisant ainsi une modification de la valeur phonique. Plus précisément,

son absence ôte la prononciation aux consonnes qui le précèdent.

(7) \* *il a rencontré une {troup}\_ {troupe} de chevaliers.* (L2-02-010)

- Erreur d'orthographe lexicale sans altération de la valeur phonique. Cette catégorie ressemble également les erreurs relevant de l'orthographe lexicale mais il s'agit des erreurs qui ne modifient pas la prononciation des mots. Autrement dit, la forme produite par le scripteur est une des variantes possibles graphiquement mais non retenue par la norme orthographique (R.Honvault, 1995 : 61). Par exemple, l'erreur suivante concerne l'adjonction d'une lettre finale, qui n'est pas prononcée et n'influe non plus la prononciation du phonogramme voisin.

(8) \* *Elle a eu un {désire}\_ {désir} que Perceval peut parler avec le roi.* (L2-01-017)

Le type suivant regroupe les erreurs d'orthographe grammaticale, où la marque propre à la catégorie grammaticale n'est pas présentée ou est mal présentée. Les erreurs relèvent de deux domaines : le problème de l'accord, par exemple, \**le matins* (pour *le matin*), et celui de la conjugaison verbale, par exemple : \**la fille a rit* (pour *la fille a ri*).

### 3.3.2. Analyse quantitative à travers *le Trameur*

Notre travail a pour objectif d'étudier les comportements orthographiques des apprenants chinois et d'avoir une première vue de leurs lacunes et de leurs points forts lors de leur apprentissage. Pour ce faire, nous relevons les erreurs, les hésitations, les options manifestées chez les apprenants dans les copies recueillies. Dans la partie précédente, nous avons exposé les principes de classement des erreurs en huit types. Néanmoins, si nous souhaitons appréhender le système des erreurs, il faut observer la



répartition des erreurs en différents types et le poids relatif de chaque type. Ainsi, nous nous donnons des éléments « pour comprendre quels sont les secteurs faibles et les secteurs mieux maîtrisés de l'orthographe » (Manesse & Cogis, 2007 : 88). Afin de faire cette analyse quantitative, les textes écrits par les apprenants ont été soumis à différentes analyses automatiques proposées par Le Trameur ainsi qu'à certain nombre de calculs manuels. Dans cette partie, nous exposerons les analyses automatiques que l'on peut effectuer à l'aide du Trameur.

### 3.3.2.1. Extraire les éléments qui nous intéressent

Tout d'abord, *Le Trameur* nous permet d'extraire les formes qui nous intéressent d'une manière automatique. Si l'on prend l'exemple de l'accord du nombre, nous pouvons trouver tous les noms dont le contexte demande un accord du pluriel. Pour ce faire, il faut faire une recherche dans le secteur de sélection. Notons que ce secteur nous permet d'exploiter les données dans toutes quatre couches d'annotation. Afin de trouver toutes les occurrences de noms au pluriel, les consignes suivantes sont nécessaires :

- `.*[^s]$` dans le crochet de « lemme », qui nous permet de trouver tous les lemmes qui ne se terminent pas par un *-s*.
- NOM dans le crochet de « catégorie », qui indique que la recherche sera faite dans la catégorie de nom.
- `.*s$` dans le crochet de « a-00004 »<sup>24</sup>, qui sert à trouver toutes les formes se terminant par un *-s*.

---

<sup>24</sup> Dans notre corpus, la quatrième couche d'annotation concerne les formes finales ou corrigées. Plus précisément, si la forme produite par l'apprenant, qui se trouve dans la première couche d'annotation, est correcte, la forme de la première couche coïncide avec celle de la quatrième. En revanche, si la forme de la première couche est erronée, la quatrième présente donc la forme corrigée.

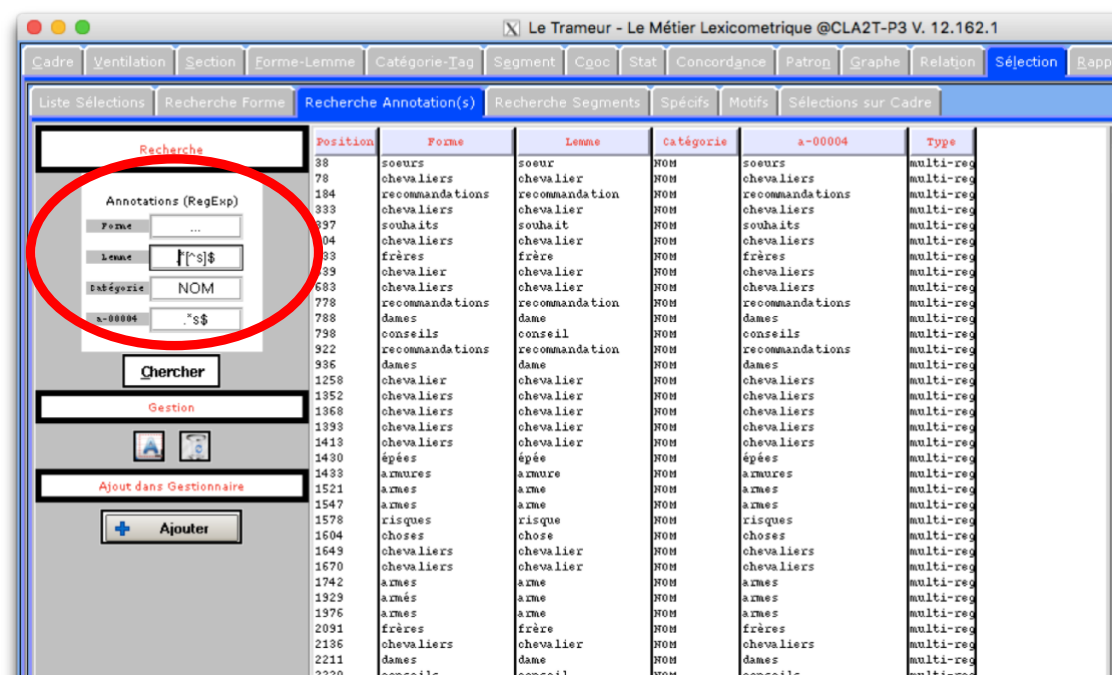


Fig. 3 : Lancement de l'enquête sur les noms au pluriel

Avec une telle opération, *le Trameur* nous présente une liste de mots, soit tous les noms qui demandent un accord du pluriel<sup>25</sup>. Ensuite, en ajoutant ces occurrences dans la « liste sélections », nous calculons leur nombre total.

<sup>25</sup> Dans ce cas nécessite une vérification, puisqu'il est tout à fait possible que dans notre corpus existent des mots se terminant par un -s et se trouvant dans un contexte du pluriel.

Le Trameur - Le Métier Lexicométrique @CLA2T-P3 V. 12.162.1

Cadre Ventilation Section Forme-Lemme Catégorie-Tag Segment Cocc Stat Concordance Patron Graphe Relation Sélection Rapport

Liste Sélections Recherche Forme Recherche Annotation(s) Recherche Segments Spécifs Motifs Sélections sur Cadre

Gestion		Position	Forme	Lemme	Catégorie	a-00004	Type
		38	soeurs	soeur	NOM	soeurs	multi-regexp (.
		78	chevaliers	chevalier	NOM	chevaliers	multi-regexp (.
		184	recommandations	recommandation	NOM	recommandations	multi-regexp (.
		333	chevaliers	chevalier	NOM	chevaliers	multi-regexp (.
		397	souhais	souhait	NOM	souhais	multi-regexp (.
		404	chevaliers	chevalier	NOM	chevaliers	multi-regexp (.
		633	frères	frère	NOM	frères	multi-regexp (.
		639	chevalier	chevalier	NOM	chevaliers	multi-regexp (.
		683	chevaliers	chevalier	NOM	chevaliers	multi-regexp (.
		778	recommandations	recommandation	NOM	recommandations	multi-regexp (.
		788	dames	dame	NOM	dames	multi-regexp (.
		798	conseils	conseil	NOM	conseils	multi-regexp (.
		922	recommandations	recommandation	NOM	recommandations	multi-regexp (.
		936	dames	dame	NOM	dames	multi-regexp (.
		1258	chevalier	chevalier	NOM	chevaliers	multi-regexp (.
		1352	chevaliers	chevalier	NOM	chevaliers	multi-regexp (.
		1368	chevaliers	chevalier	NOM	chevaliers	multi-regexp (.
		1393	chevaliers	chevalier	NOM	chevaliers	multi-regexp (.
		1413	chevaliers	chevalier	NOM	chevaliers	multi-regexp (.
		1430	épées	épée	NOM	épées	multi-regexp (.
		1433	armures	armure	NOM	armures	multi-regexp (.
		1521	armes	arme	NOM	armes	multi-regexp (.
		1547	armes	arme	NOM	armes	multi-regexp (.
		1578	risques	risque	NOM	risques	multi-regexp (.
		1604	choses	chose	NOM	choses	multi-regexp (.
		1549	chevaliers	chevalier	NOM	chevaliers	multi-regexp (.
		1670	armes	arme	NOM	armes	multi-regexp (.
		1742	armes	arme	NOM	armes	multi-regexp (.
		1929	armes	arme	NOM	armes	multi-regexp (.
		1976	armes	arme	NOM	armes	multi-regexp (.
		2091	frères	frère	NOM	frères	multi-regexp (.
		2136	chevaliers	chevalier	NOM	chevaliers	multi-regexp (.
		2211	dames	dame	NOM	dames	multi-regexp (.
		2220	conseils	conseil	NOM	conseils	multi-regexp (.
		2226	hommes	homme	NOM	hommes	multi-regexp (.
		2488	chevaliers	chevalier	NOM	chevaliers	multi-regexp (.
		2540	chevaliers	chevalier	NOM	chevaliers	multi-regexp (.
		2758	recommandations	recommandation	NOM	recommandations	multi-regexp (.
		2822	ans	an	NOM	ans	multi-regexp (.
		2929	armes	arme	NOM	armes	multi-regexp (.
		3206	frères	frère	NOM	frères	multi-regexp (.
		3239	chevaliers	chevalier	NOM	chevaliers	multi-regexp (.
		3401	armes	arme	NOM	armes	multi-regexp (.
		3464	armes	arme	NOM	armes	multi-regexp (.
		3585	chevaliers	chevalier	NOM	chevaliers	multi-regexp (.
		3610	épées	épée	NOM	épées	multi-regexp (.
		3614	armures	armure	NOM	armures	multi-regexp (.

Sélection par Nom

Nombre d'éléments dans le gestionnaire: 715

Fig. 4 : Occurrences de noms au pluriel

Comme le visualisent les deux figures suivantes (Fig. 5 et Fig. 6), nous pouvons également examiner le contexte de ces formes nominales ainsi que leur distribution en fonction de chaque apprenant ou chaque groupe d'apprenants<sup>26</sup>.

<sup>26</sup> La graphie affichée dans la deuxième image permet de visualiser le nombre de formes nominales au pluriel produit par chaque apprenant.

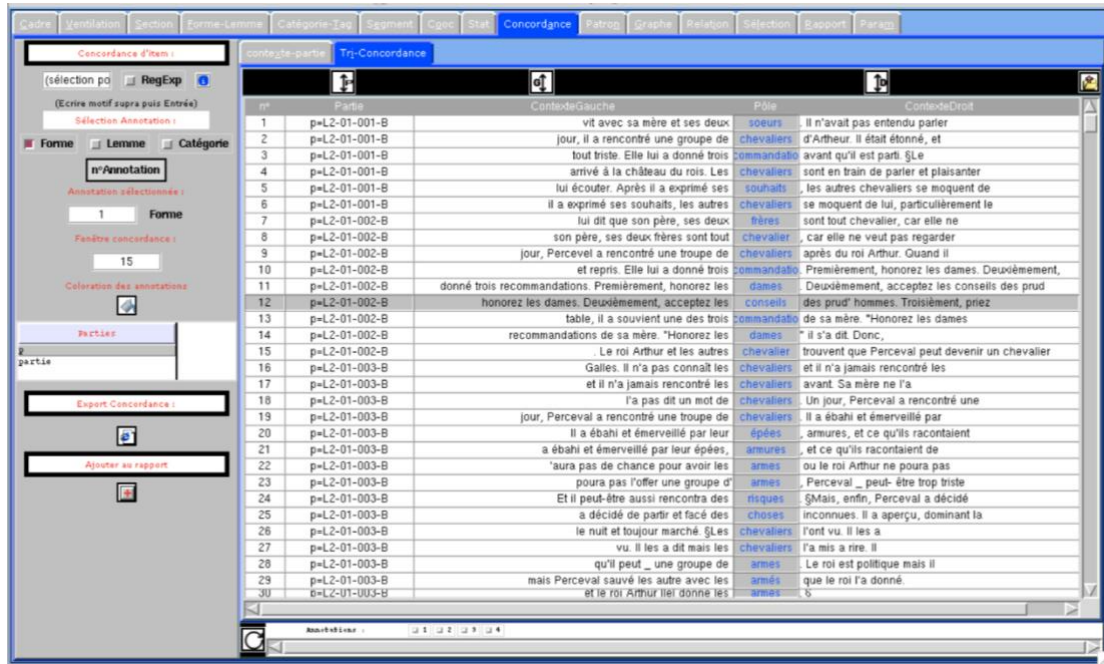


Fig. 5 : Contextes gauche et droit de noms au pluriel

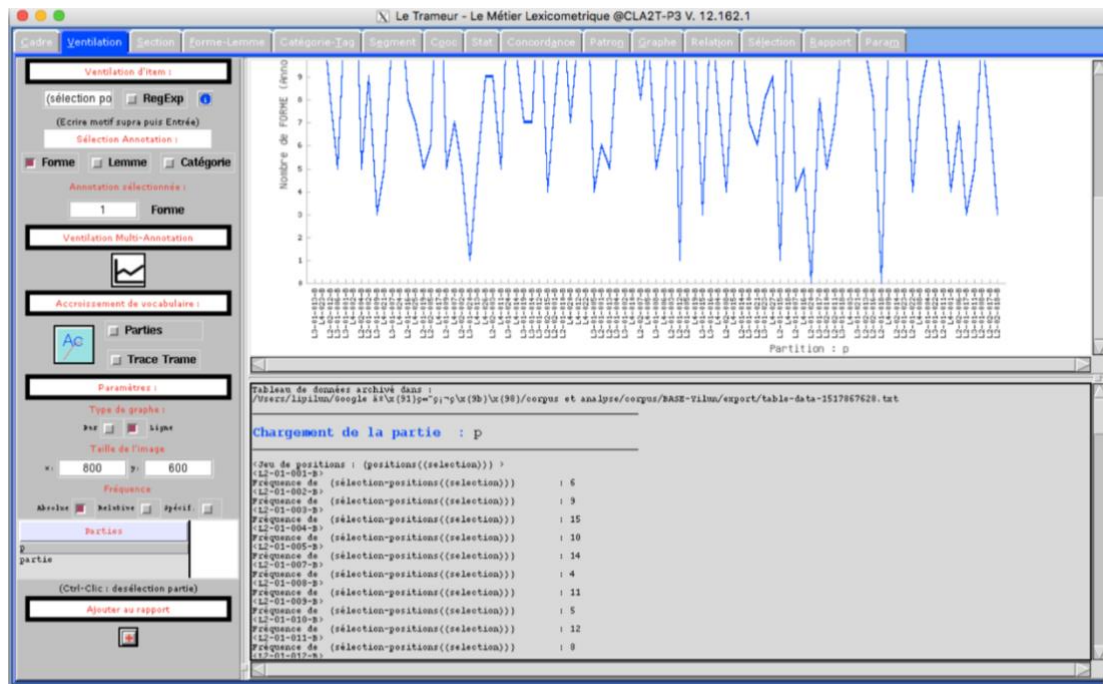


Fig. 6 : Distribution de noms au pluriel en fonction de niveau linguistique

### 3.3.2.2. Deux manières de repérer les formes non-normatives

*Le Trameur* nous permet également de repérer les formes non-normatives de manière

automatique. Si nous reprenons l'exemple de l'accord du nombre, nous pouvons trouver toutes les omissions de la marque du pluriel –s, en faisant appel à une recherche dans le secteur « sélection ». Les consignes seront suivantes :

- .\*[^s]\$ dans le crochet de « forme », qui nous permet de trouver tous les formes nominales produites par les apprenants qui ne se terminent pas par un –s.
- NOM dans le crochet de « catégorie », qui indique que la recherche sera faite dans la catégorie de nom.
- .\*s\$ dans le crochet de « a-00004 », qui sert à trouver toutes les formes se terminant par un –s.

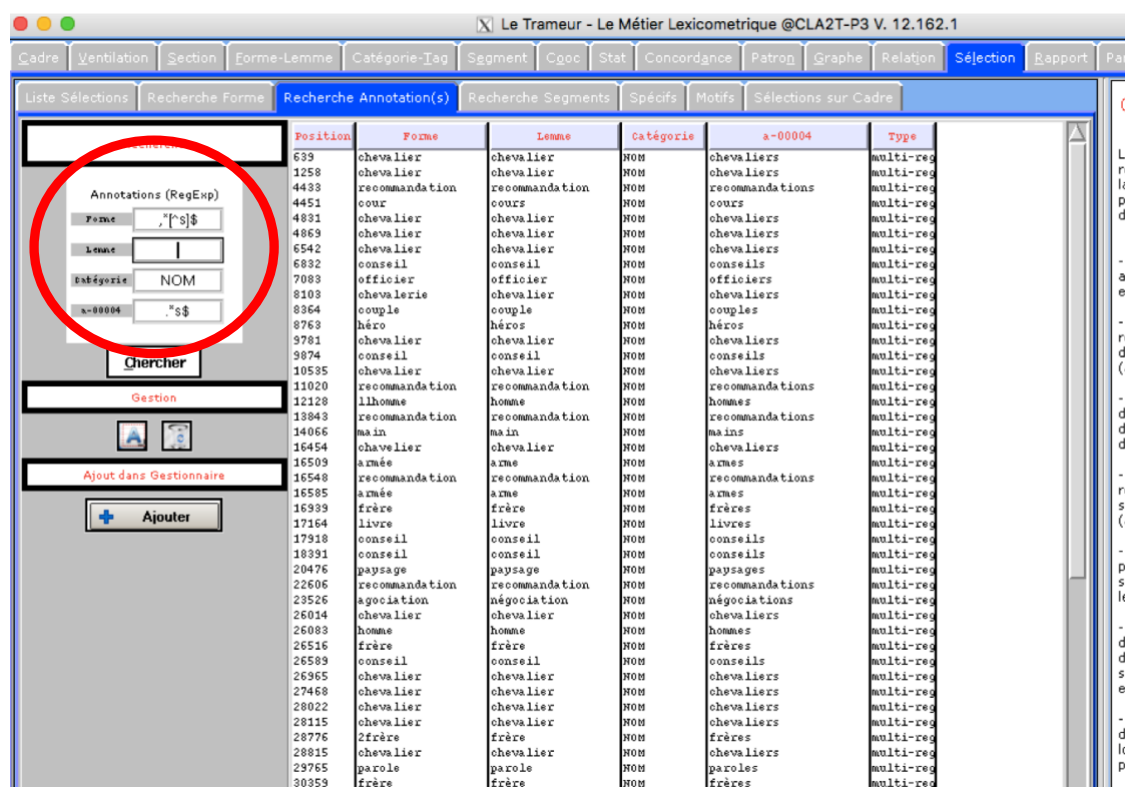


Fig. 7 : Lancement de la recherche de formes non-normatives

Une telle opération nous donnera une liste des occurrences dont la forme de la première couche d'annotation ne coïncide pas avec celle de la quatrième. De plus, la différence porte sur l'absence du morphème –s. Autrement dit, il s'agit des erreurs relevant de l'omission de la marque du pluriel –s. Remarquons qu'un travail de

nettoyage manuel sera nécessaire, puisqu'il existe des noms au singulier se terminant par un *-s* dans la liste de formes sélectionnées :

- (9) a. \* *ses deux frères sont {chevalier}\_ {chevaliers}*. (L2-01-002)  
b. \* *pour la première {foi}\_ {fois}* (L4-021)

En croisant le nombre de noms au pluriel et celui d'omission, nous pouvons déduire le taux de réussite de la morphologie nominale du nombre avec un simple calcul. Dans le secteur de « ventilation », nous pouvons également examiner la distribution des erreurs de l'omission en fonction de chaque apprenant ou chaque groupe d'apprenant. Parallèlement, une telle opération pourrait être effectuée pour d'autres catégories grammaticales, telles que déterminant, adjectif.

*Le Trameur* nous offre une autre possibilité de trouver les formes non-normatives sans préciser les caractéristiques de formes erronées. Nous nous concentrons cette fois sur la catégorie verbale, et souhaitons extraire toutes les formes verbales erronées produites par les apprenants dans ce corpus. Tout d'abord, comme l'illustre la figure 8, on crée une nouvelle couche d'annotation concaténant la forme produite par l'apprenant et la forme normative sous la forme X\_Y. Autrement dit, la forme de la première couche d'annotation et celle de la quatrième seront concaténées dans une nouvelle couche d'annotation, la cinquième.

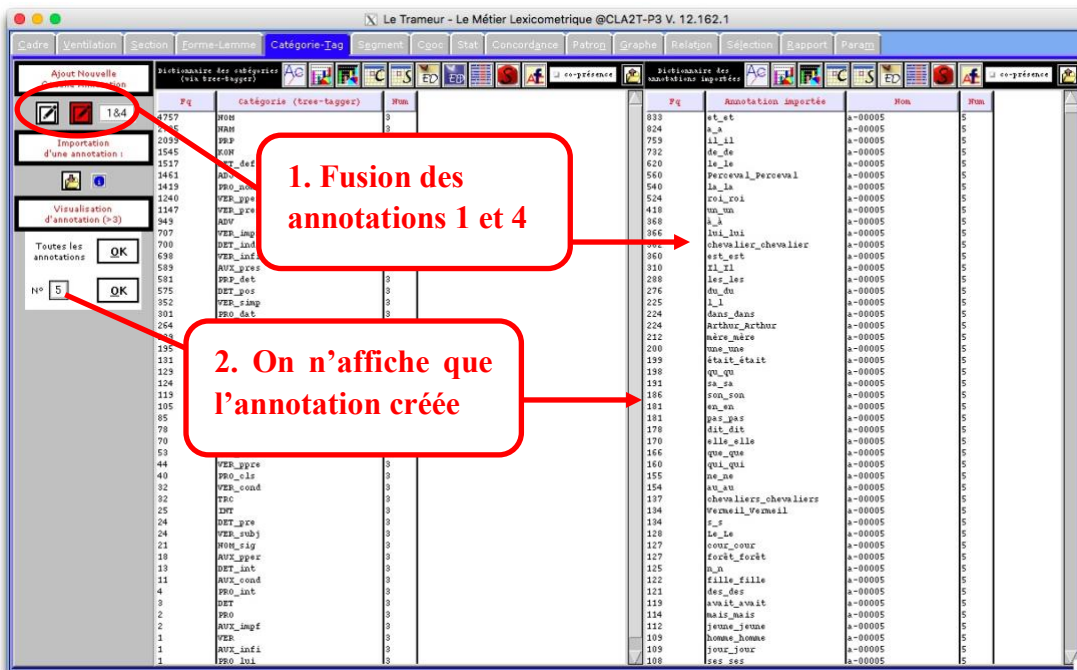


Fig. 8 : Création de la cinquième couche d'annotation

Deuxièmement, on cherche les formes verbales correctes : celles pour lesquelles  $X=Y$  dans la cinquième couche d'annotation. Et pour faire cette recherche, la consigne sera :  $(w+)\_1$ .

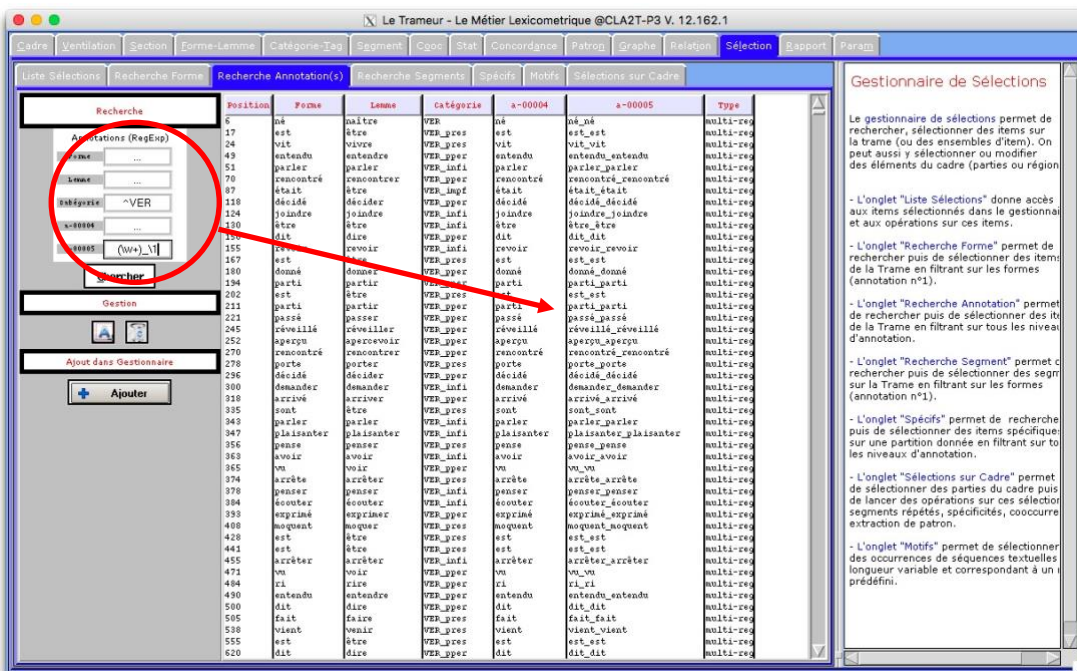


Fig. 9 : Recherche des formes normatives



A partir l'étape précédente, on a obtenu une liste de formes verbales correctes. On les sélectionne et les ajoute dans le secteur « liste sélections », où on crée une nouvelle annotation avec la valeur 0 pour ces formes verbes correctes.

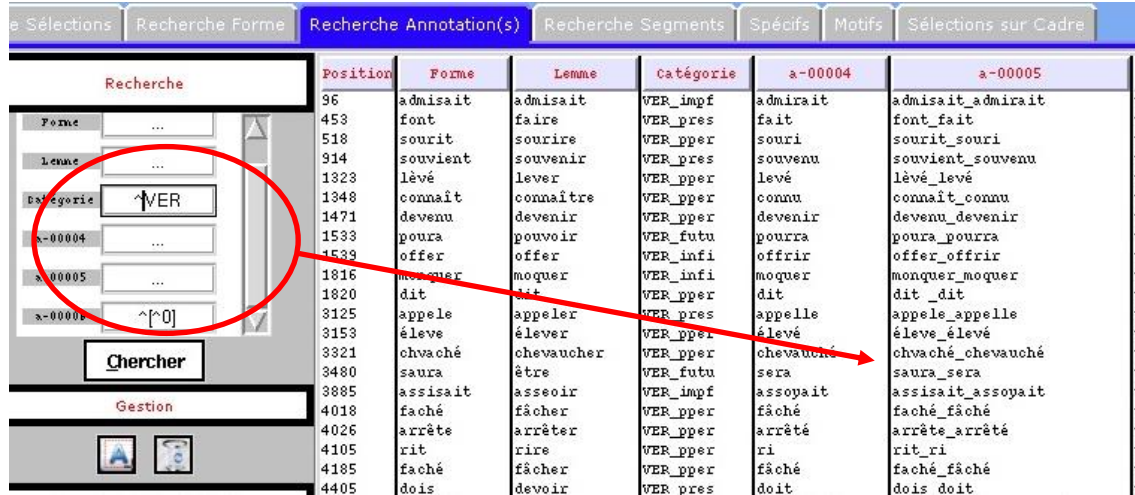


Fig. 10 : Création de la nouvelle annotation avec la valeur 0

Ainsi est créée la sixième couche d'annotation a-00006, dans laquelle les formes verbales normatives sont marquées avec la valeur 0. Dans l'étape suivante, afin d'extraire toutes les formes verbales non-normatives, on cherche les occurrences telles que l'annotation dans la sixième couche n'est pas 0. Et le consigne sera : `^[^0]` dans le crochet « a-00006 ».

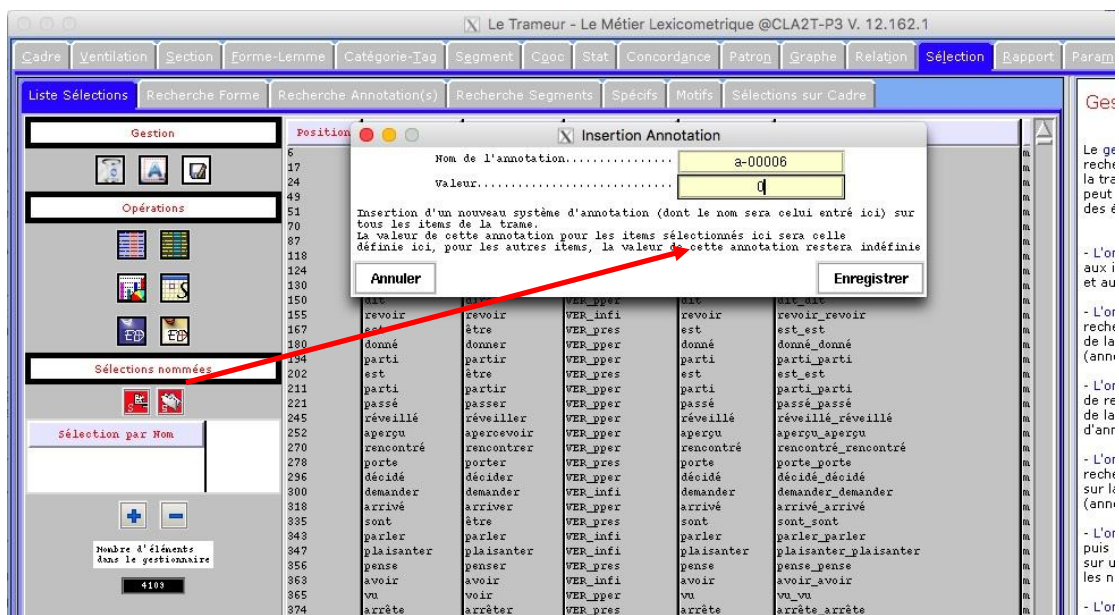


Fig. 11 : Création de la sixième couche d'annotation



Finalement on a obtenu la liste de formes verbales erronées. Comme le montre l'image suivante, on pourrait également exporter ces formes verbales et leurs concordances (contexte gauche et contexte droite) dans un fichier au format Excel, afin de faire d'autres opérations de calcul.

n°	Partie	ContexteGauche	Pôle	ContexteDroit
1	p=L2-01-003-B	le roi Arthur ne pourra pas l'	offer	une groupe d'armes, Perceval
2	p=L2-01-023-B	paroles ironique. Le roi Arthur s'est	fâcher	de sa grossièreté, en réclamant de
3	p=L2-01-023-B	d'or. Bien sûr, il a	laisser	la coupe d'or un chevalier
4	p=L2-02-010-B	pays de Galles. Sa mère ne lui	laisser	voir des livre de chevalier et
5	p=L2-02-011-B	les dame, suivre les conseil et	prendre	Dieu. § Donc Perceval est parti.
6	p=L2-02-011-B	a dit au chevaliers et a	demander	leur armes. Ecoutant ça, des chevaliers
7	p=L2-02-013-B	mère était très triste mais ella a	laisser	son fils suivre [son] sa rêve.
8	p=L2-02-014-B	que elle ne veut pas laisser Perceval	meurer	au combat comme son père et ses
9	p=L2-02-015-B	dit que le roi Arthur ne pas	prendre	son service et ne pas donner
10	p=L3-01-005-B	Keu était très jaloux et il se	moquer	de Perceval, le roi [éta] se
11	p=L3-01-007-B	la cour du moi. § Perceval dit :	Donner	moi armes, car je veux être
12	p=L3-01-023-B	Arthur, jaloux de roi et s'est	moquer	de roi. [Pour faire sortir_d] Dans la salle,
13	p=L3-01-023-B	Arthur, jaloux de roi et s'est	moquer	de roi. [Pour faire sortir_d] Dans la salle,
14	p=L4-018-R	une armure vermeille. Natüusement Perceval lui	demander	ses armes [d] au roi. la

Fig. 12 : Liste des formes non-normatives

Pour extraire les formes verbales non-normatives, les recherches plus affinées seront également possibles avec *Le Trameur*. L'image suivante montre la recherche des formes verbales erronées dont l'infinitif se termine par -er. Avec une telle opération, on se concentre sur le premier groupe de verbes.

Position	Forme	Lemme	Catégorie	a-00004	a-00005
1323	lèvé	lever	VER_pper	levé	lèvé_levé
1539	offer	offer	VER_infi	offrir	offer_offrir
1816	monquer	moquer	VER_infi	moquer	monquer_moquer
3125	appelle	appeler	VER_pres	appelle	appelle_appelle
3153	éleve	élever	VER_pper	élevé	éleve_élevé
3321	chvaché	chevaucher	VER_pper	chevauché	chvaché_chevauché
4018	fâché	fâcher	VER_pper	fâché	fâché_fâché
4026	arrête	arrêter	VER_pper	arrêté	arrête_arrêté
4185	fâché	fâcher	VER_pper	fâché	fâché_fâché
4482	communicaté	communiquer	VER_pper	communiqué	communicaté_communiqué
4942	honorez	honorer	VER_infi	honorer	honorez_honorer
5014	chevanché	chevaucher	VER_pper	chevauché	chevanché_chevauché
5090	demandé	demander	VER_infi	demander	demandé_demander
5314	ranger	venger	VER_infi	ranger_venger	ranger_venger
5557	pleuvait	pleurer	VER_impf	pleurait	pleuvait_pleurait
5681	parlaient	parler	VER_pper	parlé	parlaient_parlé
8160	allé	aller	VER_infi	aller	allé_aller

Fig. 13: Recherche des formes verbales erronées dont l'infinitif se termine par -er

La recherche pourrait également être effectuée pour une sous-catégorie de verbe, telle que les formes verbales en imparfait :

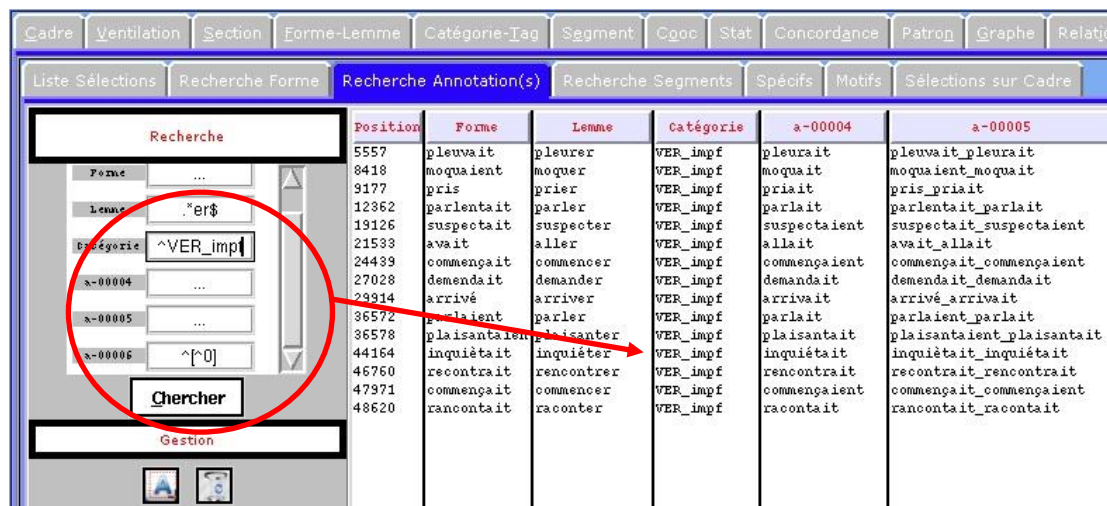


Fig. 14 : Recherche des formes verbales erronées de l'imparfait

### 3.3.3. Démarche d'analyse

Notons que les opérations avec *Le Trameur* susmentionnées nous permet de faire une analyse quantitative plus efficace. Et avant d'examiner les performances orthographiques des apprenants chinois, nous nous intéressons dans un premier temps les résultats bruts. Ainsi, nous pouvons répondre à une première question : quelles sont les zones de fragilité des apprenants chinois du point de vue global de leur maîtrise de l'orthographe française ? Le tableau suivant montre la répartition des erreurs en différents types et le poids relatif de chaque type d'erreur.

Tableau 7. La répartition des erreurs orthographiques

Types d'erreur		Nombre d'erreurs	Répartition d'erreurs (pourcentage)
1	Mauvais découpage du mot	9	1%
2	Substitution de mot	92	10%
3a	Erreur orthographique à dominante	133	15%

	phonétique		
3b	Erreur phonogrammique	27	3%
3c	Erreur d'orthographe lexicale sans altérer la valeur phonique	76	9%
4a	Erreur d'orthographe grammaticale : l'accord du nombre ou du genre	260	30%
4b	Erreurs d'orthographe grammaticale : les désinences verbales	284	32%
Total		881	100%

Première remarque à partir l'observation de ce tableau : les erreurs orthographiques les plus fréquentes des étudiants chinois de notre corpus sont les erreurs d'orthographe grammaticales, ainsi que les erreurs orthographiques à dominante phonétique.

Remarquons que les données se laissent certes traduire en chiffres, par le nombre des erreurs, mais comme l'indiquent Manesse *et al.*, « le sens de ces chiffres se donne dans la mise à jour des relations entre eux » (Manesse & Cogis, 2007 : 79). Ce sont les relations qui nous font comprendre les lacunes et les points forts chez les apprenants chinois lors de leur apprentissage de l'orthographe française. Donc dans la section suivante, la partie empirique, les erreurs orthographiques les plus fréquentes chez les apprenants chinois seront décrites et examinées de près. Afin de donner du sens aux données, nous nous attacherons à une analyse quantitative et qualitative.

La partie empirique se divise en trois parties. Dans la section 4.1, nous analysons les erreurs orthographiques à dominante phonétique, en les comparant avec d'autres erreurs d'orthographe lexicale, à savoir, les erreurs phonogrammiques et les erreurs d'orthographe lexicale sans altérer la valeur phonique. Nous aborderons ensuite le problème d'accord (l'accord en nombre et en genre) dans la section 4.2. La partie empirique se terminera par la discussion des erreurs orthographiques relevant des terminaisons verbales.

Cette partie empirique se veut la fois descriptive et comparative. D'un côté, pour chaque section, les données seront examinées de façon synchronique, afin de proposer une esquisse de la performance orthographique des apprenants chinois. De l'autre côté, pour mieux connaître le développement orthographique en français L2, les résultats obtenus par les apprenants chinois à des niveaux linguistiques différents seront comparés. En outre, les résultats seront commentés et les scénarios psycholinguistiques possibles pour interpréter les erreurs orthographiques seront également présentés, en faisant référence au développement langagier des apprenants L2, ainsi qu'aux divers facteurs en jeu dans le processus de la production écrite. Avec l'association des éléments de la recherche qualitative et quantitative, nous souhaitons mettre en lumière les traitements orthographiques utilisés par les apprenants chinois et leurs tendances interlangagiers spécifiques.

## IV. PARTIE EMPIRIQUE

### 4. 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]	各 [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] ; [ɔ̃] prononcé [ɔŋ] ; [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* [mãge].
- [ʒ] est prononcé comme les consonnes affriquées [tʃ] ou [dʒ] : *bonjour, je ...* [põtʃux 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>ẽ].

#### 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.

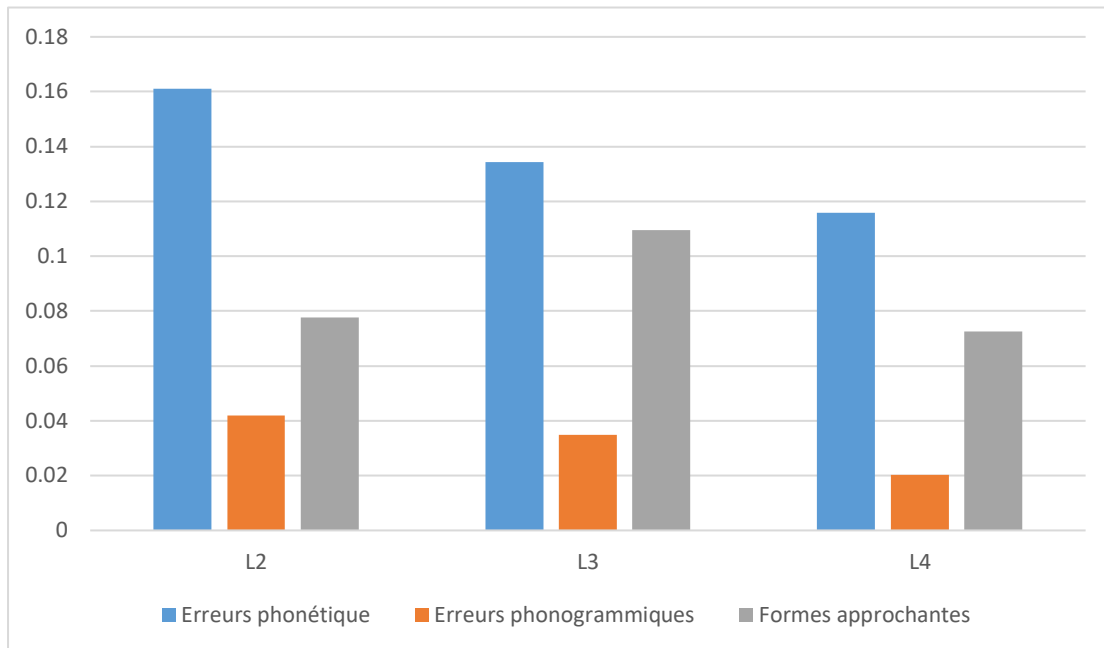
#### 4.1.6. Résultats

##### 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) ; \*ingnoraît (ignoraît) (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] : \*maledroitenant (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)
- \* conseils (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 [ɛ] 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.

#### 4.1.7.1.1. Différents modèles théoriques

Le concept de surdit  phonologique en langue  trang re a  t  introduit par le linguiste Polivanov dans les ann es 30. Selon lui, les phon mes et les autres repr sentations phonologiques de la langue maternelle sont  troitement li s   notre activit  perceptive. M me quand nous percevons des mots d'une langue ayant un syst me phonologique tout diff rent, nous sommes enclins   les d composer en des repr sentations phonologiques propres   la langue maternelle (Polivanov, 1931).

Quelques ann es plus tard, Troubetzkoy, le p re de la phonologie, illustre ce concept de surdit  phonologique par sa m taphore c l bre du crible phonologique :

« Le syst me phonologique d'une langue est semblable   un crible   travers lequel passe tout ce qui est dit. Seules restent dans le crible les marques phoniques pertinentes pour individualiser les phon mes. Tout le reste tombe dans un autre crible o  restent les marques phoniques ayant une valeur d'appel : plus bas se trouve encore un crible o  sont tri s les traits phoniques caract risant l'expression du sujet parlant. Chaque homme s'habitue d s l'enfance   analyser ainsi ce qui est dit et cette analyse se fait d'une fa on tout   fait automatique et inconsciente. Mais en outre le syst me des cribles, qui rend cette analyse possible, est construit diff remment dans chaque langue. L'homme s'approprie le syst me de sa langue maternelle. Mais s'il entend parler une autre langue, il emploie involontairement pour l'analyse de ce qu'il entend le "crible phonologique" de sa langue maternelle qui lui est familier. Et comme ce crible ne convient pas pour la langue  trang re entendue, il se produit de nombreuses erreurs et incompr hensions. Les sons de la langue  trang re re oivent une interpr tation phonologiquement inexacte, puisqu'on les fait passer par le "crible phonologique" de sa propre langue. » (Troubetzkoy, 1939 : 54)

Le concept du crible phonologique nous permet de comprendre les raisons d'une perception d fectueuse en langue  trang re, cependant il se fonde sur les produits finis

des traitements perceptifs. Il faut attendre le début des années 60, pour que la psycholinguistique et la psychologie cognitive prennent le relai et proposent de nouvelles explications de ce phénomène. Le concept de la perception catégorielle est ainsi introduit par ces disciplines pour expliquer la surdit  phonologique. Le processus de perception catégorielle commence tr s t t. Chaque enfant cat gorise progressivement ce qu'il entend fr quemment, c'est- -dire qu'il regroupe les unit s sonores qu'il conna t dans une m me cat gorie, et il d limite graduellement des fronti res entre elles afin de les discriminer. Ces cat gories sonores, limit es en nombre, contribuent au processus de communication. Ainsi, chaque individu acquiert la ma trise du syst me sonore de sa langue maternelle.

Les travaux de Kuhl (Kuhl & Iverson, 1995) compl tent cette notion de perception catégorielle, en d veloppant une th orie traitant de l'effet aimant de la langue maternelle (Native Language Magnet). Dans chaque cat gorie des  l ments sonores, il existe un  l ment le plus repr sentatif, qui est appel  prototype. Il attire vers lui la perception des sons qui sont situ s   proximit . Lorsqu'on apprend une langue  trang re, les phon mes de la langue maternelle agissent comme des aimants, qui attirent vers eux des sons qui ont pourtant des propri t s acoustiques diff rentes. Autrement dit, le syst me phonologique de la langue maternelle constitue un filtre qui devient un handicap pour le processus de perception des phon mes de la langue  trang re, car il nous emp che de percevoir des diff rences qui ne sont pas significatives en langue maternelle mais qui peuvent  tre significatives en langue  trang re (Anne-Flore, 2012). Selon Kuhl, ces interf rences entre langue maternelle et langue  trang re sont responsables de notre difficult     tendre notre carte de perception phonologique afin d'y int grer les  l ments sonores distinctifs en langue  trang re. Cependant, Kuhl consid re qu'un apprenant en langue  trang re peut parvenir   cr er de nouvelles fronti res phonologiques suite   un entraînement intensif.

Flege (1992, 2003)  labore un mod le de perception et de production de la parole, pour expliquer notamment les difficult s perceptives rencontr es par les apprenants adultes lors de leur apprentissage d'une langue  trang re. Le Mod le d'acquisition de la parole (Speech Learning Model ou SLM) est bas  sur deux  l ments principaux : le premier postule que pour les apprenants en langue  trang re, il est impossible de s parer



le système phonétique de leur langue maternelle et celui de la langue cible. La seconde stipule que les mécanismes qui nous permettent d'apprendre une langue restent intacts toute notre vie. Cependant, au fur et à mesure que les catégories pour la langue maternelle se développent, il est de plus en plus difficile de créer de nouvelles catégories qui correspondraient aux unités sonores de la langue étrangère. Ainsi, pour les enfants, dont les catégories phonologiques sont encore en construction, il est plus facile d'ajouter de nouvelles catégories. Cela explique, dans un certain sens, les différences concernant les performances en perception phonologique entre les enfants et les adultes. Flege indique également que, lors de l'apprentissage d'une langue étrangère, les sons les plus proches et les plus éloignés de la langue maternelle seraient les mieux maîtrisés. Pour les sons les plus proches, les apprenants peuvent les assimiler facilement aux catégories préexistantes, alors que pour les seconds, qui n'appartiennent pas aux catégories perceptives de la langue maternelle, les apprenants sont obligés de créer de nouvelles catégories. Ajoutons également que les différences perceptives entre la langue maternelle et la langue étrangère sont traitées sur une base du continuum. Selon le modèle de Flege, toutes les modifications des catégories existantes ainsi que les créations de nouvelles catégories amèneraient une réorganisation du continuum. Quant à la production, Flege constate qu'il existe une corrélation significative entre la perception et la production, mais il n'affirme pas que toutes les erreurs de production en L2 résultent des éléments perceptuels.

Escudero (2005) introduit un nouveau modèle synthétique des théories antérieures : le modèle de perception linguistique d'une langue seconde (L2LP). Selon ce modèle plus souple et plus complet, l'apprenant en langue étrangère passe plusieurs étapes pour créer de nouvelles catégories phonologiques. Plus précisément, après avoir changé sa carte perceptive, l'apprenant devient plus sensible aux différences sonores qui ne sont pas distinctives dans sa langue maternelle mais qui le sont en langue cible. Ainsi, il est capable d'établir graduellement de nouvelles catégories. Escudero explique que ce processus commence par un état initial, passe par un stade d'apprentissage et de développement, et atteint enfin un stade final. A l'état initial, l'apprenant utilise les catégories phonologiques de sa langue maternelle, et une mauvaise perception phonologique provient donc d'un classement erroné des unités sonores de la langue étrangère dans les catégories de la

langue maternelle. Au fur et à mesure du contact avec la langue cible, l'apprenant commence à modifier sa carte perceptive et développe de nouvelles catégories phonologiques pour arriver au stade final. Ainsi, l'acquisition de la phonologie d'une langue étrangère entre dans un contexte dynamique.

Grâce à ces théories concernant l'acquisition de la phonologie d'une langue étrangère, nous pouvons voir que tout individu est conditionné par les spécificités phonétiques originales de sa langue maternelle. Pour un apprenant adulte chinois du français L2, ses difficultés de perception phonologique pourraient s'expliquer par le filtre phonologique qu'il met en place très jeune au contact de la langue chinoise. Ce filtre a tendance, notamment au début de l'apprentissage, à réduire considérablement les capacités de discrimination et de catégorisation phonologiques des apprenants chinois. Cette fausse appréciation des phonèmes du français peut également se manifester au niveau de la production. 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.

#### 4.1.7.1.2. Difficultés lors de la production graphique

Partant de ce constat au niveau oral, nous vérifierons dans la partie suivante si les différences entre les systèmes phonologiques jouent également un rôle dans la production écrite. Autrement dit, la question de recherche est de savoir si les erreurs orthographiques à dominante phonétique sont dues à des difficultés au niveau phonique liées aux différences existantes entre le système phonologique du français et celui du chinois. En revenant à notre corpus d'écrits, nous constatons que l'effet probable de la dissemblance existant entre les deux systèmes phonologiques en jeu se manifeste dans la production orthographique des apprenants.

Pour la confusion des voyelles, il existe les erreurs concernant la nasalisation ou la dénasalisation :

- (52) a. \* Keu, le sénéchal, jaloux, aigre et {prèt}\_ {prêt} à se {monquer}\_ {moquer}  
a dit. (L2-01-003)

- b. \* {Aperceval}\_ {Perceval} +R[combattre]+R+//combat//+R avec un autre chevalier qui le {raconte}\_ {rencontre} dans {le}\_ {la} forêt quand il part sa maison. (L3-002)
- c. \* Le roi observait le jeune homme inconnu et {rancontait}\_ {racontait} son [chagrin] ennemi qui [sa] +R[a]+R+//avait//+R saisi la coupe du roi et répandu sur la reine [vin] le vin qu'elle avait refusé. (L4-022)
- d. \* Il est {audance}\_ {audacieux} et menaçant. (L4-026)
- e. \* le roi ~~lui~~ donna son {explanation}\_ {explication} de son {ignanrance}\_ {ignorance} impolie et lui interrogea +R[son]+R+//sa//+R raison de venir à la cour. (L4-027)
- f. \* Perceval {recontra}\_ {rencontra} un groupe de chevaliers. (L4-016, L4-19, L4-25, L4-27)

Les erreurs concernent la consonne [R], qui n'existe pas dans le système phonologique du chinois :

- (53) a. \* il {admisait}\_ {admirait} ça trop. (L2-01-001)
- b. \* Mais [qu] il y a [eut] eu des [rit] {rites}\_ {rires}. (L3-018)

Les difficultés relèvent de l'opposition de voisement de consonnes :

- (54) \* Perceval n'a pas connu que{c\_QUOTE\_a}\_ {ç\_QUOTE\_a} été une {blaque}\_ {blague}. (L2-01-003)

Les erreurs concernent la non réalisation de groupes consonantiques :

- (55) a. \* Un jour qu'il chassait dans la forêt, il rencontrait une troupe de {chevalier}\_ {chevaliers}, et il ébahi puis {émereillé}\_ {émerveillé} par leurs épées. (L3-012)
- b. \* Le roi dit à lui que s'il {renconte}\_ {rencontre} devant la porte et qui s'en

va avec ta coupe, il {veux}\_ {veut} être le chevalier {Vermeils}\_ {Vermeil}.

(L3-016)

c. \* suivez les conseils des [prud]'hommes qui [est] sont {monte}\_ {montrés} honorable. (L4-008)

d. \* il a {renconté}\_ {rencontré} {du}\_ {le} roi ~~Arthur~~ et il ~~lui~~ demandait {de\_être}\_ {d\_QUOTE\_être} son chevalier. (L4-009)

e. \* il n'y a [q] une seul[e] idée dans le cœur de Perceval que [l] leur {ressembler}\_ {ressembler} et devenir un membre entre eux, comme un chevalier. (L4-026)

La forme erronée provient de la non-réalisation de syllabes fermées :

(56) \* +R[Au]+R+//En//+R {contrait}\_ {contraire} de sa +R[mére]+R+//mère//+R, ce jeune homme s'est dépêché de prendre {sa}\_ {son} chemin. (L2-01-023)

Cependant, tous les cas que nous venons de citer ne constituent pas de tendances orthographiques, et il ne s'agit que des erreurs sporadiques. Notons que l'interférence entre le système phonologique de l'apprenant et celui du français peut expliquer certains faits mais pas tous. De plus, contrairement aux recherches de Gonac'h & Mortamet (2011) et de Katoozian (2013), où les scripteurs afghans et iraniens rencontrent les difficultés spécifiquement vocaliques, et que les scripteurs hispano-américains sont confrontés particulièrement les difficultés consonantiques, les erreurs orthographiques à dominante phonétique trouvées dans notre corpus ne portent pas particulièrement sur les consonnes, même le système consonantique du français et celui du chinois font preuve de peu de similarités. Notons que la plupart de confusions de consonnes ou de voyelles ne pourraient pas s'expliquer, nous semble-t-il, par les dissemblances existant entre les deux systèmes phonétiques.

Pour la forme erronée suivante, le scripteur a remplacé le phonème [o] par [ø]. Le dernier phonème n'existe pas dans le système vocalique du chinois, et pose souvent une difficulté pour les apprenants chinois lors de la production orale.

(57) \* Il [a] est [arriver] arrivé au {+R[château]+R+//châteaux//+R}\_ {château}. (L2-02-014)

L'erreur suivante montrent une confusion de la voyelle orale [ɛ] avec la voyelle orale [i]. Cependant cette discrimination existe dans le système vocalique du chinois.

(58) \* Tant elle {crignait}\_ {craignait} qu'il ne meure au combat... (L3-015, L3-016)

De plus, de nombreuses erreurs orthographiques à dominante phonétique portent sur l'omission ou l'adjonction de lettres ou de syllabes, qui sont également difficiles à expliquer directement par les différences entre les deux systèmes phonétiques. Par exemple :

(59) a. \* Perceval a dit qu'il [ve] voudrait {deveniaire}\_ {devenir} Vermeil. (L2-02-012)

b. \* A la {diffère}\_ {différence} +R[des]+R+//de//+R  
+R[les\_autres]+R+//la\_pluspart//+R, il y avait une ~~belle~~ jeune fille qui  
~~n\_QUOTE\_a\_pas~~ ri depuis de six ans. (L2-01-014)

Ainsi il faut être prudent face à une telle justification pour expliquer les erreurs orthographiques à dominante phonétique. Comme Katoozian (2013) l'a évoqué, on peut s'interroger aussi sur l'absence de l'influence, due aux dissemblances entre les deux systèmes phonologiques français et chinois, chez les étudiants de notre corpus qui ont produit des formes correctes. Il semble que, chez eux, les dissemblances entre les deux systèmes phonologiques ne jouent aucun rôle, tandis que ces étudiants ont le même système phonologique que leurs camarades qui ont commis les erreurs phonétiques. Selon Katoozian (2013), ce paradoxe s'explique par le fait que les erreurs orthographiques à dominante phonétique relèvent des erreurs de la langue, et prouve que celui qui les a produites ne connaissait pas les mots en question. La même justification est également

évoquée par Manesse et *al.*, pour expliquer les erreurs de ce type observées chez les élèves français. Selon les chercheurs, le mot concerné est « visiblement inconnu de l'élève qui tente d'en restituer les sons qu'il semble percevoir pour la première fois » (2007 : 73). Notons que les deux recherches sont basées sur la dictée, où les apprenants sont obligés de transmettre le matériau phonique en matériau graphique, même s'il s'agit des mots inconnus. Cependant, les étudiants de notre corpus ont été confrontés à la rédaction, tâche d'écriture où les étudiants utilisent leurs connaissances linguistiques préalables pour produire leur propre texte. De ce fait, les erreurs orthographiques à dominante phonétique relèvent de l'inaccessibilité de l'image auditive des mots dans le cas de la dictée et de la rédaction. Cependant, en rédaction, cette inaccessibilité ne peut pas s'expliquer par le fait que les mots en question sont inconnus pour les apprenants, puisque les apprenants ont recours à leurs connaissances linguistiques préalables lors de la rédaction.

#### **4.1.7.2. Développement lexical des apprenants L2**

Notre hypothèse est que les erreurs orthographiques à dominante phonétique observées dans notre corpus sont liées au développement lexical du français chez les étudiants chinois. Pour ceux qui ont commis ce type d'erreur, les mots en question n'étaient pas inconnus, mais se trouvaient au stade initial du développement. Notons que pour les mots se situant au stade initial du développement, l'entrée lexicale ne contient que les informations formelles, orthographe et prononciation, ainsi qu'un pointeur qui sert de lien entre le mot en langue étrangère et son équivalent en langue maternelle.

Il faut souligner que même les informations phonologiques et graphiques stockées à l'entrée lexicale ne sont pas toujours justes et accessibles pour les apprenants chinois. Pour les étudiants de notre corpus, qui apprennent le français dans une université chinoise, ils ne sont pas en contact permanent avec l'image auditive correcte du mot, et « l'image qu'ils ont des mots français devient pour ainsi dire plus ou moins floue » (Gak, 1962 : 71). De plus, il est tout à fait possible que les informations phonologiques stockées dans leur lexique mental soient incorrectes, ce qui s'explique par la fausse appréciation des phonèmes français. Comme nous l'avons expliqué auparavant, il s'agit d'une difficulté

de perception phonologique très fréquente chez les apprenants allophones, dont le système phonologique de la langue maternelle se différencie largement de celui de la langue cible. Ainsi, il est fort probable que, lors de la production écrite, les apprenants ne possèdent pas les informations phonologiques correctes concernant les mots à produire, tandis que les apprenants français connaissent déjà bien la forme parlée des mots à produire. Contrairement aux apprenants chinois, les apprenants français sont en mesure de prévenir les erreurs orthographiques altérant l'image acoustique du mot à l'aide de la forme orale déjà bien maîtrisée. Chez les étudiants de notre corpus se dessine nettement une telle influence due aux informations phonologiques incorrectes. On voit notamment la confusion des consonnes qui se distinguent par le voisement, \* *blaque* à la place de *blague* ; la dénasalisation d'une voyelle nasale, \* *racontre* à la place de *rencontre* ; la non-réalisation des groupes consonantiques, \* *renconte* à la place de *rencontre*, \* *monte* à la place de *montre*, etc.

Les performances orthographiques des étudiants chinois de notre corpus sont également influencées, nous semble-t-il, par les habitudes scolaires et l'enseignement dispensé à l'université. Les cours proposés dans les départements universitaires sont majoritairement assumés par les professeurs chinois, en utilisant les manuels chinois. Selon l'observation de Cuet et Marguerie (2007), les pratiques de classe dans les universités chinoises correspondent à ce que l'on appelle méthode traditionnelle ou encore méthode grammaire-traduction, qui consiste à privilégier la grammaire, le vocabulaire et la traduction dans l'enseignement. Quant à l'enseignement du nouveau vocabulaire, il est présenté souvent « sous forme de listes que l'étudiant devra apprendre par cœur » (Cuet & Marguerie, 2007).

Pour les étudiants chinois, formés dans une culture ancrée dans l'écrit, l'apprentissage du vocabulaire est centré sur la mémoire visuelle ou auditive, l'apprentissage par cœur (la forme orthographique, le sens traduit en chinois, la catégorie grammaticale, etc.) et le recopie par écrit. Et tout cela favorise une réceptivité pour l'écrit plutôt que pour l'oral, qui est aussi remarqué par les enseignants de FLE en France, proposant souvent une approche communicative dans leurs cours (Robert, 2002). Ainsi, il est fort probable que les étudiants chinois ont tendance à prendre des notes ou à

visualiser la graphie des mots pour apprendre le sens et l'orthographe. Et avec la difficulté de perception phonologique, due aux dissemblances existant entre les deux systèmes phonologiques, ce besoin de visualisation serait renforcé chez les étudiants chinois.

Nous pouvons, par conséquent, préciser l'organisation lexicale au stade initial pour les étudiants chinois de notre corpus. Comme Jiang (2000) l'évoque, une entrée lexicale en langue étrangère ne contient, au début, que les informations formelles du mot concerné, qui sont, plus précisément, orthographiques et phonologiques. Pour les étudiants de notre corpus, l'apprentissage du français dépend principalement de l'enseignement en classe, ce qui introduit ainsi une pauvreté de l'input en termes de quantité et de qualité. Étant donné ce contact insuffisant avec l'image auditive correcte du mot et les dissemblances existant entre les deux systèmes phonologiques, il est fort probable que les étudiants soient confrontés à de grandes difficultés au niveau de perception et de production phonologique, et que les informations phonologiques stockées à l'entrée lexicale chez eux soient gravement discréditées. Cependant, avec la culture écrite venant de la tradition chinoise, l'enseignement basé sur la grammaire et la traduction, ainsi que l'insécurité linguistique ressentie par les apprenants dans le domaine de la phonologie, les apprenants chinois sont probablement plus à l'aise pour retenir les informations orthographiques. De plus, en étudiant le français en Chine, les étudiants peuvent plus facilement s'imprégner des contextes écrits que des contextes oraux, ce qui contribue également au stockage à l'entrée lexicale des informations orthographiques. Ainsi, au stade initial du développement lexical, les informations orthographiques présentent plus de stabilité et de crédibilité que les informations phonologiques.

Lors de la rédaction, tâche d'écriture choisie par notre étude, les scripteurs produisent, en terme général, les mots qu'ils connaissent. En priori, les étudiants récupèrent directement les formes orthographiques stockées à l'entrée lexicale. Cette récupération nécessite cependant que la forme orthographique des mots ait été mémorisée sous un format stable, qu'elle soit facilement mobilisée et que le scripteur ait un bon degré de certitude (Fayol & Jaffré, 2008 : 141). Si la forme des mots n'est que partiellement connue par le scripteur, ou si le scripteur est mentalement surchargé pendant la tâche d'écriture, la mobilisation de la forme orthographique serait difficile et l'utilisation de



cette stratégie subirait des risques redoutables.

Dans ce cas, le scripteur pourrait faire référence à d'autres stratégies possibles afin de compenser des risques. L'une des stratégies de secours, c'est l'emploi des correspondances entre phonèmes et graphèmes. Ces correspondances permettent au scripteur d'aboutir à une forme orthographique plausible s'il est en mesure de retrouver la forme phonologique correcte du mot. Comme nous l'avons évoqué plus haut, les informations phonologiques stockées dans l'entrée lexicale chez les étudiants chinois sont souvent dégradées, et sous un format instable. C'est pourquoi nous pouvons trouver les formes erronées, telles que \* *blaque* à la place de *blague* ; \* *racontre* à la place de *rencontre* ; \* *renconte* à la place de *rencontre*, \* *monte* à la place de *montre*, etc.

Le scripteur pourrait également s'appuyer sur une analogie orthographique lorsque la remémoration de la forme orthographique des mots n'est pas disponible. Il s'agit de se référer à une forme orthographique bien connue par le scripteur pour en produire une autre qui entretient avec la première une relation de ressemblance formelle ou du point de vue de sens (Fayol & Jaffré, 2008 : 141).

Dans l'exemple (60), le scripteur a produit la forme verbale du mot *espérer* en se référant à la forme orthographique du verbe *expecter*.

(60) \*Sa mère {n\_QUOTE\_especte}\_ {n\_QUOTE\_espère} pas que ... (L3-002)

L'utilisation de cette stratégie est également observée dans l'exemple (61) :

- (61) a. \* {plaisant}\_ {plaisantant} (L3-009) qui a comme référent la forme *plaisant* du verbe *plaire* ;
- b. \* {jusquà\_TIRET\_là}\_ {jusque\_TIRET\_là} qui a comme référent la forme contractée *jusqu'à* ;
- c. \* pour la forme erronée {volontir}\_ {volontaire} (L4-019), le choix de la voyelle *i* pourrait s'expliquer par l'analogie avec l'adverbe *volontiers*.
- d. \* {derrier}\_ {derrière} (L4-023) qui a pour référent l'adjectif *dernier*.

Cette stratégie concerne également la référence qui est faite à la famille morphologique. La forme nominale erronée \* {meures}\_ {morts} (L2-02-015) pourrait être construite à partir des formes verbales du subjonctif, telles que *je meure, tu meures, il meure...*

Les scripteurs dans l'exemple (62) se réfèrent également à la forme orthographique de l'anglais pour produire son équivalent en français :

- (62) \* {adventures}\_ {aventures} (L2-01-013) ;  
\* {honez}\_ {honorez} (L2-02-003) ;  
\* {réal}\_ {réel} (L4-017) ;  
\* {montation}\_ {montagnes} (L4-009) ;  
\* {riviers}\_ {rivières} (L4-009) ;  
\* {l\_QUOTE\_offer} {l\_QUOTE\_offrir} (L2-01-003) ;  
\* {palace}\_ {palais} (L4-022) ;  
\* {explanation}\_ {explication} (L4-027)

Lorsque la forme orthographique des mots à produire ne peut pas être récupérée directement, le scripteur peut recourir aux règles. Les règles, si elles existent, permettent sûrement au scripteur de trouver la forme orthographique correcte. Rappelons cependant que « peu de phénomènes sont formalisables par des règles simples et applicables de manière générale » (Fayol & Jaffré, 2008 : 141), et la généralisation des règles n'est pas sans risque.

Avec la connaissance insuffisante de la formation des adverbes en *-ment*, le scripteur dans l'exemple (63) a construit la forme adverbiale de l'adjectif *gentil* en ajoutant *-mment*. Il a ignoré que les doubles consonnes *mm* n'apparaissent que dans le suffixe *-amment* ou *-emment*, qui sert à construire l'adverbe à partir des adjectifs se terminant par *-ant* ou *-ent*. Le scripteur a généralisé les règles sans tenir compte de ce que l'adverbe *gentiment* est une forme irrégulière.

- (63) \* Le roi, cependant, a salué {gentilmment}\_ {gentiment} Perceval et lui expliqué

son chagrin [et]. (L4-023)

Le scripteur dans l'exemple (64) a remarqué le suffixe *-ie*, qui est généralement employé pour former des noms désignant un état, une qualité ou une condition. En essayant de construire un nom désignant la qualité de chevalier, le scripteur a ajouté directement le suffixe *-ie* au nom *chevalier*.

(64) \* Il y a six ans que la fille ne rit pas, mais quand elle rit : ça montrait qu'elle rit un homme qui a le prix de toute {chevalerie}\_ {chevalerie}. (L4-011)

Il nous semble que le scripteur a rencontré le nom *chevalerie* pour la première fois dans le texte de source lors de la tâche d'écriture. Et une seule rencontre ne suffit pas à assurer la mémorisation de la forme orthographique ou de l'image auditive du mot. De ce fait, il a fait référence à ses connaissances préalables, notamment les règles implicites concernant la formation des mots. Nous pouvons également trouver dans notre corpus les formes erronées, telles que \*{chevaleure}\_ {chevalerie} (L2-02-015) ; \* {chavalaire}\_ {chevalier} (L2-02-015, L2-02-016).

Dans l'exemple (65), le scripteur a construit le verbe \* *communicater* à partir de sa forme nominale *communication*.

(65) \*il {s\_QUOTE\_a}\_ {s\_QUOTE\_est} s'a {communicaté}\_ {communiqué} avec le roi. (L2-01-007)

De nombreuses études sur l'apprentissage de l'orthographe lexicale confirment l'influence positive des facteurs visuels sur l'acquisition de l'orthographe lexicale, notamment « le traitement visuel simultané de l'ensemble des lettres du mot pendant la lecture » (Chaves & al., 2012). La prise d'information simultanée du mot entier favorise généralement la mémorisation de son orthographe, toutefois, cette stratégie n'est pas sans risque lors de la production écrite. Les formes erronées ci-dessous nous donnent

l'impression que les scripteurs ont mémorisé visuellement l'image orthographique du mot entier, et que la forme orthographique du mot mémorisée n'a pas été stable ou bien mobilisée, ce qui a pour conséquence la permutation ou l'inversement des lettres et le déplacement des accents.

- (66) a. \* Pendant les {prèmes}\_ {premières} années de leur vie, sa mère n'ose pas lui parler de la chevalerie, car elle ne [e] supporte plus de perdre encore un de sa famille-son père et ses deux frères sont tous morts dans l'aventure chevalière. (L4-023)
- b. \*Perceval [innogre] {innogra}\_ {ignora} {pus}\_ {puis} et [p] il était parti directement. (L3-022)
- c. \* Ils ont passé une nuit dans {un}\_ {une} {fôrt}\_ {forêt}. (L2-01-018)

Si la forme orthographique et phonologique de mots mémorisée par le scripteur ne présente pas de stabilité, la production orthographique d'un mot ou d'une syllabe pourrait subir l'influence de son voisin précédent.

- (67) a. \* Dans la cour de roi Arthur, tout le monde plaisantait et {parlentait}\_ {parlait}. (L2-01-023)
- b. \* Les hommes là sont assis, parlant et {plaisaitant}\_ {plaisantant}. Le roi Arthur est placé +R[en]+R+//au//+R haut bout de la table, restait pensif et muet. (L3-013)

#### 4.1.8. Conclusions

L'objectif de ce chapitre était de décrire l'acquisition de l'orthographe des mots du lexique chez les apprenants chinois. Pour ce faire, nous avons comparé les performances des apprenants chinois avec celles des élèves français, en analysant les erreurs lexicales de trois catégories différentes, à savoir les erreurs orthographiques à dominante

phonétique, les erreurs phonogrammiques, ainsi que les erreurs lexicales sans incidence phonique. A partir de ces analyses, un certain nombre de tendances intéressantes ont pu être dégagées.

A la différence des élèves français, les erreurs orthographiques à dominante phonétique sont nombreuses chez les apprenants chinois. De plus, ce type d'erreurs ne diminue pas en nombre au cours de la scolarité. En outre, les erreurs phonogrammiques et les erreurs lexicales sans incidence phonique sont beaucoup moins lourdes en poids relatif. Néanmoins, les erreurs phonogrammiques diminuent en nombre en fonction du niveau scolaire des apprenants, alors que les erreurs lexicales sans incidence phonique persistent et même augmentent légèrement en nombre au cours de la scolarité.

Ainsi, nous constatons que les erreurs orthographiques à dominante phonétique constituent une tendance orthographique spécifique aux apprenants du français L2. Néanmoins, la véritable cause des erreurs de ce type reste difficile à déterminer. La dissemblance entre le système phonologique du chinois et celui du français peut quasiment expliquer certaines graphies erronées mais pas toutes. Il nous semble plus pertinent d'examiner ces erreurs en faisant référence au développement lexical des apprenants chinois du français L2.

Sans contact permanent avec l'image auditive correcte du mot, et avec les dissemblances existant entre les deux systèmes phonologiques, il est fort probable que les informations phonologiques stockées à l'entrée lexicale chez les étudiants chinois soient gravement discréditées. En revanche, avec la culture écrite venant de la tradition chinoise, l'enseignement basé sur la grammaire et la traduction, ainsi que l'insécurité linguistique ressentie par les apprenants dans le domaine de phonologie, les apprenants chinois sont probablement plus à l'aise à retenir les informations orthographiques. Autrement dit, au stade initial du développement lexical, les informations orthographiques présentent plus de stabilité et de crédibilité que les informations phonologiques. Lors de la production écrite, les étudiants récupèrent en priori directement les formes orthographiques stockées à l'entrée lexicale. Si la forme des mots n'est que partiellement connue par le scripteur, ou si le scripteur est mentalement surchargé pendant la tâche d'écriture, la mobilisation de la forme orthographique serait alors difficile et l'utilisation de cette stratégie subirait

des risques redoutables. Dans ce cas, le scripteur doit avoir recours à d'autres stratégies possibles afin de compenser des risques, telles que l'emploi des correspondances entre phonèmes et graphèmes, de l'analogie, des règles apprises implicitement.

Notons que l'emploi de ces stratégies aboutissent à des solutions qui ne rencontrent pas toujours la graphie normée. Nous trouvons ainsi dans notre corpus de nombreuses erreurs orthographiques à dominante phonétique. Sans accéder à l'image phonique correcte de mots, les apprenants ont proposé les graphies erronées représentant une grande diversité. L'omniprésence et la persistance des erreurs de type phonique nous font penser qu'une bonne maîtrise du système phonologique du français est essentielle pour une bonne acquisition de l'orthographe. Comme Ameer-amokrane (2007) l'indique, si certaines confusions au niveau phonique peuvent ne pas gêner la communication orale, étant donné que le contexte de l'énoncé sert à lever l'ambiguïté, elles seront presque toujours source de difficulté orthographique. Les analyses que nous avons faites précédemment nous conduisent à affirmer la nécessité d'assurer la correction phonétique pour faire progresser les apprenants chinois en orthographe.

## 4.2. Orthographe et syntaxe : le problème des accords

### 4.2.1. Syntaxe de l'accord : le genre et le nombre

En français, les mots de certaines catégories grammaticales sont dotés de deux ou plusieurs formes. Ces catégories concernent en effet les déterminants, les noms, les adjectifs et les verbes. Les formes de ces mots varient en fonction des catégories morphologiques, telles que le nombre, le genre, les personnes, etc.

Notons que tous les mots ne sont pas déterminés par les mêmes catégories : les verbes varient en fonction de la personne, du mode et du temps ; les noms ont un seul genre (ils sont masculins ou féminins), mais la plupart du temps les deux marques du nombre, celle du singulier et celle du pluriel ; les adjectifs s'accordent à la fois en genre et en nombre. Mais évidemment « dans un énoncé donné, une seule forme est sélectionnée » (Cogis, 2007 : 98). Par exemple, selon ce que l'on veut signifier, on parlera donc d'un chevalier, ou *des chevaliers*, de *la forêt*, et *des forêts*.

Il n'est pas superflu de remarquer que le nombre du nom est déterminé par le choix du locuteur, et qu'en revanche, le genre du nom est fixé d'avance par la langue. Comme tout nom est « pourvu d'un genre inhérent, masculin ou féminin » (Riegel, Pellat & Rioul, 1994), le genre du nom ne relève pas d'une décision du locuteur (*un chevalier / une arme*). Quant à l'adjectif, il s'accorde au nom auquel il se rapporte (*une armure vermeille*). Il en va de même pour le nombre du verbe, qui est décidé par le nombre du sujet auquel le verbe se rapporte. Ainsi existe-il « une relation étroite entre la forme que prend le mot dans un énoncé et les relations syntaxiques qu'il entretient avec d'autres mots » (Cogis, 2007 : 98).

Si l'on parle de la mise en place d'une relation syntaxique dans un énoncé, on entre dans le domaine de l'accord. L'accord est défini comme la mise en cohérence des marques syntaxiques de plusieurs signes à l'intérieur d'un texte. En d'autres termes, lorsque deux signes différents d'un texte ont « la même caractérisation syntaxique », ils sont en accord (Weinrich, 1989). En particulier dans la grammaire française, on parle d'accord lorsque deux signes ont une caractérisation cohérente en genre et en nombre.

Et nous nous concentrons dans ce chapitre sur le phénomène de l'accord en français.

#### 4.2.1.1. Orthographe et marques morphologiques

En français, lorsqu'un mot figure dans un énoncé, il est nécessairement « affecté d'une valeur grammaticale » (Cogis, 2007 : 98). C'est-à-dire que le mot pur n'existe pas dans le contexte de l'énoncé. En étudiant le français, il faut en effet appréhender clairement le rapport mot/formes. Et comme Cogis l'indique, tout cela n'est pas propre à l'écrit (2007 : 98). En revanche, c'est l'apprentissage de l'écrit qui permet aux apprenants de mieux comprendre le rapport mot / formes, ainsi que la relation syntaxique établie entre les mots de différentes catégories grammaticales.

Remarquons qu'un mot français se présente sous la forme d'« une suite de lettres qui lui est propre et qui, à l'exception de certains homophones, le différencie des autres » (Cogis, 2007 : 99). Néanmoins, les mots variables ont en commun certaines lettres. Il s'agit des marques morphologiques, indiquant les catégories de genre et de nombre. Ainsi, tous les noms au pluriel se terminent par un *-s* ou un *-x* ; tous les adjectifs au féminin prennent la marque morphologique *-e* ; le pluriel de verbes (3<sup>e</sup> personnes du pluriel du présent de l'indicatif) est marqué par *-nt*. Or, en français, ces marques morphologiques, et donc les accords, ne se prononcent pas (Jaffré & Fayol, 2016 ; Fayol, 2008). C'est-à-dire que, d'une manière générale, ces marques servant à indiquer la catégorie de genre ou de nombre ne sont pas en relation avec la chaîne orale, et ne transcrivent pas un élément phonique. A la différence d'autres langues, qui ont soit des marques morphosyntaxiques très limitées comme l'anglais, soit des marques correspondant à des variations phoniques comme l'italien (Cogis, 2007 : 99), le français présente les marques morphologiques silencieuses très riches, et cela rend leur transposition à l'écrit moins aisée.

**Tableau 24.** Exemples de morphologie silencieuse et audible dans le syntagme nominal et le syntagme verbal (*Légende : souligné = morphologie audible ; en caractères gras = morphologie silencieuse*)

A      Le petit garçon se promène dans la forêt.



- B Les petits garçons se promènent sur les dunes.
- C Les jeunes enfants jouent dans le jardin.
- D Le joli vélo bleu est garé devant l'école.
- E La jolie voiture bleue est garée devant l'école.
- F Les jolies voitures bleues sont garées devant l'école.

Afin d'illustrer la morphologie silencieuse et audible, nous reprenons les exemples utilisés par Ågren (2005) dans son travail. Au sein d'un syntagme nominal, l'accord entre le noyau nominal et ses compléments, à savoir déterminants et adjectifs, doit être marqué par le genre et le nombre. Si nous comparons A et B, nous constatons que le pluriel n'est audible que sur le déterminant (*les petits garçons, les dunes*). Pour les noms et les adjectifs, le pluriel est marqué par un -s muet. Le syntagme nominal *les jeunes enfants* dans la phrase C présente la seule exception : étant donné que le nom *enfant* commence par une voyelle, il existe une liaison entre ce nom et l'adjectif préposé *jeune*, et le -s servant à marquer le pluriel de l'adjectif devient donc audible. Quant à l'accord en genre, la difficulté morphologique reste la même si nous comparons D et E : les déterminants sont les seuls portant une marque de genre audible. La complexité du système d'accord se voit clairement dans la phrase F : neuf marques morphologiques expriment le nombre et le genre, dont deux seulement sont en relation avec la chaîne orale. Il en va de même dans le syntagme verbal. Les marques indiquant la 3<sup>e</sup> personne du pluriel restent souvent inaudibles (*les petits garçons se promènent ; les jeunes enfants jouent*).

#### **4.2.1.2. Acquisition d'un sous-système linguistique spécifique à l'écrit**

A partir des exemples susmentionnés, nous constatons que le français a un sous-système linguistique spécifique à l'écrit pour marquer la catégorie de genre et de nombre. Et les marques morphologiques silencieuses constituent des zones orthographiques fragiles, difficilement maîtrisées par les apprenants du français. Pour bien orthographier, l'apprenant doit connaître toutes les formes d'un mot donné, ainsi que les marques

graphiques propres à chaque catégorie des mots. Et pour la mise en place de ces connaissances morphographiques, l'apprenant doit avoir conscience des relations syntaxiques que ce mot entretient avec les autres. Ainsi la pratique de l'accord orthographique se trouve-t-elle au croisement de deux niveaux, à savoir la morphologie et la syntaxe. Pour faire un accord dans la production orthographique, une analyse morphosyntaxique permanente serait nécessaire, puisqu'on peut rarement se fier à l'oral pour trouver la marque congruente.

Confronté à cette complexité au plan orthographique, tout scripteur aura un long trajet à parcourir avant d'arriver à la maîtrise complète des traitements morphosyntaxiques impliqués (voir, par exemple, Brissaud & Jaffré, 2003 ; Fayol & Largy, 1992 ; Jaffré & David, 1999), et ceci est sans doute d'autant plus vrai lorsqu'il s'agit de s'approprier ce système français dans une langue étrangère qui est typologiquement distincte de la langue française. De ce fait, nous examinerons dans ce chapitre la pratique de l'accord orthographique des apprenants chinois. En analysant respectivement leurs performances de l'accord en nombre et en genre, nous souhaitons savoir comment les apprenants chinois répondent à ces difficultés multidimensionnelles dans une production écrite réelle.

#### 4.2.2. Accord du nombre

La maîtrise de la morphologie du nombre joue un rôle particulièrement vital pour la compréhension et la production du français écrit, puisque le français écrit est caractérisé par un système binaire pour marquer l'opposition sémantique singularité / pluralité le plus souvent inaudible à l'oral (Dubois, 1965). Ainsi, les apprenants, français ou allophones, doivent apprendre ce système d'opposition spécifique de l'écrit : « -s pour le pluriel nominal et adjectival ; -nt pour le pluriel verbal » (Fayol, 2003). L'accord non marqué phonétiquement du nombre correspond à l'une des difficultés majeures de l'apprentissage de l'orthographe du français, que les apprenants doivent surmonter. Cette difficulté est également observée dans notre corpus des étudiants chinois.

### **4.2.2.1. Questions de recherche**

En mettant l'accent sur les formes erronées relevant de l'accord du nombre, ce chapitre consacrera à décrire le développement de la morphologie du nombre du français écrit pour les apprenants chinois. Les questions principales auxquelles nous souhaitons répondre en ce qui concerne l'acquisition de la morphologie écrite du nombre sont donc les suivantes :

- Est-il possible, à partir des études sur le développement de la morphologie du nombre du français écrit en L1, de supposer un développement semblable dans le domaine de L2 ?
- Quelles seraient les différences et les similitudes entre le développement de L1 et de L2 en ce qui concerne le nombre ?

Pour répondre à ces questions concernant l'acquisition de la morphologie du nombre, nous devons établir tout d'abord le sens de l'opposition en nombre, puis préciser ce que les études de L1 apportent à l'acquisition de la morphologie écrite du nombre en français. Ensuite on décrira un modèle d'apprentissage de la morphologie du nombre, basant sur l'opposition entre apprentissage d'instances et apprentissage de règles. Enfin nous présenterons le profil de l'accord en nombre nominal et verbal observé chez les étudiants chinois de notre corpus, en reposant sur les formes erronées qu'ils ont produites. Nous proposerons également les hypothèses susceptibles pour expliquer leur profil.

### **4.2.2.2. Expression morphologique du nombre**

#### **4.2.2.2.1. Distinction sémantique et syntaxique en français**

Avant d'en venir à la production écrite de la morphologie, plusieurs notions concernant la catégorie grammaticale du nombre méritent d'être soulevées. Notons que le concept de nombre connaît en français une distinction à la fois sémantique et syntaxique.

Selon la distinction sémantique, l'opposition singulier/pluriel fournit des instructions de guidage pour la compréhension correcte du texte. Plus précisément, le singulier permet au lecteur/auditeur d'« interpréter le sens d'un mot en tant qu'ensemble », tandis que le pluriel donne pour consigne « le sens d'un mot en tant qu'éléments disjoints » (Weinrich, 1989 : 46). A l'inverse du genre, qui est prédéterminé par la contrainte lexicale, le nombre est choisi librement pour chaque signe en fonction des exigences du texte. Pour l'accord en nombre, il faut signaler qu'à l'intérieur de la phrase, il existe une différence sémantique entre les différentes classes grammaticales. Pour la catégorie des noms, les marques du nombre changent leur contenu conceptuel. Il n'en est pas de même avec les autres composants du groupe nominal ou verbal. Les marques du nombre ne changent pas le sens sémantique des adjectifs ou des verbes. A ce propos, il est intéressant de mentionner que, selon les études du français L1 (voir, par exemple, Totereau, Barrouillet & Fayol 1998 ; Fayol, 2003), cette différence entre le nom, d'un côté, et l'adjectif et le verbe, de l'autre, constitue un des facteurs essentiels dont l'acquisition et la mise en œuvre des accords en nombre semblent dépendre.

En ce qui concerne la distinction syntaxique, l'opposition en nombre existe en français de façon régulière et sous forme de marque flexionnelle, pour les différents éléments de la phrase. Dans la majorité des cas, le pluriel est assuré par l'ajout d'un –s dans le syntagme nominal (adjectifs et noms)<sup>41</sup>. De même, le pluriel s'exprime par l'ajout du morphème –nt dans le syntagme verbal (verbes à la troisième personne)<sup>42</sup>. L'opposition en nombre existe également pour les déterminants. Cependant, à la différence des adjectifs et des noms, les marques du nombre des déterminants introduisent une modification de l'image phonologique. En outre, les déterminants au pluriel sont très fréquents dans l'input des apprenants de français.

---

<sup>41</sup> Selon New et al. (2004), 2% des noms et des adjectifs prennent un –x au pluriel.

<sup>42</sup> Un petit groupe de verbes marquent le nombre avec un pluriel audible en [ø], tels qu'*ils sont/ont/vont/ont*. De plus, certains verbes irréguliers marquent le pluriel avec un changement de radical et l'ajout du morphème –nt (Hedbor & al., 2006).

#### 4.2.2.2. Le nombre en chinois et en anglais

Avant d'analyser la production morphologique des apprenants chinois, il nous semble essentiel de souligner l'influence possible de l'expression du nombre en chinois et d'autres langues étrangères sur leur production en français, puisque l'apprentissage d'une langue étrangère peut être influencé par d'autres langues déjà acquises par l'apprenant.

Comme nous l'avons indiqué plus haut, la variation morphologique est presque zéro en chinois. Contrairement au français, les mots chinois restent invariables, quelle que soit leur position syntaxique. En ce qui concerne l'opposition singulier/pluriel, pour le chinois standard, les marques morphologiques de pluriel sur les noms sont inexistantes si l'on écarte le suffixe *-men* (们), « qu'on trouve sur les pronoms personnels, sur certains noms collectifs animés humains de façon optionnelle, ainsi que plus rarement sur des noms propres dissyllabiques » Lam (2006) :

- 我/我们  
wo/women  
Je/nous
  
- 老师/老师们  
laoshi/laoshimen  
professeur/professeur-men  
professeur<sup>43</sup>/des professeurs

---

<sup>43</sup> Notons que le déterminant défini est absent en chinois. Les notions de définitude et d'indéfinitude sont « compensées par d'autres stratégies (position du groupe nominal dans la phrase, présence d'un démonstratif, présence d'un numéral/indéfini avec un classifieur, etc.) ». Un nom nu peut avoir des référents différents selon le contexte :

老师

laoshi

professeur

le professeur, les professeurs, des professeurs, un professeur

- 小强们

Xiao Qiang men

Nom propre –men

Le groupe/la bande de Xiao Qiang

Cela dit, Lam (2006) a évoqué également que l'emploi de ce suffixe collectif –*men* est soumis à de nombreuses contraintes, et qu'il est par conséquent très irrégulier et limité en chinois. Ainsi, il est de mise de penser que chez les apprenants chinois, dont la langue maternelle présente une variation morphographique minimale, la mise en place de l'accord en nombre, phénomène qui leur est inconnu, pourrait être problématique.

Étant donné que l'anglais est la première langue étrangère des apprenants chinois de notre corpus, il est pertinent de rendre compte d'un transfert éventuel des connaissances de l'anglais vers le français. De ce fait, il n'est pas sans importance de rappeler que l'anglais marque le nombre nominal aussi bien à l'oral qu'à l'écrit, avec le même morphème qu'en français : le *-s*. Cependant, les verbes et les adjectifs ne sont pas accordés en nombre.

#### **4.2.2.3. L'acquisition de la morphologie écrite du nombre en Français**

##### **L1**

Avant d'analyser les productions morphologiques du nombre des apprenants chinois, il nous semble pertinent de décrire de façon sommaire l'état des recherches dans le domaine de l'acquisition de la morphologie écrite du nombre en français langue maternelle. Cette relecture devrait contribuer à notre compréhension de la tâche s'imposant à un apprenant du français L2.

Les enfants français, même depuis très jeunes, disposent déjà de connaissances sur l'écrit, ce que Jaffré appelle « une intelligence orthographique » (1992 : 28). La mise en œuvre de ces connaissances dans la production réelle concerne, en revanche, un processus difficile à gérer et à automatiser. En touchant aux difficultés

multidimensionnelles, la catégorie du nombre linguistique du français écrit pourrait « bien être l'un des domaines linguistiques les plus longs à maîtriser » (Jaffré & David, 1999 : 7).

#### **4.2.2.3.1. L'accord en nombre : la logique de la pluralité vs la difficulté de l'inaudible**

Comme il a été mentionné plus haut, l'expression morphologique du nombre connaît un avantage par rapport à celle du genre, puisqu'il existe « un lien entre l'expression linguistique et la pluralité dans le sens » (Ågren, 2005). C'est ce que Jaffré et David appellent « la logique de la pluralité » (1999). Ainsi, le nombre connaît un fondement sémantique simple et transparent, et cela devrait faciliter l'acquisition de son expression morphologique au niveau linguistique.

Remarquons encore que la difficulté de la représentation morphologique du nombre au niveau de l'écrit réside dans le décalage entre l'oral et l'écrit, ainsi que le coût cognitif qui en résulte. Néanmoins, cette difficulté pourrait être dans un sens compensée par la régularité de l'expression morphologique du nombre au niveau graphique : le *-s* pour les noms et le *-nt* pour les verbes (la 3<sup>e</sup> personne du pluriel). Cela explique probablement le fait que les enfants, dès le début de première primaire (CP), connaissent déjà la marque du pluriel et savent l'interpréter (Fayol, 2003). Jaffré et David (1999 : 15) confirment également dans leur étude que dès l'âge de six/sept ans, certains enfants savent justifier la marque du pluriel *-s* : « ...y en avait beaucoup alors on met un 's' », et qu'un peu plus tard, ils peuvent valider le singulier sans « s » : « pasque y'a pas plusieurs chocolats ». Or, bien que les enfants disposent d'une connaissance, ils ne la mettent pas en œuvre. Selon un corpus contenant des écrits produits par des enfants du cycle II (maternelle, CP et CE1), Jaffré et David constatent que seulement 10% des noms dans un contexte au pluriel sont bien accordés en nombre. Fayol (2003) propose une hypothèse pour expliquer ce fait : « les enfants de ce niveau ne disposent pas d'une capacité attentionnelle suffisante pour gérer l'application de l'accord nominal ». En menant une étude auprès des enfants de CP et CE1, Largy (2001) confirme cette hypothèse : les enfants sont capables de détecter les erreurs d'accord en nombre et de les corriger, alors même qu'ils les commettent eux-mêmes très souvent.

Dans la sphère nominale, Fayol nous indique un autre facteur facilitant l'apprentissage de l'expression morphologique du nombre : « la présence d'un -s en fin de mot constitue un indice fort et fiable du la pluralité<sup>44</sup> » (2003b : 48). Et la probabilité de commettre des erreurs d'interpréter la présence d'un -s final est minime. Cela est également confirmé par la recherche de Totereau et *al.* (1997), qui montre que le repérage et l'interprétation d'un -s final chez des enfants français sont à la fois précoces et sûrs. En revanche, le -e final comme indice du féminin reste beaucoup moins fort et fiable que le -s du pluriel, puisqu'il existe en français de nombreux noms masculins se terminant en -e.

Or, il en va tout autrement pour la sphère verbale puisque la marque du pluriel -s, omniprésente dans la sphère nominale, sera remplacé par -nt pour les verbes. Comme l'indiquent Jaffré et David (1999 : 18) : « Cette dissymétrie est à l'origine de bien des difficultés dans la maîtrise de l'orthographe du nombre ». Dans les écrits d'enfants, les chercheurs observent que les enfants surgénéralisent le -s aux verbes, en utilisant le -s à la fois pour les noms et pour les verbes (Fayol, 2003). Au fur et à mesure de la scolarité, les enfants arrivent de mieux en mieux à identifier les différentes catégories grammaticales et ainsi, à mieux opérationnaliser la morphologie du nombre. Néanmoins, selon les travaux de Brissaud et Jaffré (2003), les erreurs relevant de l'accord en nombre persistent jusqu'à la fin du lycée et même au-delà.

#### **4.2.2.3.2. Une perspective développementale sur l'acquisition de la morphologie écrite du nombre en français L1**

C'est seulement au cours des dernières décennies que les chercheurs ont commencé à envisager la théorisation de l'apprentissage de la morphologie écrite du nombre. En reposant sur une étude fine des formes erronées produites par les apprenants français, les recherches ont fait apparaître que les performances de la morphologie écrite du nombre suivaient une évolution complexe (Fayol, 2003).

---

<sup>44</sup> Selon la base informatisée Brulex (Content, Mousty & Radeau, 1990 cités dans Fayol, 2003) sur 19 384 mots, seuls 504 se terminent par un -s au singulier, tels qu'*une souris, un tapis*.



Lors de la première phase, les enfants français de CP ne marquent le pluriel ni pour les noms, ni pour les verbes. Ils produisent le plus souvent les mots sous leur forme neutre. Cependant, leur performance en compréhension est plus précoce qu'en production. Autrement dit, pendant cette période, les enfants connaissent la marque du pluriel nominal et peuvent l'interpréter, mais la mise en œuvre de cette règle leur est problématique. Selon Fayol, ce fait pourrait s'expliquer par une capacité attentionnelle insuffisante des enfants de ce niveau. Lors de la production écrite, leur attention est principalement centrée sur, d'une part, « la détermination de l'orthographe du mot » et, d'autre part, « la réalisation graphique » (Fayol, 2003). Comme la transcription graphique mobilise déjà la plupart de l'attention des enfants, ils ne sont plus en mesure d'appliquer l'accord nominal.

Durant une deuxième phase (surtout en CE1 et CE2), les enfants sont capables de mettre en œuvre la marque pluriel –s, mais ils l'emploient à la fois pour les noms (correctement) et pour les verbes (erronément) (Fayol, 2003). Cette surgénéralisation du morphème –s (pluriel nominal) aux verbes s'explique probablement par le fait que les enfants de ce niveau ne différencient pas encore les catégories syntaxiques.

Dans une troisième phase, les enfants savent utiliser la flexion –nt pour les verbes. Pourtant, les chercheurs ont observé une surgénéralisation de cette flexion –nt aux noms. Notons que cette compétition entre les deux marques pluriels –s et –ent se produit notamment lorsque les noms ont un homophone verbal (Largy, Fayol & Lemaire, 1996). Les erreurs de substitutions de –s à –nt sont également observées par les chercheurs, qui affectent notamment « les verbes ayant un homophone nominal plus fréquent » (Fayol, 2003). Étant donné que l'application d'une règle ne dépend pas de la fréquence des items, les erreurs de substitution ne pourraient pas ainsi être dues au recours à la règle de l'accord en nombre. Il est fort probable que ces formes erronées s'expliquent par la récupération soit des associations entre un mot et une flexion (morphèmes –s ou –nt) soit des instances composées activées globalement, par exemple, (*ils*) *ferment* ou (*les*) *fermes*, lors de la réalisation graphique.

#### 4.2.2.3.3. Vers un modèle d'apprentissage : apprentissage d'instances et apprentissage de règle

L'évolution présentée ci-dessous nous conduit à envisager l'opposition entre l'apprentissage d'instances et l'apprentissage de règle, problème beaucoup discuté par les chercheurs qui veulent construire un modèle d'apprentissage de la morphologie du nombre (voir, par exemple, Largy & Cousin, 2004 ; Cousin, Largy & Fayol, 2004). Ce débat théorique mené par les psycholinguistes en France concerne la relation complexe « entre application de règles morphologiques, ce qui correspond à des processus combinatoires, et récupération d'instances toutes faites » (Hedbor & Ågren, 2006).

Selon Anderson (1993, 1995), les savoirs enseignés et mémorisés verbalement sont initialement stockés chez les apprenants sous la forme des connaissances déclaratives. La pratique conduirait par la suite à une procéduralisation des connaissances. Au stade initial de la procéduralisation, le coût de l'application de la règle serait très élevé. Chez les apprenants-débutants, cette application pourrait également être en concurrence avec la transcription graphique lors de la mobilisation attentionnelle et aussi sensible à la difficulté des tâches d'écriture. De ce fait, au début de la procéduralisation de la règle d'accord en nombre, les apprenants pourraient omettre de mettre les marques du nombre, mais avec les effets de la pratique, une diminution rapide de ces erreurs d'omission pourrait être observée (Thévenin, Totereau, Fayol & Jarousse, 1999). Ainsi, les apprenants automatisent progressivement les procédures d'accord. Cependant, il nous semble essentiel de faire mentionner qu'il existe toujours les situations suffisamment coûteuses en attention, où la mobilisation de la procédure d'accord est perturbée. Ainsi apparaissent les oublis des marques du nombre, y compris chez les adultes (Fayol, Largy & Lemaire, 1994).

En ce qui concerne la morphologie écrite du nombre, la difficulté relève également de l'existence des deux marques pluriels, morphèmes *-s* et *-nt*, dont « l'utilisation repose sur des conditions qui sont à la fois partiellement communes (la pluralité) et différentes (nom ou adjectif : *-s* ; verbe : *-nt*) » (Fayol, 2003). Pour les psycholinguistes,

le travail principal se trouve dans l'identification des processus qui permettent aux apprenants d'associer correctement les marques du pluriel et catégories syntaxiques. En effet, comme le souligne Fayol (2003), « rien n'assure a priori que les processus mis en œuvre soient des correspondants mentaux des règles enseignées ».

De plus, plusieurs études nous suggèrent que l'acquisition de la morphologie écrite du nombre ne s'appuie pas uniquement sur l'apprentissage des règles d'accord, mais qu'elle est également sensible à la fréquence des « instances spécifiques d'associations entre certaines racines et certains morphèmes » (Fayol, 2003). Cet effet de fréquence par le mot *asperge*, qui est plus fréquemment rencontré comme nom que comme verbe. De ce fait, il indique que le mot *asperge* serait plus facilement associé à la marque du pluriel nominal *-s* que celui du pluriel verbal *-nt*. Ainsi, la production de la morphologie écrite du nombre concerne non seulement l'automatisation de règles mais également « la mémorisation d'associations entre morphèmes prêtes à servir » (Fayol, 2003).

Étant donné que les scripteurs disposent de deux processus différents pour le traitement des accords en nombre, il est pertinent de penser qu'il y aurait compétition entre l'application d'une règle et la récupération directe d'instances stockées en mémoire. Ici nous pouvons reprendre les erreurs de substitution dont nous avons parlé plus haut. Lorsque le scripteur doit écrire une phrase telle que *L'épicier sort les cagettes et il les asperge*, il se trouve dans une situation où mobilisation de règles et récupération d'instances conduisent à des réalisations morphologiques différentes, comme le mot *asperge* pourrait être à la fois nom et verbe. L'occurrence d'effet d'homophonie et de fréquence des homophones (le mot *asperge* est plus fréquemment rencontré comme nom que comme verbe) nous suggère que le recours à la récupération d'instances gagne, au moins dans certains cas, lors de la réalisation d'accords en nombre. Ce processus permet au scripteur d'effectuer la plupart des accords sans recourir aux procédures coûteuses en attention et fragiles. Mais il faut insister sur le fait que ces dernières procédures sont toujours disponibles au scripteur, et que le scripteur pourrait les mobiliser comme procédure de secours.

A part les discussions théoriques décrites au-dessus, il existe également deux autres éléments qui pourraient jouer un rôle lors de la production morphologique écrite du nombre.

Tout d'abord, c'est l'effet du caractère sémantique. Comme nous l'avons indiqué plus haut, l'accord nominal en nombre est sémantiquement motivé, tandis que l'accord verbal est purement formel. Est-ce que l'apparition tardive de l'accord verbal observée chez les apprenants, dont nous avons parlé plus haut, est due à l'absence de motivation sémantique ? Ou est-ce qu'elle s'explique plutôt par la fréquence de la marque du pluriel (le morphème *-s* est plus fréquent que le morphème *-nt*) ? L'acquisition de l'accord adjectival pourrait probablement mettre en lumière cette question. L'accord adjectival en nombre, comme celui du verbe, est purement formel. De ce fait, si c'est l'effet sémantique qui intervient, l'accord adjectival en nombre devrait être acquis plus tardif que celui du nom. En revanche, si c'est la fréquence de la marque du pluriel qui joue un rôle, le pluriel adjectival devrait apparaître plus précocement que le pluriel verbal, puisque dans la graphie, le pluriel est marqué, pour la plupart des adjectifs comme pour les noms, par le morphème *-s*. Les recherches de Thévenin et *al.* (1999) nous suggèrent que chez les enfants français, le pluriel adjectival est significativement plus précoce que le pluriel verbal, mais aussi plus tardif que le pluriel nominal. La précocité d'utilisation du morphème *-s* pour accorder les adjectifs plaide effectivement en faveur de la fréquence et la similitude de la marque du pluriel, mais comme les verbes, l'absence de la motivation sémantique retarde en même temps cette utilisation du morphème *-s* pour les adjectifs.

L'information syntaxique constitue le deuxième élément qui pourrait influencer la gestion des accords. En ce qui concerne l'accord de l'adjectif, Pacton et *al.* ont mené une recherche sur les erreurs de substitution de *-nt* à *-s*. Leur travail a montré que ce type d'erreur est moins fréquent pour les adjectifs qui sont mis en position post verbal (par exemple : *Les bouteilles vendues du marchand sont vides.*) que ceux qui se trouvent en position post nominal (par exemple : *Les bouteilles vides du marchand sont vendues.*). Ainsi, il est de mise de penser que l'information syntaxique favorise la réduction des erreurs de substitution de *-nt* à *-s*. Cependant, il faut noter que ce type

d'erreur survient alors même que les participants ne confondent pas les deux catégories syntaxiques, à savoir le verbe et l'adjectif, ce qui donne lieu à une interprétation en faveur des erreurs de récupération, au lieu de catégorisation.

En suivant cette logique, Fayol (2003) a comparé l'accord en nombre des adjectifs et ceux des noms pour tester « l'effet de proximité de l'item par rapport au déterminant ». Étant donné que l'opposition du nombre est directement audible pour les déterminants du français, il est fort probable que l'item qui les suit immédiatement serait mieux accordé que les items suivants. Si l'on prend l'exemple de la phrase *Les musiques douces accompagnent le repas*, le nom *musiques* serait fréquemment accordé par les scripteurs plus que l'adjectif *douces* puisque le nom *musiques* est placé immédiatement après le déterminant *les*. Les données recueillies auprès des enfants de 8 à 11 ans ont montré, d'une part, que les noms étaient plus fréquemment accordés que les adjectifs et, d'autre part, que l'item se situant immédiatement le déterminant était mieux accordé que l'item suivant, observation qui nous renseigne à la fois sur l'effet de la motivation sémantique et sur celui de la dimension syntaxique.

A partir des discussions théoriques présentées plus haut, nous examinerons par la suite la production écrite de l'accord en nombre observée chez les étudiants chinois de notre corpus. Les trois groupes d'étudiants des niveaux différents seront comparés d'un point de vue du développement. Et cette partie empirique sera divisée en quatre sections : celle du nom, du verbe, de l'adjectif et du déterminant. Les deux morphèmes du nombre, -s pour les éléments des syntagmes nominaux et -nt pour les verbes, se trouveront au centre de notre analyse. La discussion théorique sera intégrée dans notre analyse des résultats empiriques, et nous émettrons également les hypothèses possibles pour expliquer le profil développemental de l'accord en nombre ainsi que les tendances de la mise en œuvre de cet accord observé chez les étudiants chinois de notre corpus.

#### **4.2.2.4. La morphologie écrite du nombre nominal**

##### **4.2.2.4.1. Résultats**

Dans les parties précédentes, nous avons cité plusieurs études qui décrivent et analysent l'acquisition et la mise en œuvre de la morphologie écrite du nombre nominal des enfants français. Ces études plaident en faveur d'une compréhension et d'une production précoce de l'accord nominal en nombre, mais elles suggèrent également une évolution complexe que suit l'acquisition de ce domaine. En suivant une méthodologie différente de celle des recherches mentionnées auparavant, nous ne pourrions nous prononcer que sur le traitement de l'accord en production écrite libre en français L2. Cependant, il nous semble possible de faire un certain nombre de parallèles entre le domaine du français L1 et celui du français L2 en ce qui concerne la morphologie écrite du nombre nominal.

Le tableau suivant illustre bien le fait que le marquage du pluriel nominal est précocement maîtrisé par les apprenants chinois, puisque la plupart des noms apparaissant dans des contextes au pluriel, le nombre est marqué par le morphème *-s* ou *-x*, et que cela vaut pour les trois stades. Déjà au stade initial, celui où se trouvent les étudiants qui ont suivi une année d'études de français et représentent un niveau linguistique élémentaire, le pourcentage d'accord correct est 89.6%.

Aux stade 2 et stade 3, le pourcentage correct du marquage du nombre augmente, d'une manière assez faible, jusqu'à respectivement 90.2% et 91%. Ainsi, les taux de réussite sont pratiquement les mêmes au stade 3 qu'au stade 1 et au stade 2. Autrement dit, le profil développemental reste très faible entre le niveau initial et le niveau assez avancé. L'hypothèse susceptible d'expliquer ce fait est que, après deux ans ou trois ans des études de français, les étudiants de L3 et de L4 sont en mesure de produire le texte avec une structure syntaxique compliquée et un lexique assez riche. L'essentiel de leur attention est sans doute capté par, d'une part, la détermination du contenu du texte et, d'autre part, la détermination de l'orthographe du mot et sa réalisation graphique. Comme ces deux éléments mobilisent la plus grande part de l'attention, les étudiants

ne disposent pas à un certain moment d'une capacité attentionnelle suffisante pour mettre en œuvre les marques du pluriel.

**Tableau 25.** L'accord nominal en nombre dans la totalité des contextes au pluriel

Stade	Contextes nominaux pluriels	Mots types	Occurrences marquées avec -s/-x	Pourcentage de noms marqués
Licence 2	271	42	243	89.6%
Licence 3	195	39	176	90.2%
Licence 4	269	41	245	91%

Cette évolution restreinte dans le domaine de l'accord nominal du nombre est également observée chez les apprenants suédois en français L2 (voir, par exemple, Hedbor & Ågren, 2006 ; Ågren, 2005). Les chercheurs ont recueilli les données auprès des apprenants suédois de quatre niveaux différents : 1<sup>ère</sup>, 2<sup>e</sup>, 4<sup>e</sup> et 5<sup>e</sup> année de français. De plus, les écrits d'un groupe des lycéens français au troisième ont été choisis comme groupe de contrôle. En ce qui concerne le taux de réussite, le stade 3 ne se distingue pas du stade 1 et du stade 2. Un changement a eu lieu au stade 4 avec un taux de réussite 96%, qui est déjà très proche de celui du groupe contrôle français. Autrement dit, avec 5 ans d'études de français, les élèves suédois ont un accord en nombre nominal quasi correct. Mais d'une manière globale, l'évolution reste toujours restreinte entre le stade 1 et le stade 4. Notons que le pourcentage d'accords corrects des apprenants chinois de la Licence 4 n'est que 91%, encore loin de celui du groupe contrôle français.

Ågren a également constaté une augmentation nette des occurrences au pluriel et notamment des noms types entre le stade 1 et le stade 4. Quant aux étudiants chinois de notre corpus, les contextes pluriels augmentent en fonction de niveaux différents mais les noms types restent toujours restreints. Cela s'explique probablement par la nature de notre tâche d'écriture, qui concerne la reconstitution d'un texte de source.

**Tableau 26.** L'accord nominal en nombre dans la totalité des contextes au pluriel chez les apprenants

suédois (Ågren, 2005)

Stade	Contextes nominaux pluriels	Mots types	Occurrences marqués avec -s/-x	Pourcentage de noms marqués
1	65	33	58	89%
2	224	89	198	88%
3	317	99	294	93%
4	282	93	272	96%
Contrôle	318	100	307	97%

Un autre élément qu'il faut faire mentionner ici, c'est les différences individuelles existant à tous les stades acquisitionnels. Comme l'indique le tableau ci-dessous, il existe à chaque stade un nombre important d'étudiants qui n'ont jamais omis le morphème du pluriel, et qui ne l'ont omis qu'une seule fois.

Pour le groupe de la Licence 2, l'écart type prononcé (20%) met en lumière les différences individuelles assez importantes à ce niveau. Nous constatons que 23 étudiants sur 40 n'ont jamais omis -s ou -x dans un contexte au pluriel, ainsi que 9 étudiants ne l'ont omis qu'une seule fois. Néanmoins, les performances des étudiants L2-01-011 et L2-02-007 attirent également notre attention. Les deux scripteurs ont fait respectivement 4 omissions sur 7 contextes et 4 omissions sur 8 contextes. De plus, la variation du nombre de contextes au pluriel est grande. Aucune occurrence au pluriel n'est trouvée dans les textes des scripteurs L2-01-018 et L2-01-020, alors que le scripteur L2-01-03 au même niveau en produit 16.

Les mêmes phénomènes sont également observés chez les scripteurs de la Licence 3 et de la Licence 4, qui nous indique les différences individuelles présentées à ces deux niveaux. Néanmoins, bien que la diminution soit modeste, il est à noter que l'écart type diminue avec les stades acquisitionnels.



**Tableau 27.** Différences individuelles des omissions du pluriel nominal

(Légende : *appr. 0 omis.* : Le nombre d'apprenants sans omission ; *1 omis.* : le nombre d'apprenants avec une seule omission ; *2-3 omis.* : le nombre d'apprenants avec deux ou trois omissions, *3<* : Le nombre d'apprenants avec plus de trois omissions, % correct moyenne : la variation de pourcentage correct parmi les apprenants d'un même stade ; écart type : mesure la variation à l'intérieur de chaque stade)

Stade	Nombre de contextes	Appr. 0 omis	% Appr. 0 omis	Appr. 1 omis	Appr. 2-3 omis	Appr. 3 <	% correct moyenne	Écart types
Licence 2 (n= 40)	271	23	57%	9	4	2	89% (33%-100%)	20%
Licence 3 (n= 23)	195	9	39%	8	6	0	89% (67%-100%)	13%
Licence 4 (n=28)	269	15	52%	8	3	2	91%	15%

#### 4.2.2.4.2. Discussion

Au travers d'une analyse quantitative, nous constatons que la morphologie écrite du nombre nominal est très prononcée dès le stade initial où le taux de réussite est déjà 90%. Cela met en évidence la production précoce du pluriel nominal chez les apprenants chinois. Cependant son développement entre les stades acquisitionnels n'est pas très marqué dans notre corpus. Et bien qu'en diminuant, la variation individuelle importante se présente à tous les stades. Toutefois, pour interpréter la nature des commissions de marques du pluriel commises par les scripteurs, il nous faut regarder les données de plus près.

Le tableau 28 illustre les mots non accordés produits par les étudiants de la Licence 2, ainsi que leurs contextes gauche et droite. Les 40 étudiants de L2 ont produit 269 formes nominales dans un contexte pluriel. Parmi elles, 242 ont été bien accordées en fonction du nombre. Observons que 5 omissions de la marque du pluriel sur 27 sont précédées d'un numéral. A première vue, l'on pourrait croire que ces 5 omissions pourraient être interprétées par le contexte (présence d'un numéral). Cependant, en regardant les données de tout le groupe, il s'avère que les apprenants de ce niveau ont aussi produit 45 occurrences où un nom précédé d'un numéral est bien accordé en nombre : le nom *an* précédé du numéral *six* est bien accordé dans toutes les 17 occurrences ; 16 occurrences sur 17 sont bien accordées pour le nom *frère* qui est précédé du numéral *deux* ; 12 occurrences sur 16 sont mises au pluriel pour le nom *recommandation* qui se situe après le numéral *trois*. Ainsi, il nous semble que ces 5 omissions relèvent d'un problème de vocabulaire. Par rapport au nom *recommandation*, les noms *an* et *frère* sont très fréquents et vus souvent sous leur forme du pluriel.

**Tableau 28.** Erreurs relevant de l'accord du nombre nominal commises par les apprenants de L2 et leur contexte

Scripteur	Contexte gauche	Mot non-accordé	Contexte droite
L2-01-002	ses deux frères sont  {tout}_{tous}	{chevalier}_{chevaliers}	

L2-01-002	les autres	{chevalier}_ {chevaliers}	trouvent que...
L2-01-011	les	{chevalier}_ {chevaliers}	
L2-01-019	les	{chevalier}_ {chevaliers}	{vermail}_ {vermeils}
L2-01-008	{un}_ {une} troupe de	{chevalier}_ {chevaliers}	
L2-01-008	{Ce}_ {Cette} troupe de	{chevalier}_ {chevaliers}	{étaient}_ {était}
L2-02-007	une troupe de	{chevalier}_ {chevaliers}	
L2-01-017	un groupe de	{chevalier}_ {chevaliers}	
L2-01-007	trois	{recommandation}_ {recommandations}	
L2-01-021	trois	{recommandation}_ {recommandations}	utiles
L2-02-002	trois	{recommandation}_ {recommandations}	
L2-02-011	trois	{recommandation}_ {recommandations}	
L2-02-008	Son père et deux	{frère}_ {frères}	ont allées...
L2-01-011	les	{conseil}_ {conseils}	de {pru_homme}_ {pru_QUOTE_hommes}.
L2-01-017	les	{conseil}_ {conseils}	
L2-02-010	les	{conseil}_ {conseils}	des {prudhommes}_ {prud_QUOTE_hommes}
L2-02-011	les	{conseil}_ {conseils}	des prud'hommes
L2-02-007	des	{armée}_ {armes}	
L2-02-007	les	{armée}_ {armes}	du chevalier vermeil
L2-01-011	des	{officier}_ {officiers}	étaient...
L2-01-014	ces	{couple}_ {couples}	ne que riraient

L2-02-009	des	{livre}_{livres}	de chevalier
L2-02-003	des	{feu}_{feux}	
L2-01-002	sa nouvelle	{armures}_{armure}	
L2-01-002	le	{pieds}_{pied}	
L2-01-005	un	{mots}_{mot}	de chevalier
L2-01-020	quelque	{choses}_{chose}	
L2-01-016	comme	{d_QUOTE_habitudes}_ {d_QUOTE_habitude}	

Un autre mot qui attire notre attention, c'est *chevalier*, nom au contexte pluriel le plus utilisé par les apprenants de L2 : 66 occurrences trouvées dans les textes écrits. Parmi elles, 8 omissions ont été commises par 6 apprenants. Plus précisément, chez les scripteurs L2-01-002 et L2-01-008, le nom *chevalier* revient non accordé à deux reprises sans apparaître sous sa forme correcte. Notons que le scripteur L2-01-008 a oublié d'accorder le mot *chevalier* lorsque ce nom se trouvait après le déterminant composé *une troupe de*. Ce type d'omission n'est pas un cas singulier : deux scripteurs L2-02-007 et L2-01-017 ont également omis de mettre la marque du pluriel *s* pour le mot *chevalier*, qui se situait après le déterminant composé *une troupe de* ou *un groupe de*. Ainsi, nous avons trouvé au total 4 omissions sur 24 lorsqu'il s'agit de l'accord du nom *chevalier*, et que ce nom se trouve après *une troupe de* ou *un groupe de*. Autrement dit, la moitié des omissions du morphème *s* pour le nom *chevalier* au contexte pluriel est liée à la présence du déterminant composé *une troupe de* ou *un groupe de*. Sans le déterminant le plus fréquent *les*, la présence du déterminant composé *une troupe de* ou *un groupe de*, même sémantiquement motivée, ne permet pas à certains scripteurs de garder la vigilance d'accorder le nom en termes de nombre.

Un phénomène particulier aux scripteurs de la Licence 2, c'est que certains apprenants ont choisi la forme du pluriel de noms, tandis que la forme du singulier est demandée par le contexte. Ici, il s'agit principalement des mots qui apparaissent souvent sous leur forme du pluriel : *mot*, *pied*, *chose*, etc. L'hypothèse la plus probable

susceptible d'expliquer ce fait est que ces scripteurs ne disposaient pas d'une capacité attentionnelle suffisante lors de la production écrite pour gérer la mise en œuvre de l'accord nominal, et que les formes les plus fréquentes ont ainsi été suremployées.

Pour les 23 apprenants de L3, ils ont produit au total 195 formes nominales qui se trouvaient au contexte pluriel. Nous pouvons constater que la proportion des noms au pluriel augmente, qui s'explique par le fait que les textes deviennent plus longs à ce stade. Cependant, le nombre de mots types reste au même niveau qu'au stade précédent. Parmi les formes nominales produites dans un contexte pluriel, 19 n'ont pas été bien accordées en termes de nombre. Et la plupart des erreurs sont interprétables à partir de l'analyse insuffisante du contexte.

Nous avons constaté la même difficulté liée au nom *chevalier* lors qu'il avait pour contexte gauche le déterminant composé *une troupe de* chez les apprenants de L3 que chez leurs camarades de L2. 22 occurrences concernant le syntagme nominal *une(la) troupe de chevalier(s)* ont été trouvées, et 8 ont été privées du morphème *s* pour le nom *chevalier*. Cependant, nous n'avons trouvé qu'une seule omission du morphème *s* pour le nom *chevalier* (sur 26) lorsqu'il se trouvait après les déterminants, tels que l'article indéfini *des*, l'article défini *les* et le déterminant possessif *ses*. Il est fort probable que certains apprenants ont été affrontés à une difficulté particulière quand le nom *chevalier* se trouvait après le déterminant composé *une troupe de*, contexte gauche inhabituel. Donc leurs ressources attentionnelles étaient surchargées par cette difficulté et cela leur empêchait de faire une analyse métalinguistique suffisante, qui aurait dû permettre la mise en œuvre de l'accord du nombre.

La difficulté liée à la présence de l'adjectif numéral revient également à ce stade. Étant donné l'absence du déterminant habituel, ainsi qu'une analyse insuffisante du contexte gauche, certains locuteurs ont omis de mettre le morphème *s* pour les noms qui se trouvaient après les numéraux *deux* et *trois*. Lors de la production des syntagmes nominaux, ils ne disposaient pas, nous semble-t-il, d'une capacité attentionnelle suffisante pour gérer l'application de l'accord nominal.

**Tableau 29.** Erreurs relevant de l'accord du nombre nominal commises par les apprenants de L3 et leur

contexte

<b>Scripteur</b>	<b>Contexte gauche</b>	<b>Mot non-accordé</b>	<b>Contexte droite</b>
L3-007	une troupe de	{chevalier}_ {chevaliers}	
L3-009	une troupe de	{chevalier}_ {chevaliers}	
L3-010	une troupe de	{chevalier}_ {chevaliers}	
L3-011	une troupe de	{chevalier}_ {chevaliers}	
L3-011	la troupe de	{chevalier}_ {chevaliers}	
L3-012	une troupe de	{chevalier}_ {chevaliers}	
L3-016	{un}_ {une} troupe de	{chevalier}_ {chevaliers}	
L3-017	{un}_ {une} troupe de	{chevalier}_ {chevaliers}	
L3-001	trois	{recommandation}_ {recommandations}	
L3-008	trois	{conseil}_ {conseils}	
L3-012	deux	{frère}_ {frères}	{meuraient}_ {mouraient}
L3-015	ses deux	{frère}_ {frères}	
L3-016	ses deux	{frère}_ {frères}	
L3-004	des	{autre}_ {autres}	{commençait}_ {commençaient}
L3-005	ses	{adieu}_ {adieux}	
L3-007	les conseil des	{prud_QUOTE_homme}_ {prud_QUOTE_hommes}	
L3-013	ses	{parole}_ {paroles}	
L3-015	les	{chevalier}_ {chevaliers}	Vermeil
L3-002	quelque	{choses}_ {chose}	

En ce qui concerne les performances des apprenants de L4, nous reconnaissons les tendances commentées aux stades précédentes : la difficulté liée particulièrement à la présence des locutions *une troupe de*, *un groupe de*, et *beaucoup de*, qui se trouvaient avant les noms à accorder comme contexte gauche. Notons que ces locutions marquent toutes la pluralité au niveau sémantique, cependant, cette indication ne permet pas à certains scripteurs de bien gérer l'application de l'accord nominal. La deuxième tendance concerne la présence des adjectifs numéraux comme contexte gauche du nom à accorder, dont nous avons déjà beaucoup parlé plus haut.

A un niveau individuel, le scripteur L4-019 attire notre attention, qui semble en difficulté pour un nom spécifique, *\* les chevalier* ou *\* ses chevalier*, qui l'a mené à une erreur d'accord à trois reprises. Le scripteur L4-009 a produit au total 13 formes nominales au contexte pluriel, et il a bien accordé les autres noms après les déterminants *les* ou *ses*, tels que *les dames*, *ses armes*. Ainsi, les formes erronées relèvent probablement d'un problème de vocabulaire. Le mot *chevalier*, nom que le scripteur n'avait pas sans doute connu avant la tâche d'écriture, brouillait la mise en œuvre de l'accord du nombre.

**Tableau 30.** Erreurs relevant de l'accord du nombre nominal commises par les apprenants de L4 et leur contexte

Scripteur	Contexte gauche	Mot non-accordé	Contexte droite
L4-004	une troupe de	{chevalier}_ {chevaliers}	
L4-009	une troupe de	{chevalier}_ {chevaliers}	
L4-025	une troupe de	{chevalier}_ {chevaliers}	
L4-018	le groupe de	{chevalier}_ {chevaliers}	du Roi Arthur
L4-018	autour du troupe de	{chevalier}_ {chevaliers}	
L4-028	{une}_ {un} groupe de	{chevalier}_ {chevaliers}	
L4-006	beaucoup de	{difficulté}_ {difficultés}	
L4-005	trois	{recommandation}_ {recommandations}	

L4-009	trois	{recommandation}_ {recommandations}	
L4-018	ses deux	{frère}_ {frères}	
L4-019	ses deux	{frère}_ {frères}	
L4-003	les conseil des	{prud_QUOTE_homme}_ {prud_QUOTE_hommes}	
L4-024	les conseil des	{prud_QUOTE_homme}_ {prud_QUOTE_hommes}	
L4-012	les	{chevalier}_ {chevaliers}	
L4-016	les histoires des	{chevalier}_ {chevaliers}	
L4-019	les	{chevalier}_ {chevaliers}	
L4-019	ses	{chevalier}_ {chevaliers}	
L4-019	ses	{chevalier}_ {chevaliers}	
L4-004	{Tous}_ {Toutes} les	{chose}_ {choses}	+R[avaient]+R+//étaient //+R {nouveau}_ {nouvelles} pour lui
L4-004	ses	{mot}_ {mots}	
L4-005	les	{conseil}_ {conseils}	des prud'hommes
L4-009	les	{montation}_ {montagnes},	
L4-018	à cause des	{combat}_ {combats}	
L4-023	les	{épée}_ {épées}	

Comme nous l'avons indiqué plus haut, chez les étudiants chinois de notre corpus, l'accord nominal en nombre est déjà très prononcé dès le stade initial. Selon nous, plusieurs faits peuvent expliquer cette acquisition précoce.

Tout d'abord, il faut prendre en compte la structure du groupe nominal observé dans les textes produits par les apprenants chinois. Le tableau ci-dessous nous montre les pourcentages de réussite de l'accord nominal du nombre selon les différents



contextes. Pour tous les niveaux, le taux de réussite est assez élevé avec la présence du déterminant au pluriel. Néanmoins, se situant après l’adjectif numéral ou le déterminant composé tel que *une troupe de*, le nom n’est pas aussi bien accordé en nombre qu’après le déterminant au pluriel.

**Tableau 31.** Taux de réussite de l’accord nominal du nombre selon les différents contextes

Stade	Déterminant + nom		Numéral + nom		Déterminant composé + nom (ex : <i>une troupe de</i> )	
L2	173/191	90%	43/48	89%	28/32	87%
L3	135/141	96%	25/30	83%	16/24	67%
L4	192/205	94%	42/46	91%	10/17	59%
Total	500/537	93%	110/124	89%	54/73	74%

Notons que la structure dominante choisie par les apprenants est sous la forme de « déterminant au pluriel + nom au pluriel ». Selon les différents niveaux d’études, les groupes nominaux ainsi construits occupent respectivement 70% (191 occurrences sur 271), 72% (141 occurrences sur 195) et 76% (205 occurrences sur 268) sur toutes les formes nominales au contexte pluriel. Tous les deux éléments dans cette structure exigent le marquage morphologique du nombre, mais contrairement au marquage du nom, celui du déterminant introduit une modification de l’image phonologique. De plus, les déterminants au pluriel sont très fréquents dans l’input des apprenants de français. Ainsi, il est pertinent de penser que, lors de la mise en œuvre de l’accord du nombre, l’essentiel de l’attention de l’apprenant est capté par celle du nom. Ce fait pourrait probablement faciliter l’accord morphologique du nom et expliquer le taux de réussite assez élevé dès le stade initial. Et si l’apprenant commence à produire des groupes nominaux plus complexes, il serait confronté à plus de difficultés à gérer l’application de l’accord nominal du nombre.

Bien que l’emploi des adjectifs soit relativement restreint dans ce corpus, il est toujours intéressant d’examiner l’application de l’accord du nombre dans les groupes

nominaux qui contiennent un ou plusieurs adjectifs. Le tableau 32 nous indique les pourcentages corrects élevés des noms accordés au pluriel dans des groupes nominaux complexes. Signalons que pour tous les niveaux, le pourcentage de réussite est plus élevé dans un groupe nominal complexe que dans un groupe nominal simple. Cependant, dans certains groupes nominaux, la marque du nombre nominal était bien mise en place, alors que l'adjectif n'était pas accordé en terme du nombre. Ces observations illustrent bien que dans un groupe nominal, le morphème *-s* est marqué pour le nom « malgré la présence d'un adjectif » (antéposé ou postposé) et « indépendamment de la présence ou l'absence du *-s* sur cet adjectif » (Ågren, 2005 : 75).

**Tableau 32.** Accord nominal du nombre dans un groupe nominal contenant un ou plusieurs adjectifs

Stade	Déterminant + nom (sans adjectif)		Adjectif + nom			
			Adjectif antéposé		Adjectif postposé	
L2	145/161	90%	9/10	90%	19/20	95%
L3	120/126	95%	4/4	100%	11/11	100%
L4	170/183	93%	9/9	100%	13/13	100%
Total	435/470	92%	22/23	96%	43/44	98%

Un autre élément qui pourrait faciliter la mise en place de l'accord nominal du nom, c'est le facteur sémantique. Contrairement aux adjectifs ou aux verbes, les noms connaissent un grand avantage sémantique en ce qui concerne l'accord en nombre. Comme nous l'avons indiqué antérieurement, les marques du nombre changent le contenu conceptuel des noms. Il n'en est pas de même avec les autres composants du groupe nominal ou verbal. Les études sur l'acquisition de l'accord du nombre en français L2 nous montrent que cette particularité du nom est cruciale pour l'acquisition et que « les marques redondantes sont plus vulnérables que les marques porteuses d'un contenu conceptuel » (Ågren, 2005 : 75). Nous reviendrons à ce facteur sémantique dans les analyses de l'accord adjectival et verbal du nombre.

Pour expliquer l'accord précoce observé chez les apprenants chinois, il nous semble indispensable de prendre en compte leurs connaissances linguistiques préalables. Comme nous l'avons indiqué auparavant, la langue chinoise ne connaît pas la variation morphologique. Par contre, l'anglais, qui est la première langue étrangère des apprenants chinois de notre corpus, marque le nombre nominal aussi bien à l'oral qu'à l'écrit, avec le même morphème qu'en français : le *-s*. Ainsi, pour les apprenants, l'accord nominal du nombre ne relève pas d'une connaissance complètement nouvelle, et il est pertinent de rendre compte d'un transfert éventuel de cette connaissance de l'anglais vers le français.

Le dernier élément que nous aimerons signaler pour comprendre l'acquisition précoce de l'accord nominal du nombre concerne l'enseignement reçu par les apprenants. Notons que les apprenants chinois reçoivent au cours de français les explications grammaticales spécifiquement dédiées à cet apprentissage. Autrement dit, les marques morphologiques du nombre et leur fonctionnement font l'objet d'un enseignement explicite. Cet enseignement fait appel à une analyse métalinguistique permettant aux apprenants de signaler les marques morphologiques du nombre, leurs co-occurrences, ainsi que leurs associations avec les catégories syntaxiques. A part l'enseignement de règles explicites, les apprenants passent beaucoup de temps à réaliser les exercices, au cours desquels ils ont à catégoriser correctement les items et choisir les flexions congruentes selon le contexte morphosyntaxique. De plus, au moment d'apprendre à écrire en français, à l'âge de 18 ans, les apprenants chinois, disposent d'une capacité cognitive plus élevée que les enfants français. L'enseignement explicite et la capacité cognitive plus élevée facilitent tous les deux une acquisition précoce de la morphologie écrite du nombre.

#### **4.2.2.5. La morphologie écrite du nombre adjectival**

Nous présenterons dans cette partie la performance de l'accord adjectival du nombre des apprenants de notre corpus. Selon le tableau suivant, nous pouvons constater que, malgré le nombre important de formes adjectivales, les apprenants n'ont pas produit

beaucoup d'adjectifs dans un contexte au pluriel en écrivant leur texte. Cependant, il est toujours possible de repérer certaines tendances dans ce domaine liées à leur traitement de la morphologie écrite du nombre adjectival. Nous analyserons les adjectifs au contexte pluriel en les groupant en trois catégories différentes selon leur placement par rapport au nom à qualifier : adjectif antéposé, postposé et position prédicative.

**Tableau 33.** Nombre des adjectifs produits selon les différents niveaux

Stade	Adj. produits au total	Adj. produits dans le contexte pluriel
Licence 2 (n=40)	615	38
Licence 3 (n=23)	375	25
Licence 4 (n=28)	472	30

#### 4.2.2.5.1. Résultats

Comme ce que nous avons observé pour l'accord nominal du nombre, la morphologie écrite du nombre adjectival est également très prononcée dès le stade initial, où le pourcentage de réussite est déjà 86%. Nous constatons que cet accord en nombre est relativement bien maîtrisé par les apprenants de trois niveaux : les pourcentages corrects pour L3 et L4 sont respectivement 96% et 90%. Néanmoins, son développement de la L2 à la L4 n'est pas très marqué, comme ce qui est montré dans le tableau ci-dessous. Comme généralement chaque apprenant n'ont produit qu'une ou deux formes adjectivales au contexte pluriel<sup>45</sup>, la variation individuelle ne sera plus discutée dans cette partie.

Contrairement à l'observation tirée de notre corpus, Ågren (2005) a constaté un

---

<sup>45</sup> Pour tous les niveaux acquisitionnels, il existe les apprenants qui n'ont produit aucun adjectif au contexte pluriel. Et le nombre maximal des formes adjectivales produites par personne est trois.

vaste développement dans ce domaine chez les apprenants suédois de français. Au stade initial, il est très rare que les apprenants fassent l'accord du nombre entre le nom et l'adjectif, et le pourcentage de réussite est moins de 20%. Mais au stade avancé, les apprenants semblent bien maîtriser cet accord en nombre, et le pourcentage correct est de presque 80%. Notons que même au niveau avancé, le pourcentage de réussite est encore loin de celui du groupe contrôle, qui est 91%. L'étude d'Ågren met aussi en évidence le fait que l'accord adjectival se fait plus tardivement que l'accord nominal.

#### **4.2.2.5.2. Le placement de l'adjectif par rapport au nom à qualifier**

Après une analyse quantitative générale, nous examinerons dans cette partie la relation entre le placement de l'adjectif par rapport au nom à qualifier et la réalisation de l'accord de cet adjectif en nombre. Fayol (2003) relève dans son étude l'importance du placement de l'adjectif pour son accord. L'adjectif se trouvant à la position prénominale devrait être le plus favorable pour l'accord en nombre, étant donné « le rôle du déterminant pour l'accord en nombre de la phrase » ainsi que « l'impact de la distance des constituants par rapport au déterminant » (Ågren, 2005).

Pienemann (1998), quant à lui, nous renseigne sur le rôle de la distance syntaxique sur l'accord. Selon sa théorie, l'accord des éléments qui se trouvent à l'intérieur d'un même syntagme serait plus facile à assurer que l'accord entre deux syntagmes différents, où il faut traverser une frontière syntaxique. Plus précisément, cette théorie parle en faveur des deux placements pré-nominal et post-nominale de l'adjectif pour l'accord, puisqu'ils se situent à l'intérieur du syntagme nominal et ainsi proche de l'élément qui impose l'accord. En revanche, le placement prédicatif de l'adjectif est moins favorable pour un accord en prenant en compte la distance syntaxique à traverser entre deux syntagmes différents. Cette théorie est confirmée par Bartning (2000) dans son travail sur l'acquisition de l'accord en genre, en indiquant que l'adjectif se situant à la position prédicative s'accorde moins bien que ceux aux positions pré ou post-nominales. Néanmoins, la différence entre les deux positions pré et post-nominale, qui se situent au même niveau hiérarchique, n'est pas évoquée dans ces études.

Les données de notre corpus dans ce domaine parlent également en faveur du

placement antéposé de l'adjectif pour l'accord en nombre. En se trouvant juste après le déterminant, où le nombre est audible, les adjectifs antéposés ont été mieux accordés que les adjectifs postposés pour tous les stades acquisitionnels. Cependant, il faut tenir compte du fait que les occurrences des adjectifs du pluriel ne sont pas nombreuses dans notre corpus, ce qui rend les données incertaines. Quant aux adjectifs prédicatifs, leur emploi semble particulièrement stable pour tous les trois niveaux. Nous n'avons trouvé qu'une seule forme erronée dans les textes de L2 et de L4. Ainsi, il est imprudent de dire que la position prédicative est difficilement maîtrisable pour les apprenants de notre corpus. En ce qui concerne la performance des apprenants de L3, toutes les occurrences produites à la position prédicative concernent un même adjectif *assis*, dont la forme ne modifie pas le contexte du pluriel. De ce fait, il nous semble pertinent d'exclure pour l'instant les adjectifs prédicatifs dans notre discussion.

**Tableau 34.** Taux de réussite de l'accord adjectival du nombre selon les différents contextes

Stade	Adj. postposé		Adj. antéposé		Adj. prédicative		Total
L2	16/20	80%	9/10	90%	7/8	87%	84%
L3	12/13	92%	4/4	100%	8/8	100%	96%
L4	12/14	86%	9/9	100%	6/7	86%	90%
Total	40/47	85%	22/23	96%	21/23	91%	89%

A partir de l'observation des données de notre corpus, nous constatons que les adjectifs au placement antéposé ont été mieux accordés que les adjectifs postposés. Il semble que la position à proximité du déterminant, où le nombre est audible favorise l'accord silencieux en nombre de l'adjectif. Néanmoins, il faut tout d'abord indiquer que les données de notre corpus ne sont pas statistiquement significatives pour montrer l'importance du placement de l'adjectif. De plus, il existe deux autres éléments qui pourraient expliquer l'avantage du placement pré-nominal de l'adjectif pour l'accord du nombre.

Premièrement, il ne s'agit pas des mêmes types d'adjectifs qui sont employés dans

les deux positions épithètes. En position antéposée dominant les adjectifs courts et fréquents, tels que *bon(bonne)*, *autre*, tandis qu'en position postposée, on trouve notamment les adjectifs comme intéressant, clair, inconnu, etc., qui sont moins fréquents que ceux en antéposition. Selon Ågren (2005), ce fait pourrait jouer un rôle pour l'accord adjectival en nombre, car il est possible que les adjectifs extrêmement fréquents en position antéposée « bénéficient d'un caractère presque figé ». A la lumière de cette hypothèse, la chercheuse a également noté que les scripteurs français dans le groupe contrôle employaient beaucoup plus d'adjectifs antéposés que les adjectifs postposés, et que c'était l'inverse pour les apprenants suédois de français de son corpus. Cependant, cette hypothèse demande encore les vérifications et approfondissements dans de futures recherches.

Le deuxième élément que nous voudrions signaler ici concerne l'influence possible mais négative de la langue maternelle, le chinois et de la première langue étrangère, l'anglais. Comme nous l'avons indiqué plus haut, en chinois, l'adjectif ne modifie pas la forme selon les contextes différents et comme l'épithète, il se situe toujours avant le nom. Et dans ce domaine, l'anglais connaît la même situation. De ce fait, la position post-nominale de l'adjectif correspond à une position marquée, qui est ainsi plus coûteux en attention lors de la production réelle. En ce qui concerne l'accord adjectival en nombre, il est pertinent de penser que pour un apprenant chinois de français, l'essentiel de son attention est probablement capté par le placement de l'adjectif en postposition, et qu'il ne dispose pas de capacité attentionnelle suffisante pour mettre en œuvre l'accord en nombre.

#### **4.2.2.6. La morphologie écrite du nombre verbal**

Nous consacrons cette partie à analyser le traitement de l'accord du nombre dans le syntagme verbal. Notons que cette analyse ne rendra compte que des formes verbales de la 3<sup>ème</sup> personne du pluriel. Comme l'indique Dubois (1967 : 35), « ...l'étude de l'opposition de marque est limitée à ce que l'on appelle la 3<sup>e</sup> personne ; on ne doit pas en effet considérer nous et vous comme les pluriels de *je* et de *tu*, mais comme d'autres

personnes ». Dans le syntagme verbal, le morphème silencieux –nt sert à marquer le nombre des verbes du premier groupe (ceux qui se terminent en –er), ainsi que d’une grande partie des verbes irréguliers, accompagnés d’un changement de radical (Ågrent, 2005). Pour les verbes *avoir, être, faire, aller*, les formes de 3PP au temps présent ont un statut assez particulier, puisque, d’un côté, leur fréquence d’occurrence est très élevée, et l’autre côté, leur marque du nombre n’est plus silencieuse.

#### 4.2.2.6.1. Résultats

Le tableau ci-dessous illustre les performances des étudiants chinois de trois niveaux différents. Comme ce qui se passe pour l’accord nominal et adjectival, le développement acquisitionnel en ce qui concerne l’accord verbal n’est pas très prononcé du niveau initial au niveau avancé. Nous constatons que l’accord du nombre est déjà bien maîtrisé par les étudiants de L2, dont le pourcentage correct est 90%. Notons que pour les étudiants de ce niveau, l’accord du nombre est exprimé aussi bien au temps du présent qu’aux temps du passé (passé composé, imparfait). Néanmoins, le pourcentage de réussite baisse jusqu’à 76% pour les étudiants de L3, et remonte finalement à 83% pour ceux de L4. Il faut indiquer qu’à partir de L3, le temps du passé simple est ajouté au texte de source, qui pourrait probablement s’expliquer cette baisse de performance.

**Tableau 35.** Performance de l’accord du nombre dans le syntagme verbal

(Légende : **Omission**, la forme verbale de la 3<sup>ème</sup> personne du singulier est mise au contexte du pluriel ; **radical incorrect**, verbe accordé en –nt mais avec un radical incorrect ; **sur-généralisation**, le morphème marquant le nombre est employé dans un contexte inapproprié ; **forme hybride**, deux types d’erreurs sont cumulés sous une même forme verbale)

	Contexte pluriel	Formes erronées	Taux de réussite	Omission	Radical incorrect	Sur-généralisation	Forme hybride
L2	59	6	90%	4	1	0	1
L3	46	11	76%	7	2	1	1



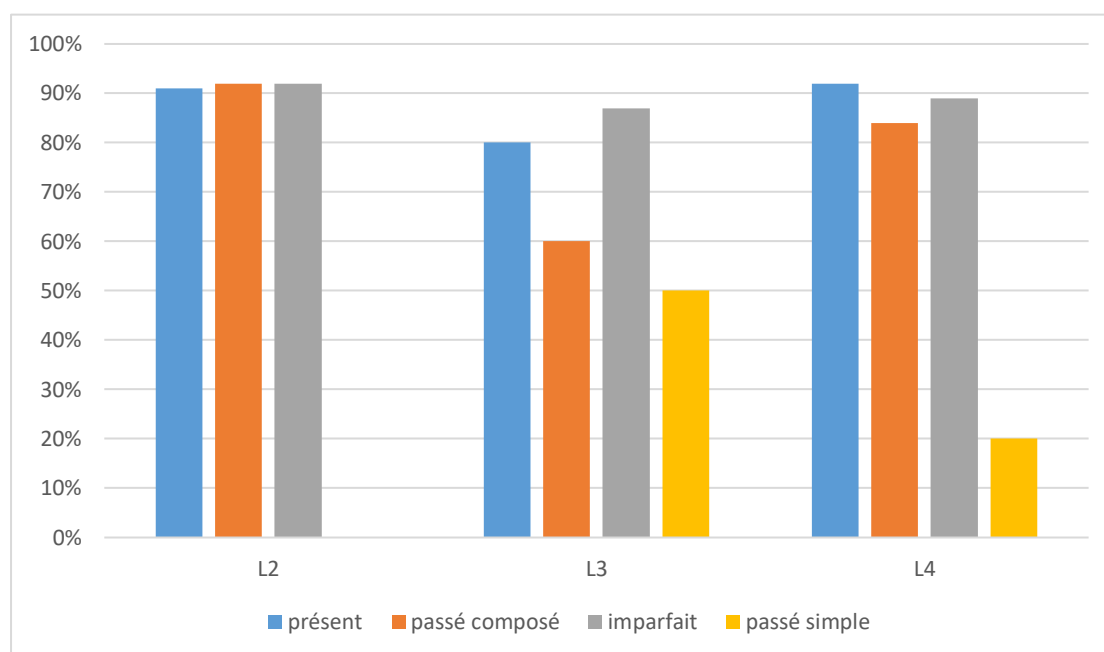
L4	58	10	83%	10	0	0	0
----	----	----	-----	----	---	---	---

Quant aux types d'erreurs commises par les étudiants, nous classons quatre types principaux. Tout d'abord, il s'agit de l'omission du morphème qui sert à marquer le pluriel. Autrement dit, les étudiants produisent la forme verbale de la 3PS alors que le contexte demande la forme du pluriel. Le deuxième type d'erreurs concerne un mauvais choix du radical verbal. Ici, il s'agit notamment des verbes irréguliers. Les formes erronées produites par les étudiants sont bien accordées en nombre avec le morphème –nt, cependant le choix du radical verbal n'est pas correct. Par exemple, *les autres chevaliers \*riraient quand le roi restait muet*. Le troisième type d'erreurs relève d'une sur-généralisation. Les étudiants choisissent un morphème inapproprié pour exprimer le nombre verbal. Par exemple, *elle parla si fort que tous \*rires*. Finalement, la forme hybride cumule deux types d'erreur différents. Par exemple, *ils ne \*serrent guère adroit*.

Nous constatons que pour les scripteurs de L2, la plupart des erreurs portent sur l'omission de la marque indiquant le pluriel : 4 formes erronées sur 6 (67%). Quant aux deux erreurs qui restent, l'une concerne le mauvais choix du radical verbal : *ces couples ne \*riraient pour ses paroles naïf*, où le scripteur a choisi le radical *rir-* pour le temps imparfait ; l'autre erreur relève d'une forme hybride : *ils ne \*serrent guère adroit*, où le scripteur a choisi un mauvais radical verbal *serr-* et un morphème inapproprié *–ent* pour marquer le pluriel. Pour les scripteurs de L3, 7 formes erronées sur 11 commises par eux concernent l'omission de la marque du pluriel. Deux erreurs portent sur le mauvais choix du radical : *les autres chevaliers \*riraient quand le roi restait muet* ; *tous les chevaliers se \*mitent à rire*. Bien que les deux formes verbales soient bien accordées en nombre, les deux scripteurs ont omis de choisir le radical congruent pour les verbes irréguliers *rire* et *mettre*. La seule erreur relevant de la sur-généralisation est trouvée dans la production de ce niveau : *les autres chevaliers \*riraient quand le roi restait muet*. Au lieu d'utiliser le morphème *–ent* pour indiquer le pluriel du verbe *rire* au passé simple, le scripteur a choisi le morphème *-s* qui est employé généralement pour marquer le nombre nominal ou adjectival. L'erreur qui reste concerne une forme

hybride : *les résultats \*surprise le roi*. Au lieu d'employer le verbe *surprendre*, le scripteur a choisi la forme nominale *surprise*, qui est notamment beaucoup plus fréquente que la forme verbale. De plus, le scripteur a omis de respecter l'accord en nombre imposé par le sujet *les résultats*. En ce qui concerne la performance des scripteurs de L3, ils ont commis 10 erreurs au total, et toutes les erreurs concernent l'omission de la marque du pluriel.

Comme il s'agit de l'emploi des temps assez variés dans les textes produits par les étudiants chinois, nous présenterons par la suite le traitement de l'accord du nombre en fonction des temps différents. Au niveau de L2, le pourcentage correct est de plus de 90% pour tous les trois temps principaux, comme ce qui est montré dans le tableau suivant. Néanmoins, en L3, avec l'ajout du temps passé simple, la baisse du pourcentage de réussite concerne tous les temps employés par les scripteurs. Quant à la performance de L4, le pourcentage de réussite pour le temps du présent, du passé composé, et de l'imparfait remonte jusqu'à presque 90%, tandis que le pourcentage correct pour le temps du passé simple n'est que 20%. Il semble pertinent de penser que les formes verbales du passé simple posent un vrai problème aux étudiants chinois.



**Fig. 16 :** Pourcentage de réussite en fonction des temps différents

Pour conclure, les résultats montrés au-dessus nous permet de constater que l'accord verbal en nombre est acquis et maîtrisé assez précocement par les étudiants chinois de notre corpus. Cependant, son développement reste plutôt faible de la L2 à la L4. Une baisse de performance est même observée en L3 avec l'ajout des formes verbales du passé simple, difficulté particulière à laquelle les étudiants ont été confrontée lors de la production écrite. En ce qui concerne le développement de l'accord verbal en nombre, ce que nous avons observé chez les apprenants chinois est différent de la performance des apprenants français, qui est caractérisée par un développement graduel. Plus précisément, chez les enfants francophones, l'accord verbal en nombre apparaissent plus tardivement que l'accord nominal. Comme Fayol (2003) l'indique, d'une manière générale, l'enfant du CP ne marque pas le pluriel ni pour le nom ni pour les verbes, en employant les mots sous leur forme neutre. A partir du CE1, l'enfant commence à utiliser le morphème *-s* pour le nom, et le surgénéraliser au verbe. Finalement, l'enfant sait employer la flexion *-nt* pour le verbe. Le caractère graduel et tardif de l'accord verbal en nombre est également remarqué par Ågren (2005), lorsqu'elle analyse le traitement de l'accord verbal mené par les apprenants suédois de français.

Dans ce domaine, une autre caractéristique qui distingue les apprenants chinois des apprenants francophones concerne les erreurs de substitution. Lors de l'acquisition de l'accord verbal en nombre, les apprenants français sont confrontés souvent à cette difficulté issue de la compétition entre les flexions *-s* et *-nt*. Pour les enfants français, les erreurs de substitution apparaissent tout d'abord puisqu'ils utilisent la flexion *-s* à la fois pour les noms et pour les verbes. Après, ils utilisent *-nt* pour les verbes mais ils « tendent à en surgénéraliser l'emploi à quelques noms, notamment lorsque ces noms ont un homophone verbal » (Fayol, 2003 ; Largy, Fayol & Lemaire, 1996). Ainsi de nouvelles erreurs de substitution apparaissent et on peut trouver également ce type d'erreur chez les adultes (Fayol, Largy & Lemaire, 1994 ; Largy, Fayol & Lemaire, 1996). Lors de la production écrite, comme Fayol (2003) l'indique, il semble que « le pluriel notionnel active des flexions concurrentes *-s* et *-nt*, induisant des interférences ». Et le scripteur fait son choix entre les deux marques en fonction de plusieurs facteurs.

Tout d'abord, il s'agit de l'ambiguïté lexicale. Plus précisément, les noms ayant un homophone verbal (ou l'inverse) sont plus sensibles aux effets d'interférence que les autres. Le deuxième élément qui pourrait jouer un rôle concerne les fréquences relatives des items. Fayol (2003) nous signale que « les mots homophones sont plus souvent infléchis avec la marque du plus fréquent d'entre eux, nom ou verbe ». Cependant, ce type d'erreur est rarement trouvé dans la production des apprenants chinois. Dans notre corpus, nous ne trouvons qu'une seule forme relevant de la sur-généralisation du morphème *-s* : *Elle parla si fort que tous \*rires*, où le scripteur a utilisé le morphème *-s* pour marquer le pluriel du verbe *rire*. Une autre forme erronée issue de la sur-généralisation montre la compétition entre les flexions *-ent* et *-ont* : *ils ne \*serrent guère adroit*, où le scripteur a choisi un morphème inapproprié *-ent* pour marquer le pluriel du verbe être au futur simple.

A part l'apparition précoce de l'accord verbal en nombre, la performance des apprenants chinois présente également certaines caractéristiques particulières en comparant avec celle des autres apprenants du français L2. Ågren (2005) analyse le traitement de l'accord verbal mené par les apprenants suédois de français en fonction des trois types différents des verbes.

Tout d'abord, il s'agit des verbes réguliers, en se terminant en *-er*. Pour les verbes de cette catégorie, la chercheuse remarque que, à part la forme verbale à la 3PS, les apprenants produisent également l'infinitif et la forme du participe passé à la place d'un verbe conjugué à la 3PP. Ce type d'erreur présente 37% au stade initial et persiste encore au stade avancé. Néanmoins, nous n'avons pas trouvé les formes erronées de ce type de notre corpus.

Le deuxième type de verbe traité dans l'étude de Ågren (2005) concerne les verbes irréguliers qui marquent le nombre non seulement par l'ajout de la flexion *-nt* mais aussi par le changement de radical, comme *il doit – il doivent*. Ågren nous signale que, dans les textes des apprenants suédois, les verbes de ce type apparaissent parfois sous une forme non finie, accordée avec le morphème *-nt* mais sans changement de radical. Chez les apprenants suédois, les erreurs de ce type sont assez fréquentes, notamment

pour les débutants (55% au stade 2 ; 38% au stade 3 et 17% au stade 4<sup>46</sup>). Il semble que les apprenants ajoutent le morphème –ent à une racine erronée construite soit sur la forme de la 3PS (*ils \*faitent*), soit sur la forme de l’infinitif (*ils \*rarent*). Selon la chercheuse, ce type d’erreur est issue de la fréquence moins importante des verbes irréguliers avec un changement de radical. En revanche, pour les francophones, il est probable que les formes verbales irrégulières à la 3PP sont gardées individuellement dans leur mémoire. Avec l’aide de l’oral, les erreurs sous la forme infinie ne sont pas trouvées chez les francophones. De plus, l’indice venant de l’oral (*il doit/ils doivent*) aide aux scripteurs francophones à mieux maîtriser l’accord en nombre silencieux pour ce groupe de verbes. Quant aux apprenants chinois, la tendance de produire une forme infinie est également trouvée dans notre corpus. Nous avons trouvé au total 4 formes erronées à la L2 et L3. Mais en termes de proportion, ce type d’erreur est beaucoup moins important dans notre corpus.

La dernière catégorie verbale concerne les verbes irréguliers fréquent *sont, ont, vont et font*, pour lesquels l’opposition entre singulier et pluriel est audible. Selon la recherche de Ågren (2005), les verbes *sont/ont/vont/font* sont très fréquents, voire dominants dans les textes écrits par les apprenants suédois. De plus, ce groupe de verbes est employé plus tôt que les verbes réguliers et de façon plus correcte. Notons également qu’un autre genre de suremploi est signalé par la chercheuse, qui pourrait être issu du modèle de ces verbes irréguliers mais extrêmement fréquents, tel qu’*ils \*prendons*. Les formes verbales *sont/ont/vont/font* sont également très fréquents dans notre corpus. Elles sont employées soit comme verbe au temps présent, soit comme auxiliaire du temps composé. Mais pour ce groupe de verbes, toutes les erreurs relèvent de l’omission de la marque du pluriel.

---

<sup>46</sup> Comme les verbes irréguliers avec un changement de radical n’apparaissent que comme cas isolés au stade 1, aucune forme infinie n’est trouvée dans ce niveau.

#### 4.2.2.6.2. Discussion

L'analyse comparative là-dessus met en évidence le fait que, chez les apprenants chinois, l'accord verbal en nombre est employé très tôt et de façon assez correcte dès le niveau initial. Cependant, le profil développemental n'est pas très prononcé du stade initial au stade avancé. En observant les germanophones du français L2 au milieu scolaire, Granget (2005) note également l'apparition très tôt de l'usage abondant des marques de l'accord pluriel de ces apprenants. Selon elle, plusieurs facteurs parlent en faveur de cet accord précoce. Premièrement, ce phénomène est sans doute issu de « l'environnement de l'apprentissage et la focalisation massive sur la morphologie verbale écrite en salle de classe » (Ågren, 2005). Deuxièmement, elle discute le rôle possible de la langue maternelle des apprenants, c'est-à-dire l'allemand, qui fait, morphologiquement, la distinction entre la 3PS et la 3PP pour le verbe.

En ce qui concerne les apprenants chinois de notre corpus, si l'on prend en compte le rôle de l'enseignement pour eux et le rôle de l'écrit dans l'enseignement, il n'est pas étonnant de constater l'usage précoce de l'accord verbal en nombre. Plus précisément, les flexions verbales et leur fonctionnement font l'objet d'un enseignement explicite. Dès le premier cours les conjugaisons verbales sont introduites sous la forme de tableau, comme l'illustre le tableau ci-dessous. Les verbes sont classés aux trois groupes. Par exemple, pour les verbes du 1<sup>er</sup> groupe, la règle de conjugaison est formulée ainsi : Ajouter au radical les terminaisons *-e, -es, -e, -ons, -ez et -ent*. Et le radical est la partie qui reste lorsqu'on enlève *-er*. Par la suite de leur apprentissage, les tableaux de ce genre, notamment pour les verbes du 3<sup>ème</sup> groupe, sont presque omniprésents et les apprenants ont à passer beaucoup de temps à les mémoriser et à réaliser de nombreux exercices là-dessus. Pour les apprenants français, le morphème *-s* sert à marquer le pluriel du nom et de l'adjectif, et le morphème *-nt* pour le pluriel du verbe. Cependant, pour les apprenants chinois de notre corpus, une telle formulation de la règle est rarement mentionnée, et la forme verbale de la 3PP est toujours introduite sous la forme de tableau des conjugaisons. De ce fait, il n'est pas surprenant de constater que les erreurs de substitution (usage du morphème *-s* pour marquer la 3PP du verbe) sont très

rare chez les apprenants chinois, car il semble que, chez eux, le lien analogique entre le pluriel du nom et celui du verbe n'est pas établi.

Être	
Je suis	Nous sommes
Tu es	Vous êtes
Il/elle est	Ils/elles sont

Un autre facteur qui pourrait faciliter l'usage de l'accord verbal du nombre concerne les conditions d'acquisition des apprenants allophones. Les enfants français apprennent à écrire en partant « d'un input oral massif et d'un système linguistique oral bien en place » (Ågren, 2005). Néanmoins, les apprenants chinois, comme nous l'avons indiqué plus haut, apprennent à parler et à écrire simultanément. De ce fait, ils sont moins affectés par le côté silencieux du français écrit que les enfants français et plus marqués par la représentation visuelle des verbes (Ågren, 2005).

Signalons également qu'avec l'ajout du temps de passé simple, une baisse de performance est observée à la L2. Et pour tous les niveaux acquisitionnels, la plupart des formes erronées concernent l'omission de la marque du pluriel. Selon nous, il semble que les apprenants disposent de la connaissance (accorder le verbe selon le nombre du sujet), mais qu'ils ne la mettent pas en œuvre. L'hypothèse la plus probable susceptible d'expliquer ce fait est que, lors de la production écrite, les apprenants ne disposent pas d'une capacité attentionnelle suffisante pour faire une analyse syntaxique permettant d'identifier le nombre du sujet. L'essentiel de leur attention est probablement capté par, d'une part, la détermination du contenu à produire et, d'autre part, la transcription graphique. Notons que la réalisation graphique est encore très coûteuse en attention chez les apprenants chinois, qui ont commis plus d'erreurs en produisant les formes verbales aux temps du passé simple et du passé composé, qu'en produisant celles du présent et de l'imparfait. Ainsi il est pertinent de penser qu'une analyse syntaxique suffisante qui sert à déterminer l'accord verbal en nombre ne serait pas bien faite si le scripteur est surchargé mentalement lors de la production écrite.

#### 4.2.2.7. Conclusion : Acquisition de l'accord en nombre des trois catégories des mots

La figure 17 montre le profit développemental de l'accord en nombre des trois catégories des mots : nom, verbe et adjectif. Comme nous l'avons mentionné auparavant, pour tous les trois catégories, l'usage de l'accord en nombre apparaît dès le niveau initial, et le profit développemental reste assez faible de la L2 à la L4. Contrairement aux apprenants français ainsi qu'aux apprenants suédois de l'étude d'Ågren (2005), les apprenants chinois de notre corpus manifestent une tendance particulière en ce qui concerne l'acquisition de la morphologie écrite du nombre.

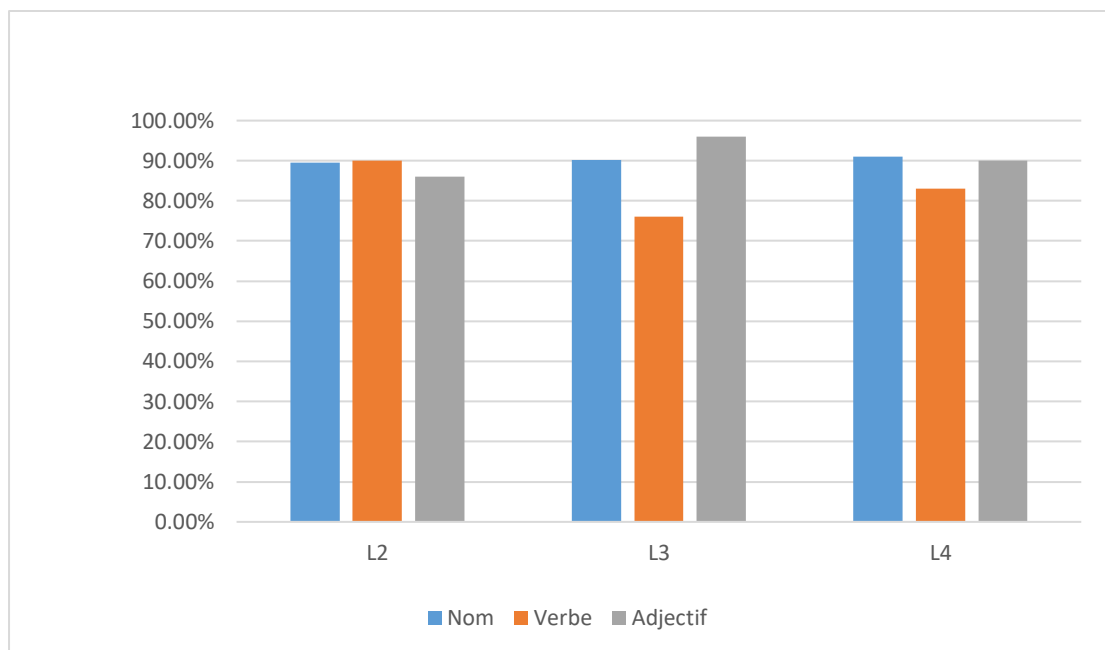


Fig. 17 : Pourcentage de réussite de l'accord du nombre

Il nous semble que le modèle de la représentation lexicale proposé par Jiang (2000) nous permet de voir d'une façon plus claire ce traitement de la morphologie écrite du nombre. Jiang envisage le traitement morphologique en langue maternelle dans un cadre de processus automatique. Avant la réalisation morphologique, les informations morphologiques sont déjà codées dans le message préverbal et la sélection de la flexion est sous le contrôle direct des éléments du message. De ce fait, même si les informations



morphologiques ne sont généralement pas considérées comme partie intégrale du lemme, les informations sémantiques contribuent elles aussi à la réalisation morphologique. Il s'agit donc d'un processus automatique de mise en correspondance entre le contenu du message à produire et les informations spécifiques concernant la morphologie associée à un item lexical.

Il est également proposé que le choix morphologique fait partie de l'encodage grammatical et que la flexion forme une partie intrinsèque du système syntaxique. En d'autres termes, lorsqu'une structure syntaxique est donnée, les informations morphologiques sont activées.

Notons également que les associations possibles entre un item lexical et les affixes relevant de sa morphologie flexionnelle sont présentées dans une même entrée (Levelt, 1989). Une fois cette entrée lexicale est ouverte, toutes ces formes lexicales deviennent automatiquement disponibles pour le scripteur.

Selon Jiang (2000), à partir de ces observations, nous pourrions envisager un processus en une seule étape de la production morphologique en langue maternelle. Si nous voulons produire le message *They left*, deux entrées lexicales sont ouvertes, une d'entre elles est celle de *leave*, avec toutes formes flexionnelles acceptables comme *leave, leaves, left, leaving*. Parmi toutes ces formes, *left* est la forme la plus intensément activée, et par la suite choisie puisqu'elle se conforme le mieux au message préverbal, l'action de quitter qui se produit dans le passé. Il s'agit d'un processus en une seule étape puisque la forme flexionnelle est choisie directement sans une médiation de la racine *leave* ou une analyse linguistique consciente.

Il n'est pas superflu d'indiquer qu'un tel processus automatique nécessite une forte intégration des informations morphologiques dans l'entrée lexicale. Autrement dit, les différentes flexions d'un item lexical doivent s'intégrer dans son entrée lexicale et ainsi deviennent automatiquement disponibles lors de la production. Ce processus automatique du traitement morphologique fait une partie importante de la compétence lexicale. Cependant, dans l'apprentissage de langue étrangère, au moins pour les scripteurs chinois dans notre corpus, qui apprennent la langue française en s'appuyant principalement sur l'enseignement en classe, les informations morphologiques ne sont

pas intégrées dans les entrées lexicales, dont nous avons parlé dans la partie précédente, mais sont apprises et stockées hors du système de l'entrée lexicale, en tant que les connaissances explicites. Cette différence concernant la représentation lexicale aurait des conséquences significatives sur le processus du traitement morphologique. Selon Jiang (2000), la sélection de la flexion pertinente devient ainsi un processus conscient et strictement séquentiel, en impliquant deux étapes nécessaires : sélection de la racine liée à l'item lexical à produire, et sélection de la flexion sur la base des connaissances morphologiques du scripteur.

Si nous reprenons l'exemple du mot *leave*, ses variations morphologiques comme *leave*, *leaves*, *left*, *leaving* seraient conservées respectivement dans une entrée séparée. Toutefois, il est concevable qu'il existe des liens entre ces entrées, qui auraient été établis pendant le processus d'acquisition de ces mots. Pour produire *They left*, l'entrée de l'item *leave* serait ouverte, cependant elle ne comporte que la racine, *leave*. La deuxième étape implique donc une sélection d'un affixe, pertinent du point de vue morphologique, selon les connaissances de l'apprenant.

Une telle production linguistique strictement séquentielle est susceptible d'être un processus conscient, qui nécessite des ressources mentales. Lorsque ce travail de sélection entre en conflit avec d'autres activités qui consomment aussi de l'énergie mentale, des erreurs morphologiques seraient produites.

Jiang ajoute également que l'enseignement de la langue étrangère se concentre particulièrement sur la racine lorsqu'un nouveau mot est introduit. Ce qui est représenté dans l'entrée lexicale, c'est donc souvent la racine, liée à des variations lexicales relevant de sa morphologie flexionnelle. Ces liens seraient probablement directionnels, de la racine à ses flexions, plutôt que l'inverse. Comme nous en avons déjà parlé, le choix de la forme pertinente du point de vue morphologique pourrait commencer par la sélection de la racine, et la flexion serait par la suite choisie selon les connaissances morphologiques du scripteur. Cela signifie que des erreurs morphologiques auraient plus de chances d'être produites lorsque le contexte linguistique demande une flexion. C'est exactement ce que nous avons observé dans la production des étudiants chinois : la plupart des erreurs, concernant la morphologie écrite du nombre, porte sur l'omission

de la flexion du pluriel. Ces formes erronées pourraient s'expliquer par le fait que les apprenants disposent de ressources mentales limitées pour que leurs connaissances morphologiques contribuent à la réalisation morphologique pertinente.

D'après nos observations, les performances des étudiants chinois pourraient s'expliquer, nous semble-t-il, par le fait qu'à la suite d'un enseignement explicite ainsi qu'une grande quantité d'exercices, les étudiants chinois sont, majoritairement, bien conscients de la morphologie écrite du nombre impliquée par la langue française et maîtrisent assez bien cette production linguistique strictement séquentielle. De ce fait, pour toutes les trois catégories, l'usage de l'accord en nombre apparaît dès le niveau initial. Néanmoins, une telle production strictement séquentielle demande toujours une analyse linguistique consciente. Et le coût de la réalisation morphologique est très élevé chez les scripteurs chinois, et entre en conflit avec d'autres activités qui mobilisent également de l'énergie mentale. Et le fait que le profit développemental reste assez faible de la L2 à la L4, en ce qui concerne l'usage de l'accord en nombre, nous renseigne sur les informations morphologiques stockées toujours au dehors de l'entrée lexicale pour ces trois groupes d'étudiants chinois.

Les bonnes connaissances des règles morphologiques des apprenants en langue étrangère sont aussi indiquées par Jiang (2000). Selon lui, les apprenants sont souvent conscients des erreurs qu'ils ont produites lors de la révision et sont capables d'énoncer les règles en ce qui concerne les flexions à produire. Jiang souligne que c'est un argument fort pour montrer la distinction entre les informations lexicales intégrées dans les entrées lexicales et celles qui existent hors de ce système. Ainsi que nous l'avons déjà dit, les informations lexicales qui ne sont pas obtenues par l'extraction du contexte existent au dehors des entrées lexicales, et contribuent donc au processus conscient, au lieu de processus automatique.

#### 4.2.3. L'accord du genre

Comrie (1999) constate que pour les langues, le genre est assigné en fonction de deux grands principes : les principes sémantiques et les principes formels. En anglais, par

exemple, des noms sont attribués à un genre selon leur sens, c'est-à-dire suivant un principe sémantique. Et les principes formels jouent un rôle beaucoup plus important dans les langues romanes, telles que l'espagnol, le français. Dans la plupart des noms français, le genre n'est pas sémantiquement motivé. Il relève au contraire de l'arbitraire du signe. Et en français, la combinaison des principes sémantique et formel rend le système d'attribution du genre opaque. Ainsi, il n'est pas surprenant que ce problème d'attribution du genre reste l'une des grandes difficultés qui gênent les apprenants de français.

A notre connaissance, alors même que de nombreuses recherches sur l'utilisation et l'acquisition de la morphologie du genre ont été effectuées en français L1 et en français L2, il existe peu de données en ce qui concerne les apprenants en français L2 et leur traitement de la morphologie du genre à l'écrit. Si les chercheurs essaient de déterminer à partir de quel âge les enfants français font systématiquement l'accord du genre à l'écrit et si le traitement de cet accord est régi par des contraintes linguistiques et cognitives, les questions similaires doivent être répondues pour la production écrite du français L2. Cette partie de notre travail a ainsi pour objectif de faire une analyse des données sur l'utilisation de la morphologie du genre afin d'apporter des éléments de réponse à ces questions.

#### **4.2.3.1. Le genre en français**

Le genre du français trouve son fondement dans le nom. Le nom est toujours pourvu d'un genre inhérent, masculin ou féminin. Et cette caractéristique lui est « fixée d'avance par le code de la langue et répertoriée dans le dictionnaire » (Weinrich, 1989 : 36). Au niveau de l'énoncé, des éléments d'autres classes grammaticales, tels que le déterminant, l'adjectif, s'accordent au nom auquel ils se réfèrent. Ainsi se construit l'accord en genre.

Étant donné que les noms sont habituellement précédés de leur déterminant en français, c'est le déterminant qui donne lieu le mieux à une discrimination perceptive pour le genre du nom. Comme on peut le voir dans les tableaux 36 et 37, non seulement l'article

simple mais aussi l'article spécifique ont des formes dont le genre est marqué. Néanmoins, l'opposition du genre est neutralisée au pluriel dans presque toutes les formes du déterminant. Autrement dit, on ne trouve que la marque du genre dans les formes du singulier.

**Tableau 36.** Articles simples

	défini		indéfini		partitif	
	singulier	pluriel	singulier	pluriel	singulier	pluriel
masculin	le	les	un	des	du	des
féminin	la	les	une	des	de la	des

**Tableau 37.** Articles spécifiques

	singulier	pluriel
masculin	ce	ces
	son	ses
féminin	cette	ces
	sa	ses

A la différence des noms, le genre des adjectifs n'est pas préétabli par le code de la langue, mais à sélectionner en fonction du contexte. Plus précisément, l'adjectifs porte la même marque du genre que le nom avec lequel il entretient un rapport de détermination. Entre le nom et l'adjectif, il y a deux possibilités principales d'accord en genre, à savoir la qualification et la prédication. Alors que le déterminant précède toujours le nom à déterminer, l'adjectif qualificatif peut être placé avant (antéposition) ou après le nom (postposition)<sup>47</sup>. Et le genre du nom détermine la marque du genre de l'adjectif. Une autre possibilité d'accord en genre entre le nom et l'adjectif est la prédication. Par l'intermédiaire du verbe prädicatif, on attribue à un sujet un adjectif

---

<sup>47</sup> Notons qu'un nombre relativement petit d'adjectifs peut être antéposé, mais les adjectifs les plus fréquemment utilisés sont parmi eux (Blanche-Benveniste, 1990 : 195).

comme prédicant. Ainsi, l'accord en genre de l'adjectif se fait avec le sujet<sup>48</sup>.

Il n'est pas superflu de rappeler qu'à l'oral, la marque du genre n'est audible que pour un tiers des adjectifs français. Dans la graphie, cependant, l'opposition du genre est conservée pour la moitié des adjectifs (Riegel, Pella & Rioul, 1994 : 359). Riegel et *al.* distinguent dans la marque du genre trois catégories d'adjectifs :

- (i) Adjectifs dont le genre n'est pas marqué
- (ii) Adjectifs dont le genre n'est marqué que à la graphie
- (iii) Adjectifs dont le genre est marqué à l'oral et à la graphie

Selon Blanche-Benveniste (1990), la relation forme-fonction pour le genre des adjectifs est relativement transparente : la forme du masculin est, pour la plupart des adjectifs, plus court d'un phonème que la correspondante du féminin. De ce fait, il est « conseillé, dans l'apprentissage et la mémorisation des formes des adjectifs, de partir du féminin pour obtenir le masculin par troncation » (Weinrich, 1989 : 41).

#### **4.2.3.2. Le traitement et l'acquisition de la morphologie du genre en français L1**

Chez des enfants francophones, l'utilisation des marques du genre est relativement « précoce, automatique et irrépessible à l'oral » (Bourdin, Leuwers & Bourbon, 2011). Roulet (2007) a marqué l'utilisation de déterminants variables dont le genre est marqué dès l'âge de deux ans. Les études de Valois & *al.* (2009) et de Trudeau & *al.* (2008) révèlent l'émergence d'adjectifs variables au féminin entre les âges de deux et trois ans. Cependant, les recherches sur l'accord en genre de l'adjectif effectué dans la production écrite nous montrent que les élèves français commettent encore des erreurs d'accord en fin de secondaire (voir, par exemple, Bourdin & *al.*, 2004 ; Fayol, 2003 ; Mariol & *al.*,

---

<sup>48</sup> A l'aide du verbe *rendre*, les prédicants-adjectifs peuvent être imputés à l'objet. Par exemple : *Qui a rendu cette fleur si belle ?* (Weinrich, 1989 : 40).

2004). Et dans ces recherches, les erreurs les plus fréquentes concernent l'omission de la marque du genre *-e*. Totereau et *al.* (1997) indiquent que la morphologie écrite du genre est acquise en deuxième année primaire. En étudiant l'accord en genre de l'adjectif dans les tâches de complètement et de dictée, Fayol (2003) montre que les élèves, dès le CE1, sont en mesure de repérer la marque du genre dans une tâche de sélection, mais échouent à la mettre en œuvre dans une tâche de production. Ce tâtonnement des élèves est également observé par Bourdin et *al.* (2011). En proposant une tâche de production écrite de phrases aux enfants de l'école primaire, les chercheurs confirment que la morphologie écrite du genre est loin d'être totalement maîtrisée, même au CM2.

Dans cette recherche, Bourdin & *al.* (2011) ont également étudié l'impact des contraintes linguistiques et cognitives sur la mise en œuvre à l'écrit de l'accord en genre de l'adjectif.

En ce qui concerne les contraintes linguistiques, comme nous l'avons indiqué auparavant, pour de nombreux adjectifs français, la marque du genre est inaudible, ce qui rend plus difficile la réalisation écrite de cet accord. Les données de Bourdin et *al.* (2011) sont conformes à cette hypothèse, en montrant que les élèves obtiennent de meilleures performances avec les adjectifs dont la marque du féminin est audible.

Selon les chercheurs, l'utilisation de l'accord en genre de l'adjectif dépend également de contraintes cognitives. Les données montrent que les erreurs commises par les élèves relèvent essentiellement de l'omission de la marque du genre. Selon Fayol (2003), ce type d'erreur pourrait être interprété dans le cadre d'un apprentissage procédural (Anderson, 1983). Les élèves peuvent savoir que l'adjectif s'accorde en genre avec le nom à déterminer (savoirs déclaratifs), mais ils ne peuvent pas mettre en œuvre cet accord lorsqu'ils rencontrent un adjectif au féminin (savoir procéduraux) parce que les élèves de ce niveau ne disposent pas, lors d'une réalisation graphique, d'une capacité attentionnelle suffisante pour gérer l'application de l'accord du genre. En suivant cette conception, Fayol indique aussi que ces erreurs d'omission devraient disparaître chez les élèves avec l'automatisation de l'accord du genre. Cette hypothèse est confirmée par les données de Bourdin & *al.* (2011). Les chercheurs constatent une

diminution des erreurs d'omission, d'une part, chez les élèves plus âgés et d'autre part, dans les tâches de production moins coûteuses en ressources attentionnelles.

#### **4.2.3.3. Le traitement et l'acquisition de la morphologie du genre en français L2**

Bien que toutes les recherches mentionnées auparavant se rapportent uniquement à la production des apprenants français, elles nous offrent les domaines de recherche intéressants concernant la production des apprenants allophones.

L'utilisation et le développement de la morphologie écrite du genre seront examinés dans cette partie de notre travail. Comme notre corpus couvre trois groupes d'apprenants de niveaux linguistiques différents, nous concentrons nos analyses sur les variétés diachroniques et synchroniques produites par les apprenants sur la distribution et/ou l'accord du genre. Plus spécifiquement, nous aborderons les questions suivantes :

- Comment les apprenants chinois de trois niveaux différents marquent le genre pour les déterminants et les adjectifs en français écrit ?
- Quel domaine est le plus difficile, déterminants ou adjectifs ? Quelles contraintes linguistiques et/ou syntaxiques sont liées aux taux de réussite de l'accord du genre ?
- Quels sont les scénarios psycholinguistiques possibles pour interpréter les erreurs concernant la morphologie du genre ?

#### **4.2.3.4. Résultats**

Nous présenterons dans cette partie la performance des apprenants chinois de trois niveaux linguistiques différents. Leur réalisation morphologique du genre sera examinée en fonction de deux catégories grammaticales, à savoir le déterminant et l'adjectif.



#### 4.2.3.4.1. Le marquage du genre pour les déterminants

Le tableau ci-dessous montre les résultats de l'enquête sur le traitement de l'attribution ou l'accord du genre pour les déterminants définis, indéfinis, et possessifs, effectué respectivement par trois groupes d'étudiants<sup>49</sup>.

**Tableau 38.** Taux de réussite de l'accord du genre pour les déterminants

(Légende : DD = déterminant défini ; DIn = déterminant indéfini ; DPo = déterminant possessif)

	DD	DIn	DPo	Total
Licence 2	398/415 = 95.90%	271/284 = 95.42%	154/159 = 96.85%	823/858 =95.92%
Licence 3	274/294 = 93.19%	144/155 = 92.90%	82/85 = 96.47%	500/534 =93.63%
Licence 4	399/432 = 92.36%	205/220 = 93.18%	143/146 = 97.94%	747/798 =93.60%
Total	1073/1143 = 93.87%	618/657 = 94.06%	379/390 = 97.17%	2070/2190 =94.52%

Si les données sont examinées d'une manière développementale, nous constatons que le marquage du genre dans les déterminants est précocement maîtrisé par les apprenants chinois, puisque pour la plupart du temps, les apprenants sont capables de choisir le genre congruent pour les déterminants. Déjà au stade initial, où les étudiants ont suivi une année d'études du français et représentent un niveau linguistique élémentaire, le pourcentage d'accord correct est 95.92%. Le résultat le plus frappant concerne les étudiants de la Licence 3 et de la Licence 4, dont le niveau de langue est plus avancé. Chez ces deux groupes d'apprenants, le pourcentage correct du marquage du genre diminue, d'une manière assez faible, jusqu'à respectivement 93.63% et 93.60%. Ainsi, il nous semble difficile de déterminer le profil développemental des

<sup>49</sup> Lors de l'inventaire, les formes élidées *l'*, ainsi que les formes contractées *au*, *du*, ont été exclues.

apprenants chinois, dont la performance concernant l'attribution ou l'accord du genre pour les déterminant est caractérisée par une acquisition précoce et une évolution très restreinte, même un recul. Cependant, il faut noter que pour le taux de réussite, il n'existe pas de différence statistiquement significative entre Licence 2 et Licence 3, ou entre Licence 2 et Licence 4. La diminution légère du pourcentage de réussite s'explique probablement par les constructions nominales plus compliquées et le choix lexical plus riche pour les noms faits par les apprenants ayant un niveau linguistique plus avancé. De ce fait, il nous semble pertinent de dire que le profil développemental reste très faible entre le niveau initial et le niveau assez avancé.

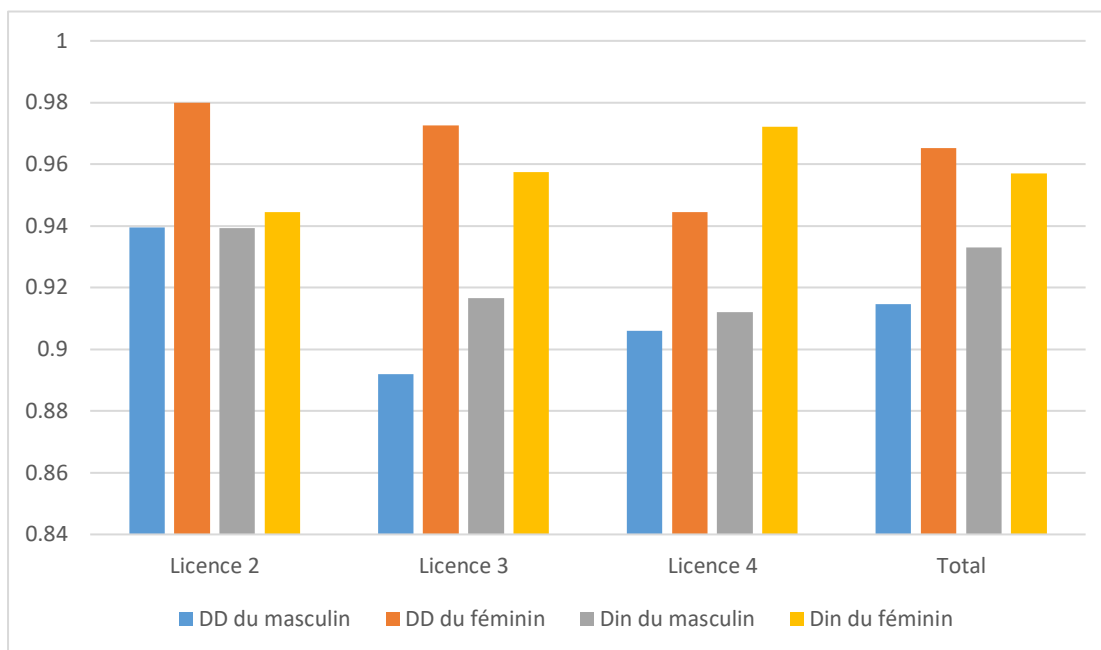
Si nous examinons l'acquisition de la morphologie du genre pour les trois catégories de déterminants, les résultats sont conformes à notre observation globale. Pour chaque catégorie de déterminants, les étudiants au stade initial démontrent déjà une bonne maîtrise concernant le marquage du genre. Cependant, une diminution légère est observée pour toutes les trois catégories de déterminants à partir de la Licence 3. Nous constatons ainsi que les apprenants commettent encore les erreurs concernant l'attribution ou l'accord du genre, et que le nombre de ces erreurs ne diminue pas significativement en fonction de la durée de l'apprentissage du français.

En ce qui concerne les variétés synchroniques, les déterminants indéfinis ne posent pas de difficulté particulière aux apprenants chinois, par rapport aux autres catégories de déterminants. Cela ne correspond pas aux résultats de Bartning (2000) et de Chini (1995). Les deux études travaillent respectivement sur l'acquisition de la morphologie du genre avec la production orale en français L2 et en italien L2. Les données de Bartning (2000) montrent que l'accord de déterminants indéfinis est le plus difficile à acquérir pour les apprenants suédois, notamment pour les apprenants de niveau avancé. Chini (1995) présente l'ordre d'acquisition suivant pour l'accord en genre en italien L2 : pronoms personnels > déterminants définis > déterminants indéfinis > adjectifs attributifs > adjectifs prédicatifs > participes passés. Cet ordre d'acquisition n'est pas observé dans notre corpus et pour les apprenants de trois niveaux différents, les déterminants indéfinis sont aussi bien accordés que les déterminants définis.

Il faut également noter que les déterminants possessifs, bien que d'une manière

légère, sont les mieux maîtrisés parmi les trois catégories de déterminants. Cela s'explique probablement par le fait que, dans notre corpus, les déterminants possessifs introduisent pour la plupart du temps les noms tels que *fil*s, *mère*, *père*, *frère*, dont le genre est sémantiquement motivé. Cette caractéristique des noms plaide vraisemblablement en faveur de l'accord de leur déterminant. Ce résultat est conforme à ceux de Bartning (2000). En analysant le marquage du genre utilisé par les apprenants suédois (en français L2) de deux niveaux différents (apprenants avancés et pré-avancés), le chercheur constate que les déterminants possessifs sont mieux maîtrisés que les déterminants définis et indéfinis (97% VS 93% / 83% pour les apprenants avancés ; 80% VS 74% / 72% pour les apprenants pré-avancés).

Le phénomène de généralisation des formes masculines existe dans notre corpus, bien qu'avec les proportions légèrement différentes pour les trois catégories de déterminants. Notons que 51 sur 70 déterminants définis avec une marque du genre incorrecte sont les formes masculines (72.85%), et que 30 sur 39 déterminants indéfinis qui ne s'accordent pas correctement avec le nom sont les formes masculines (76.92%). Quant aux déterminants possessifs, cette proportion est tombée à 63.63%. La figure 18 permet de visualiser les performances des apprenants de niveaux différents. Nous constatons que, pour les apprenants de trois niveaux, la forme *le* est la forme par défaut utilisée dans le contexte qui demande un déterminant défini. Toutefois, la forme masculine du déterminant indéfini n'est pas sur-généralisée qu'à partir de la Licence 3. Autrement dit, les apprenants de la Licence 2 n'ont pas encore commencé à utiliser la stratégie de sur-généralisation du genre masculin pour les déterminants indéfinis.



**Fig. 18** : Réalisation de la morphologie du genre selon les différents types de déterminant

#### 4.2.3.4.2. Le marquage du genre pour les adjectifs

Cette partie porte sur l'acquisition de l'accord adjectival dans les quatre positions de l'adjectif français : postposition (*il portait une armure vermeille*) ; antéposition (*il avait posé la coupe sur une grosse pierre*) ; position prédicative (*ma douleur est grande*) ; apposition (*triste et désespérée, sa mère n'a pas eu...*).

Tout d'abord, si l'on examine les données d'une manière synchronique, pour chaque groupe, les différences de taux de réussite entre les positions attributives et les positions prédicatives ne sont pas statistiquement significatives. Cette observation est contraire aux résultats de Bartning (2000), selon lesquels l'accord adjectival en position attributive se développe et est maîtrisé plus tard que celui en position prédicative. Notons que les adjectifs en apposition sont beaucoup moins fréquents dans la production des apprenants. Bien que l'échantillon soit très petit, les résultats du tableau 39 démontrent que l'accord adjectival en apposition est acquis plus tard que l'accord des autres positions. Plus précisément, pour les apprenants de la Licence 2 et de la Licence 3, l'accord adjectival en appositions pose plus de difficultés que celui des autres positions.

**Tableau 39.** Taux de réussite de l'accord du genre pour les adjectifs

(Légende : *Postposé = adjectif postposé ; antéposé = adjectif antéposé ; prédicatif = adjectif  
prédicatif*)

	Postposé	Antéposé	Prédicatif	Apposition	Total
Licence 2	136/138 = 98.55%	163/164 = 99.39%	116/116 = 100%	10/11 =90.91%	425/429 =99.06%
Licence 3	77/80 = 96.25%	105/110 = 95.45%	73/74 = 98.64%	11/13 =84.61%	266/277 =96.02%
Licence 4	104/108 = 96.29%	114/119 = 95.79%	85/92 = 92.39%	11/11 =100%	314/330 =95.15%
Total	317/326 = 97.23%	382/393 = 97.2%	274/282 = 97.16%	32/35 =91.42%	1005/1036 =97%

Les données montrent également qu'il existe peu de développement sur l'axe diachronique. Nous constatons que le taux de réussite diminue du niveau initial au niveau avancé, en ce qui concerne l'accord adjectival en position attributive et prédicative. Les adjectifs en apposition constituent une exception si l'on examine le taux de réussite de leur accord de la Licence 2 à la Licence 4. Toutefois, le profil développemental n'est pas mis en évidence, et le taux de réussite de l'accord adjectival baisse en fonction du niveau, bien que d'une façon non significative.

#### **4.2.3.5. Variation intra-individuelle : ontogénèse des erreurs de genre**

Selon les chercheurs Dewaele et Véronique (2001), les modificateurs portant les marques du genre incorrectes peuvent donner lieu à deux interprétations complètement différentes. Tout d'abord, la tête nominale pourrait être assignée au genre erroné, et les modificateurs sont par conséquent accordés erronément. Une autre possibilité concerne les erreurs d'accord de genre. Plus précisément, la tête nominale est assignée au genre correct, mais cette information du genre n'est pas fournie aux modificateurs. Ainsi est produite une erreur d'accord.

A partir de ces deux principes, nous discuterons dans cette partie des facteurs divers en jeu dans l'attribution et l'accord du genre observés dans notre corpus. Les données qualitatives et quantitatives seront présentées afin d'illustrer les interprétations psycholinguistiques pour comprendre les erreurs de genre portées. Dewaele et Véronique (2001) supposent que les erreurs de genre pourraient résulter d'une variété de causes venant de différents stades du processus de production langagière. Notons que les scénarios suivants offrent les hypothèses pour comprendre ce qui se passe dans le processus de la production langagière et les éléments entraînant les erreurs de genre spécifiques. Cependant ils sont très spéculatifs, et il existe toujours d'autres explications possibles.

#### 4.2.3.5.1. Erreurs concernant l'attribution du genre

Dans cette catégorie, la tête nominale, comme nous avons discuté auparavant, est assignée au genre incorrect et ses modificateurs portent ainsi les marques du genre erronées. Nous constatons que, dans de nombreux cas, les apprenants semblent avoir ignoré les indices du genre : soit ces indices sont présentés dans les terminaisons morpho-phonologiques, soit ils sont motivés par les informations sémantiques, notamment le sexe biologique de noms. Les deux scripteurs ci-dessous nous offrent les illustrations frappantes de ces deux cas. Tout d'abord, dans l'exemple (68a), le mot *chevalerie* est un nom féminin, comme la plupart des noms avec la terminaison *-ie*<sup>50</sup>. Deuxièmement, contrairement à la plupart des noms dont le genre relève de l'arbitraire du signe, le genre d'un certain nombre de noms est sémantiquement motivé. Autrement dit, on peut déduire le genre de ces noms à partir de leur sens. Il s'agit particulièrement des êtres vivants que différencie le sexe. C'est le cas du nom *mère*, dont le genre est mal prédit par le scripteur dans l'exemple (68b).

- (68) a. \* Sa mère ne l'avait jamais dit {le}\_ {la} chevalerie. (L3-020)  
b. \* {Son}\_ {Sa} mère s'est inquiétée de son départ. (L2-02-011)

---

<sup>50</sup> Le suffixe *-ie* sert à former des noms désignant un état, une qualité ou une condition.

L'utilisation constante de la marque du genre erronée, comme les déterminants attachés au nom *cour* dans l'exemple (69), produites par le même scripteur :

- (69) a. \* Avant qu'il allait {au}\_ {à\_la} cour du roi Arthur, il n'avait [pas] entendu parler aucune chose à propose de la chevalerie. (L3-003)
- b. \* Il décida de suivre des chevaliers [d] {au}\_ {à\_la} cour du roi pour lui demander des {amées}\_ {armes}. (L3-003)
- c. \* Dans {le}\_ {la} cour du roi, il y eut des chevaliers, Perceval dit {son}\_ {sa} {démandation}\_ {demande} au roi. (L3-003)

Dans l'exemple (70), les scripteurs, en ignorant le genre des noms, ont omis de choisir le genre congruent pour les déterminant et les adjectifs qualificatifs :

- (70) a. \* Et il a rencontré un homme qui porte {un}\_ {une} arme très {fort}\_ {forte}. (L1-01-001)
- b. \* il est revenu à la cour de roi avec {une}\_ {un} {grande}\_ {grand} succès. (L1-01-014)

L'information diacritique incorrecte du genre est stockée à l'entrée des noms concernés entraîne l'activation et la sélection du nœud de genre erroné. La règle de l'accord a été appliquée, ce qu'on voit notamment dans l'exemple, mais portait les informations incorrectes du genre. Ainsi, il est pertinent de penser que les erreurs dans l'exemple (70) concernent le problème de l'attribution du genre, mais pas celui de l'accord du genre.

Notons que les erreurs concernant l'attribution du genre pourraient être également de nature temporaire. Dans l'exemple (71), trois occurrences du nom *coupe* ont été trouvées dans le texte produit par le même scripteur. Le scripteur a utilisé les déterminants du genre féminin dans les premiers énoncés, et choisi un déterminant du genre masculin pour le nom *coupe* dans le dernier énoncé. Ici il est pertinent de penser

que le choix du genre masculin est dû à une erreur momentanée, ou un lapsus, et que ce nom est encodé comme un féminin dans le lexique mental du scripteur.

- (71) a. \* Il réclama de //combattre\_avec\_Vermeil\_et// ramener la coupe pour le roi. (L4-026)
- b. \* Perceval lui dit qu'il avait envie de demander une coupe d'or du roi et devenir le chevalier. (L4-026)
- c. \* Perceval tua Vermeil et [rap] rattrapa {le}\_{la} coupe [d] en or. (L4-026)

Cette catégorie d'erreur est la plus importante, avec 119 occurrences, qui représentent 77.8% du nombre total d'erreurs de genre trouvées dans notre corpus.

#### 4.2.3.5.2. Erreurs concernant l'accord en genre

Dans certains cas, les formes erronées produites par les scripteurs donnent lieu à une interprétation en faveur d'erreurs d'accord de genre. La tête nominale ne déclenche pas un accord soit avec des déterminants ou des adjectifs dans le groupe nominal, soit avec des adjectifs dans le groupe verbal.

Dans l'exemple (72), l'adjectif *mécontent* ne s'accorde ni en genre ni en nombre avec le nom *connaissance* qu'il qualifie.

- (72) \* Après les connaissances {mécontent}\_{mécontentes}, Perceval rencontra quelques challenges. (L3-021)

Nous avons identifié 12 occurrences de ce type d'erreur (7.8% au total).

Certaines erreurs portées par les adjectifs résultent, nous semble-t-il, d'une connaissance insuffisante des formes orthographiques d'adjectifs.

- (73) \* Il portait une armure vermeil neuve. (L2-02-003)

Notons que dans cet exemple, tous les modificateurs du nom *armure* portait la



marque congruente du genre, sauf l'adjectif qualificatif *vermeil*. Cela s'explique probablement par le fait que l'adjectif *vermeil* était un mot nouveau pour le scripteur durant la production écrite, et qu'il ne connaissait pas la forme graphique du genre féminin.

Dans l'exemple (74), le scripteur a bien prédit le genre du nom *roi* et choisi par conséquent le genre congruent pour l'article défini. Mais il a omis de choisir la marque du genre appropriée pour l'adjectif *surpris*, sur lequel le genre du sujet rejaillit aussi par le verbe prédicatif *être*. Selon nous, l'existence du mot *surprise*, comme nom, pourrait probablement amoindrir la vigilance du scripteur sur l'accord du genre.

(74) \* Le roi était {surprise}\_{surpris} que cette fille n'avait ~pas~ {rit}\_{ri} avant. (L1-01-017)

6 occurrences ont été identifiées dans cette sous-catégorie (4% au total).

Selon les chercheurs Dewaele et Véronique (2001), le scripteur pourrait également commettre les erreurs d'accord en raison d'une désactivation prématurée du nœud de genre. Dans ce scénario, les chercheurs supposent que l'information correcte concernant le genre est bien attachée au lemme. Et le nœud de genre est activé et sélectionné lors de la production langagière, mais cette information n'est pas fournie systématiquement à tous les modificateurs durant le processus. Cela pourrait s'expliquer par une désactivation du nœud de genre avant que tous les modificateurs n'aient été « servis ».

Cette hypothèse pourrait être utilisée pour expliquer l'exemple (75), où les déterminants et les adjectifs ne gardent pas la même marque du genre avec leur tête nominale féminine :

- (75) a. \* Il portait {un}\_{une} armure vermeille toute neuve. (L2-02-013)  
b. \* Il a rendu compte d' {un}\_{une} fille assise à la table. (L4-025)

La forme du féminin des adjectifs *vermeil* et *neuf* dans l'exemple (75a) ou *assis*

dans (75b) montre que les deux scripteurs étaient conscients du genre des deux noms concernés et qu'ils ont réussi à choisir le genre congruent pour les adjectifs. Cependant, ils ont omis de choisir le genre approprié pour les déterminants au début du groupe nominal, ou ils ont désactivé le nœud de genre prématurément.

Cette sous-catégorie contient 7 occurrences, représentant 4.5% du nombre total d'erreurs.

Une autre sous-catégorie des erreurs d'accord de genre concerne *Apparent Gender Errors*. Pour comprendre les erreurs de cette sous-catégorie, il faudrait revenir au processus où le message préverbal est produit. Le message préverbal active le lemme nominal et ses informations diacritiques, y compris celle du genre, qui permet au locuteur de choisir un déterminant congruent. Cependant, ce lemme nominal ne peut pas être récupéré et le locuteur doit ainsi activer un autre lemme pour le remplacer. Si le genre du deuxième lemme est différent de celui du premier, deux déterminants de genre différents pourraient précéder le nom, qui créait donc l'impression d'une erreur d'accord de genre. Dewaele et Véronique (2001) ont trouvé plusieurs occurrences de ce type dans leur corpus oral. Par exemple, *Euh il y a une, un centre médical à Bruxelles et euh ce centre donnait...* Selon les chercheurs, cette forme erronée pourrait s'expliquer par une tentative avortée de récupérer un autre nom féminin, tels que *clinique* ou *maison de santé*.

Bien que cette hypothèse soit utilisée par Dewarle et Véronique pour expliquer les erreurs trouvées dans le corpus oral, elle pourrait nous renseigner sur certains facteurs en jeu dans l'exemple suivant :

Le transfert venant d'homophones pourrait encore compliquer le choix du genre pour les apprenants chinois de français.

(76) a. \* Finalement Perceval a quitté le cours de roi. (L2-02-014)

b. \* Perceval est arrivé dans dans le cours de roi Arthur. (L2-02-014)

Lors du processus de la production, le message préverbal demandait un nom indiquant le lieu résidant par le roi, cependant le scripteur n'a pas été capable de

récupérer le lexème *cour*, et il a activé son homophone *cours* pour le remplacer. Ainsi, ce type d'erreur ne peut pas être traité comme erreur purement de genre.

(77) \* Il est allé {au}\_ {à\_la} cour du roi Arthur. (L2-01-006)

Dans notre corpus, nous constatons que de nombreuses erreurs concernent la prédiction du genre du nom *cour*. Cela pourrait s'expliquer par une mauvaise attribution du genre. Mais ici, une deuxième explication est possible. C'est que le processus de l'accord en genre pourrait être compliqué par l'existence de son homophone *cours*. Le fait que l'un est féminin et l'autre masculin perturbe probablement le choix du scripteur.

En suivant cette idée, nous pouvons fournir une deuxième explication pour comprendre les erreurs de genre concernant le nom *coupe* :

(78) \* Perceval a [suivr] suivi {le}\_ {la} troupe et il a obtenu {un}\_ {une} coupe en or dans sa main droite. (L2-02-013)

Le nom *coupe* entraîne de nombreuses erreurs d'accord de genre dans notre corpus, comme ce qu'a fait le scripteur dans l'exemple (78). Il a choisi le déterminant masculin alors que le nom *coupe* exige un déterminant féminin. Selon nous, ce mauvais choix du déterminant pourrait provenir d'un effet d'analogie, ou de transfert. La coupe d'or, objet lié au monde des chevaliers, pourrait être inconnu pour les étudiants chinois. Il est probable qu'en essayant de récupérer ce mot inconnu lors de la production du texte, ils ont fait référence aux mots comme *coup*, ou *couple*, qui sont plus fréquents dans leur première année des études du français, et qui ont une forme phonétique et graphique semblable que celle de *coupe*, et que par analogie, ils ont donné un déterminant masculin au nom *coupe*.

Il existe également les cas où le lemme cherché par le scripteur est déjà activé, avec ou sans l'information du genre correcte, mais le scripteur n'est pas capable de récupérer ce lexème cible, ou ne peut récupérer que partiellement ce lexème. Ainsi le locuteur peut s'appuyer sur ses ressources linguistiques préalables et créer une

invention lexicale. Dans l'exemple suivant, nous voyons que le nom *lever* n'est que partiellement récupéré par le scripteur. La forme produite par le scripteur a la tige du lexème cible combinée avec un suffixe non-cible, influencé probablement par un autre nom *levée*.

(79) \* Il marchait {jusqu\_QUOTE\_à\_la}\_ {jusqu\_QUOTE\_au} {levé}\_ {lever} du soleil à la mer.

7 occurrences ont été identifiées dans cette sous-catégorie, ce qui représente 4.5% du nombre total d'erreurs de genre.

#### **4.2.3.6. Discussion**

##### **4.2.3.6.1. Taux de réussite et limites de constituants**

Pour expliquer l'acquisition de l'accord en genre de L2, une perspective théorique importante est la théorie de processabilité (la TP) de Pienemann (1988), qui met en lumière quelles formes langagières en L2 sont processables à quels niveaux de développement langagier. Le développement grammatical (morphologique et syntaxique) est régi par des contraintes psycholinguistiques, organisées dans une hiérarchie, et basées sur la processabilité de structures grammaticales. Ainsi, l'apprenant L2 ne peut restructurer que les connaissances morphologiques et syntaxiques qui lui sont processables à son niveau actuel. Notons que chaque niveau de cette hiérarchie acquisitionnelle constitue une condition préalable pour entrer dans le niveau suivant (Lindström, 2013).

Bartning (2000) nous indique qu'un concept clé de ce système développemental est l'échange d'informations grammaticales aux différents niveaux structurels et développemental. Selon ce modèle développemental, l'apprenant identifie et acquiert tout d'abord les mots de la langue cible au niveau initial. Le niveau suivant consiste à classer et catégoriser les lexèmes selon leurs informations, tels que la définitude, le nombre, le cas, etc. C'est le niveau de la morphologie lexicale. Le niveau suivant

permet à l'apprenant d'échanger les informations grammaticales entre les éléments se situant dans un même syntagme, par exemple, entre la tête nominale et ses modificateurs tel que déterminants, adjectifs, dans un syntagme nominal. Lorsque l'apprenant est capable de traiter les informations morphosyntaxiques, il pourrait entrer au niveau suivant, qui implique l'échange d'informations grammaticales entre deux syntagmes différents. Enfin serait établie la relation hiérarchique entre les propositions principales et subordonnées, et l'échange d'informations entre les propositions de niveaux différents serait possible à ce stade acquisitionnel.

**Tableau 40.** Théorie de la processabilité de Piennemann appliquée à l'acquisition de l'accord adjectival en français (cité par Bartning 2001)

Stade	Procédure de processabilité	Échange d'informations	Marquage d'accord	Exemple
5	Procédure de proposition subordonnée	Propositions principale et subordonnée	Accord entre propositions principale et subordonnée	La maison qui est verte
4	Procédure inter-syntagmatique	Phrase simple	Accord d'adjectifs attributs	La maison est verte
3	Procédure intra-syntagmatique	Syntagme	Accord de déterminants et d'adjectifs épithètes	La/une maison verte
2	Procédure de catégorie	Catégorie	Attribution du genre et du nombre	Vert, -e ; Maison, -s
1	Mot			Vert maison

Ainsi, selon ce modèle développemental, l'accord adjectival situant dans un groupe nominal serait acquis avant celui qui se trouve dans une position prédicative (*la pomme sucrée* VS *la pomme est sucrée*). Notons que la théorie de Piennemann concerne

l'émergence de la première apparition d'un phénomène grammaticale, alors que nos données représentent notamment les stades de développement ultérieurs. Comme nous nous intéressons également au traitement d'un phénomène grammatical chez l'apprenant après l'émergence de la première occurrence, il nous semble intéressant d'étendre la théorie de Pienemann aux stades de développement ultérieurs et d'examiner une structure donnée à la lumière du modèle hiérarchique. Sachant que cette théorie concerne également un ordre hiérarchique des procédures d'automatisation. Comme Dewaele et Véronique (2001) l'indiquent, les structures maîtrisées plus tôt par l'apprenant devrait être automatisées plus facilement.

Les données du présent travail montrent que, pour chaque groupe d'apprenants, les différences de taux de réussite entre les positions attributives et les positions prédicatives ne sont pas statistiquement significatives. Cela implique que, en ce qui concerne l'accord adjectival, la question de savoir si l'information grammaticale a besoin de traverser les frontières posées par les différents constituants n'est plus pertinente dans l'interlangue de ces apprenants. Autrement dit, l'échange d'informations grammaticales entre les constituants différents ne pose plus le problème aux apprenants. L'étude de Dewaele et Véronique (2001) montre les résultats similaires en examinant le traitement de l'accord adjectival en français chez un groupe d'apprenants néerlandais au niveau avancé.

Cependant, les adjectifs en apposition représentent un profil développemental différent. Les résultats montrent que l'accord adjectival en apposition pose plus de difficultés que celui des autres positions aux apprenants de la Licence 2 et de la Licence 3. Bien que l'échantillon soit très petit, cela nous indique que l'accord adjectival en apposition est acquis et automatisé plus tard que l'accord des autres positions, qui est conforme au modèle hiérarchique de Pienemann.

Cependant, lorsque nous étendons la théorie de Pienemann à la maîtrise et l'automatisation de l'accord adjectival, c'est la forme féminine qui obtient le taux de réussite le plus bas. Les données montrent que plus de 80% des erreurs s'expliquent par l'omission de la marque du genre féminin. Les apprenants ont évidemment du mal à produire et automatiser l'inflexion de la forme féminine. Comme Bartning (2001)

l'indique, cette difficulté avec la forme morphologique semble être plus prévisible et systématique qu'avec l'échange d'informations grammaticales à travers les frontières posées par de différents constituants.

#### 4.2.3.6.2. Taux de réussite et nature du modificateur

Dans la partie précédente, nous avons examiné respectivement les données de l'accord des déterminants et de l'accord adjectival. Notons qu'une différence significative est constatée si l'on compare les deux types de modificateurs. C'est-à-dire le marquage du genre pour les déterminants pose plus de problèmes que celui pour les adjectifs. Si dans notre corpus les adjectifs présentent un taux de réussite plus élevé que les déterminants en ce qui concerne l'accord en genre, cette observation est contraire aux résultats de Dewaele & Véronique (2001) et de Bartning (2000).

A partir d'une analyse de l'interlangue française des locuteurs néerlandais, Dewaele et Véronique (2001) constatent que le taux de réussite pour l'accord en genre des déterminants était plus élevé que celui des adjectifs. Cette observation est partiellement confirmée dans l'étude de Bartning (2000). Chez les apprenants avancés, le taux de réussite pour les déterminants est plus élevé que pour les adjectifs (Det =90% VS Adj =81%), alors que c'est l'inverse chez les apprenants pré-avancés (Det =74% VS Adj =80%). Bartning suggère que « This is probably due to the random use of the gender distinction on the determiner at earlier stages » (2000 : 235). Toutefois, il n'en est pas de même pour les apprenants de notre corpus.

**Tableau 41.** Accord en genre pour les déterminants et les adjectifs

	Licence 2	Licence 3	Licence 4	Total
Déterminant	823/858 =95.92%	500/534 =93.63%	747/798 =93.60%	2070/2190 =94.52%
Adjectif	425/429 =99.06%	266/277 =96.02%	314/330 =95.15%	1005/1036 =97%

Les apprenants chinois des trois niveaux ont commis plus d'erreurs d'accord de

genre sur les déterminants que sur les adjectifs. Bien que l'accord en genre reste toujours problématique pour les déterminants, il nous semble indispensable d'indiquer que, le taux de réussite atteint 95.92% au stade initial. Pour les apprenants ayant un an d'apprentissage de français, ils démontrent une bonne maîtrise de l'accord en genre. Ainsi, l'hypothèse proposée par Bartning (2000), selon nous, ne permet pas à expliquer les performances des apprenants chinois. Comme ce que nous avons énoncé auparavant, Bartning (2000) suggère que le taux de réussite plus bas des déterminants par rapport aux adjectifs observés chez les apprenants pré-avancés résulte d'une utilisation aléatoire de l'opposition du genre pour les déterminants aux stades initiaux, et que ce phénomène disparaît chez les apprenants avancés, qui ont eu une meilleure performance pour les déterminants que pour les adjectifs. Il n'en est pas de même avec les apprenants chinois. Ils ont une bonne maîtrise de l'accord en genre des déterminants, cependant le taux de réussite n'augmentait pas avec l'avancement scolaire. Et les déterminants leur posent toujours plus de problèmes que les adjectifs.

Ces résultats sont également contraires aux observations tirées des études de l'acquisition L1. A partir d'une tâche de complètement réalisée à l'oral, Roulet et Jacobowicz (2006) montrent que les enfants français de 4 et de 6 ans commentaient plus d'erreurs d'accord de genre sur les adjectifs que sur les déterminants. Cela fait écho à la recherche de Tucker et al. (1977), qui indiquent également dans leur travail l'importance des déterminants pour la caractérisation du genre. Autrement dit, pour les apprenants L1, le genre de noms est principalement acquis par leur déterminant. En analysant le traitement de l'accord en genre chez les apprenants d'espagnol L2, Grüter et ses collègues (2011) proposent une théorie différente pour comprendre comment les apprenants acquièrent l'accord en genre d'une langue étrangère.

Les chercheurs suggèrent que cette difficulté liée à l'accord en genre des déterminants s'explique probablement par la représentation lexicale des noms. Pour les langues qui demandent le marquage du genre, tous les noms d'un genre grammatical donné sont liés à un nœud de genre spécifiant leur genre grammatical (Schriefers & Jescheniak, 1999 : 577). Autrement dit, les informations du genre sont stockées, comme caractéristique intrinsèque, à l'entrée de chaque nom au lexique mental. Pour les noms



du français, les indices phonologiques et sémantiques ne sont pas suffisants pour établir leur appartenance du genre. Ainsi, afin de détecter le genre de noms, les apprenants doivent avoir recours au lien de cooccurrence entre noms et leurs modificateurs marqués par le genre, notamment entre noms et leur déterminant. Le calcul de ce lien de cooccurrence et des probabilités de transition se révèle être un mécanisme essentiel dans l'apprentissage langagier chez de jeunes enfants (Saffran et *al.*, 1996). Étant donné que les probabilités de transition entre déterminants et noms (en arrière) sont généralement élevées (Pelucchi, Hay & Saffran, 2009), il n'est pas surprenant que les jeunes enfants traitent parfois le groupe déterminant-nom comme un bloc non-segmentable (Carroll, 1989). Selon Grüter et ses collègues (2011), ce type d'erreur concernant la non-segmentation illustre l'association étroite entre noms et déterminants dans le lexique mental L1. Les chercheurs indiquent également que l'association étroite entre déterminants ayant une marque du genre et noms est susceptible d'être lexicalement spécifique. Il semble pertinent de supposer que, avec l'enrichissement du vocabulaire, une association entre noms et indices du genre plus abstraits serait établie. Étant donné les fortes associations initiales entre noms et déterminants, les associations entre noms et leur nœud de genre, apparaissant dans le lexique mental L1 plus mature, pourraient être aussi fortes.

Quant aux apprenants L2, selon les chercheurs, calculer les informations distributives n'est pas le seul ou le moyen le plus efficace pour apprendre les noms et établir leur catégorie du genre. Les apprenants L2 au-delà de la petite enfance ont accès à beaucoup d'informations métalinguistiques de l'input pour identifier et apprendre les noms. Ainsi, il est peu probable que les apprenants L2 s'appuient, autant que les apprenants L2, sur la cooccurrence des éléments, tels que noms et déterminants, pour établir la catégorie du genre de noms. Ainsi, il est pertinent de penser que les associations fortes entre noms et déterminants émergées dans le lexique mental L1 ne se développeraient pas autant chez les apprenants L2. Les résultats tirés de notre corpus montrent que les apprenants chinois ont atteint une bonne maîtrise de l'accord en genre dès le niveau de Licence 2. Cela plaide en faveur de l'établissement de la catégorie du genre dans la représentation lexicale de noms chez ces apprenants. Toutefois, les

associations entre les noms et les informations de leur catégorie du genre ne seraient pas aussi fortes dans le lexique mental L2 que dans celui de L1. Ainsi persistent les erreurs de ce type tout au long de la scolarité.

#### 4.2.3.6.3. Sur-généralisation et effets de L1

Selon les données de notre corpus, la sur-généralisation du genre masculin est assez fréquente dans l'interlangue des apprenants. Comme nous l'avons indiqué auparavant, la forme *le* est la forme par défaut utilisée dans le contexte qui demande un déterminant défini, alors que la forme masculine du déterminant indéfini n'est pas sur-généralisée qu'à partir de la Licence 3. Cette observation qui implique le choix de la forme masculine au lieu de la forme féminine confirme les résultats de Bartning (2000), qui a constaté la sur-généralisation du genre masculin pour les déterminants définis et les adjectifs chez les apprenants au niveau avancé. Cependant, il n'en va pas de même pour les apprenants au niveau pré-avancé. Bartning suggère que les apprenants pré-avancés n'ont pas encore commencé à utiliser la stratégie de sur-généralisation du genre masculin, et elle soutient que la forme du masculin est acquise avec celle du féminin.

A partir de l'observation de l'interlangue du français produit par les locuteurs anglais, Carroll (1989 : 573) présente une théorie intéressante pour expliquer la sur-généralisation du genre masculin, ainsi que la difficulté à maîtriser l'accord de genre. Selon elle, il existe un transfert dans l'acquisition de système du genre français par les apprenants anglophones. Les francophones apprennent des déterminants comme une partie de noms, et l'association forte entre déterminant et nom leur permet de retrouver facilement les caractéristiques morphosyntaxiques de noms connus. Et il n'en va pas de même pour les apprenants anglophones. Ils transfèrent la catégorie de noms de leur langue maternelle, qui ne dispose pas de la caractéristique inhérente du genre, à l'acquisition de noms français. Ainsi lors de l'apprentissage du français, les anglophones apprennent des déterminants comme unité indépendante, et utilisent d'autres informations linguistiques pour identifier le genre de nouveaux mots. Carroll (1989 : 581) indique également que les apprenants pourraient développer un système de genre raisonnablement correct, mais qu'ils n'acquièrent jamais la compétence des

locuteurs natifs. Lorsque les apprenants allophones ne sont plus capables de retrouver le genre de noms, ils pourraient recourir à la stratégie de sur-généralisation, bien qu'il existe des risques. D'une manière générale, la forme du masculin est également la forme non-marquée en français. Selon Wise (1997), la forme du masculin peut être utilisée comme un générique, pour inclure le genre du féminin. Il s'agit d'une forme polysémique qui fonctionne à la fois au niveau lexical et grammatical.

#### **4.2.3.7. Conclusion**

L'objectif de ce travail est de décrire l'utilisation de la morphologie écrite du genre chez les apprenants chinois du français L2. A partir de l'analyse synchronique et diachronique des données, un certain nombre de tendances intéressantes ont pu être dégagées. Après un an d'études du français, les apprenants chinois démontrent une maîtrise précoce de la morphologie écrite du genre. Le taux de réussite est de plus de 90% pour les déterminants et les adjectifs. Cela nous montre que les connaissances sur les marques du genre sont bien maîtrisées par les apprenants et ces derniers sont en mesure de les mettre en œuvre dans une production écrite pour la plupart du temps. L'hypothèse la plus probable susceptible d'expliquer ce fait est que les apprenants chinois sont plus âgés et disposent d'une capacité cognitive assez avancée. Contrairement aux enfants français, qui cherchent à faire des tentatives de catégorisation lors de leur construction du système écrit, les apprenants chinois ont déjà construit le système écrit et peuvent transférer leurs connaissances linguistiques préalables en étudiant une langue étrangère. A l'aide d'un enseignement explicite et d'un entraînement intensif, ils sont en mesure de faire une analyse morphosyntaxique suffisante pour les syntagmes nominaux et verbaux, et puis de gérer l'application des règles de l'accord du genre.

Néanmoins, les erreurs persistent et le taux de réussite ne diminue pas significativement en fonction de niveau linguistique. Quant aux graphies erronées commises par les apprenants, elles sont hétérogènes et les explications sont multiples. L'erreur peut résulter d'une mauvaise attribution du genre : la tête nominale est

assignée au genre incorrect et ses modificateurs portent ainsi les marques du genre erronées. Rappelons que l'attribution du genre congruent à la tête nominale pose une difficulté majeure aux apprenants chinois et que les erreurs de ce type représentent 77.8% du nombre total d'erreurs trouvées dans le corpus. Les graphies erronées produites par les apprenants peuvent également s'expliquer par un mauvais accord de genre : la tête nominale ne déclenche pas un accord soit avec des déterminants ou des adjectifs dans le groupe nominal, soit avec des adjectifs dans le groupe verbal. Ce type d'erreur est beaucoup moins important dans notre corpus, représentant un peu plus de 20% du nombre total d'erreurs de genre. Un autre élément à remarquer dans notre corpus concerne le fait que le marquage du genre pour les déterminants pose plus de problèmes que celui pour les adjectifs.

Grüter et ses collègues (2011) proposent une théorie différente pour comprendre comment les apprenants acquièrent la morphologie du genre d'une langue étrangère. D'après leur étude, les associations entre les noms et les informations de leur catégorie du genre ne seraient pas aussi fortes dans le lexique mental en L2 que dans celui de la L1. Les apprenants chinois se trouvent ainsi confrontés à la difficulté d'attribuer le genre congruent à la tête nominale et de sélectionner le déterminant approprié. Cette hypothèse liée aux manières d'intégrer des informations morphologiques dans la représentation lexicale des apprenants allophones explique, nous semble-il, le nombre important des erreurs relevant de l'attribution du genre dans notre corpus, ainsi que la difficulté posée par les déterminants.

A la lumière de ces résultats, il apparaît souhaitable de formuler de nouvelles propositions didactiques pour un enseignement plus efficace de la morphologie du genre. Au lieu de mettre tout l'accent de l'enseignement sur l'automatisation de l'accord du genre, il faut également aider les apprenants à intégrer plus efficacement les informations morphologiques de mots dans leur représentation lexicale. Pour les apprenants chinois du français L2, une association étroite entre noms et déterminants tant au niveau de l'oral que de l'écrit pourrait faciliter leur réalisation de la morphologie écrite du genre.

### **4.3. Orthographe et morphologie : le problème des verbes**

La maîtrise de la morphologie joue un rôle important pour la compréhension et pour la production écrite du français, dont le système est riche de flexions et caractérisé par un décalage entre la morphologie de l'oral et celle de l'écrit (Jaffré, 1998). Au sein de la zone de la morphologie écrite, les formes verbales relèvent de la classe morphosyntaxique la plus variable. Ainsi leur réalisation écrite demande-t-elle aux scripteurs « le plus lourd travail d'apprentissage et l'effort de contrôle le plus grand » (Brissaud & Sandon, 1999).

En français, l'apprentissage de la morphologie verbale écrite est un parcours semé d'embûches. Si le problème est posé pour des apprenants natifs, il ne peut que l'être de manière encore plus ardue pour les apprenants du français L2. Le traitement de la morphologie verbale constitue par conséquent un objet d'étude fructueux tant en L1 qu'en L2.

Dans ce qui suit, nous présenterons d'abord les caractéristiques essentielles de la morphologie verbale du français, ainsi que les études principales sur l'acquisition de la morphologie verbale en français L2. Dans la deuxième partie, nous examinerons la réalisation de la morphologie verbale, en classant les graphies erronées sous deux ensembles : les erreurs issues de la morphologie verbale sans altérer la valeur phonique et les formes erronées altérant la valeur phonique de verbes.

#### **4.3.1. La morphologie verbale du français**

« Catégorie grammaticale complexe qui incorpore plusieurs marques de flexion (temps, mode, personne, nombre, genre) » (Escoubas-Benveniste & Di Domenico, 2017), le verbe constitue un aspect fondamental dans l'acquisition du français, L1 et L2. A l'écrit, la zone du verbe relève de la classe morphosyntaxique la plus variable.

Selon Catach (1995), la morphologie verbale est particulièrement concernée par la zone des morphogrammes au sein du plurisystème de l'orthographe française. Situés après le noyau central des phonogrammes, les morphogrammes sont directement

chargés de la représentation d'unités significatives (Jaffré, 2003) et répondent, tout particulièrement pour le verbe, aux besoins de la facilité de lecture, ainsi que de la clarté de lecture-écriture. Les morphogrammes signalent « les contenus grammaticaux des paradigmes fermés de désinences » (Escoubas-Benveniste & Di Domenico, 2017). Notons que cette « sémiographie » possède l'image graphique multiple. Par exemple, il existe plusieurs morphogrammes différents pour marquer la personne au singulier et ils varient en fonction des groupes de verbes : *-e, -es, -e* pour les verbes en *-er* ; *-s, -s, -t* ou *-s, -s, ø* ou encore *-x, -x, -t* pour les deux autres groupes de verbes. Les marques du pluriel sont systématiques et régulières dans le secteur de la morphologie verbale écrite : *-ons, -ez, -nt*. L'essentiel de la distinction graphique (les marques du singulier et de la 3PP) relève dans cette zone d'une morphologie sans correspondance phonique (Sandon & Brissaud, 1999).

Ce décalage entre la morphologie de l'écrit et celle de l'oral constitue un autre défi pour l'apprentissage de la morphologie verbale écrite. Ågren explique cette morphologie silencieuse de manière suivante (2005 : 30) :

« D'un point de vue morphologique, les marques de genre, de nombre, de personne, de temps et de mode diffèrent entre l'oral et l'écrit. De façon générale, le code écrit peut être caractérisé par la redondance des marques grammaticales (les petits enfants jouent). Par contre, l'oral se caractérise par l'économie morphologique et utilise moins de marques grammaticales que l'écrit. [...] Le français écrit se caractérise par rapport au français oral entre autres par une morphologie silencieuse qui touche les homophones verbaux et les accords en nombre et en genre. Ces caractéristiques de l'écrit sont considérées par de nombreux chercheurs comme des zones orthographiques fragiles, difficilement maîtrisées par les apprenants du français L1 et même les adultes » (Ågren, 2005 : 30).

En bref, tous les sons de mots français possèdent une image graphique à l'écrit ; en revanche, certaines graphies ne se prononcent jamais et d'autres graphies peuvent se

prononcer ou non en fonction de leur statut grammatical et des contextes<sup>51</sup>. Les morphogrammes flexionnels muets introduisent un autre problème de la morphologie verbale, homophonie.

L'homophonie systématique du français touche particulièrement la zone du verbe. Pouradier Duteil (1997) remarque que :

« Contrairement à une idée très répandue, le système verbal du français est complexe certes, mais très régulier. Pendant de longues générations on s'est contenté de décrire le verbe du seul point de vue de la graphie [...]. Cela revenait à décrire la langue sans tenir compte de son organisation phonique. » (Pouradier Duteil, 1997 : 7)

Elle signale que, dans le secteur de la morphologie verbale, la 1PP et la 2PP sont marquées à l'oral par respectivement les voyelles [ɔ̃] (-ons) et [e] (-ez) et que ces voyelles se retrouvent à tous les temps du verbe, sauf à celui du passé simple. En revanche, les autres personnes ne sont pas marquées à l'oral<sup>52</sup>.

Selon Bassano et *al.* (2001 : 32), « l'homophonie constitue au départ pour l'enfant un facteur de facilitation : une forme verbale sera d'autant plus facile à percevoir, repérer et reproduire qu'elle se trouve répétée de façon identique dans l'input ». Nouveau (2017) indique également que l'homophonie devrait faciliter considérablement l'apprentissage des conjugaisons en français L2. Néanmoins, l'hétérographie des formes fléchies homophones représentent une difficulté majeure à l'écrit pour les apprenants du français (Voir par exemple Jaffré, 2003 ; Brissaud et *al.*, 2006). En outre, la difficulté de transcrire les formes verbales homophones s'accroît pour la flexion du participe passé, dont la transcription demande « l'analyse de phénomènes morphologiques discontinus et plus ou moins distants, la connaissance des

---

<sup>51</sup> L'instabilité phonique des morphogrammes se traduit par l'apparition intermittente de nouveaux sons consonantiques, par exemple, -s [z] : *penses-y* (Escoubas-Benveniste & Di Domenico, 2017).

<sup>52</sup> L'accord de la 3PP est marqué par le changement du radical.

règles d'accord, la mise en œuvre des procédures visuographiques d'auto-contrôle (*il(s)...-é(s) ; elle(s)...-é(e)(s)*) » (Escoubas-Benveniste & Di Domenico, 2017).

Après avoir décrit les caractéristiques essentielles de la morphologie verbale écrite du français, nous présenterons dans la partie suivante les recherches principales sur l'acquisition de la morphologie verbale en L2.

#### 4.3.2. L'acquisition de la morphologie verbale en L2

Dans le domaine du français L2, la question de l'acquisition de la morphologie verbale intéresse aussi bien les chercheurs que les praticiens de l'enseignement de langue. Cet intérêt est motivé d'une part par le fait que le verbe est le noyau de la phrase et pose ainsi l'un des fondements nécessaires à l'acquisition du français, d'autre part parce que son appropriation présente des poches de résistance même à un niveau avancé. Michot (2015) explique le rôle majeur de la morphologie verbale dans l'acquisition L2 de manière suivante :

« L'appropriation de la morphologie verbale est de loin le domaine le plus étudié en acquisition de la morphosyntaxe en L2, ce qui s'explique par le rôle central que le verbe joue dans le processus de grammaticalisation de l'interlangue des apprenants, aussi bien sur le plan morphologique que syntaxique » (Michot, 2015 : 24).

Dans ce domaine, deux séries de travaux nous semblent particulièrement intéressants pour comprendre l'acquisition de la morphologie verbale en français L2 : la première est dominée par une discussion sur la finitude verbale ; la seconde porte principalement sur la construction du profil développemental d'apprenants au cours de leur acquisition de la morphologie verbale.



### 4.3.2.1. La finitude verbale

Dans ce travail, la finitude verbale concerne particulièrement la présence ou l'absence de toute « une série de marques morphologiques permettant de situer la forme verbale dans le temps et dans l'espace » (Michot, 2013). Véronique (2009) explique également cette notion :

« Un verbe fléchi ou conjugué est dit disposer d'une finitude (finiteness) morphologique. Est finie toute forme verbale conjuguée qui permet un contraste de temps, de mode etc. à l'aide de sa désinence, les autres formes verbales (participiales ou infinitives) sont dites non finies » (Véronique, 2009 : 369).

La finitude verbale fait l'objet de travaux importants en acquisition (voir par exemple : Klein & Perdue, 1997 ; Bartning & Schlyter, 2004). Cet intérêt est motivé par son rôle important pour la construction phrastique des énoncés des apprenants. Selon Perdue et ses collègues (2002), c'est autour du verbe fléchi que s'organisent d'autres éléments grammaticaux tels que la négation, les adverbes, ainsi que les pronoms. De plus, l'appréhension de la finitude permet aux apprenants d'apprendre le système de marquer le temps, l'aspect et le mode (TAM). La finitude verbale constitue ainsi un aspect fondamental dans le développement de l'interlangue des apprenants L2.

Klein et Perdue (1997) étudient l'interlangue des apprenants adultes L2 se situant à un stade initial et ils remarquent un phénomène, qu'ils appellent « la variété de base » (the basic variety) dans la production morphologique verbale de ce groupe d'apprenants :

« There is no inflection in the basic variety, hence no marking of case, number, gender, tense, aspect, agreement by morphology. Thus, lexical items typically occur in one invariant form. It corresponds to the stem, the infinitive or the nominative in the target language [...] » (Klein & Perdue

1997 : 322).

Donc le défi de l'acquisition de la morphologie verbale n'est pas seulement de connaître les différentes formes flexionnelles d'un paradigme morphologique, mais surtout de réussir la combinaison entre forme morphologique et contexte morphosyntaxique (Thomas, 2017). Selon Prévost et White (2000), l'acquisition de la morphologie en L1 est toujours réussie :

« Once morphosyntactic and semantic features have been mapped to a particular form, that association is permanent, preventing non-finite forms from surfacing in finite positions, as well as agreement mismatches. (...) In other words, once the more fully specified forms are acquired, they replace underspecified forms, which can no longer surface. Hence, one sees a developmental progression in L1 acquisition » (Prévost & White, 2000).

Néanmoins, cela n'est que partiellement vrai pour l'acquisition en L2. Prévost et White (1999, 2000) ont mené un travail sur deux apprenants adultes de français et deux apprenants d'allemand, en étudiant leurs productions orales spontanées. Les données montrent que les formes non-finies alternent avec les formes finies tout au long de la période de l'observation en français L2 et en allemand L2 et que la surutilisation des formes non-finies ne disparaissent jamais. L'absence du progrès dans le domaine de l'acquisition de la morphologie flexionnelle est également observée dans les études longitudinales. En étudiant le développement langagier d'un locuteur chinois de l'anglais L2 sur une période de 8.5 ans, Lardiere (1998) montre que dans l'état final de la grammaire anglaise, la production morphologique (en ce qui concerne l'accord et le temps) reste toujours variable et ne correspond pas toujours à la norme.

Bien que cette difficulté des apprenants L2 liée à la morphologie verbale soit largement reconnue, il n'y pas de consensus sur la nature de cette absence de flexion verbale. De manière générale, il existe deux approches différentes pour expliquer l'utilisation variable de la morphologie flexionnelle par des apprenants L2. La première

suppose que de telles difficultés se situent au niveau de la représentation lexicale ou la compétence de l'apprenant L2, tandis que la deuxième approche attribue ces difficultés à un certain manque au niveau du traitement ou du contrôle.

Selon the competence deficit approach (CDA), les erreurs morphologiques commises par les apprenants L2 sont le signe d'une acquisition incomplète des connaissances morphologiques (Jiang, 2004). Les apprenants peuvent connaître les règles sur les morphèmes flexionnels, mais comme Hale (1988) l'évoque, cette compréhension consciente et intellectuelle ne fait pas partie de leur compétence linguistique intégrée. En suivant cette approche, the Impaired Representation Hypothesis (IRH) suppose que l'absence de la flexion verbale s'explique par l'absence des catégories fonctionnelles de la part de l'apprenant. Plus précisément, la distinction entre les différentes catégories (noms, verbes, pronoms, etc.) n'est pas encore maîtrisée par les apprenants L2 (voir, par exemple, Meisel, 1997 ; Eubank, 1993, 1994 ; Beck, 1998).

The performance deficiency approach (PDA) s'y oppose, en revanche, en considérant les erreurs morphologiques en L2 comme les difficultés d'accéder, de récupérer ou de contrôler les connaissances morphologiques qui ont déjà été internalisées chez les apprenants. S'inscrivant dans le même courant, Prévost et White (2000) expliquent l'absence de flexion verbale en L2 par la théorie de Missing Surface Inflection (MSI) : les apprenants d'une L2 sont conscients des catégories flexionnelles et les marques d'accord et ils peuvent, en revanche, éprouver de la difficulté à réaliser la morphologie superficielle. Si l'apprenant produit une forme d'infinitif dans un contexte pour la forme du présent, il s'agit d'une erreur de substitution. Cet infinitif correspond à une forme par défaut et présente les caractéristiques des formes verbales finies.

Selon Jiang (2004), la différence entre la CDA et la PDA réside dans le fait que la première approche suppose que les apprenants L2 font plus qu'ils ne savent vraiment à l'aide de connaissances explicites et qu'en revanche, selon la seconde approche, les apprenants L2 savent plus qu'ils ne font, en raison de leur contrôle limité sur l'interlangue, qui est en train de développer. La différence entre les deux approches peut

être clairement démontrée par deux manières d'interpréter la performance du même sujet (P), décrit pour la première fois par Krashen et Pon (1975). Le sujet P, locuteur natif de chinois, a vécu aux États-Unis pendant plus de 10 ans et a très bien réussi à l'université. Cependant, elle a souvent commis les erreurs relevant des morphèmes flexionnels dans la production orale spontanée. Une fois que ces erreurs lui étaient présentées, elle disposait des connaissances linguistiques pour les corriger. En outre, elle était bonne scriptrice et ne faisait presque pas d'erreur à l'écrit. Selon Krashen (1981), le sujet P est capable d'utiliser les connaissances apprises pour surveiller et modifier la production et ses erreurs dans la production orale suggèrent que les morphèmes flexionnels n'ont pas encore été acquis. Sarwood Smith (1986) s'y oppose en proposant que le sujet P corrige par sa compétence acquise en L2 et ses erreurs dans la production orale implique un problème de contrôle.

#### **4.3.2.2. Les itinéraires acquisitionnels**

La deuxième série de travaux concerne la construction d'un modèle de l'acquisition morphologique. L'une des études significatives dans ce domaine est celle de Bartning et Schlyter (2004), qui présente un modèle de l'acquisition du français par les adultes suédois. Dans ce modèle, les chercheuses tracent l'échelle du développement d'un certain nombre de traits caractéristiques : la morphologie verbale, la négation, les pronoms objets, le genre ainsi que la subordination. En ce qui concerne la morphologie verbale, les chercheuses étudient d'abord une opposition fondamentale chez les apprenants suédois, « celles des formes finies courtes (*étudie*) » d'une part et « les formes non-finies (*étudier/étudié*) » d'autre part (Bartning et Schlyter, 2004). A partir de deux corpus oraux, les chercheuses peuvent observer un développement net : les apprenants emploient des formes invariables dans différents contextes, formes finies courtes (type *parle*) ou formes non-finies (type *parler* ou *parlé*) ; dans une deuxième étape, les apprenants peuvent faire une distinction de telle sorte que les formes finies courtes sont employées dans les contexte finis et les formes non-finies dans les contextes non finis ; lors d'une troisième étape, les formes verbales sont utilisées en

accord avec la langue cible et l'apprenant a atteint un niveau assez avancé, presque à la perfection (Bartning & Schlyter, 2004).

Néanmoins, la notion de « forme finie courte » nous semble ambiguë, définie par les deux chercheuses comme des verbes conjugués de l'indicatif présent au singulier (Bartning & Schlyter, 2004). La distinction entre formes finies courtes et formes non-finies est claire pour des verbes tels que *courir* ou *parler* (*court* vs *courir/couru* ; *parle* vs *parler/parlé*), tandis qu'elle l'est moins pour des verbes tels que *finir* ou *dire* (*finit* vs *finir/fini* ; *dit* vs *dire/dit*). En outre, Bartning et Schlyter ne font pas la distinction entre les différents types de verbes. Les verbes se distinguent l'un de l'autre en fonction du nombre de leurs bases orales<sup>53</sup>. Ainsi la conjugaison de verbes tels que *parler*, ayant une base orale, ne pose pas le même problème aux apprenants L2 que celle de verbes tels que *prendre*, ayant 4 formes orales.

En deuxième lieu, les chercheuses étudient dans quelle mesure les apprenants marquent le pluriel dans l'accord sujet-verbe (*nous parlons, ils sont, ils prennent*). Les chercheuses constatent que l'acquisition de l'accord sujet-verbe de tous les verbes confondus concerne « un processus long et lent » (Bartning et Schlyter, 2004). Le scénario développemental pour l'accord sujet-verbe proposé par les deux chercheuses est le suivant (Bartning et Schlyter, 2004) :

1. L'opposition au singulier est bien maîtrisée pour les copules et auxiliaires (*je suis/il est ; j'ai/il a*) ;
2. L'accord de la 1PP *nous V-ons* est productif<sup>54</sup> ;
3. L'accord de la 3PP *ils V-ont* (*ils sont/ont/vont/ont*) est productif ;
4. L'accord de la 3PP des verbes lexicaux irréguliers (*ils prennent*) est productif ;
5. L'accord de la 3PP des verbes lexicaux irréguliers est correct.

---

<sup>53</sup> Cela implique les différentes formes que peuvent revêtir le radical d'un verbe (Nouveau, 2017).

<sup>54</sup> Dans cette étude, les chercheuses définissent que l'accord est productif lorsqu'au moins 75% des contextes sont marqués.

En ce qui concerne l'écrit, Ågren (2005) souhaite vérifier si le profil développemental parcouru par les apprenants suédois en français L2 à l'écrit est le même que celui proposé par Bartning et Schlyter (2004). En mettant l'accent sur l'accord en nombre du français écrit, la chercheuse décrit le développement morphologique observé dans un corpus écrit. Selon elle, il existe un profil développemental net entre les apprenants à un stade initial et ceux à un stade avancé<sup>55</sup>. En ce qui concerne la morphologie verbale, son étude traite des formes de la 3PS et de la 3PP. Elle indique que, par rapport au français L2 parlé, la production écrite des apprenants montre également une émergence relativement précoce de l'accord verbal morphologique (Ågren, 2005 : 140). Cependant, cet accord verbal connaît « un développement étendu dans le temps » chez les apprenants suédois (Ågren, 2005 : 275).

A partir de ces travaux, on constate un intérêt particulier des linguistes et des didacticiens pour l'acquisition des formes verbales et son développement : Housen, Kemps et Pierrard (2006, 2009) appliquent l'échelle proposée par Bartning et Schlyter (2004) pour savoir si elle correspond à la description des apprenants dans des contextes instructionnels différents ; en s'intéressant particulièrement aux apprenants avancés finnophones, Paloheimo (2014) examine l'acquisition de la morphologie verbale à partir d'un corpus écrit ; afin d'affiner l'itinéraire acquisitionnel proposé par Bartning et Schlyter (2004), Michot (2014) aborde une étude sur l'acquisition de l'accord sujet-verbe en examinant les différents types de verbes produits ; un numéro thématique très récent de dimension internationale (Bulletin Vals-Asla, 2017) porte spécifiquement sur l'acquisition et la production des formes verbales à l'oral et leur emploi en contexte.

Bien que l'acquisition de la morphologie verbale en français L2 constitue un objet d'étude fructueux dans différents pays, elle n'a pas encore été examinée de manière systématique chez les apprenants chinois en français L2. Les études sur la morphologie verbale que nous avons mentionnées ci-dessus ont été menées auprès des apprenants scandinaves ou germaniques du français L2, dont la langue maternelle présente de

---

<sup>55</sup> Le nombre des erreurs diminue significativement en fonction du niveau linguistiques des apprenants suédois, du stade 1 au stade 4 (Ågren, 2005 : 128).

constantes variations morphologiques. Par conséquent, nous souhaitons contribuer à une meilleure compréhension de l'acquisition de la morphologie verbale, en examinant la production écrite d'un groupe d'apprenants, jusqu'alors peu connu, qui apprennent à écrire en français dans une salle de classe en Chine, et qui possèdent une langue maternelle représentant les variations morphologiques minimales.

#### **4.3.3. Les apprenants chinois face aux problèmes de morphologie verbale**

Les formes verbales produites par l'apprenant du français constituent un des paradigmes qui se prêtent le mieux à l'observation de l'utilisation et l'acquisition de la morphologie flexionnelle écrite de cette langue. Aussi l'objectif de cette partie est-il d'examiner comment les apprenants chinois, ayant une langue maternelle issue de la famille linguistique lointaine du français, réalisent la morphologie verbale écrite dans leur production écrite. Nous souhaitons également vérifier s'il existe les tendances interlangagières propres aux apprenants chinois du français L2, dont la langue maternelle présente une variation morphologique minimale.

Pour ce faire, les formes erronées produites par les apprenants chinois sont regroupées sous deux ensembles : le premier concerne les erreurs issues de la morphologie verbale sans altérer la valeur phonique ; le second porte sur les formes erronées altérant la valeur phonique de verbes.

##### **4.3.3.1. Erreurs issues de la morphologie verbale sans altération de la valeur phonique**

Nous traiterons tout d'abord les erreurs issues de la morphologie verbale qui n'altèrent pas la valeur phonique. Pour la plupart du temps, il s'agit de l'orthographe de formes homophones. Parmi tous les problèmes rencontrés dans l'apprentissage de l'orthographe, le traitement de l'homophonie attire particulièrement l'attention des chercheurs, puisque l'homophonie constitue un phénomène important dans le cas de l'orthographe du français. Afin de dresser une image aussi complète que possible pour

ce phénomène linguistique du français écrit, nous présenterons tout d'abord les notions essentielles, pour comprendre pourquoi la distinction des homophones verbaux pose une difficulté particulière en français écrit. Nous mènerons ensuite une analyse des données de notre corpus dans une perspective développementale et essaieront de caractériser l'interlangue des apprenants en nous basant sur une analyse des graphies erronées. Les scénarios psycholinguistiques possibles pour interpréter ces erreurs seront également présentés, en faisant référence aux caractéristiques spécifiques du traitement morphologique des apprenants de L2.

#### **4.3.3.1.1. L'homophonie en français**

Si la morphologie du français est motivée par des structures linguistiques originales, issues d'une influence étymologique, elle a également fait l'objet de tractations, de débats, résultant d'une attitude normative (Jaffré & Brissaud, 2006). Comme Jaffré l'indique (2006), ce n'est pas un hasard si la morphologie française est souvent liée à l'homophonie. Et il serait réducteur de ramener la morphologie des formes homophones à un accident de l'histoire (Jaffré & Brissaud, 2006). Donc il nous semble important de remettre le problème de l'homophonie en perspective historique dans la présente étude, avant d'entrer en analyse empirique.

##### **4.3.3.1.1.1. Une vue historique de la question de l'homophonie**

La distinction des homophones en langue française « ne fait pas partie d'un projet d'emblée systématique » (Jaffré, 2006). Même le latin, qui maintient une relation étroite entre l'écrit et l'oral, n'échappe pas à une tentative sémiographique, consistant à recourir à des lettres étymologiques pour distinguer des homophones (Desbordes, 1990).

Au cours des XII<sup>e</sup> – XIII<sup>e</sup> siècles, époque où on observe les premières manifestations de l'orthographe française, l'hétérographie semble « se frayer un chemin au coup par coup, dans le sillage complexe de la polyvalence phonographique et de l'étymologie » (Jaffré, 2006). Dans les époques suivantes, on assiste à un débat entre les tenants de l'ancienne orthographe et ceux de la nouvelle orthographe. Les



défenseurs de l'ancienne orthographe jouent un rôle décisif dans l'histoire, en préconisant les lettres superflues et les consonnes finales, même quand elles ne sont pas étymologiques. Ce courant est suivi par l'Académie. D'une façon générale, si l'origine des formes homophones est différente, ce sont le plus souvent « les vestiges de la prononciation disparue qui servent à établir des différences graphiques » (*bon vs bond, chêne vs chaîne*) ; s'il s'agit de la même origine, c'est « l'évolution sémantique qui rend nécessaire la distinction graphique » (*compte vs conte, dessin vs dessein*) (Jaffré, 2006).

Cependant, ce fonds étymologique ne se limite pas à la logographie. On le trouve également à la morphographie afin d'étoffer les désinences des mots. A partir du XI<sup>e</sup> siècle, on voit une chute des consonnes finales d'origine latine à l'oral, en revanche, ces consonnes sont restituées à l'écrit où elles jouent un rôle lexical et grammatical. Signalons que les rapports morphologiques de genre, de nombre et de dérivation contribuent le plus à ces changements progressifs. Grâce au fait qu'ils comportent des éléments historiques spécifiques, les homophones hétérographes de l'orthographe du français font preuve d'une diversité importante.

#### **4.3.3.1.1.2. Comment gérer les homophones en français ?**

L'opposition graphique des signes peut porter sur des phonogrammes (*ancree vs encre, pain vs pin*), ou sur des lettres non phonographiques (*voie vs voix*) ; elle peut opposer des éléments du lexique (*cent vs sang*), de la grammaire (*a vs à, ou vs où*) (Jaffré, 2006). D'autre part, l'hétérographie peut aussi être partielle, c'est-à-dire que la distinction graphique affecte seulement une partie fonctionnelle du mot. C'est notamment ce qui se passe pour les désinences verbales (*aimer vs aimé*) et les catégories grammaticales (nombre : *livre vs livres* ; genre : *continental vs continentale*). Il s'agit d'un des domaines les plus complexes de l'apprentissage de l'orthographe du français, parce que les formes « ne s'opposent pas cette fois sur la base d'un sens lexical ou d'une fonction grammaticale mais ont une partie sémantique commune et s'opposent pour l'essentiel par leur fonctions grammaticales » (Jaffré, 2006). Il s'agit donc de la gestion des structures bidimensionnelles pour le scripteur. En outre, ces distinctions graphiques

n'ont pas de contrepartie phonologique.

De ce fait, les scripteurs doivent faire référence à plusieurs moyens dans la production écrite pour lever l'ambiguïté des homophones hétérographes. Jaffré (2006) propose trois moyens essentiels et souvent complémentaires : les contextes, les catégories linguistiques et les réseaux morphologiques. Tout d'abord, le recours au contexte est très souvent considéré comme un moyen efficace pour différencier ce qui se ressemble. Dans une phrase telle que « *Je préfère le [pẽ] du boulanger de St-Michel* », le mot *boulangier* nous permet de sélectionner directement la forme graphique *pain*. Cependant, dans une phrase telle que « *Je me demande pourquoi le [pẽ] a brûlé* », le contexte ne semble pas être suffisant pour lever l'ambiguïté de termes homophones. La référence aux catégories d'appartenance des homophones est aussi un moyen utile pour distinguer divers cas de figure. Si l'on prend l'exemple des homophones *ou* et *où*, on voit que les deux mots appartiennent aux catégories morphologiques différentes, conjonction et adverbe (ou pronom relatif). Si l'on regarde l'autre couple, celui des homophones *pain* et *pin*, on voit que, bien qu'ils appartiennent à une même catégorie, celle du nom, ces deux mots ont un sémantisme totalement différent. Enfin, on pourrait se servir des réseaux morphologiques pour lever l'ambiguïté des homophones hétérographes. Autrement dit, les homophones hétérographes appartiennent à des réseaux dérivationnels, ce qui nous permet de justifier certaines de leurs configurations orthographiques. Si l'on prend l'exemple des homophones en [sã] : tel que *cent* (*centaine*), *sang* (*sanguin*), *sent* (*sentir*), on voit que ces analogies constituent une option forte de l'orthographe du français. Néanmoins, il existe des mots qui appartiennent à une même série et qui présentent des formes orthographiques différentes, tels que *souffler* et *boursouffler*.

Il s'ensuit que pour bien gérer les homophones hétérographes dans l'orthographe du français, les scripteurs doivent recourir à plusieurs stratégies et les mobiliser de manière différenciée en fonction des situations.

#### **4.3.3.1.1.3. Acquisition de l'orthographe des homophones**

Si nous prenons en compte l'importance et la complexité des homophones

hétérographes dans l'orthographe du français, il nous semble particulièrement crucial de se demander comment les apprenants traitent les homophones lors de l'apprentissage de l'orthographe. Notons que « les travaux sur l'acquisition et l'apprentissage de l'orthographe ont maintes fois souligné - et continuent de le faire - à quel point cette distinction constitue très vraisemblablement la source majeure des problèmes orthographiques » (Jaffré, 2006). Les données empiriques attestent que les erreurs d'homophonie occupent une place particulière dans la production écrite des enfants : elles sont nombreuses, et d'autant plus, elles apparaissent dès le début de la scolarité et résistent jusqu'au collège voire au lycée (Ros Dupont, 2000).

Dans le domaine du français L2, au contraire, le traitement des homophones hétérographes reçoit peu d'attention des chercheurs, surtout en ce qui concerne les apprenants chinois. Pour étudier le rapport des étudiants natifs et allophones à l'orthographe française, Makassikis et Pellat (2011) font une analyse de copies et une expérimentation orthographique sur les homonymes. A travers l'analyse de copies, les chercheurs constatent qu'une répartition des erreurs orthographiques est à peu près similaire dans les deux groupes : les deux zones à haut risque sont, en français L1 comme en français L2, les phonogrammes et les morphogrammes. Cependant, les résultats relèvent une petite différence : les étudiants étrangers commettent moins d'erreurs sur les logogrammes et les lettres étymologiques et historiques que les étudiants français.

Les chercheurs font ainsi l'hypothèse que les étudiants étrangers sont plus forts dans l'écriture des homonymes en situation de production. Ils supposent que cette différence résulte de la maîtrise de la langue maternelle des étudiants allophones, à laquelle ils peuvent faire la référence : « ils ont tendance à associer les lettres étymologiques et historiques du français, généralement inaudibles ou insuffisamment audibles, à leurs équivalents souvent audibles dans leur langue, ce qui aide à les écrire correctement » (Makassikis & Pellat, 2011). De même, pour les logogrammes, « le passage par l'équivalent en langue maternelle ou par une autre langue connue des apprenants leur permet de distinguer clairement ce qui en français est identique phonétiquement » (Makassikis & Pellat, 2011).

En revanche, dans l'expérimentation sur les homonymes, les chercheurs constatent que les étudiants francophones sont plus performants dans l'écriture des homonymes dans un exercice à trous. Selon les chercheurs, ces performances orthographiques amoindries s'expliquaient par « la fréquence de rencontre et d'utilisation des mots. Les chercheurs remarquent également un faible pourcentage de réussite lorsque les mots sont rarement vus et employés. Néanmoins, le pourcentage de réussite s'accroissait quand les mots sont souvent rencontrés et utilisés par les étudiants. Puisque les étudiants étrangers sont moins exposés aux mots français et ont moins d'occasion de les employer dans la vie réelle, ils sont moins forts dans les performances orthographiques que les étudiants natifs.

La recherche de Makassikis et Pellat nous donne quelques éléments de réflexion sur le traitement des homophones chez les apprenants du français L2. Si l'apprentissage de l'orthographe des homophones ne pose pas les mêmes problèmes en français L2 qu'en français L1, les différences entre les populations des apprenants ne sont pas précisées dans cette recherche. Il faut noter que le développement de l'écrit en L2 se rapporte à la distance entre la langue maternelle et la langue cible<sup>56</sup> (Daigle, Mathieu & Montésinos-Gelet, 2008). Et chaque population d'apprenant « affronte, en fonction des spécificités de sa langue maternelle, des défis différents » lors de l'apprentissage du français (Gerolimich & Stabarin, 2017). De surcroît, le classement des erreurs des apprenants du français L1 et du français L2, bien que d'un grand intérêt, ne nous semble pas suffisant, pour répondre aux questions sur les modes d'acquisition de l'orthographe des homophones.

#### 4.3.3.1.2. Acquisition de la morphologie écrite des verbes

Le français dispose d'un système flexionnel riche caractérisé par un décalage entre la morphologie de l'oral et celle de l'écrit (Jaffré, 1998). Comme l'écrit du français

---

<sup>56</sup> Les chercheurs ont mené une étude sur l'apprentissage de correspondances entre phonèmes et phonogrammes auprès de trois groupes d'apprenants en début d'apprentissage du français L2 : des anglophones, des hispanophones et des sinophones. Soumis à une dictée de mots, les anglophones étaient les plus performants, tandis que les sinophones ont obtenu les scores les plus faibles. Selon les chercheurs, les performances différentes s'expliquent probablement par les effets de la distance L1-L2 sur les apprentissages en L2.

distingue des formes verbales qui sont homophones à l'oral, «la transcription par les correspondances phonographiques les plus fréquentes ne garantit pas la justesse morphographique » (Lefrançois, Chevrot, & Brissaud, 2006). Le souci historique de distinguer à l'écrit ce que l'oral confond (Andrieux-Reix & Pellat, 2006), dont nous avons parlé plus haut, confronte les scripteurs d'aujourd'hui à un problème crucial.

Les formes verbales homophones sont en outre un lieu de difficulté majeure. Si nous reprenons l'exemple de Brissaud et *al.* (2006), face à une forme phonologique comme [truve], le scripteur dispose de neuf formes graphiques fréquentes : *trouver*, *trouvé*, *trouvés*, *trouvée*, *trouvées*, *trouvais*, *trouvait*, *trouvaient*, *trouvez*. Pour produire la bonne forme verbale, le scripteur est confronté à un traitement morphologique des multiples facettes, qui appelle les oppositions temporelles (imparfait vs passé composé), modales (infinitif vs participe passé) et catégorielles (personnes, genre et nombre) (David, Guyon & Brissaud, 2006). Selon Brissaud et *al.* (2006), ce traitement morphologique nécessite l'opération d'une double sélection : celle d'un morphonogramme de mode-temps (-ai-, -é-, -er...) puis celle d'un ou deux morphogrammes sans correspondant phonique (-e, -s, -t, -ent) renvoyant aux catégories de la personne, du genre ou du nombre, si le morphonogramme choisi leur permet.

Devant cette opération compliquée au niveau orthographique, tout apprenant aura un long trajet à parcourir avant d'arriver à la maîtrise complète des formes verbales homophones. Depuis une dizaine d'années, l'homophonie verbale fait l'objet de travaux importants en français L1. Brissaud et ses collègues ont conduit une série d'études transversales sur la distinction graphique des homophones verbaux, ainsi que sur les traitements menés par les apprenants francophones (les études concernant les apprenants français : Brissaud, 1998 ; Brissaud & Sandon, 1999 ; Brissaud & Chevrot, 2000 ; les études concernant les apprenants français et québécois : Chevrot, Brissaud & Lefrançois, 2003 ; Lefrançois, Brissaud & Chevrot, 2006). Avec une attention particulière aux formes verbales en /E/, *Langue Française* consacre un numéro à ces aspects de l'orthographe du français, en conjuguant les efforts de champs différents : linguistique, psycholinguistique, neurolinguistique et didactique. Néanmoins, l'acquisition de la morphologie verbale écrite reçoit peu d'attention de la part des

chercheurs dans le domaine du français L2. À notre connaissance, il existe très peu de données en ce qui concerne les apprenants du français L2 et leur traitement de l'homophonie verbale à l'écrit.

A partir d'une lecture des travaux importants menés dans les deux domaines, le français L1 et L2, nous souhaitons éclairer les notions essentielles pour comprendre la réalisation et l'apprentissage de formes verbales homophones, ainsi que les différences entre les apprenants francophones et les apprenants allophones qui pourraient jouer un rôle dans le processus d'apprentissage de ce phénomène morphologique.

#### **4.3.3.1.2.1. L'acquisition de la morphologie verbale écrite en français L1**

En français, la transcription par les correspondances entre phonèmes et graphèmes ne garantit pas la justesse morphographique. Selon Brissaud et al. (2000), deux phénomènes empêchent ce passage de l'oral à l'écrit : d'abord, l'existence de morphogrammes muets ; ensuite, la polyvalence graphique, qui oblige le scripteur à faire une sélection parmi les phonogrammes concurrents. Rappelons que le dernier phénomène est vital dans le cas des homophones en /E/. Comme nous l'avons indiqué auparavant, face à une forme phonologique telle que [truve], le scripteur dispose au moins de neuf formes graphiques fréquentes. Ainsi, lors de l'acquisition de la morphologie écrite, les apprenants se trouvent-ils confrontés à de nombreuses difficultés, liées à l'opacité de la langue française. Pour comprendre ces difficultés des apprenants, deux séries de travaux importants ont été menées en français L1.

La première série concerne la morphologie écrite du nombre, à savoir les marques du pluriel nominal *-s*, et du pluriel verbal *-nt*. En reposant sur une étude fine des formes erronées produites par les apprenants français, les recherches ont fait apparaître que les performances de la morphologie écrite du nombre suivaient une évolution complexe (Fayol, 2003). Lors de la première phase, les enfants français de CP ne marquent le pluriel ni pour les noms, ni pour les verbes. Ils produisent le plus souvent les mots sous leur forme neutre (par exemple, le singulier pour le nom ; la forme de la 3PS pour le verbe). Cependant, leur performance en compréhension est plus précoce qu'en production. Autrement dit, pendant cette période, les enfants connaissent la marque du

pluriel pour le nom et savent également l'interpréter, mais la mise en œuvre de cette règle leur est problématique. Selon Fayol, ce fait pourrait s'expliquer par une capacité attentionnelle insuffisante des enfants de ce niveau. Lors de la production écrite, leur attention est principalement centrée d'une part, sur la détermination de l'orthographe du mot et d'autre part, sur la réalisation graphique. Comme la transcription graphique mobilise déjà la plus grande partie de leur attention, les enfants français ne sont plus en mesure d'appliquer l'accord nominal. Durant une deuxième phase (surtout en CE1 et CE2), les enfants sont capables de mettre en œuvre la marque pluriel *-s*, mais ils l'emploient à la fois pour les noms (correctement) et pour les verbes (erronément) (Fayol, 2003). Cette sur-généralisation du morphème *-s* (pluriel nominal) aux verbes s'explique probablement par le fait que les enfants de ce niveau ne différencient pas encore les catégories syntaxiques. Dans une troisième phase, les enfants savent utiliser la flexion *-nt* pour les verbes. Pourtant, les chercheurs ont observé une sur-généralisation de cette flexion *-nt* aux noms. Tout se passe comme si la notion du pluriel activait la compétition entre les deux marques du pluriel, *-s* et *-nt*. Et plusieurs facteurs pourraient jouer un rôle lors de la sélection de la flexion (Fayol, 2003) :

« L'ambiguïté lexicale, les noms et les verbes ayant des homophones respectivement verbaux ou nominaux (par exemple : *timbre/fouille*, etc.) sont plus sensibles aux effets d'interférence que les autres ; les fréquences relatives des items, les mots homophones sont plus souvent infléchis avec la marque du plus fréquent d'entre eux, nom ou verbe » (Fayol, 2003).

Étant donné que l'application d'une règle ne dépend pas de la fréquence des items (Fayol, 2003), les erreurs de substitution ne pourraient pas ainsi être dues au recours à la règle de l'accord en nombre. Selon Fayol (2003), il est fort probable que ces formes erronées s'expliquent par la récupération soit des associations entre un mot et une flexion (morphèmes *-s* ou *-nt*) soit des instances composées activées globalement, par exemple, *(ils) ferment* ou *(les) fermes*, lors de la réalisation graphique. Une telle opération est beaucoup plus économique en attention cognitive. De plus,

l'enseignement explicite de la grammaire fournit aux élèves les moyens de contrôle, avec lesquels les élèves sont capables de corriger les erreurs issues de l'activation automatique.

La deuxième série de travaux concerne les formes verbales homophones en /E/. L'objectif de ces recherches est d'examiner comment les élèves français s'y prennent pour sélectionner une graphie quand une dizaine de formes verbales est disponible. En observant les erreurs produites par les élèves de CM2 et les collégiens, Brissaud et *al.* mettent en lumière des tendances récurrentes dans la production des finales verbales en /E/ : tendance à accorder le verbe avec l'élément qui le précède, soit avec le sujet de la phrase, soit avec le mot interposé entre le sujet et le verbe ; tendance à utiliser la forme en *-é* au lieu de la forme en *-er* ; propension à ne pas accorder le participe passé employé avec l'auxiliaire avoir.

Notons que ces tendances aboutissent à des solutions qui ne rencontrent pas toujours la graphie normée (Brissaud & *al.*, 2006). Et certaines tendances s'interagissent en synergie, d'autres s'interagissent en contradiction. Afin de vérifier ces tendances repérées chez les élèves français, et d'en préciser leur évolution développementale, Brissaud et *al.* conduisent une autre étude auprès des élèves de l'école primaire en France et au Québec. A l'aide d'une épreuve se composant de courtes phrases lacunaires de structure « sujet-verbe-complément d'objet direct », les chercheurs tracent le profil développemental de l'utilisation des formes verbes en *-é* et en *-er* : au CE2, la forme en *-er* est majoritairement employée par les élèves. Selon Chevrot et *al.* (2003), les élèves de ce niveau de scolarité utilisent la forme en *-er* comme forme prototypique du verbe ; ensuite, le répertoire des formes verbales se diversifie ; quant aux formes en *-é*, si les élèves au CE2 les sélectionnent comme un des morphogrammes représentant le son /E/, les élèves de niveau plus élevé les sélectionnent puisque les formes en *-é* permettent de marquer l'accord. Plus précisément, la forme prototypique du verbe ne serait plus la forme en *-er* à ce moment-là. Et les élèves ont tendance à faire un choix morphographique en sélectionnant une forme variable, et la forme en *-é* est un bon candidat avec sa possibilité de marquage du nombre ou du genre.



#### **4.3.3.1.2.2. L'acquisition de la morphologie verbale écrite en français L2**

Confronté à la complexité au niveau morphologique, l'apprenant du français doit s'approprier un sous-système spécifique à l'écrit (Hedbor & Ågren, 2006). Néanmoins, la situation d'apprentissage de l'écrit chez les apprenants du français L2 diffère sur plusieurs points par rapport à celle des enfants français. Tout d'abord, au moment d'apprendre à écrire en français, les apprenants L2 sont plus âgés, et présentent ainsi un niveau plus élevé de la capacité cognitive. Deuxièmement, 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. Par conséquent, la procédure phono-centrée (Guyon, 2003) qui est une façon d'écrire presque uniquement centrée sur l'oral, observée dans les premiers écrits des enfants, semble absente chez les apprenants L2 (Hedbor & Ågren, 2006).

Depuis une décennie, les chercheurs dans le domaine du français L2 commencent à prendre en considération le français écrit, et plus spécialement l'acquisition de la morphologie. L'une des études significatives sur l'acquisition de la morphologie écrite en français L2 est celle d'Ågren (2008). En mettant l'accent sur l'accord en nombre du français écrit, la chercheuse décrit le développement morphologique observé dans le corpus écrit constitué par les apprenants suédois du français L2. Selon elle, il existe un profil développemental net entre les apprenants à un stade initial et ceux à un stade avancé. En ce qui concerne la morphologie verbale, son étude traite des formes de la 3PS et de la 3PP. Elle indique que, par rapport au français L2 parlé, la production écrite des apprenants démontre également une émergence relativement précoce de l'accord verbal morphologique (Ågren, 2005 : 140). Cependant, cet accord verbal connaît « un développement étendu dans le temps » chez les apprenants suédois (Ågren, 2005 : 275).

Granget (2005) conduit plusieurs études semblables chez les apprenants germanophones du français L2. En menant une analyse de l'accord sujet-verbe, la chercheuse relie le développement du marquage de l'accord verbal à l'enseignement en milieu scolaire germanique. A partir d'un corpus de récits écrits, elle constate un taux d'accord très élevé : 95% des verbes fléchis sont accordés avec leur sujet. Selon elle, cela s'explique probablement par un enseignement précoce et intensif des conjugaisons

verbales. De plus, un transfert entre langues est également soulevé : la notion de l'accord est présente en allemand et cela rend l'accord sujet-verbe en français accessible aux apprenants.

Dans le domaine du français L2, la morphologie de l'accord, et plus précisément de l'accord du nombre, fait l'objet de nombreux travaux, notamment chez les chercheurs scandinaves et germaniques. En revanche, en ce qui concerne l'homophonie verbale, il reste peu de données empiriques, et l'acquisition de cette morphologie a reçu peu d'attention de la part des chercheurs. Cependant, comme nous l'avons indiqué plus haut, l'apprenant L2 aborde simultanément le français oral et écrit. Ainsi nous semble-t-il extrêmement intéressant de savoir comment et quand les apprenants L2 mettent en œuvre les homophones verbaux. De surcroît, la plupart des études sur la morphologie de l'accord ont été menées auprès des apprenants scandinaves ou germaniques du français L2, dont la langue maternelle présente de constantes variations morphologiques. Par conséquent, nous souhaitons contribuer à une meilleure compréhension d'un groupe d'apprenants, jusqu'alors peu connu, qui apprennent à écrire en français dans une salle de classe en Chine, et qui possèdent une langue maternelle représentant les variations morphologiques minimales. Pour ce faire, nous nous concentrerons tout d'abord sur les formes verbales homophones en /E/, et examinerons par la suite les performances orthographiques d'autres verbes.

#### 4.3.3.1.3. Les formes verbales homophones en /E/

##### 4.3.3.1.3.1. Questions de recherche

Le questionnement principal de cette sous-partie est de savoir comment les apprenants chinois du français L2 mettent en œuvre les formes verbales homophones en /E/ dans la production écrite. Rappelons que la complexité de cette acquisition provient du fait de la double sélection qu'impliquent les finales, dont nous avons parlé plus haut. Il s'agit en effet des « multiples facettes morphologiques d'une même réalité phonique » (Sandon & Brissaud, 1999). Une série de difficultés importantes et hétérogènes à comprendre et à maîtriser le marquage de ces finales verbales en /E/ est constatée chez les élèves français (Brissaud & Chevrot, 2000) : les erreurs produites par les élèves

pourraient relever de l'extension de la marque du nombre nominal *-s* aux verbes, où les élèves étendent à la catégorie du verbe la marque typique de la catégorie du nom ; les erreurs pourraient être syntaxiques et les élèves généralisent le domaine d'application de l'accord du participe passé avec le sujet ; les formes erronées pourraient également être phonographiques, et les élèves substituent la finale en *-é* à *-ai-* ou *-er*, qui pourraient s'expliquer par une sélection du graphème le plus probable correspondant au phonème /E/.

Si les erreurs concernant les homophones en /E/ persistent quantitativement chez les élèves français jusqu'à un âge avancé, nous nous demanderons quel scénario développemental les apprenants chinois présentent dans leur production écrite. Notre travail consiste également à savoir si les difficultés ou tendances observées chez les élèves français sont présentées chez les apprenants chinois, et si à partir des formes erronées produites par les apprenants chinois, une certaine régularisation à l'écrit pourrait être trouvée.

#### 4.3.3.1.3.2. Résultats

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.

Le tableau 42 permet de visualiser l'évolution, du niveau initial au niveau assez avancé, du pourcentage d'erreurs quand la forme de l'infinitif ou celle du participe passé est respectivement attendue.

**Tableau 42.** Évolution du pourcentage de réussite en fonction du niveau des apprenants

	La cible <i>-er</i>			La cible <i>-é</i>		
	Occurrences totale	Formes erronées	Pourcentage de réussite	Occurrences totale	Formes erronées	Pourcentage de réussite
L1	116	10	91.38%	375	17	95.56%
L2	90	2	97.78%	89	8	91.01%

L3	96	1	98.96%	134	3	97.01%
Total	302	13	95.69%	598	28	95.31%

Pour la cible *-er*, nous constatons que les formes verbales sont précocement maîtrisées par les apprenants chinois, étant donné que, la plupart du temps, les apprenants sont capables de choisir les finales verbales congruentes. Déjà au stade initial, où les apprenants ont suivi une année d'études de français et représentent un niveau linguistique élémentaire, le pourcentage de réussite est de plus de 90%. La figure nous permet également de saisir l'amélioration continue des performances des apprenants en fonction de leur niveau de langue. Pour les apprenants de L4, le pourcentage de réussite est de presque de 99%.

Quant à la cible *-é*, l'image des performances des apprenants est presque identique à celle qui se dégage du corpus de la cible *-er*. Nous constatons une maîtrise précoce au stade initial. Le pourcentage de réussite atteint jusqu'à 97% pour les apprenants de L4, bien qu'il diminue légèrement en L3.

Les tableaux 43 et 44 permettent de suivre en détail l'évolution en fonction de niveau de langue des différentes formes erronées produites par les apprenants. L'analyse des types d'erreurs, conduite à partir de ces deux tableaux, permet de faire trois remarques :

- Les finales verbales sont choisies dans le respect de la morphographie (Sandon & Brissaud, 1999). Nous constatons que les morphogrammes choisis par les apprenants, même erronés, sont toujours des morphogrammes plausibles en finale verbale et les morphogrammes finaux complémentaires, marques de personne ou de nombre, ne sont jamais l'objet d'une erreur d'association.
- Il nous semble que les apprenants sont confrontés à plus de difficultés lorsque la cible est *-é*. Si l'on compare les erreurs commises de deux catégories, le nombre de graphies erronées différentes est plus limité quand la cible est *-er*. De plus, lors de la sélection du morphogramme, les apprenants ont commis plus d'erreurs de

substitution (substituer *-é* à *-ai-*, *-er*, *-ez*) quand la cible est *-é*, et ce type d'erreurs persiste jusqu'à un niveau assez avancé.

- Quant au problème de l'accord, les omissions de marques du nombre ou du genre sont plus nombreuses que les adjonctions. L'erreur consistant à insérer une marque du nombre ou du genre dans le cas où la langue ne requiert aucun accord est en effet très rare dans notre corpus. Nous n'avons observé que deux occurrences : \* *La guerre a commencée* ; \* *Sa mère lui a élevée*.

**Tableau 43.** Formes erronées produites par les apprenants de trois niveaux quand la cible est *-er*

Nb	La cible <i>-er</i>			
	<i>-é</i>	<i>-ez</i>	autre	total
L2	3	3	4	10
L3	0	2	0	2
L4	0	1	0	1

**Tableau 44.** Formes erronées produites par les apprenants de trois niveaux quand la cible est *-é*

(Légende : *-ai-* = les formes verbales de l'imparfait ; 3PS = les formes verbales de la 3<sup>e</sup> personne du singulier ; omission = omission des marques du nombre ou du genre ; adjonction = adjonction des marques du nombre ou du genre)

Nb	La cible <i>-é</i>							
	<i>-er</i>	<i>-ez</i>	<i>-ai-</i>	3PS	omission	adjonction	autre	total
L2	5	1	3	6	1	1	0	17
L3	1	0	0	1	5	1	0	8
L4	2	0	0	1	0	0	1	4

Nous observons donc que les règles de sélection et d'association morphographiques semblent bien maîtrisées chez les apprenants chinois. Dès le stade initial, les apprenants ne sélectionnent que des morphogrammes de finales verbales et les morphogrammes qui leur sont associés. Les associations impossibles telles que -

*ers*, *-éent* ne sont pas trouvées dans notre corpus. Cela est conforme aux résultats obtenus par les collégiens francophones. Selon Sandon et Brissaud (1999), « les règles de sélection et d'association morphologiques semblent bien maîtrisées au collège ». En outre, les tendances lourdes observées chez les élèves et les collégiens français ne se dessinent pas nettement chez les apprenants chinois. A partir de CE2, la forme en *-é* devient la forme prototypique du verbe chez les élèves français, et les élèves ont également une forte propension à marquer cette forme en genre et en nombre. Néanmoins, les erreurs de substitution sont très hétérogènes chez les apprenants chinois, notamment chez les apprenants du niveau initial. Et il n'existe pas de forme prototypique du verbe. De plus, si les élèves français ont tendance à accorder la forme verbale avec le sujet de la phrase, ou avec l'élément qui précède le verbe, ce souci de faire l'accord et de garder la congruence syntaxique n'est pas observé dans notre corpus. La difficulté des apprenants chinois consiste davantage à marquer le nombre ou le genre dans les cas où la langue requerrait un accord.

#### **4.3.3.1.3.3. Analyse des erreurs**

La première analyse des données que nous avons faite dans la partie précédente a permis de décrire le scénario développemental des apprenants chinois face aux homophones en /E/, ainsi que les tendances générales se dessinant dans la production écrite des formes verbales. L'analyse détaillée des graphies erronées dans cette partie permet d'envisager la lecture d'erreurs comme celle de la trace de l'activité cognitive des apprenants.

##### 4.3.3.1.3.3.1. Erreurs non-spécifiques

L'analyse des erreurs nous montre que la plupart des formes erronées relèvent d'une procédure de substitution (69% pour les erreurs à l'infinitif ; 68% pour les erreurs au participe passé). En analysant ce type d'erreur, nous examinerons respectivement les performances des trois groupes d'apprenants.

Dans les écrits des apprenants de L2, nous avons identifié 12 formes erronées de cette catégorie.

L'infinifitif des verbes est attendu dans les phrases suivantes, alors que les scripteurs ont choisi la forme verbale de la 2PP du présent.

- (80) a. \* La première c'était que {honorez}\_ {honorer} les dames. (L2-01-008)  
b. \* {donnez}\_ {donner} honneur en ce siècle et bonne fin. (L2-02-006)  
c. \* priez Dieu de le {donnez}\_ {donner} honneur en ce siècle et bonne fin.  
(L2-02-015)

Le même phénomène est observé dans les phrases suivantes. Après les verbes auxiliaires *aller* et *faire*, ou la préposition *de*, la forme infinitive des verbes est attendue. Mais ici les scripteurs ont choisi le participe passé, qui sert en règle générale à construire le temps passé composé.

- (81) a. \* il a dit {que\_il}\_ {qu\_QUOTE\_il} allait {demandé}\_ {demander} ses armes. (L2-01-008)  
b. \* Ça fait {étonné}\_ {étonner} à Keu. (L2-01-011) (erreur lexicale ?)  
c. \* il a décidé de partir et {allé}\_ {aller} à la cour {de\_la}\_ {du} roi Arthur.  
(L2-01-014)

Les apprenants ont également produit les erreurs de substitution dans le contexte où la forme verbale du participe passé est demandée. Les formes erronées ci-dessous ont été produites pour construire le temps passé composé, qui est formé par le présent du verbe *être* ou *avoir* associé au rétro-participe. Mais lors de la production, les scripteurs ont offert une combinaison impossible entre le présent de l'auxiliaire et la forme infinitive du verbe concernant pour signaler le temps composé.

- (82) a. \* Le roi Arthur s'est {fâcher}\_ {fâché} de sa grossièreté, en {réclimant}\_ {réclamant} de +R[avec]+R+//s\_QUOTE\_il\_y\_a//+R assez de temps Perceval +R[allais]+R+//allait//+R devenir un bon chevalier avec des bonnes manières. (L2-01-023)

- b. \* il a {laisser}\_ {laissé} la coupe d'or à un chevalier en dehors et lui ~a~ dit de +R[transféré]+R+//transférer//+R sa reconnaissance {à\_la}\_ {au} roi Arthur. (L2-01-023)
- c. \* Après avoir {communiquer}\_ {communiqué} avec un chevalier à la porte, [μxxxμ] il a {ve}\_ {vu} le roi perceval. (L2-02-002)
- d. \* mais ella a {laisser}\_ {laissé} son fils suivre [son] {sa}\_ {son} rêve. (L2-02-012)
- e. \* Il a dit {au}\_ {aux} chevaliers et ~les~ a {demander}\_ {demandé} {leur}\_ {leurs} armes. (L2-02-010)

Le scripteur dans l'exemple (83) a présenté une autre combinaison impossible entre le présent du verbe auxiliaire *avoir* et la forme de la 2PP du présent du verbe *participer* pour signaler le temps passé composé.

- (83) \* La première fois qu'il a {participez}\_ {participé} à la cour du roi Arthur, Perceval était un jeune homme, grand, beau de visage, mais à demi sauvage. (L2-01-015)

Notons que la maîtrise des formes verbales homophones en /E/ se développe assez rapidement chez les apprenants chinois. En licence 3, seulement 4 formes erronées ont été trouvées.

Dans l'exemple (84), les deux scripteurs ont choisi les formes d'infinitif, qui ne correspondaient pas aux contraintes syntaxiques :

- (84) a. \* Perceval dit : {Donner}\_ {Donnez} moi armes, car je veux être le chevalier Vermeil. (L3-006)
- b. \* Le sénéchal, frère de lait du roi Arthur, jaloux de roi et s'est {moquer}\_ {moqué} de roi. (L3-022)

Les scripteurs suivants ont produit les formes verbales d'impératif, tandis que



celles d’infinitif étaient attendues :

- (85) \* Sa mère pleure et lui dit {de\_honorez}\_ {d\_QUOTE\_honorer} les dames et  
~~de~~ elle lui +R[demanda]+R+//donna//+R [3] trois  
{recommandations}\_ {recommandations} : Honorez les dames ; suivez les  
conseils des prud’hommes et priez Dieu de vous {donnez}\_ {donner} bonneur.  
(L3-013)

Pour les apprenants de la Licence 4, nous n’avons trouvé que deux formes erronées relevant de la substitution. Cela montre que chez ces apprenants, les formes verbales homophones en /E/ est bien maîtrisées par les apprenants de ce niveau. Les formes fournies par les scripteurs suivants sont plausibles au point de vue phonologique. Cependant, elles ne correspondent pas à leur contexte syntaxique.

- (86) a. \* Dans son {innocent}\_ {innocence}, Perceval s’est {déclarer}\_ {déclaré}  
aller demander au roi ces armes qu’il avait rêvées et [les\_ob] s’est montré sûr  
qu’il les obtiendrait. (L4-023)
- b. \* priez Dieu de vous {donnez}\_ {donner} honneur en ce siècle et [le] bonne  
fin. (L4-024)

En résumant, nous constatons que la maîtrise des formes verbales homophones en /E/ se développe assez rapidement, et que le nombre d’erreurs diminue significativement selon le niveau de langue des apprenants. Parmi les erreurs de cette catégorie, nous avons trouvé que 6 erreurs étaient issues du même extrait du texte de ressource :

*Elle soupira et reprit : « Écoutez toutefois. Voici mes trois recommandations : honorez les dames, suivez les conseils des prud’hommes1 et priez Dieu de vous donner honneur en ce siècle et bonne fin. »*

Les 6 formes erronées sont suivantes :

- (87) a. \* La première c'était que {honorez}\_ {honorer} les dames. (L2-01-008)  
b. \* {donnez}\_ {donner} honneur en ce siècle et bonne fin. (L2-02-006)  
c. \* priez Dieu de le {donnez}\_ {donner} honneur en ce siècle et bonne fin.  
(L2-02-015)  
d. \* Sa mère pleure et lui dit {de\_honorez}\_ {d\_QUOTE\_honorer} les dames.  
(L3-013)  
e. \* Honorez les dames ; suivez les conseils des prud'hommes et priez Dieu de  
vous {donnez}\_ {donner} bonneur. (L3-013)  
f. \* priez Dieu de vous {donnez}\_ {donner} honneur en ce siècle et [le] bonne  
fin. (L4-024)

Ainsi une hypothèse autre que l'acquisition de l'inflexion verbale pourrait être donnée pour expliquer les 6 erreurs concernées. Lors de la reconstitution du conte, les apprenants ont transformé le discours direct ci-dessus en discours indirect, en faisant référence aux constructions, telles que *dire à quelqu'un de faire quelque chose*. Et ces constructions demandent cependant les formes verbales d'infinitif. Il nous semble que, lors du passage du discours direct au discours indirect, les apprenants ne disposaient pas de capacité attentionnelle suffisante et ont copié directement les formes verbales à l'impératif du texte de source. Autrement dit, les apprenants étaient en mesure de choisir de bonnes marques d'accord en fonction du contexte. Mais le coût de la transformation du discours direct en discours indirect était très élevé, et le problème se posait finalement au niveau de la réalisation morphologique.

Remarquons que les erreurs de substitution sont également fréquentes chez les élèves français, et à partir des analyses de graphies erronées et de commentaires métagraphiques des élèves, les chercheurs ont établi un scénario évolutionnaire (Sandon & Brissaud, 1999) : au CE2, la forme en *-er* est majoritairement employée par les élèves. Cela s'explique probablement par le fait que la forme en *-er* est utilisée

comme la forme prototypique du verbe. Ainsi la sélection de la forme en *-er* montre-t-elle que, lors de la production écrite, l'élève a reconnu l'élément comme relevant de la classe du verbe. Ensuite le remplacement se fait plus nettement en faveur de la marque du participe qu'en faveur de la marque de l'infinitif. Cette sélection pourrait s'établir sur « des critères de probabilité phonographique » (Brissaud & Chevrot, 2000), en tenant compte du fait que le graphème *-é* est l'option la plus probable correspondant à /E/. Au collège, la sélection de *-é* est davantage un choix morphographique qu'un choix phonographique. A ce moment-là, le verbe est perçu comme un élément variable. Après avoir compris la solidarité syntagmatique existant en français entre le verbe et son sujet, les élèves français sont amenés à sélectionner la forme *-é* avec sa possibilité de marquer le nombre et le genre.

Néanmoins, ce scénario développemental n'est pas remarqué dans notre corpus. La cause principale de ces différences serait sans doute qu'en entrant en langue française, les apprenants chinois atteignent déjà un niveau assez élevé dans leurs capacités cognitives. Un enseignement précoce et très explicite de la conjugaison verbale permet aux apprenants chinois de maîtriser les règles et les catégories grammaticales abstraites. Et après un entraînement intensif, les apprenants sont capables de faire une analyse morphosyntaxique et de sélectionner la finale verbale congruente lors d'une production écrite. Cependant, les erreurs surgissent quand l'apprenant ne dispose pas d'une capacité attentionnelle suffisante pour gérer l'application des règles. Avec le souci de correspondance phonie-graphie, l'apprenant choisit ainsi un des graphèmes correspondant au phonème /E/, qui conduit probablement à une forme erronée. Notons également que le nombre d'erreurs de substitution diminue en fonction du niveau de langue. L'hypothèse avancée est qu'avec un niveau de langue plus élevé, la mobilisation et la mise en œuvre des règles deviennent plus faciles pour les apprenants.

#### 4.3.3.1.3.3.2. Erreurs d'utilisateur du dictionnaire

Contrairement à ce qui se passe pour les élèves français, les erreurs relevant de la substitution de *-er* à *-é* semblent plus nombreuses que celles de substitution de *-é* à *-er*

chez les apprenants chinois. Remarquons qu'une autre hypothèse est possible pour expliquer ce type d'erreur, mise à part la confusion issue des différents graphèmes correspondant à /E/.

Jiang (2000) propose d'envisager le traitement morphologique en L1 et en L2 sous deux modèles différents. En L1, avant la réalisation morphologique, les informations morphologiques sont déjà codées dans le message préverbal et la sélection de la flexion est sous le contrôle direct des éléments du message. Il s'agit d'un processus automatique de mise en correspondance entre le contenu du message à produire et les informations spécifiques concernant la morphologie associée à un item lexical. Jiang indique que le choix morphologique fait partie de l'encodage grammatical et que la flexion forme une partie intrinsèque du système syntaxique. En d'autres termes, lorsqu'une structure syntaxique est donnée, les informations morphologiques concernées sont activées. Notons que les associations possibles entre un item lexical et les affixes relevant de sa morphologie flexionnelle sont présentées dans une même entrée (Levelt, 1989). Une fois cette entrée lexicale ouverte, toutes les variations morphologiques deviennent automatiquement disponibles pour le scripteur.

Ainsi pouvons-nous envisager un processus en une seule étape lors de la production morphologique en L1. Si le scripteur veut produire le message *They left*, deux entrées lexicales sont ouvertes, l'une d'entre elles est celle du verbe *leave*, avec toutes les formes flexionnelles acceptables comme *leave*, *leaves*, *left*, *leaving*. Parmi ces formes, *left* est la forme la plus intensément activée, et par la suite choisie par le scripteur puisqu'elle se conforme le mieux au message pré-verbal : l'action de quitter qui se produit dans le passé. Il s'agit d'un processus en une seule étape, étant donné que la forme flexionnelle est choisie directement sans médiation de la racine *leave* ou analyse linguistique consciente.

Il n'est pas superflu d'indiquer qu'un tel processus automatique nécessite une forte intégration des informations morphologiques dans l'entrée lexicale. Autrement dit, différentes flexions d'un item lexical doivent s'intégrer dans son entrée lexicale et ainsi deviennent-elles automatiquement disponibles lors de la production. Ce processus automatique du traitement morphologique fait partie intégrante de la compétence

lexicale. Cependant, dans l'apprentissage L2, au moins pour les apprenants chinois de notre corpus, qui apprennent le français en s'appuyant principalement sur l'enseignement en classe, les informations morphologiques ne sont pas intégrées dans les entrées lexicales, mais sont apprises et stockées hors du système de l'entrée lexicale, en tant que connaissances explicites. Cette différence concernant la représentation lexicale aurait des conséquences significatives sur le processus du traitement morphologique. Selon Jiang (2000), la sélection de la flexion pertinente devient ainsi un processus conscient et strictement séquentiel, en impliquant deux étapes nécessaires : la sélection de la racine liée à l'item lexical à produire, et la sélection de la flexion sur la base des connaissances morphologiques du scripteur.

Si nous reprenons l'exemple du verbe *leave*, ses variations morphologiques telles que *leave*, *leaves*, *left*, *leaving* sont conservées respectivement dans une entrée séparée. Toutefois, il est concevable d'imaginer qu'il existe des liens entre ces entrées, établis pendant le processus d'acquisition du verbe. Pour produire *They left*, l'entrée de l'item *leave* serait ouverte, cependant elle ne comporte que la racine, *leave*. La deuxième étape implique donc une sélection d'un affixe, pertinent du point de vue morphologique, selon les connaissances de l'apprenant.

Une telle production linguistique strictement séquentielle est susceptible d'être un processus conscient qui nécessite des ressources mentales. Lorsque ce travail de sélection entre en conflit avec d'autres activités qui consomment aussi de l'énergie mentale, des erreurs morphologiques se produisent. Pour expliquer les erreurs de substitution de *-er* à *-é*, notre hypothèse est que lors de la production écrite, le traitement morphologique séquentiel entre en conflit avec d'autres activités cognitives. Par conséquent, l'apprenant est incapable de faire une analyse morphosyntaxique suffisante, et choisit ainsi la forme verbale de l'infinitif, forme de lemme. Dans notre corpus, nous avons également trouvé l'occurrence suivante, provenant de l'inaccomplissement du traitement morphologique :

(88) \* Elle pleure et {laisser}\_ {laisse} partir Perceval. (L3-025)

Jiang ajoute également que l'enseignement de L2 se concentre particulièrement sur la racine lorsqu'un nouveau mot est introduit. Le nouveau vocabulaire est souvent présenté sous forme de liste, et comme entrée de dictionnaire. Pour l'apprenant de L2, ce qui est présenté dans une entrée lexicale, c'est donc souvent la racine, avec les liens à ses variations flexionnelles. Ces liens seraient probablement directionnels, de la racine à ses flexions, plutôt que l'inverse. Comme nous en avons déjà parlé, le traitement morphologique séquentiel commence par la sélection de la racine puis la sélection de la flexion. Cela signifie que les erreurs morphologiques auraient plus de chances d'être produites lorsque le contexte linguistique demande une flexion.

#### 4.3.3.1.3.3. Erreurs de sur-généralisation

Comme nous l'avons indiqué plus haut, la sur-généralisation de l'accord avec le sujet est une tendance récurrente chez les élèves français, mise au jour dans plusieurs études (Voir notamment Brissaud & Chevrot, 2000 ; Totereau & Brissaud, 2006). Les chercheurs indiquent que la propension à sélectionner une forme verbale en -é accordée en nombre ou en genre avec le sujet augmente de la 3<sup>e</sup> à la 4<sup>e</sup> année de primaire, et se stabilise puis diminue entre la 3<sup>e</sup> et la 4<sup>e</sup> année du secondaire (Brissaud, Chevrot & Lefrançois, 2006). Rappelons également que la substitution de la flexion -é munie des marques -s ou -e à la forme -er est probablement issue de la propension à réaliser l'accord du verbe avec le sujet de la phrase, comme si la phonographie était analysée à nouveau à la lumière de la morphonographie (Chevrot & al., 2003). Afin de comprendre le mécanisme derrière cette généralisation de l'accord, Brissaud et al.(2006) nous présentent plusieurs possibilités : soit l'accord provient de l'application délibérée d'une procédure combinatoire, comme ce que disent les enfants dans les entretiens métagraphiques : « ils sont plusieurs c'est -és » (Brissaud & Cogis, 2002) ; Soit le groupe nominal jouant la fonction du sujet tel que *les garçons* active la forme -és ; La troisième possibilité consiste à la mise en place d'un schéma général de type *pluriel \_pluriel*, à force de rencontre de *ils \_és*, *ils \_nt*, *les \_s* dans l'environnement écrit. Cependant, dans l'état actuel, les données empiriques ne permettent pas aux chercheurs de trancher entre ces hypothèses.

Dans notre corpus, les graphies erronées relevant de la sur-généralisation de l'accord sont très rares, et nous n'avons trouvé que deux occurrences (pour les apprenants de L1, une forme erronée pour 371 occurrences ; pour les apprenants de L2, une forme erronée pour 77 occurrences), où les apprenants ont généralisé à l'auxiliaire *avoir* la règle d'accord du genre qu'ils maîtrisaient pour *être* :

- (89) a. \* La guerre a {commencée}\_ {commencé}. (L1-16)  
b. \* Sa mère lui a {élevée}\_ {élevé}. (L2-09)

Selon nous, la cause principale de ces différences serait que lors du traitement morphologique, la solidarité syntagmatique existant en français entre le verbe et son sujet n'est pas prise en compte chez les apprenants chinois. Notons qu'en L1, les informations sémantiques contribuent, elles aussi, à la réalisation morphologique (Largy, Fayol & Lemaire, 1996). Avant la réalisation morphologique, les informations sémantiques du sujet, telles que le nombre du pluriel, sont déjà codées dans le message préverbal. Après avoir compris la solidarité syntagmatique entre le verbe et son sujet, l'élève français éprouve le besoin de faire varier la forme verbale selon les informations sémantiques du sujet lors de la sélection de la flexion morphologique. Néanmoins, il est peu probable que les informations sémantiques du sujet comme celle du nombre soient codées dans le message préverbal de l'apprenant chinois, étant donné que la morphologie du nombre n'existe pas en langue chinoise (Jiang, 2007). Ainsi l'apprenant chinois met-il en œuvre l'accord sujet-verbe en suivant strictement les règles enseignées explicitement : accord pluriel de l'imparfait, du participe passé avec *être* et du participe passé avec *avoir* (accord avec objet antéposé). Quand l'apprenant est mentalement surchargé et n'est plus en mesure de gérer l'application de ces règles, les erreurs surgissent et elles portent essentiellement sur l'omission de marques d'accord.

En outre, il est important de remarquer que les apprenants de L2 ne s'imprègnent pas suffisamment des contextes langagiers, notamment les apprenants dépendant principalement de l'enseignement en classe. Étant donné cette pauvreté de l'input en

termes de quantité et de qualité, le schéma général de type *pluriel \_pluriel*, dont nous venons de parler, a peu de chance d'être mis en place chez les apprenants chinois.

#### **4.3.3.1.3.4. Discussions**

L'objectif de cette sous-partie était de décrire le développement de la production des formes verbales homophones en /E/ chez les apprenants chinois du français L2. Malgré le caractère préliminaire de notre étude, un certain nombre de tendances intéressantes a pu être dégagé. Après un an d'études du français, les apprenants chinois prouvent une maîtrise précoce des formes verbales en /E/. Le taux de réussite est de plus de 90% pour la cible *-er* et la cible *-é*. Et le nombre d'erreurs diminue en fonction du niveau de langue. L'hypothèse la plus probable susceptible d'expliquer ce fait est que les apprenants chinois sont plus âgés et disposent d'une capacité cognitive assez avancée. Contrairement aux enfants français, qui cherchent à faire des tentatives de catégorisation lors de leur construction du système écrit, les apprenants chinois ont déjà construit le système écrit et peuvent transférer leurs connaissances linguistiques préalables en étudiant une langue étrangère. A l'aide d'un enseignement explicite et d'un entraînement intensif, ils sont en mesure de faire une analyse morphosyntaxique suffisante lors de la sélection des formes verbales, et puis de gérer l'application des règles de conjugaison pour les verbes du premier groupe. Néanmoins, les erreurs persistent dans la production des apprenants, étant donné que l'application des règles entre en conflit avec d'autres activités mentales, mobilisant la plus grande partie de l'attention des apprenants.

Quant aux graphies erronées commises par les apprenants, elles sont hétérogènes, et les explications sont multiples. Soit l'erreur résulte d'une analyse morphosyntaxique insuffisante et l'apprenant fait la sélection parmi toutes les morphogrammes correspondant à /E/ (notons que sans une maîtrise préalable du français oral, les apprenants pourraient choisir une finale altérant l'image phonétique de la forme verbale : erreur de substitution de la forme 3PS à la cible *-é*). Soit l'erreur s'explique par un accomplissement du traitement morphologique séquentiel, spécifique chez les



apprenants de L2. Cela conduit les apprenants à choisir la cible *-er*, forme de l'infinitif, tandis que le contexte linguistique demande une forme verbale conjuguée. Étant donné les caractéristiques de ce traitement morphologique conscient, l'évolution de compétition entre la cible *-er* et *-é*, ainsi que la propension à faire l'accord sujet-verbe, observées chez les élèves français, ne se présentent pas dans la production des apprenants chinois. En mettant l'analyse des erreurs en relation avec la situation d'acquisition de la morphologie des apprenants, nous constatons que les multiples éléments doivent être pris en compte afin de comprendre les formes verbales en /E/ produites par un groupe d'apprenants de L2, et qu'une erreur peut en cacher une autre lorsque les apprenants ayant une langue maternelle lointaine essaient de s'intégrer au système morphologique du français.

#### 4.3.3.1.4. D'autres formes verbales homophones

Dans notre corpus, les erreurs relevant de la morphologie verbale sans altérer la valeur phonique concernent également les verbes autres que ceux du premier groupe. Dans cette sous-partie, les formes erronées sont regroupées sous deux ensembles en fonction de leur nature : le premier concerne les lettres finales muettes ; le second relève d'un non-accord sujet-verbe. Nous présenterons par la suite les analyses quantitative et qualitative de ces graphies erronées.

##### 4.3.3.1.4.1. Lettres finales muettes

Les erreurs de cette catégorie concernent l'adjonction, l'omission ou un mauvais choix de lettres finales muettes. Dans les écrits produits par les apprenants de L2, nous avons trouvé 7 formes erronées, qui représentent 22% du nombre total d'erreurs (22) concernant les terminaisons verbales sans altérer la valeur phonique.

Les erreurs dans l'exemple (90) sont portées sur l'ajout d'une lettre finale supplémentaire *t*. Il nous apparaît que ces adjonctions ne se font nullement au hasard et que deux explications possibles pourraient être données ici : soit ces lettres supplémentaires consistent à transmettre au participe passé les marques du sujet de la 3PS qui précède le verbe conjugué ; soit elles proviennent d'un effet d'analogie. Plus

précisément, elles pourraient être influencées par la forme verbale d'indicatif de la 3PS, *elle sourit (elle rit)*, qui a une fréquence importante dans la lecture des apprenants chinois.

- (90) a. \* la femme n'avait pas {sourit}\_ {sourit} pendant longtemps. (L2-01-001)  
b. \* elle a {rit}\_ {ri} à nouveau. (L2-01-006)  
c. \* Le roi était {surprise}\_ {surpris} que cette fille n'avait ~pas~ {rit}\_ {ri} avant. (L2-01-017)

Également, l'adjonction de la lettre finale *s* dans les phrases suivantes provient, selon nous, d'un effet d'analogie. En ajoutant un *s* final, les scripteurs ont pris probablement la forme d'indicatif de la 1PS du verbe *rire*.

- (91) a. \* Perceval a vu une belle fille sur la table et la fille a {ris}\_ {ri} pour lui. (L2-02-012)  
b. \* Elle n'a pas {ris}\_ {ri} depuis elle était 6 ans. (L2-02-012)

Il nous semble que la forme erronée suivante vient également d'un effet d'analogie. Le scripteur a confondu probablement le participe passé du verbe *promettre* avec soit la forme d'indicatif de la 3<sup>e</sup> personne du verbe *permettre*, soit le participe passé des verbes en *-ire*, par exemple, *dire/dit*.

- (92) \* Keu, le prud'homme, lui a {promit}\_ {promis} [que] qu'il peut avoir tout. (L2-02-004)

En oubliant la terminaison *e*, le scripteur a produit une forme verbale impossible, qui est probablement contaminée par la forme nominal *cri*.

- (93) \* il {cri}\_ {crie} à Perceval. (L2-02-015)

Dans cette catégorie, les apprenants de L3 ont produit 12 erreurs, ce qui représente 54% du nombre total d'erreurs (22).

Pour les formes erronées ci-dessous, il s'agit de l'oubli de la lettre finale *-t*, qui sert à indiquer la 3PS du passé simple, pour les verbes du 2<sup>ème</sup> groupe et du 3<sup>ème</sup>. Quant aux verbes *avoir*, *vouloir*, et *rire*, les formes conjuguées produites par les scripteurs, *eu*, *voulu*, et *ri*, proviennent sans doute d'un effet d'analogie. Elles pourraient être contaminées par la forme verbale de leur participe passé, qui a une fréquence importante dans la lecture des scripteurs.

- (94) a. \* Pour {realiser}\_ {réaliser} son rêve, il [doi] {eu}\_ {eut} besoin d'avoir un combat avec l'ennemi du roi (L3-003)
- b. \* Devant la troupe de {chevalier}\_ {chevaliers}, il a dit qu'il {voulu}\_ {voulut} aller demander les armes au roi. (L3-011)
- c. \* il {obtin}\_ {obtint} le respect du roi Arthur quand il revint dans [d] la cour. (L3-013)
- d. \* Il lui dit ce qu'il aurait dit aux chevaliers, Keu, sénéchal, frère de lait du roi {ri}\_ {rit} de lui. (L3-015)
- e. \* Elle {ri}\_ {rit} et elle l'admira. (L3-015)

Les scripteurs dans l'exemple (95) ont omis de prendre la terminaison congruente pour construire le temps passé simple, et par conséquent les formes conjuguées ne s'accordent pas avec leur sujet, qui est à la 3PS. Selon nous, ces formes erronées proviennent également d'un effet d'analogie : afin de construire le temps passé simple, les scripteurs ont pris directement la forme du participe passé des verbes *comprendre* et *reprendre*.

- (95) a. \* Perceval ne {compris}\_ {comprit} pas il moquait de lui. (L3-006)
- b. \* Sa mère soupira et {repris}\_ {reprit} trois recommandations. (L3-021)

Dans l'exemple (96), les scripteurs ont essayé de former la conjugaison du

subjonctif, tandis que les formes verbales ont été mal construites. Pour le scripteur L3-011, il s'agit d'un oubli de la lettre finale muette, *e*. Quant aux scripteurs L3-013 et L3-018, ils ont choisi la lettre *-t*, qui est souvent utilisée après le sujet de la 3PS.

- (96) a. \* La mère de Perceval ne l'a pas dit les choses sur le chevalier, parce qu'elle était peur que {sa}\_ {son} fils peut-être [meure] {meur}\_ {meure} comme son père et ses deux frères. (L3-011)
- b. \* Elle avait peur qu'il {meurt}\_ {meure} dans un combat comme son père et ses ~deux~ frères. (L3-013)
- c. \* Parce que sa [pe] mère a peur qu'il {+R[mort]+R+//meurt//+R}\_ {meure} au combat comme son père et ses deux frères. (L3-018)

Les erreurs dans l'exemple (97) sont portées sur l'ajout d'une lettre finale supplémentaire *-t*. Il nous apparaît que ces adjonctions ne se font nullement au hasard mais qu'elles sont issues d'un effet d'analogie. Deux explications pourraient être données : soit les graphies erronées consistent à transmettre au participe passé les marques du sujet de la 3PS qui précède le verbe conjugué ; soit les apprenants ont pris directement la forme verbale du passé simple, le temps qu'ils ont appris récemment.

- (97) a. \* Il a demandé au garde de la cour de donner la coupe d'or {à}\_ {au} roi et dire à cette belle fille qui n'avait pas {rit}\_ {ri} depuis de 6 ans mais ~qui~ riait pour lui qu'il {reviendrai}\_ {reviendrait}. (L3-010)
- b. \* Mais il avait [voudrait] {voulut}\_ {voulu} un soldat, un chevalier de roi Arthur. (L3-014)

Pour les apprenants de L4, 17 erreurs ont été identifiées dans cette catégorie, représentant 48.5% du nombre total d'erreurs (35).

Les erreurs ci-dessous sont portées sur l'ajout d'une lettre finale supplémentaire *-t*. Notons que les formes conjuguées du verbe *rire* (*sourire*) sont mal orthographiées par de nombreux scripteurs. La forme erronée \* *a rit* (*sourit*) provient sans doute de la

forme du présent (*il/elle*) *rit*, qui a une fréquence assez importante dans la lecture des scripteurs.

- (98) a. \* Le fou a {rit}\_ {ri}. (L4-002)
- b. \* La voix ~de\_ discussion~ augmenta parce que la fille n'avait pas {rit}\_ {ri} depuis 6 ans. (L4-002)
- c. \* Perceval embrassa sa mère et il est {partit}\_ {parti}. (L4-007)
- d. \* le roi ~Arthur~ s'est {sentit}\_ {senti} très dommage, et il {crois}\_ {croit} que Perceval {devenirait}\_ {deviendrait} un bon chevalier s'il [μxxxμ] reste. (L4-007)
- e. \* Et elle a rit, elle n'a pas {rit}\_ {ri} depuis 6 ans. (L4-012)
- f. \* Quand il a parlé au roi, tous les gens ont {rit}\_ {ri}. (L4-021)
- g. \* Il ~a~ {aperçut}\_ {aperçu} une belle fille qui a {sourit}\_ {sourit} aussi. (L4-021)
- h. \* Il a dit qu'elle n'avait pas {sourit}\_ {sourit} depuis +R[6]+R+//six//+R ans. (L4-021)
- i. \* Mais aujourd'hui elle a {sourit}\_ {sourit}. (L4-021)
- j. \* Elle a {sourit}\_ {sourit} au Perceval. (L4-021)
- k. \* Il a {aperçut}\_ {aperçu} en sortir un chevalier bien {équipé}\_ {équipé}. (L4-023)
- l. \* Perceval a [ap] {aperçut}\_ {aperçu} une belle fille [μxxxμ] à la table qui n'a +R[pas]+R+//jamais//+R {rit}\_ {ri} depuis six ans, mais quand cette fille s'est senti les regards de Perceval, elle a {rit}\_ {ri} de nouveau. (L4-024)

Il nous semble que cette forme erronée vient également d'un effet d'analogie. Le scripteur dans l'exemple (99) a confondu probablement le participe passé du verbe *promettre* avec la forme du passé simple de la troisième personne du singulier, *promit*, ou le participe passé des verbes en *-ire*, par exemple, *dire*.

- (99) \* Arthur a {promit}\_ {promis} de lui donner l'armure si {perceval}\_ {Perceval}

pouvait {convincre}\_{convaincre} l'ennemi du roi, qui était justement {devat}\_{devant} la porte de la cour. (L4-011)

Dans l'exemple (100), il s'agit de l'oubli de la lettre finale *t*, qui sert à indiquer la 3PS du passé simple, pour les verbes du 2<sup>ème</sup> groupe et du 3<sup>ème</sup>. Les formes conjuguées produites par les scripteurs, *parti et pu*, proviennent sans doute d'un effet d'analogie. Elles pourraient être influencées par la forme verbale de leur participe passé, qui a une fréquence importante dans la lecture des scripteurs.

- (100) a. \* Avant ~que~ Perceval l'embrassa et {parti}\_{partit}, elle lui donne trois recommandations. (L4-010)
- b. \* Sa mère ne [plu] {pu}\_{put} pas le retenir et lui raconta sa peur en pleurant. (L4-027)

Le scripteur dans l'exemple (101) a choisi une mauvaise lettre finale muette pour marquer la 3PS du présent. La forme qu'il a produite sert à signaler le subjonctif.

- (101) \* Il y a une fille qui s' {assoie}\_{assoit} autour {de\_la}\_{du} groupe ~de~ {chevalier}\_{chevaliers}. (L4-018)

Nous apercevons que les apprenants de trois niveaux différents représentent un profil de développement intéressant si l'on analyse quels types de formes verbales dont les erreurs relèvent. Pour les apprenants de L2, 6 erreurs (7 au total) portent sur les formes du passé composé. Ces six formes erronées ne concernent que deux verbes, *rire* et *promettre*, et il s'agit tous, nous semble-t-il, d'erreurs de fausse analogie. Plus précisément, lors de la production écrite, les locuteurs ont confondu la forme du passé composé des verbes *rire* et *promettre* avec la forme du présent du singulier : *rit / ris* pour la forme *ri* ; *promit* pour la forme *promis*. L'hypothèse la plus probable susceptible d'expliquer ce fait est que les apprenants ne disposaient pas, en produisant le texte, d'une capacité attentionnelle suffisante pour activer la forme verbale s'adaptant au

contexte morphosyntaxique. Ainsi ont-ils eu recours aux formes verbales qui leur paraissaient les plus fréquentes : les formes du singulier du présent de l'indicatif.

A partir de L3, l'introduction du passé simple fait apparaître une évolution complexe concernant le choix des terminaisons verbales. Nous constatons que 7 occurrences sur 12 portent sur les formes du passé simple, et que toutes ces sept erreurs pourraient s'expliquer par le suremploi du modèle des formes du passé composé. Pour les scripteurs qui ont produit ces erreurs, six d'entre eux ont pris directement la forme du participe passé des verbes, par exemple, *voulu* pour *voulut* ; quant au scripteur qui reste, il a produit la forme *obtin* pour le passé simple du verbe *obtenir*. Il nous semble que le radical a été construit en suivant les règles de conjugaisons du passé simple. Cependant, l'erreur concerne l'omission de la consonne finale *t* qui sert à établir le lien avec le sujet de la troisième personne du singulier, et cela pourrait s'expliquer par une contamination de formes conjuguées du participe passé. Pour ces scripteurs, il est fort probable que les formes conjuguées du passé simple, comme de nouvelles connaissances, ne s'intégraient pas suffisamment dans leur lexique mental. Et lors de la production écrite, la détermination de l'orthographe posait ainsi les problèmes à ces scripteurs, et par conséquent, ils ont fait référence aux formes conjuguées du participe passé, qui étaient fréquentes dans l'input et assez bien maîtrisées par les scripteurs de ce niveau.

En observant les écrits des apprenants de L4, nous constatons que de nouvelles erreurs de surgénéralisation se manifestent. 16 formes erronées sur 19 portent sur les formes du participe passé, et ces erreurs relèvent pour la plupart du temps de l'ajout de la lettre finale *-t*. Il nous semble que ces adjonctions ne se font nullement au hasard mais qu'elles consistent à transmettre au participe passé les marques du sujet de la 3PS qui précède le verbe conjugué. Le souci de marquer la 3PS provient probablement des formes de l'indicatif présent, mais aussi des formes du passé simple, qui deviennent plus fréquentes selon le niveau linguistique des apprenants, et qui sont omniprésentes dans le texte de source.

En faisant une analyse des types d'erreurs et de leur distribution, nous concluons que les erreurs telles que l'ajout ou l'omission de la lettre finale aux terminaisons

verbales suivent une évolution complexe et qu'elles ne pourraient pas être expliquées par une simple intention de faire l'accord avec le sujet. Au niveau initial, les formes verbales de l'indicatif présent sont introduites d'abord aux apprenants, et puis les formes du participe passé. Pour les apprenants de ce niveau, l'ajout de lettres finales porté par les formes du participe passé pourrait s'expliquer par l'interférence des formes de l'indicatif présent, qui sont fréquentes dans l'input et mieux maîtrisées par les apprenants. Au niveau suivant, avec l'introduction du passé simple, la détermination de l'orthographe des formes verbales devient plus compliquée, et une tendance à omettre la lettre finale des formes du passé simple est observée, erreurs induites par une fausse analogie des formes du participe passé. Au niveau avancé, les formes du passé simple sont au fur et à mesure maîtrisées par les apprenants. Toutefois, une nouvelle tendance d'ajouter une lettre finale aux formes du participe passé se manifeste, qui disparaît déjà chez les apprenants au niveau intermédiaire. Avec la forte présence du passé simple dans le texte de source, le souci de mettre une lettre finale aux formes du participe passé est probablement dû à une fausse analogie avec les formes du passé simple.

#### **4.3.3.1.4.2. Accord sujet-verbe**

Notons que les verbes de la langue française sont soumis à la conjugaison régie par le sujet. Autrement dit, il existe des formes verbales différentes en fonction du rôle communicatif et du nombre du sujet. Pour les formes erronées ci-dessous, il existe un désaccord entre les formes verbales choisies par les scripteurs et le sujet des verbes. Plus précisément, le rôle communicatif du sujet ne se reflète pas dans la conjugaison des verbes suivants.

Pour les apprenants en L2, les erreurs relevant de cette catégorie (10 occurrences) concernent 45% du nombre total (22) des erreurs relevant de la morphologie verbale et n'altérant pas la valeur phonique.

Pour construire le temps imparfait, les deux scripteurs dans l'exemple (102) ont choisi la terminaison – *ais*, qui sert à indiquer la 1PS ou 2PS, tandis que le sujet est à la 3PS.



(102) a. \* A la {court}\_ {cour}, il y {avais}\_ {avait} une fille. (L2-02-004)

b. \* Le roi a dit que si on lui {avais}\_ {avait} appris l'emploi de la lance. (L2-02-011)

La forme verbale *reviendras* ne s'accorde pas avec son sujet *il*, puisque le premier est à la 2PS alors que le dernier est à la 3PS.

(103) \* Perceval s'est en et il lui a dit qu'il [rien] {reviendras}\_ {reviendra}. (L2-02-001)

Le scripteur dans l'exemple (104) n'a pas choisi la marque flexionnelle congruente et par conséquent, la forme verbale *peut* de la 3PS ne s'accorde pas avec son sujet *je*, de la 1PS.

(104) \* je {peut}\_ {peux} être le chevalier vermeil. (L2-01-020)

Inversement, les deux scripteurs dans l'exemple (105) ont choisi la terminaison –x, qui sert à indiquer la 1PS ou 2PS, alors que le sujet est à la 3PS.

(105) a. \* Il [semblait] {veux}\_ {veut} partir {est}\_ {et} servir sa nation, parce que la {guère}\_ {guerre} a {commancée}\_ {commencé}. (L2-01-016)

b. \* Il {+R[a\_donné]+R+//peux\_donner//+R}\_ {peut\_donner} {bonneur}\_ {bonheur} en ce siècle et bonne fin. (L2-02-005)

Les scripteurs chinois font une confusion entre la terminaison –s et –t. Ils ont choisi les formes verbales terminant en -s et ces dernières ne s'accordent pas avec leur sujet, qui est à la 3PS.

(106) a. \* Perceval {dois}\_ {doit} en servir. (L2-01-007)

b. \* Il se {crois}\_ {croit} que s'il le demande {à}\_ {au} roi le roi ~le~ lui va

donner aussi. (L2-02-009)

c. \* belle fille {crois}\_ {croit} que c'est dommage de ne pas pouvoir savoir son prénom. (L2-02-015)

d. \* Perceval {reviens}\_ {revient}, il y a eu les exclamations et le {rois}\_ {roi} lui ~a~ exprimé désolé de l'avoir {rire}\_ {ri}. (L2-02-010)

Pour les apprenants de L3, 5 erreurs ont été identifiées dans cette catégorie, représentant 23% du nombre total d'erreurs (22).

Dans l'exemple (107), il existe un désaccord entre les formes verbales choisies par les scripteurs et leur sujet. Plus précisément, les scripteurs ont choisi la terminaison –s, ou –x, qui sert à indiquer la 1PS ou 2PS, alors que le sujet est à la 3PS.

(107) a. \* Et Perceval le {convaincs}\_ {convainc} à la fin. (L3-002)

b. \* Elle lui dit qu'elle {crois}\_ {croit} que nul chevalier ne te surpassera !  
(L3-004)

c. \* Si vous rencontra devant la porte et qui s'en va avec votre coupe. Tu {peut}\_ {peux} à la cour du moi. (L3-006)

d. \* Le roi dit à lui que s'il {renconte}\_ {rencontre} devant la porte et qui s'en va avec ta coupe, il {veux}\_ {veut} être le chevalier {Vermeils}\_ {Vermeil}.  
(L3-016)

Dans l'exemple (108), il s'agit d'un oubli de la lettre finale *t*, qui sert à indiquer la 3PS.

(108) \* Il a demandé au garde de la cour de donner la coupe d'or {à}\_ {au} roi et dire à cette belle fille qui n'avait pas {rit}\_ {ri} depuis de 6 ans mais ~qui~ riait pour lui qu'il {reviendrai}\_ {reviendrait}. (L3-010)

Dans cette catégorie, les apprenants de L4 ont produit 14 erreurs, ce qui représente 40% du nombre total d'erreurs (35).

Pour les formes erronées dans l'exemple (109), il existe un désaccord entre les formes verbales choisies par les scripteurs et leur sujet. Plus précisément, le rôle communicatif du sujet ne se reflète pas dans la conjugaison des verbes suivants. Les scripteurs ont tous choisi la terminaison *-s* ou *-x*, qui sert à indiquer la 1PS ou 2PS, tandis que le sujet des verbes est à la 3PS.

- (109) a. \* Donc elle le [fa] {fis}\_ {fit} ne rencontre pas la chevalerie et ne voir pas des chevaliers. (L4-010)
- b. \* Et puis il {fis}\_ {fit} la connaissance que le pire ennemi du roi, c'est le chevalier Vermeil +R[à]+R+//dans//+R la forêt. Il lui menaça de sa fille et du vin auparavant. (L4-010)
- c. \* on ne {sais}\_ {sait} pas son nom. (L4-003)
- d. \* le roi ~~Arthur~~ s'est {sentit}\_ {senti} très dommage, et il {crois}\_ {croit} que Perceval {devenirait}\_ {deviendrait} un bon chevalier s'il [μxxxμ] reste. (L4-007)
- e. \* Perceval, [veu] se croit qu'il {peux}\_ {peut} résoudre ce problème. (L4-013)
- f. \* il [cou] a dit au roi qu'il voulait d'être un chevalier et il {veux}\_ {veut} des armes du chevalier Vermeil. (L4-016)
- g. \* Puis il [parti] {pars}\_ {part} et reste avec {le}\_ {la} cour Vermeil. (L4-017)
- h. \* Mais personne ne {peux}\_ {peut} [éviter] rien faire face à la fatalité. (L4-018)
- i. \* Il se {sens}\_ {sent} étonné quand il voit leurs épées et armures, les choses de chevalier qu'il sait jamais. (L4-018)
- j. \* Keu, le frère lait du roi Arthur, {toujour}\_ {toujours} jaloux, aigre, [mo] lui moquait : Allons-y, {prend}\_ {prends}-les\_. (L4-019)
- k. \* il {aperçois}\_ {aperçoit} un château et un chevalier. (L4-026)
- l. \* Perceval recherche du chevalier Vermeil, il {veux}\_ {veut} combattre avec lui. (L4-026)

m. \* il l'en {veux}\_ {veut} bien, dont il {veux}\_ {veut} les demander au roi, il est {naïf}\_ {naïf}. (L4-028)

Pour les erreurs de cette catégorie, nous concluons qu'il s'agit pour la plupart du temps de la production des formes de la 3PS : les formes de la 3PS se confondent avec celles de la 1PS ou 2PS. Pour de nombreux verbes français, les formes 1PS, 2PS, et 3PS se distinguent les uns des autres à l'écrit, mais elles se confondent à l'oral. Ainsi avons-nous prévu la production des erreurs relevant de cette confusion. Cependant, pour les apprenants de trois niveaux, les erreurs ne concernent que la confusion de formes la 3PS avec celles de la 1PS ou de la 2PS (il n'existe que deux cas d'exception). Étant donné que les formes de la 3PS sont normalement d'abord acquises par les apprenants, il devrait s'agir des formes les plus accessibles et les plus maîtrisées par les apprenants dès le début de leur apprentissage et nous nous sommes attendus à plus de formes accordées. Une hypothèse susceptible d'expliquer ce fait est que la tâche de l'écriture concerne la restitution par écrit d'un conte du chevalier Perceval, et que la plupart des phrases produites par les apprenants ont été construites autour d'un sujet de la troisième personne du singulier. Indiquons également que les erreurs de cette catégorie concernent notamment les verbes qui sont fréquents dans l'input des apprenants, tels que *pouvoir*, *vouloir*, *devoir*, *croire*, *savoir*, *revenir*, etc. Le problème se pose également aux verbes avec les formes conjuguées plus compliquées, tels que *convaincre*, *apercevoir*. Toutefois, nous n'avons trouvé que deux occurrences dans ce cas.

#### 4.3.3.1.5. Conclusions

En observant les formes verbales homophones en /E/, nous constatons certaines tendances interlangagières spécifiques à ce groupe d'apprenants. Tout d'abord, il faut indiquer que les apprenants chinois possèdent une bonne maîtrise de cette homophonie verbale : le taux de réussite est de plus de 90% dès le stade initial, et le nombre d'erreurs diminue en fonction du niveau linguistique des apprenants. Rappelons que chez les enfants francophones, un scénario développemental se dessine nettement : les élèves

français, à partir de CE2<sup>57</sup>, sélectionnent la forme en –é comme forme prototypique et ont une forte propension à marquer cette forme en genre et en nombre. En revanche, nous n’avons pas trouvé de forme prototypique dans la production des apprenants chinois. En outre, le souci d’accorder la forme verbale avec le sujet de la phrase n’est pas observé dans notre corpus. D’après Totereau et *al.* (2006), marque l’accord avec le sujet même quand il n’est pas attendu, c’est « avoir compris la solidarité syntagmatique qui existe, en français, entre le verbe et son sujet ».

Chez les apprenants chinois, les erreurs relevant de l’accord sujet-verbe sont peu nombreuses. L’hypothèse la plus probable susceptible d’expliquer ce fait est que le processus de l’application de conjugaisons verbales est loin d’être automatisé pour ce groupe d’apprenants. En suivant strictement une opération séquentielle pour réaliser les terminaisons verbales, les apprenants chinois auraient peu de chances à produire ce type d’erreur de sur-généralisation.

Pour les formes verbales restantes, les erreurs concernent essentiellement d’une part, les lettres finales muettes et d’autre part, l’accord sujet-verbe. Selon nos observations, le premier type d’erreur ne pourrait pas s’expliquer par l’intention d’établir le lien syntagmatique entre le verbe et son sujet. Il existe bien évidemment les graphies erronées, telles que \* *la femme n’avait pas {sourit} \_{souri} pendant longtemps.* (L2-01-001), pour lesquelles l’ajout de la lettre finale muette pourrait s’expliquer par l’expression de la solidarité syntagmatique entre le verbe et son sujet de la 3<sup>e</sup> personne du singulier. Néanmoins, après avoir examiné toutes les occurrences regroupées dans cette catégorie, il nous semble plus pertinent d’interpréter ces erreurs par un effet d’analogie.

Tout d’abord, il faut indiquer que certaines graphies erronées dans cette catégorie ne pourraient pas s’expliquer par une sur-généralisation de l’accord entre le verbe et son sujet. Nous trouvons en effet les occurrences telles que : \* *la fille a {ris} \_{ri} pour lui.* (L2-02-012). Dans ce cas, l’ajout de la lettre finale –s ne résulte pas d’une intention de faire l’accord entre le verbe et le sujet, puisque la lettre finale –s renvoie d’une façon

---

<sup>57</sup> Chez les élèves plus jeunes, la forme prototypique est la forme en -er.

générale à un sujet de la 1<sup>ère</sup> personne et qu'il s'agit d'un sujet de la 3PS dans cet exemple. D'après notre analyse, cette forme erronée provient probablement d'un effet d'analogie. Lors de la production écrite, le scripteur a pris directement la forme d'indicatif de la 1PS du verbe *rire*.

Quand le scripteur n'arrive pas à récupérer directement la forme verbale congruente lors de la production du texte, il pourrait avoir recours à une stratégie de l'analogie orthographique : soit il prend directement une forme verbale qu'il connaît bien, et c'est ce que nous avons montré dans l'exemple plus haut ; soit il construit une forme verbale à partir d'une autre forme qu'il connaît, et cela nous permet de comprendre les graphies erronées telles que \* *il {obtin}\_ {obtint} le respect du roi Arthur quand il revint dans [d] la cour.* (L3-013). L'omission de la lettre finale *-t* pourrait être contaminée par la forme du participe passé du verbe *obtenir*.

Deuxièmement, il n'est pas inintéressant de rappeler que les graphies erronées dans cette catégorie produites par les apprenants de trois niveaux linguistiques différents représentent un scénario développemental particulier. Pour les apprenants de L2, la majorité des erreurs (6 graphies erronées sur 7) concerne les formes verbales du passé composé. En analysant ces erreurs, nous constatons que les scripteurs ont confondu la forme du passé composé des verbes avec les formes de l'indicatif présent (3PS ou 1PS). Notons que ces dernières ont une fréquence importante pour les apprenants de ce niveau.

A partir du niveau L3, 7 graphies erronées sur 12 portent sur les formes du passé simple, et que ces sept erreurs pourraient s'expliquer par le suremploi du modèle des formes du passé composé : *il {voulu}\_ {voulut} aller demander les armes au roi.* (L3-011). L'hypothèse possible pour expliquer ce fait est que les apprenants n'ont pas encore eu une bonne maîtrise de formes verbales du passé simple et qu'ils ont ainsi fait référence aux formes conjuguées du participe passé, fréquentes dans l'input et assez bien maîtrisées par les scripteurs de ce niveau.

Si les erreurs relevant des formes verbales du passé composé réduisent significativement en L3, elles reviennent dans la production des apprenants de L4. 16 formes erronées sur 19 portent sur les formes du participe passé, et ces erreurs

concernent pour la plupart du temps de l'ajout de la lettre finale *-t*. Ce souci de marquer la troisième personne du singulier provient probablement des formes verbales du passé simple, qui deviennent plus fréquentes selon le niveau linguistique des apprenants, et qui sont omniprésentes dans le texte de source.

Si l'on observe les formes verbales homophones autres que celles en /E/, le deuxième type d'erreur relève d'un non-accord sujet-verbe. Par exemple, \* *je {peut}\_ {peux} être le chevalier vermeil*. (L2-01-020). Ici, il s'agit pour la plupart du temps d'une confusion : les formes de la troisième personne du singulier se confondent avec celles de la première personne du singulier ou de la deuxième personne, ou inversement. Et cette confusion s'explique, nous semble-t-il, par une analyse insuffisance morphosyntaxique. Plus précisément, le coût de certaines activités graphiques est très élevé chez le scripteur et mobilise la plus grande part de l'attention. Ainsi, il n'arrive pas à analyser le contexte morphosyntaxique de la forme verbale à produire et omet de choisir une terminaison congruente.

En résumé, il nous semble que les formes verbales homophones sont loin d'être stables chez les apprenants chinois. Pour les formes homophones en /E/, les règles de conjugaison sont assez simples à maîtriser et en suivant une opération séquentielle, les apprenants ont une bonne performance. Néanmoins, les erreurs surgissent lorsque les apprenants ne disposent pas d'une capacité attentionnelle suffisante pour gérer l'application de règles. Pour les formes verbales homophones restantes, les difficultés sont de la même nature. Lorsque les apprenants sont incapables de récupérer directement les formes verbales lors de la production écrite, ils écrivent une forme verbale à partir d'une autre qu'ils connaissent. Les erreurs relevant de la sur-généralisation présentent un profil développemental intéressant, illustrant le fait que les apprenants s'appuient souvent à une stratégie de l'analogie lors de la production des formes verbales. Quant aux erreurs concernant l'accord sujet/verbe, ce type d'erreur se produit lorsque les apprenants ne disposent pas d'une capacité attentionnelle suffisante pour faire une analyse morphosyntaxique congruente. Notons que ce type d'erreur surgit même pour les verbes ayant une fréquence particulièrement importante dans l'input de l'apprenant, tels que *pouvoir*, *vouloir*.

Les performances des apprenants chinois en ce qui concerne la morphologie écrite des homophones verbales nous illustrent les tendances interlangagières spécifiques. En outre, l'observation des graphies erronées permet de mettre en lumière deux stratégies graphiques principalement adoptées par les apprenants chinois lorsqu'ils se trouvent confrontés à la difficulté de produire une forme verbale congruente : la stratégie de substitution et celle de sur-généralisation. En faisant référence à ces deux types de stratégie graphique, nous avons déjà proposé, dans les parties ci-dessus, les scénarios psycholinguistiques possibles pour interpréter les erreurs des homophones verbaux.

A la différence des enfants francophones, ayant un système morphologique basé sur un input majoritairement oral, les apprenants chinois reçoivent simultanément un input oral et écrit dès le début de leur apprentissage. En produisant les formes verbales à l'écrit, les apprenants chinois commettent également un grand nombre d'erreurs altérant la valeur phonique du verbe concerné. L'observation de ces formes erronées est conforme à nos attentes. Parmi les formes erronées issues de la morphologie verbale altérant la valeur phonique, nous avons identifié deux types d'erreur principaux : erreur de substitution, et celles relevant de la fausse analogie. Dans la partie suivante, nous mènerons une analyse des graphies erronées, afin d'illustrer les deux stratégies graphiques adoptées par les apprenants.

#### **4.3.3.2. Erreurs issues de la morphologie verbale altérant la valeur phonique**

##### **4.3.3.2.1. Erreurs de substitution**

Les erreurs ci-dessous montrent que le problème de la finitude verbale ne concerne pas seulement les verbes du premier groupe. Le fait que les apprenants commettent les erreurs relevant de la finitude verbale ne pourront pas s'expliquer tout simplement par la confusion des différentes formes verbales issues de la même image phonétique (par exemple : la forme *parlé* est substituée par *parler*).

Pour les apprenants de L2, 9 occurrences ont été trouvées, qui représentant 32% du nombre total d'erreurs (28).



Même avec la présence de l’auxiliaire *avoir*, les scripteurs dans l’exemple (110) ont choisi les formes de l’infinitif.

- (110) a. \* Perceval lui a fait rire, qui n’a pas {rire}\_ {ri} depuis six ans. (L2-02-002)
- b. \* Et celle qui ne [ria] rirait pas [ri] a [ri] {rire}\_ {ri}. (L2-02-007)
- c. \* Perceval {reviens}\_ {revient}, il y a eu les exclamations et le {rois}\_ {roi} lui ~a~ exprimé désolé de l’avoir {rire}\_ {ri}. (L2-02-010)
- d. \* Il y a six ans que la belle fille n’avait pas {rire}\_ {ri}. (L2-02-015)

Inversement, les formes finies ont été également trouvées dans les contextes finis. Les apprenants dans l’exemple (111) ont choisi soit les formes du participe passé, soit les formes de l’indicatif présent dans la situation où les formes de l’infinitif étaient demandées.

- (111) a. \* Il a décidé de partir et {devenu}\_ {devenir} un chevalier. (L1-01-003)
- b. \* il a l’idée de leur ressembler et {parti}\_ {partir} à la {cœur}\_ {cour} du roi. (L1-01-009)
- c. \* C’est l’heure de {parti}\_ {partir}. (L1-01-011) (erreur lexicale ?)
- d. \* mais elle ne peut pas {ri}\_ {rire} depuis 6 ans. (L2-02-016)
- e. \* Il est parti à la cour du roi Arthur pour [a] {gagne}\_ {gagner} les {armée}\_ {armes} [rêve\_de] du chevalier vermeil. (L2-02-007)

Un autre type d’erreur de substitution est également trouvé dans la production des apprenants de L2. Dans l’exemple (112), les apprenants ont choisi les formes verbales de l’indicatif présent, tandis que les contextes demandaient celles du participe passé. Et la présence des auxiliaires *avoir* et *être* n’a pas facilité le choix des terminaisons verbales faits par les apprenants.

- (112) a. \* il a {souvient}\_ {souvenu} une des trois recommandations de sa mère. (L1-01-002)

- b. \* Il n'a pas {connaît}\_ {connu} les chevaliers. (L1-01-003)
- c. \* le roi s'est {fiché}\_ {fâché} et lui a {arrête}\_ {arrêté} ses paroles. (L2-01-006)
- d. \* il [p] est {parts}\_ {parti}. (L1-01-018)
- e. \* Perceval lui a {demande}\_ {demandé} +R[sa]+R+//son//+R {armée}\_ {arme}. (L1-01-023)
- f. \* il y {avais}\_ {avait} une fille, qui n'avait pas ri depuis 6 ans, a {apparaît}\_ {apparu} son appréciation [d] pour Perceval. (L2-02-004)
- g. \* il est {arrive}\_ {arrivé} dans la salle où le roi {est}\_ {et} ses chevaliers étaient assis.
- h. \* il a {pense}\_ {pensé} que, s'il a {demandait}\_ {demandé} au roi. (L2-02-006)
- i. \* Ainsi que le roi [se] a {rend}\_ {rendu} {compt}\_ {compte} à lui. (L2-02-013)
- j. \* {Le}\_ {La} première fois qu'il a {vient}\_ {venu} à la cour du roi Arthur, Perceval était un jeune, grand, beau, noble homme mais à demi sauvage. (L2-02-016)

Dans les écrits des apprenants de L3, les erreurs concernant la finitude verbale ont été également trouvées, mais avec une présence moins importante. 5 occurrences ont été identifiées, représentant 22% du nombre total d'erreurs (23).

Pour les phrases suivantes de l'exemple (113), les formes verbales se trouvaient juste après les propositions. Au lieu de choisir les infinitifs, les scripteurs ont produit les formes de l'indicatif du présent.

- (113) a. \* il a demandé à un chevalier de {porte}\_ {porter} au roi sa coupe et saluer de sa part. (L3-001)
- b. \* pour [lu] {protège}\_ {protéger} à la mort {du}\_ {de\_la} guerre, sa mère ne lui raconte aucune histoire de chevaliers, donc il ne sait pas du tout de chevaliers. (L3-002)

Les situations inverses se sont produites également chez les apprenants de L3. Ils ont omis de réaliser l'inflexion verbale dans les contextes finis.

- (114) a. \* il se {moquer}\_ {moqua} de Perceval. (L3-004)  
b. \* comme il {comprendre}\_ {comprend} nombre de choses qu'il ne sait pas avant. (L3-016)  
c. \* Le chevalier {rire}\_ {rit}, et il [en] [vele] allait au roi avec Perceval. (L3-006)

Pour les apprenants de L4, les erreurs relevant de la finitude verbale sont encore abondamment présentes. Nous avons trouvé 16 formes erronées, qui représentent 32% du nombre total d'erreurs (50). Dans l'exemple (115), les apprenants ont choisi les formes verbales de l'infinitif en ignorant les contextes finis.

- (115) a. \* Il était une fois un homme il s'appelle Perceval qui {vivre}\_ {vit} {à\_la}\_ {au} manoir isolé de la +R[cours]+R+//cour//+R du roi Arthur [il\_est] avec sa mère, il est un homme qui un peu sauvage mais il avait un visage beau et grand. (L4-005)  
b. \* Le chevalier {rire}\_ {rit} pour ses mots. (L4-008)  
c. \* elle ne [pas] {rire}\_ {rit} pas depuis de 6ans. (L4-026)  
d. \* Elle pleure et {laisser}\_ {laisse} partir Perceval. (L4-026)  
e. \* Dans ces gens, il y avait une femme qui lui {faire}\_ {faisait} penser à sa mère. (L4-012)  
f. \* Il passait les incertaines péripéties ~~dans\_la\_rue~~ et il en {voir}\_ {vit} {de}\_ {des} rudes. (L4-014)  
g. \* Perceval lui {demander}\_ {demanda} ses armes [d] au roi, le chevalier lui se moqua. (L4-019)

Les apprenants dans l'exemple (116) ont produit les formes de l'infinitif, en

ignorant la présence de l’auxiliaire *avoir* ou *être*, qui se trouvaient au contexte immédiat. L’exemple (116d) constitue un cas d’exception. Deux formes verbales étaient juxtaposées, et ainsi la deuxième, forme erronée ne se trouvait pas juste après l’auxiliaire *avoir*.

- (116) a. \* Elle n’eut {dire}\_ {dit} rien et [Perceval] il n’eut {savoir}\_ {su} rien. (L4-005)
- b. \* il est {partir}\_ {parti} seulement pour cette fille. (L4-007)
- c. \* Il n’a même pas {connaître}\_ {connu} son nom! (L4-025)
- d. \* Perceval lui salua, et le roi lui exprima le pardon, expliqua la raison : «... et {répandre}\_ {répandu} sur la reine Guenièvre. (L4-019)

Les formes finies qui se situaient dans les contextes non finis ont été également trouvées. Dans l’exemple (117a), le scripteur a employé la forme du futur simple du verbe *dire* après la préposition *sans*. Le scripteur a fait un choix verbal dont nous ne pouvons pas trouver aucune hypothèse susceptible. Le scripteur dans l’exemple (117b) a choisi la forme verbale de l’indicatif présent après la préposition *de*. Il apparaît que la présence du pronom clitique *lui* a perturbé le choix du scripteur. Et les deux scripteurs suivants ont choisi les formes verbales du participe passé en ignorant la présence du verbe *aller* ou *pouvoir* dans le contexte gauche immédiat.

- (117) a. \* Sa mère lui élevé dans une place isolée, sans lui [les] {dira}\_ {dire} les mots sur les chevaliers. (L4-006)
- b. \* Ce sont : honorez les dames, suivez les conseils des prud’hommes. et priez Dieu de lui {donne}\_ {donner} [honn] [hono] honneur et bonne fin. (L4-013)
- c. \* Le {chevail}\_ {chevalier} dit à Perceval qu’il allait {revenu}\_ {revenir} un jour pour {le\_venge}\_ {la\_vengeance}. (L4-019)
- d. \* Perceval a vaincu finalement au cas où il n’a pu {su}\_ {savoir} comment utiliser des armes. (L4-020)

#### 4.3.3.2.2. Erreurs d'analogie

Les erreurs concernant les terminaisons verbales pourraient être issues d'une fausse analogie. Autrement dit, l'apprenant suit un autre modèle pour conjuguer une forme verbale. Dans les écrits produits par les apprenants de L2, nous avons trouvé 6 formes erronées, qui représentent 21% du nombre total d'erreurs (28).

Pour construire la forme du participe passé du verbe *rire*, le modèle des verbes du premier groupe a été suremployé dans l'exemple (118a). Nous constatons que le scripteur a enlevé la terminaison *-re* du verbe *rire*, et ajouté directement *-é*, terminaison pour les formes du participe passé des verbes du premier groupe, au radical *ri-*. Dans l'exemple (118b), le verbe *recevoir* s'est conjugué également selon un faux modèle. Le radical se construit comme « infinitif moins *-oir* », auquel se rattache le suffixe *-u*.

(118) a. \* {Il}\_ {Ils} se sont [riait] {rié}\_ {ri}. (L2-02-007)

b. \* Quand sa mère a {receu}\_ {reçu} l'information de sa partie. (L1-01-016)

Le verbe *appeler* ou *rappeler* redouble le *l* devant un *e* muet pour obtenir le son [ɛ], car, suivi de deux consonnes, le *e* se prononce [ɛ] sans l'ajout d'un accent grave. Cependant, pour les deux scripteurs suivants, cette règle de conjugaison n'a pas été activée lors de leur production. Dans l'exemple (119a), le scripteur a conjugué le verbe *appeler* en suivant le modèle des verbes en *-er* normaux. Quant au scripteur dans l'exemple (119b), il a doublé la lettre *l* sans connaître que le modèle du redoublement de consonnes ne s'applique pas à la forme du participe passé.

(119) a. \* Il y avait un jeune homme qui s' {appele}\_ {appelle} Perceval. (L2-01-005)

b. \* Il s'est {rappellé}\_ {rappelé} des conseils de sa maman. (L2-02-013)

La procédure de conjugaison est compliquée pour la situation suivante. Le scripteur a rattaché le suffixe *-ent*, qui sert notamment à marquer la troisième personne du pluriel, au radical *ser-*. De plus, pour obtenir le son [ɛ] de la voyelle *e*, le scripteur a doublé la lettre *r*. Il nous semble que le scripteur a fait preuve d'« une créativité de

conjugaison » : il s'est inspiré d'autres formes verbales pour produire une forme acceptable (Santens, 2016).

(120)\* Sa mère a peur parce [que] qu'ils ne {serrent}\_ {seront} guère {adroit}\_ {adroits}. (L2-02-006)

Dans la catégorie de la fausse analogie, les apprenants de L3 ont produit 12 erreurs, ce qui représente 52% du nombre total d'erreurs (23). Nous constatons que le nombre d'erreurs augmente avec l'introduction des formes verbales du passé simple.

En ignorant la forme correcte du passé simple, le scripteur dans l'exemple (121) a pris directement la forme du participe passé du verbe *entendre*.

(121)\* Quand le roi {entendu}\_ {entendit} son départ, il était [re] [eut\_regret] triste de le faire partir sans se revoir. (L3-003)

Pour la conjugaison du passé simple, les verbes du 3<sup>e</sup> groupe ont trois modèles de terminaisons. Le scripteur dans l'exemple (122) a conjugué les verbes rendre et abattre en suivant le faux modèle, qui s'explique probablement par une fausse analogie des formes du passé composé.

(122) a. \* Il déclarait que la personne qui la {rendut}\_ {rendit} rire allait être ~~bon~~ chevalier. (L3-023)

b. \* il {abattut}\_ {abattit} l'ennemi. (L3-023)

Pour la forme verbale erronée suivante, elle est construite comme « infinitif *suivre* moins *-re* », auquel se rattache le suffixe *-a*, terminaison au passé simple des verbes du 1<sup>er</sup> groupe.

(123)\* quand il chassa dans {le}\_ {la} forêt, il rencontra le roi ~~Arthur~~ et {son}\_ {sa} troupe qui {les}\_ {le} {suiva}\_ {suivit}. (L3-013)

Dans l'exemple (124), le scripteur a suivi la même logique en rattachant la terminaison *-a* au radical *rejoint-*, forme verbale de la 3PS de l'indicatif présent.

(124) \* il {rejointa}\_ {rejoinit} {le}\_ {la} {troupe}\_ {troupe} et repartit. (L3-013)

Le scripteur a ignoré dans le cas suivant que le modèle du redoublement de consonnes ne s'appliquait pas à la forme du passé simple.

(125) \* une belle fille {assis}\_ {assise} à côté {appella}\_ {appela} les trois [comma] recommandations. (L3-013)

Les erreurs ci-dessous s'expliquent probablement par une fausse analogie avec les formes verbales du subjonctif.

(126) a. \* ça [ser] {serve}\_ {servit} à rien. (L3-013)

b. \* son père et 2 frère [et\_me] {meuraient}\_ {mouraient} au combat. (L3-012)

Il nous semble que le scripteur dans l'exemple (127) a suivi le modèle du futur simple pour construire la forme du passé simple du verbe *voir*.

(127) \* quand il arriva, il {verre}\_ {vit} plein de choses nouvelles. (L3-003)

Quant aux erreurs suivantes concernant le verbe *rire*, les formes se conjuguèrent selon le faux modèle.

(128) a. \* Cette fille [n'avai] n'avait pas {ru}\_ {ri} que le jour où elle verrait celui qui aurait le prix de toute chevalerie. (L3-020)

b. \* Dans la salle, il y avait aussi une belle jeune fille qui ne {ru}\_ {rit} jamais depuis 6 ans. (L3-022)

- c. \* Mais elle [e] a {ru}\_ {ri} pour la première fois quand elle [le] a rencontré Perceval. (L3-022)

Pour les apprenants de L4, 21 erreurs ont été identifiées dans cette catégorie, représentant 42% du nombre total d'erreurs (50). Parmi ces erreurs, 12 formes erronées concernent les formes du passé simple. Cela nous montre que la maîtrise des formes du passé simple n'est pas encore stable chez les apprenants de ce niveau.

Dans l'exemple (129a) et (129b), les scripteurs ont rattaché la terminaison *-ait* à l'infinitif des verbes pour construire les formes du conditionnel. Afin de construire la forme de l'imparfait du verbe *repartir*, le scripteur dans l'exemple (129c) a suivi le même modèle. Dans les exemples (129d) et (129e), les scripteurs ont ignoré la forme du passé simple des verbes et par conséquent rattaché la terminaison *-a* à l'infinitif.

- (129) a. \* Elle a dit que Perceval peut-être {devenirait}\_ {deviendrait} le meilleur chevalier. (L4-004)
- b. \* le roi ~Arthur~ s'est {sentit}\_ {senti} très dommage, et il {crois}\_ {croit} que Perceval {devenirait}\_ {deviendrait} un bon chevalier s'il [μxxxμ] reste. (L4-007)
- c. \* Mais Perceval [repartirai] ignorait et {repartirait}\_ {repartait} {au}\_ {à\_la} cour du roi Arthur. (L4-008)
- d. \* Quand il {acquérira}\_ {acquerra} les manières, il devient preux ! (L4-002)
- e. \* Il {partira}\_ {partit} [son] à la direction du roi Arthur. (L4-006)

Les formes verbales suivantes relèvent des verbes du 3<sup>e</sup> groupe. Cependant, les scripteurs ont choisi la terminaison *-a* pour les verbes du 1<sup>er</sup> groupe. Soit le radical se construit comme « infinitif moins *-ir* » (pour les verbes *souvenir*, *obtenir*, *offrir*), soit comme « infinitif moins *-e* » (pour les verbes *rire*, *dire*, *combattre*), soit comme « infinitif moins *-re* » (pour le verbe *entendre*). La seule exception concerne le verbe *pleurer*, verbe du 1<sup>er</sup> groupe. Dans l'exemple (130e), le scripteur l'a conjugué selon le modèle des verbes du 2<sup>e</sup> groupe en utilisant la terminaison *-it*.



- (130) a. \* Il {souvena}\_ {souvint} le conseil de sa mère et {le}\_ {la} salua. (L4-002)
- b. \* Le roi {obtenu}\_ {obtint} les armes sans savoir le nom de Perceval. (L4-015)
- c. \* Mais elle ne pouvait pas changer la décision de son fils, alors elle {offra}\_ {offrit} trois recommandations à son fils. (L4-015)
- d. \* Dans {le}\_ {la} cour, [une] il y avait une belle fille, lui regarda et {rira}\_ {rit}. (L4-006)
- e. \* Quand il {dira}\_ {dit} à sa mère, elle était en inquiétude et {pleurit}\_ {pleura}. (L4-015)
- f. \* Le roi lui {dira}\_ {dit} que c'était le chevalier Vermeil de la forêt de Quinqueroi qui portait les armes du roi. (L4-015)
- g. \* {En\_fin}\_ {Enfin}, Perceval {combatta}\_ {combattit} avec [le\_chevallier] le chevalier Vermeil et donna les armes au roi et puis partit. (L4-015)
- h. \* Quand il {entenda}\_ {entendit} la demande de Perceval, il était en colère. (L4-015)
- i. \* Le roi {réponça}\_ {répondit} qu'il avait ~en~ eu +R[une\_coupe\_d\_QUOTE\_or]+R+//mais\_elle//+R était {violée}\_ {volée} justement par le chevalier Vermeil. (L4-027)

Pour les verbes combattre et sourire, le scripteur dans l'exemple (131) les a conjugués selon le modèle de verbes du 2<sup>e</sup> groupe. Il a directement rattaché la terminaison *-issaient* au radical.

- (131) a. \* Ils se {combattaient}\_ {combattaient} dans la forêt. (L4-019)
- b. \* Elle ne {sourissait}\_ {souriait} plus depuis [il] elle +R[était]+R+//avait//+R six ans. (L4-019)

Les formes erronées dans l'exemple (132) concernent le verbe *rire*. Pour conjuguer le verbe *rire*, les scripteurs ont tous construit le radical comme « infinitif

moins -e ».

- (132) a. \* Dans le château, le roi Arthur et ses {chevalier}\_ {chevaliers} {assisaient}\_ {assoyaient} autour d'une table, parlant et {rirant}\_ {riant}. (L4-019)
- b. \* Perceval aperçut une belle jeune fille [dis] parlant {ave}\_ {avec} la reine en {rirant}\_ {riant}. (L4-019)
- c. \* Elle {rirait}\_ {riaït}. (L4-020)
- d. \* une belle jeune fille regardait {perceval}\_ {Perceval} et [ris] {rirait}\_ {riaït}. (L4-022)

#### **4.3.3.3. Conclusion et discussion**

L'objectif de ce chapitre a été de mettre l'accent sur la production écrite de la morphologie verbale en français L2. Le défi de l'acquisition pour les apprenants chinois n'est pas seulement de connaître les différentes formes d'un paradigme morphologique donné, mais surtout de réussir à combiner la forme verbale adéquate avec le contexte morphosyntaxique (Ellis & Collins, 2009). En effectuant une analyse exploratoire et donc de nature plutôt descriptive, nous avons d'une part, examiné les formes et le développement de la morphologie verbale du français L2 écrit et d'autre part, mis l'accent sur les tendances interlangagières des apprenants chinois à des niveaux linguistiques différents. Nous souhaitons dans ce bilan, revenir aux principaux résultats que nous avons évoqués plus haut dans ces deux domaines.

En analysant la réalisation morphologique sur les divers types de verbes, nous avons identifié deux types d'erreurs principaux : erreur de substitution et erreur relevant de la sur-généralisation.

##### **4.3.3.3.1. Erreur de substitution**

Les erreurs de cette catégorie sont multiples et couvrent presque tous les types de formes verbales. Par la suite, nous examinerons respectivement la réalisation verbale

de l'infinitif, de l'indicatif présent, du passé composé et du passé simple.

**Tableau 45.** Formes erronées quand la forme de l'infinitif est attendue

*(Légende : Participe passé = la forme de l'infinitif est remplacée par la forme du participe passé ;  
2PP = la forme de l'infinitif est remplacée par la forme de l'indicatif présent de la 2<sup>e</sup> personne du  
pluriel ; la forme de l'infinitif est remplacée par la forme de l'indicatif présent de la 3<sup>e</sup> personne (ou  
1<sup>ère</sup> personne) du singulier)*

Stade	Nombre de contextes	Nombre d'erreurs	Participe passé	2PP	3PS/1PS	Nombre moyen d'erreurs par apprenant
L2 (n=40)	296	11	8	3	0	0.28
L3 (n=23)	173	4	0	2	2	0.17
L4 (n=28)	229	5	2	1	1	0.18
Total	698	20	10	6	3	0.22

Le tableau 45 nous illustre les performances des apprenants chinois en ce qui concerne les formes verbales de l'infinitif. Bien que pour les apprenants chinois, l'infinitif soit la forme de référence pour désigner un verbe, l'emploi de l'infinitif dans les contextes où il est attendu pose encore un problème, particulièrement pour les apprenants du stade initial. L'hypothèse la plus probable susceptible d'expliquer ce fait est que les apprenants ne disposaient pas d'une capacité attentionnelle suffisante pour faire une analyse morphosyntaxique. Les apprenants ont donc produit des formes du présent au singulier, de la 2PP et du participe passé dans les contextes pour un infinitif après un verbe introducteur, ainsi que dans les contextes pour un infinitif ayant la fonction de complément circonstanciel. Et ces erreurs de substitution diminuent en fonction du niveau linguistique.

- (133) a. \* Le {chevail}\_ {chevalier} dit à Perceval qu'il allait {revenu}\_ {revenir} un jour. (L4-019)
- b. \* priez Dieu de lui {donne}\_ {donner} [honn] [hono] honneur et bonne fin. (L4-013)

Le tableau 46 permet de visualiser la réalisation des formes du participe passé des apprenants de trois niveaux différents. Nous constatons que les formes du participe passé jouent un rôle important à l'écrit des apprenants : ils ont produit au total 1240 occurrences. En outre, leur performance présente un scénario développemental intéressant. Les apprenants de L2 ont substitué souvent des formes de l'infinitif et du présent au singulier à une forme du participe passé. Pour les apprenants de L3, ces erreurs de substitution diminuent significativement. Néanmoins, ce type d'erreur revient au niveau de L4 et des formes du participe passé ont été remplacé soit par celle de l'infinitif, soit par celle du passé simple. La dernière a été introduite à partir de la L2 et était omniprésente dans le texte de source. Ces résultats démontrent clairement que les formes du participe passé n'ont pas encore été stables chez les apprenants de trois niveaux différents.

**Tableau 46.** Formes erronées quand la forme du participe passé est attendue

*(Légende : Infinitif = la forme du participe passé est remplacée par la forme de l'infinitif ; 2PP = la forme du participe passé est remplacée par la forme de l'indicatif présent de la 2<sup>e</sup> personne du pluriel ; la forme du participe passé est remplacée par la forme de l'indicatif présent de la 3<sup>e</sup> personne (ou 1<sup>ère</sup> personne) du singulier ; la forme du participe passé est remplacée par la forme du passé simple)*

Stade	Nombre de contextes	Nombre d'erreurs	Infinitif	2PP	3PS/1PS	Passé simple	Nombre moyen d'erreurs par apprenant

L2 (n=40)	728	27	9	3	17	0	0.68
L3 (n=23)	212	3	1	0	0	2	0.13
L4 (n=28)	300	22	6	0	0	16	0.77
Total	1240	52	16	3	17	18	0.57

Le tableau 47 nous illustre une bonne maîtrise de formes verbales de l'indicatif présent. Aucune erreur n'a été trouvée dans la production des apprenants de L2 et de L3. Néanmoins, les trois graphies erronées produites par les apprenants de L4 sont assez surprenantes. Les deux scripteurs ont substitué la forme de l'infinitif, forme de base également fréquente, à la forme du présent.

- (134) a. \* Le chevalier {rire}\_ {rit} pour ses mots. (L4-008)  
 b. \* elle ne [pas] {rire}\_ {rit} pas depuis de 6ans. (L4-026)  
 c. \* Elle pleure et {laisser}\_ {laisse} partir Perceval. (L4-026)

**Tableau 47.** Formes erronées quand la forme de l'indicatif présent est attendue

(Légende : *Infinitif* = la forme de l'indicatif présent est remplacée par la forme de l'infinitif)

Stade	Nombre de contextes	Nombre d'erreurs	Infinitif	Nombre moyen d'erreurs par apprenant
L2 (n=40)	536	0	0	-
L3 (n=23)	213	0	0	-
L4 (n=28)	398	3	3	0.11
Total	1240	3	3	0.03

Le tableau 48 permet de visualiser la réalisation de formes du passé simple. Cela ne concerne que les apprenants de L3 et de L4, étant donné que ce temps est introduit à partir du niveau L3. Nous constatons que les formes erronées sont assez homogènes : les formes du passé simple ont été remplacées soit par celle de l’infinitif, soit par celle du participe passé.

**Tableau 48.** Formes erronées quand la forme du passé simple est attendue

(Légende : *Infinitif* = la forme du passé simple est remplacée par la forme de l’infinitif ;  
*participe passé* = la forme du passé simple est remplacée par la forme du participé passé)

Stade	Nombre de contextes	Nombre d’erreurs	Infinitif	Participe passé	Nombre moyen d’erreurs par apprenant
L3 (n=23)	179	9	3	6	0.39
L4 (n=28)	172	5	3	2	0.18
Total	351	14	6	8	0.27

Les apprenants chinois de ce corpus ont produit des erreurs typiques du français L2, telles que la confusion entre l’infinitif et le participe passé pour des verbes réguliers du premier groupe<sup>58</sup>, ou l’utilisation de la forme du présent au singulier à la place du participe passé et de l’infinitif. Il nous semble important de rappeler que la confusion entre l’infinitif et le participe passé pour des verbes réguliers ne pourrait pas s’expliquer tout simplement par le fait que l’apprenant ne fait pas la distinction entre les différentes formes verbales ayant la même image phonique. Dans notre corpus, la confusion entre l’infinitif et le participe passé pour des verbes irréguliers, ainsi que l’utilisation de la forme du présent au singulier à la place du participe passé et de l’infinitif pour des

<sup>58</sup> Ce type d’erreur est également fréquent en français L1, voir par exemple Jaffré et Brissaud, 2006.

verbes réguliers sont également observées et ces erreurs ont modifié en effet l'image phonique du verbe donné. Il nous semble que ces erreurs surgissent lorsque l'apprenant ne dispose pas d'une capacité attentionnelle suffisante dans la production écrite. Ainsi, l'apprenant n'est pas en mesure soit d'accomplir le traitement morphologique séquentiel, spécifique aux apprenants L2, soit de récupérer directement la forme verbale congruente dans la mémoire. Par conséquent il a recours à des formes par défaut, correspondant plutôt des formes de base fréquentes, comme l'infinitif, le participe passé ou la forme du présent singulier. Sans une maîtrise préalable de l'oral, l'apprenant pourrait choisir une finale altérant l'image phonique de la forme verbale. Dans notre corpus, nous avons trouvé que l'apprenant a produit la forme du passé simple, une forme conjuguée et peu fréquente à la place du passé composé. Ce choix assez surprenant s'explique probablement par le fait que le texte de source utilisé dans la présente étude comporte de nombreux usages du passé simple. Ces formes pourraient, par conséquent, être particulièrement activées puis réutilisées comme forme par défaut par les apprenants (Thomas, 2017).

#### 4.3.3.3.2. Erreur de sur-généralisation

Lorsque l'apprenant se trouve confronté à la difficulté de réaliser la forme verbale, il a recours soit à des formes par défaut, soit à des formes créatives. Dans le second cas, l'apprenant tente de créer des formes verbales en suivant le modèle de conjugaison qu'il connaît déjà. Ainsi surgissent les erreurs de sur-généralisation. Pour les apprenants de L2, 5 occurrences ont été trouvées et trois entre elles concernent des formes du participe passé. L'erreur ci-dessous relève de la sur-généralisation du modèle des verbes en *-er* :

(135) \* {Il}\_ {Ils} se sont [riait] {rié}\_ {ri}. (L2-02-007)

Dans la catégorie de la fausse analogie, les apprenants de L3 ont produit 12 erreurs, dont 8 concernent les formes verbales du passé simple. Le temps du passé simple est introduit aux apprenants chinois à partir de L3 et par conséquent, les formes verbales de cette catégorie n'étaient pas encore très présentes dans l'input des apprenants de L3.

Ces formes étaient ainsi moins productives. Cependant, les apprenants étaient créatifs et ils ont tenté de créer des formes en suivant le modèle de conjugaison qu'ils connaissaient déjà, tel que celui du passé composé, ou celui du passé simple pour les verbes en *-er* :

- (136) a. \* Il déclarait que la personne qui la {rendut}\_ {rendit} rire allait être ~bon~ chevalier. (L3-023)
- b. \* quand il chassa dans {le}\_ {la} forêt, il rencontra le roi ~Arthur~ et {son}\_ {sa} troupe qui {les}\_ {le} {suiva}\_ {suivit}. (L3-013)

Pour les apprenants de L4, nous avons identifié 21 graphies erronées de sur-généralisation et elles sont plus hétérogènes, couvrant les formes du passé simple, du conditionnel et de l'imparfait. La sur-généralisation du modèle des verbes réguliers est également observée pour ce groupe d'apprenants. Étant donné que les verbes réguliers, notamment les verbes en *-er*, sont les plus nombreux dans l'interlangue des apprenants L2, ils sont les plus accessibles et les plus productifs. Ainsi leur modèle pourrait s'appliquer aux verbes des autres modèles de conjugaison (Santens, 2016) :

- (137) a. \* Perceval {combattr}\_ {combattit} avec le chevalier Vermeil et donna les armes au roi et puis partit. (L4-015)
- b. \* Ils se {combattaient}\_ {combattaient} dans la forêt. (L4-019)

Un autre type d'erreur de sur-généralisation concerne la propension à faire l'accord sujet-verbe pour les formes verbales homophones en /E/. Cette tendance est très récurrente chez les élèves français : la propension à sélectionner une forme verbale en *-é* accordée en nombre ou en genre avec le sujet augmente de la 3<sup>e</sup> à la 4<sup>e</sup> année de primaire, et se stabilise puis diminue entre la 3<sup>e</sup> et la 4<sup>e</sup> année du secondaire (Brissaud, Chevrot & Lefrançois, 2006). Dans notre corpus, les graphies erronées relevant de la sur-généralisation de l'accord sont très rares, et nous n'avons trouvé que deux occurrences :



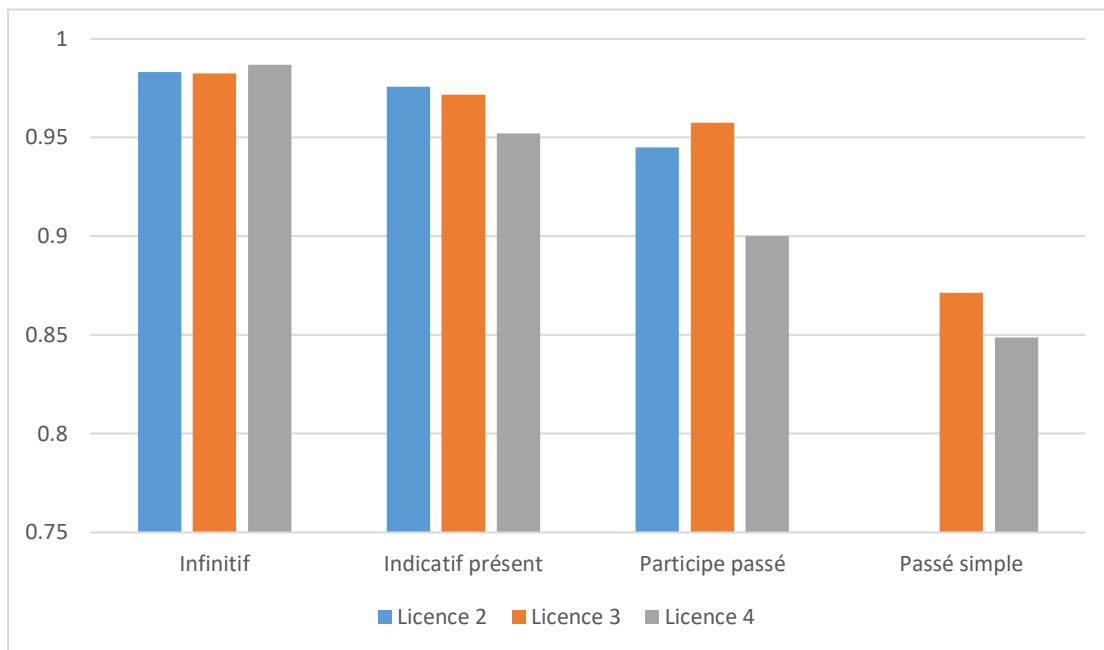
(138) a. \* La guerre a {commencée}\_ {commencé}. (L2-01-16)

b. \* Sa mère lui a {élevée}\_ {élevé}. (L3-09)

L'hypothèse la plus probable susceptible d'expliquer ce fait est que la solidarité syntagmatique existant en français entre le verbe et son sujet n'est pas prise en compte chez les apprenants chinois lorsqu'ils traitent la morphologie verbale. A la différence des enfants francophones, l'apprenant chinois met en œuvre l'accord sujet-verbe en suivant strictement les règles enseignées explicitement : accord pluriel de l'imparfait, du participe passé avec *être* et du participe passé avec *avoir* (accord avec objet antéposé). Quand l'apprenant est mentalement surchargé et n'est plus en mesure de gérer l'application de ces règles, les erreurs surgissent et elles portent essentiellement sur l'omission de marques d'accord.

#### 4.3.3.3.3. Le profil développemental des apprenants de trois niveaux linguistiques différents

La figure suivante nous permet de visualiser l'évolution du pourcentage de réussite pour les différentes catégories des formes verbales. En ce qui concerne la réalisation des formes verbales de l'infinitif, de l'indicatif présent et du participe passé, le taux de réussite est de plus de 90% pour les apprenants de tous les trois niveaux linguistiques. Cela nous montre que les connaissances de la morphologie verbale sont bien maîtrisées par les apprenants et que ces derniers sont alors en mesure de les mettre en œuvre dans une production écrite pour la plupart du temps. Nous constatons que l'infinitif est le mieux maîtrisé par rapport aux autres catégories des formes verbales. Et cela s'explique probablement par le fait que l'infinitif est « la forme de référence pour désigner un verbe » (Santens, 2016). En revanche, le passé simple correspond à la catégorie verbale la moins maîtrisée, représentant un taux de réussite de plus de 80%. Notons que les verbes du passé simple sont beaucoup moins fréquents dans l'input des apprenants et entraînent ainsi un faible degré de représentation en mémoire.



**Fig. 19** : Évolution du pourcentage de réussite pour les différentes catégories des formes verbales en fonction des niveaux linguistique différents

Dans notre corpus, la maîtrise précoce de la morphologie verbale s'explique sans doute par le fait que les apprenants chinois sont plus âgés et disposent d'une capacité cognitive assez avancée. Contrairement aux enfants français, qui cherchent à faire des tentatives de catégorisation lors de leur construction du système écrit, les apprenants chinois ont déjà construit le système écrit et peuvent transférer leurs connaissances linguistiques préalables en étudiant une langue étrangère. Avec un enseignement explicite et un entraînement intensif, ils sont en mesure de faire une analyse morphosyntaxique suffisante lors de la sélection des formes verbales. Néanmoins, les erreurs persistent dans la production des apprenants et représentent une variété importante.

La performance des apprenants chinois est liée aussi, nous semble-t-il, à la tâche de production choisie par cette étude : il s'agit d'une production écrite. Ågren (2005) explique l'écart entre les deux registres, oral et écrit, par la différence de temps de réaction :

« En écrivant, le scripteur a le temps d'élaborer son message, de s'arrêter

pour réfléchir, de se corriger ou de se compléter. Ainsi, le produit écrit final a un caractère fini et construit. À l'oral, l'élaboration et l'émission du message se font presque simultanément, ce qui résulte en reprises, hésitations ou ruptures de construction laissant des traces dans le message » (Ågren 2005 : 28).

Ainsi dans la production écrite, l'apprenant a plus de temps pour réfléchir avant de mettre en œuvre la forme verbale. Si la forme verbale est faiblement représentée en mémoire chez l'apprenant, elle ne pourra pas être récupérée directement. Afin de réaliser la forme verbale congruente, l'apprenant pourrait toujours appliquer les règles de conjugaison ou faire référence à d'autres stratégies graphiques.

Bien que les apprenants prouvent une bonne maîtrise de la morphologie verbale après un an d'étude du français, il faut indiquer que leur profil développemental n'est pas net. En d'autres termes, les formes erronées ne diminuent pas du stade initial au stade assez avancé. Les données de notre corpus ne confirment pas les résultats des autres recherches en français L2, selon lesquelles les apprenants du français L2 démontrent clairement un profil développemental évident en réalisant la morphologie verbale (voir notamment Ågren, 2005 ; Bartning & Schlyter, 2004). Il nous semble qu'après trois ans d'études de français, les formes verbales sont encore loin d'être stables chez les apprenants chinois. En d'autres termes, elles ne sont pas complètement intégrées dans le système lexical des apprenants. Et les erreurs surgissent lorsque les apprenants ne disposent pas d'une capacité attentionnelle suffisante pour récupérer directement les formes verbales ou gérer l'application de règles de conjugaison.

Rappelons un scénario développemental intéressant à partir de l'observation des graphies erronées sans altération de valeur phonique. Pour les apprenants de L2, les formes verbales de l'indicatif présent sont introduites en premier temps, puis les formes du participe passé. La majorité des erreurs concernant les formes du passé composé relèvent de l'ajout de lettres finales. Ce type d'erreur pourrait s'expliquer par l'interférence des formes de l'indicatif présent, qui sont fréquentes dans l'input et mieux maîtrisées par les apprenants. Au niveau suivant, avec l'introduction du passé

simple, la détermination de l'orthographe des formes verbales devient plus compliquée et une tendance d'omettre la lettre finale des formes du passé simple est observée : *il {voulu}\_{voulut} aller demander les armes au roi.* (L3-011). L'hypothèse possible d'expliquer ce fait est que les apprenants n'ont pas encore eu une bonne maîtrise de formes verbales du passé simple et qu'ils ont ainsi fait référence aux formes conjuguées du participe passé, qui étaient fréquentes dans l'input et assez bien maîtrisées par les scripteurs de ce niveau. Si les erreurs relevant des formes du passé composé réduisent significativement en L3, elles reviennent dans la production des apprenants de L4 : une nouvelle tendance d'ajouter une lettre finale *-t* aux formes du participe passé se manifestent. Ce souci de marquer la troisième personne du singulier provient probablement des formes verbales du passé simple, qui deviennent plus fréquentes pour les apprenants de ce niveau, et qui sont omniprésentes dans le texte de source.

Cette analyse des erreurs illustre bien le fait que les formes verbales du passé composé sont loin d'être stables dans le système lexical des apprenants chinois, bien que ces derniers soient en mesure de les mettre en œuvre de manière productive. Avec l'avancement dans l'enseignement du français, les erreurs relevant de la catégorie du passé composé diminuent en Licence 3. Néanmoins, les écrits de L4 présentent une grande diversité de formes verbales au niveau de temps et avec l'interférence des formes du passé simple, les erreurs concernant le passé composé reviennent dans la production des apprenants de L4.

Cela explique probablement pourquoi le nombre des erreurs en ce qui concerne les formes verbales ne diminue pas significativement au fil de la scolarité. À l'aide d'un enseignement explicite et d'un entraînement intensif, les apprenants chinois, qui sont plus âgés et disposent alors d'une capacité cognitive plus développée, présentent une bonne maîtrise des formes verbales dès le stade initial. Néanmoins, pour les apprenants chinois, qui apprennent la langue française en s'appuyant principalement sur l'enseignement en classe, les différentes flexions d'un verbe ne sont pas complètement intégrées dans son entrée lexicale et par conséquent, elles ne pourraient pas être automatiquement disponibles dans la production. Ainsi le processus du traitement morphologique devient-il coûteux en attention et les erreurs surgissent lorsque les

apprenants ne disposent pas d'une capacité attentionnelle suffisante. Pour les apprenants de L3 et de L4, ils pourraient au fur et à mesure produire les textes plus longs et les structures verbales plus compliqué. Cependant, le temps du passé simple est introduit dans le texte de source à partir de la Licence 3. Ainsi le coût de la sélection de terminaisons verbales devient-il plus élevé et une grande partie de l'attention de la part des apprenants est demandée. La persistance des erreurs observée dans notre corpus pourrait s'expliquer par ces divers facteurs, à savoir, la représentation lexicale et le développement langagier spécifique aux apprenants L2, ainsi que les caractéristiques du texte de source utilisé dans la présente étude.

#### **4.4. Pistes didactiques**

Les analyses dans ce chapitre empirique nous permettent de mieux comprendre comment les apprenants appréhendent le système linguistique du français par le biais de leur apprentissage des formes orthographiques. Comme Gerolimich et Stabarin (2017) l'évoquent :

« Chaque population d'apprenants affronte, en fonction des spécificités de sa langue maternelle, des défis différents, qui font ressortir les spécificités du français (liaisons, différences phono- orthographiques et morphologie silencieuse, richesse de la flexion verbale, multifonctionnalité de certaines structures, fonctions de l'imparfait, etc.). Cette approche contrastive offre un intérêt aussi bien pour la conceptualisation des phénomènes linguistiques que pour une application didactique différenciée. »

A la lumière des résultats ressortis de ce chapitre, nous souhaitons par la suite quelques nouvelles propositions didactiques pour un enseignement du français L2 qui serait plus proche de son emploi effectif et plus orienté vers les apprenants chinois.

##### **4.4.1. Erreurs orthographiques altérant la valeur phonique**

Dans le corpus du présent travail, nous avons constaté un grand nombre d'erreurs orthographiques altérant la valeur phonique. Ce type d'erreur, bien que classé et repéré parmi les erreurs orthographiques, ne concerne pas ces dernières, en raison de la différence de nature avec d'autres types d'erreurs orthographiques au sens strict : il nécessite une intervention autre que l'ordre orthographique (Katoozian, 2014). Comme Cogis (2005 : 18) le souligne dans sa recherche, ce type d'erreur nécessite une correction phonétique, ou un apprentissage plus approfondi de la prononciation. Étant donné le travail de précision suscité par l'apprentissage de la prononciation, les participants apprennent « à travailler de manière précise et à développer leurs capacités

d'observation » (Dufeu, 2008). Cette précision dans l'observation peut être transférée dans le domaine de l'orthographe. De ce fait, il est souhaitable qu'un enseignement approfondi de la phonétique française et de la phonologie soit introduit dès les niveaux élémentaires de l'apprentissage du français L2.

Néanmoins, à la situation chinoise, l'enseignement du français L2 au milieu universitaire, qui accorde depuis toujours une place plus importante à l'écrit qu'à l'oral, est dû partiellement à un héritage des traditions chinoises dans le domaine de l'enseignement de langues étrangères. Nous constatons que la transmission des connaissances langagières et les échanges verbaux dans les classes se passent essentiellement à travers un manuel. Ainsi les activités orales sont-elles souvent dépendantes d'une source écrite. Cette importance de l'écrit se marque également dans le système d'évaluation. Comme dans le test national TFS4, les candidats font tous les sujets par écrit et n'ont pas l'occasion de faire face à un jury, de répondre à ses questions ou de réagir à ses remarques. Notons que cette prééminence de l'écrit se reflète dans la difficulté de perception et de reproduction orale des apprenants chinois du français L2. Même si actuellement la compétence orale est considérée comme importante dans l'enseignement-apprentissage de la situation chinoise, peu de place est accordée à la prononciation et surtout « les moyens investis ne sont pas à la hauteur des objectifs » (Dufeu, 2008).

#### **4.4.1.1. Les lacunes des manuels**

Nous remarquons un besoin de présenter des exercices de prononciation dans les manuels de français édités en Chine. Notons que ces manuels commencent toujours pour le débutant chinois par un entraînement intensif de la phonétique et des correspondances phonèmes-graphèmes. Néanmoins, leur présentation laisse souvent supposer qu'il s'agit plus de répondre à des exigences didactiques qu'aux véritables besoins des apprenants (Dufeu, 2008).

Tout d'abord, l'examen des exercices proposés par différents manuels démontre une absence d'analyse précise des erreurs. Dans le domaine des sons, chacun est présenté selon le lieu d'articulation, selon le degré d'aperture (pour les voyelles), et selon le mode

de formation (pour les consonnes). Par exemple, les sons /p/ et /b/ sont présentés en paire. Comme consonnes labiales, ils se prononcent avec les lèvres. Puis pour produire ces consonnes occlusives, on est obligé de bloquer le passage de l'air en fermant complètement la bouche et ensuite de le relâcher. Les deux sons s'opposent l'un à l'autre selon le paramètre sourd/sonore, c'est-à-dire la présence ou l'absence de vibrations des cordes vocales. En dernier lieu sont présentées leurs correspondances en graphèmes (*En route*, 2017 : 4). Si les caractéristiques de ces deux consonnes sont bien expliquées en détail, il n'est pas tenu compte, en revanche, de leur niveau de difficulté et de leur possibilité de correction en fonction de leur environnement et de leur place dans la chaîne sonore. Soulignons que l'opposition de voisement pose des difficultés aux apprenants chinois, puisque la plupart de consonnes en chinois standard sont sourdes ou sourdes aspirées. Autrement dit, les apprenants chinois sont confrontés à la discrimination entre les deux sons au niveau de la perception et de la reproduction, notamment s'il s'agit de son /p/ se trouvant devant les voyelles. Quant aux exercices, les manuels se contentent souvent de proposer des phrases ou des mots en opposition qu'on écoute sur le CD qui accompagne le manuel. Par exemple, l'exercice ci-dessous invite l'apprenant à lire les mots suivants et à discriminer les différents sons (*En route*, 2017 : 5) :

/p/	pâle	peste	taper	pile	pape
/b/	bal	label	abbé	bile	barbe

Néanmoins, nous n'avons pas trouvé d'indication ou de véritable développement d'exercice qui favorise l'écoute et la perception de la discrimination ou d'analyse précise ciblée à cette difficulté spécifique aux apprenants chinois.

Deuxièmement, les exercices de prononciation sont souvent proposés d'une manière fort succincte. Dans le cadre d'une leçon, on présente un exercice sur un phénomène de prononciation : par exemple, la liaison de la langue française, puis on passe à la leçon suivante avec un autre thème, tel que l'enchaînement, « laissant supposer que le problème abordé précédemment est résolu » (Dufeu, 2008). Ainsi, avec cet assemblage de phénomènes de prononciation et d'exercices juxtaposés, avons-nous l'impression que la



conception de progression est absente. Sans une conception globale, les exercices de prononciation sont proposés isolément et ne permettent pas une intégration dans l'ensemble du processus d'apprentissage du français.

#### **4.4.1.2. Le manque de formation des enseignants**

A la situation chinoise, une des raisons de la primauté de l'écrit dans l'enseignement-apprentissage du français peut venir du fait que les enseignants ont été « plus préparés à travailler autour de textes et d'exercices écrits qu'à animer des activités orales » (Dufeu, 2008). Soulignons que le problème principal réside alors dans la formation des enseignants : d'une manière générale, on enseigne peu les connaissances concernant les caractéristiques prosodiques (rythme et mélodie) de la langue française et encore moins sur les moyens de les transmettre (Dufeu, 2008). Ainsi, ces enseignants du français L2 ne disposent-ils pas de connaissances théoriques suffisantes pour orienter et éclairer leur pratique pédagogique. De ce fait, les pratiques employées dans les cours de français se réduisent souvent à une présentation confinée des sons du français, faisant essentiellement appel à la méthode articulatoire. En outre, peu de place est accordée aux moyens servant à améliorer la perception des apprenants. Les enseignants ne disposent pas de pratiques nécessaires pour créer les conditions d'écoute pour que les apprenants puissent percevoir les différences entre L1 et L2.

Remarquons également qu'en Chine, l'enseignement-apprentissage du français au milieu universitaire se limite souvent au domaine de la phonologie. L'attention des enseignants est souvent orientée vers la perception et la production des sons du français, et on accorde peu de place au rythme de la langue française. Il n'est pas superflu d'indiquer que le rythme constitue la base prosodique d'une langue et que c'est en fonction de ce rythme que les sons de cette langue sont prononcés (Dufeu, 2008). Mais la transmission des caractéristiques rythmiques du français aux apprenants chinois nous semble importante, puisqu'elle facilite leur apprentissage des sons de cette langue. Néanmoins, les gestes pour corriger le rythme, la mélodie ou les sons « doivent être continuellement adaptés aux réactions des apprenants » (Dufeu, 2008). Et cela nécessite

une expérience, qui, d'une manière générale, n'est pas transmise lors de la formation des enseignants du français.

Un autre élément pour expliquer la place dominante de la langue écrite dans l'enseignement-apprentissage du français pourrait venir du fait que, dans les universités chinoises, la plupart des cours du français sont assurés par les enseignants chinois. Il est fort probable que beaucoup d'enseignants ne sont pas « clairement conscients de leurs propres erreurs de prononciation » (Dufeu, 2008). De ce fait, ils ne peuvent remarquer que partiellement les erreurs de prononciation chez les apprenants. Il est donc compréhensible que les enseignants chinois hésitent à s'engager dans le domaine de la prononciation, dans lequel ils éprouvent eux-mêmes un sentiment d'insécurité, dû au fait qu'ils ressentent leurs déficiences sans pouvoir cerner toutes les erreurs de prononciation commises par leurs apprenants.

#### **4.4.1.3. Pistes didactiques**

Étant donné le terrain didactique chinois actuel et le nombre important des erreurs orthographiques altérant la valeur phonique, il est souhaitable d'introduire un enseignement plus approfondi de la phonétique française (exercices articulatoires) et de la phonologie (exercices d'opposition) dans les niveaux élémentaires de l'apprentissage du français (Katoozian, 2014), et cet enseignement devrait être accompagné de l'initiation des enseignants à la phonétique et à la phonologie (Mahmoudian, 1976).

Les enseignants pourraient alors aider les apprenants à prendre conscience du système phonologique de la langue chinoise, à travers une description parallèle et contrastive des systèmes phonologiques du français et du chinois. Cela pourrait permettre aux apprenants chinois de remarquer les dissemblances existant entre les deux systèmes phonologiques et de prévoir les zones à risque (Katoozian, 2014). A partir des analyses précises des erreurs de prononciation, on devrait créer les conditions nécessaires, qui habituent les apprenants chinois « à la fois à entendre et émettre correctement les phonèmes du français » (Katoozian, 2014).

Katoozian (2014) indique également que la pratique régulière de la dictée

traditionnelle enregistrée par une voix native joue un rôle positif pour améliorer la perception des apprenants L2. Les recherches menées auprès des apprenants de l'anglais L2 démontrent clairement l'effet positif et significatif de la pratique de la dictée sur l'amélioration du niveau de la compréhension orale (voir, par exemple, Kiyani & Shiramiry 2002 ; Farshid & Farshid 2010).

L'enseignement-apprentissage sur la discrimination auditive, tout comme sur l'articulation pourrait se faire à travers les exercices spécifiques, ciblés sur les difficultés particulières confrontées par les apprenants. Par exemple, Rey-von Allemen (1982 : 296-301) propose une série d'exercices d'opposition, qui pourraient être appliqués avant ou après la dictée. Il s'agit de la perception des oppositions entre les consonnes (/v/~b/ : *viens vite / bien vite*), les semi-voyelles (/l~/j/ : *vous allez à pied / vous alliez à pied*), les voyelles (/a~/ã/ : *vous m'attendez ? / vous m'entendez ?*) et le rythme (/a~/aa/ : *tu l'as pris ? / tu l'as appris ?*) (Katoozian, 2014). Étant donné la spécificité des deux systèmes phonologiques en présence, les exercices d'opposition pourraient viser d'un côté la distinction des voyelles moyennes [e] [ɛ] [o] [ɔ], du fait de l'existence des voyelles diphtongues en chinois : [ei] , [ai] , [ou] , [au], et de l'autre, l'opposition de voisement, étant donné que le chinois dispose peu de consonnes sonores, etc.

Finalement, le contact permanent avec la forme orale du français joue également un rôle significatif. A ce sujet, Amokrane (2011) suggère de « multiplier les occasions de production spontanée », afin de « permettre aux apprenants de mettre leurs connaissances phonétiques en pratique ». Pour les gestes de correction, Champagne-Muzar *et al.* (1998) indiquent qu'« une correction immédiate et ponctuelle risque d'interrompre la transmission d'un message, nuisant ainsi à la spontanéité d'expression de l'apprenant ». De ce fait, il est souhaitable que les enseignants notent les erreurs de prononciation commises par les apprenants, de manière à y revenir juste après la production spontanée, et que les enseignants interviennent immédiatement après et demandent aux apprenants de reprendre les éléments qui leur ont posé problème.

#### 4.4.2. Erreurs orthographiques relevant de la morphologie verbale

L'acquisition de la morphologie verbale constitue l'un des fondements à la communication en français L2 et un nombre important d'erreurs orthographiques relevant de la morphologie verbale a été identifié dans notre corpus. Cette partie a pour objectif d'élaborer quelques pistes de réflexion afin de faciliter l'apprentissage de la morphologie verbale.

##### 4.4.2.1. Oral et écrit pour l'apprentissage de conjugaisons verbales

Si les verbes du premier groupe présentent une morphologie relativement régulière et par conséquent facile à acquérir, la diversité des paradigmes des verbes du deuxième groupe et du troisième groupe pose des difficultés aux apprenants chinois du travail présent. A part les erreurs relevant des terminaisons verbales, les graphies erronées se situent également au niveau de la réalisation « des différentes formes que peut revêtir le radical (ou base lexicale) du verbe conjugué » (Nouveau, 2017). Plus précisément, nous constatons que les apprenants respectent l'accord de personne et la désinence du temps. Les erreurs découlent, en revanche, du choix erroné de la base lexicale.

Dans les cours de français à la situation chinoise, la didactique de la conjugaison s'appuie généralement sur la graphie et les tableaux classiques de conjugaison. Nouveau (2017) commente cette pratique pédagogique d'une façon suivante :

« Privilégiant le français écrit au détriment du français parlé, ils (les tableaux classiques de conjugaison) visualisent les flexions graphiques des personnes, négligent les bases lexicales orales dans la formation des temps et confèrent peu d'importance aux véritables indicateurs de la personne, les pronoms personnels, éléments antéposés aux verbes. » (Nouveau, 2017)

Selon Nouveau, il serait souhaitable de « tirer parti de la morphologie verbale » afin de démontrer le décalage entre la graphie et la phonie. Cela permettrait, par conséquent, aux

apprenants de conscientiser les graphèmes muets. L'examen des lettres muettes dans les formes verbales met en relief l'homophonie, ainsi que l'association étroite entre le verbe et le pronom clitique, ce dernier servant de véritable indicateur de la personne.

#### 4.4.2.2. Classification de verbes

Les erreurs produites par les apprenants chinois pourraient également résulter d'une systématisation incohérente des classes verbales. Dans l'enseignement du français L2 à la situation chinoise, une approche traditionnelle est souvent utilisée, qui offre une classification verbale tripartite : 1<sup>er</sup>, 2<sup>e</sup> et 3<sup>e</sup> groupes en fonction de la désinence infinitive de verbes. Toutefois, comme Riegel et *al.* (1997, 468) le remarquent, « il ne suffit pas de connaître l'infinitif du verbe pour élaborer automatiquement une forme verbale en ajoutant au radical une désinence : la régularité apparente des infinitifs masque les multiples variations des radicaux de nombreux verbes français ».

Cependant, il existe un deuxième type de classification verbale, initié par Dubois (1967) dans sa *Grammaire structurale du français*, en considérant les variations du radical verbal comme indices majeurs de systématisation. En suivant l'idée de Dubois, Gérolimich et Stabarin (2007 : 155) proposent une classification en 5 groupes verbaux qui illustre à la fois le nombre de bases lexicales et la distribution de ces bases en fonction de personne :

**Tableau 49.** Classification verbale de Gerolimich et Stabarin (cité dans Nouveau, 2017)

<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>
<b>1 base</b>	<b>2 bases</b>	<b>2 bases</b>	<b>2 bases</b>	<b>3 bases</b>
	Alternance vocalique ou base 2 allongée par une semi-consonne	Base 2 allongée par une consonne épenthétique ou latente	Alternance vocalique et base 2 allongée	
- parler - ouvrir	- appeler, jeter - essayer	- finir - partir	- craindre - savoir	- prendre - venir

	- autres : <i>voir</i> , <i>distraindre</i> , <i>etc.</i>	- <i>connaître</i> - autres : <i>lire</i> , <i>etc.</i>	- autres : <i>hair</i>	- <i>vouloir</i> - autres : <i>boire</i> , <i>etc.</i>
[parl]	[apɛl]/ [apəl]	[fini]/[finis]	[krɛ̃]/ [krɛ̃ɲ] [sɛ]/[sav]	[prɑ̃]/ [prən]/[prɛn]
[uvr]	[esɛ]/[esej] [vwa]/[vwaj]	[par]/[part]		[vjɛ̃]/ [vən]/ [vjɛn]
P1-2-3-6 : Base	P1-2-3-6 : Base 1	P1-2-3 : Base 1	P1-2-3 : Base 1	P1-2-3 : Base 1
P4 : Base+[ɔ̃]	P4 : Base 2+[ɔ̃]	P4 : Base 2+[ɔ̃]	P4 : Base 2+[ɔ̃]	P4 : Base 2+[ɔ̃]
P5 : Base+[e]	P5 : Base 2+[e]	P5 : Base 2+[e]	P5 : Base 2+[e]	P5 : Base 2+[e]
		P6 : Base 2	P6 : Base 2	P6 : Base 3

Selon Nouveau (2017), une telle systématisation met en relief deux constantes importantes des paradigmes verbaux :

« D'une part que pour quatre personnes du paradigme, les personnes du singulier et la troisième du pluriel (P1,2,3,6), la forme verbale correspond systématiquement à la ou les base(s) nue(s) du verbe (désinence muette), et d'autre part, que seules deux désinences sont prononcées ([ɔ̃] à la 1<sup>ère</sup> personne pluriel (P4) et [e] à la 2<sup>e</sup> personne pluriel (P5). » (Nouveau, 2017)

Une telle représentation des paradigmes verbaux n'est pas sans intérêt pour l'enseignement de conjugaison. Pour chaque paradigme verbal, l'enseignant pourrait visualiser les formes orales transcrites en API, tout en les liant aux formes orthographiques correspondantes. Cette pratique permettrait de mettre en exergue « l'importance de la base lexicale en tant que squelette verbal » (Nouveau, 2017).

#### 4.4.2.3. Utilisation des corpus dans l'apprentissage de la morphologie verbale

En français, de nombreux corpus pédagogiques sont créés, s'accordant spécialement aux apprenants L2, tels que *Sacodeyl*, *Fleuron*, *Lancom*, etc. Ces ressources peuvent être adaptées par les enseignants de français en Chine, afin de répondre à des besoins pédagogiques spécifiques.

Les apprenants chinois sont souvent confrontés pendant leur apprentissage du français à une pauvreté de l'input en termes de quantité et de qualité. Le corpus peut alors leur proposer une alternative à l'enseignement traditionnel en classe, avec l'accès libre et direct aux données authentiques, massives, structurées et multicontextuelles. Les enseignants peuvent également puiser dans les ressources afin de renouveler la didactique de la conjugaison.

Généralement, les corpus pédagogiques offrent une structure de consultation de leurs données sonores, ainsi que des transcriptions orthographiques. Nouveau (2017) décrit la consultation de données de façon suivante :

« La consultation de corpus numérisés permet l'observation de l'homophonie des formes écrites. Les moteurs de recherche génèrent des listes d'occurrences ciblant des besoins spécifiques. Il suffit de lancer des requêtes telles que *je finis*, *tu finis*, *(il) finit*, etc. dans le moteur de recherche ... pour extraire des énoncés sonores produits par une variété de locuteurs dans diverses situations conversationnelles. » (Nouveau, 2017)

Une telle consultation de données permet aux apprenants de conscientiser l'écart entre les formes orales et écrites des paradigmes verbaux, d'observer l'homophonie des formes verbales. Avec la consultation du corpus, on peut également démontrer aux apprenants le rôle majeur du pronom clitique dans les constructions verbales. Comme l'exemple 139 l'illustre, des exercices peuvent être basés sur les paires verbales

homophones P3 et P6. En écoutant les énoncés sonores, les apprenants peuvent noter la liaison obligatoire en [z] qui apporte l'information du nombre et différencie la forme verbale P3 [iluvr] de la forme P6 [ilzuvr]. De plus, les apprenants peuvent par la suite documenter ce phénomène linguistique et construire leurs propres corpus de formes verbales.

(139) a. Il découvre un trésor. / Ils découvrent un trésor.

b. Il ouvre la porte. / Ils ouvrent la porte.



## IV. CONCLUSION

### 5. Bilan et perspectives

L'objectif de cette recherche a été de mettre l'accent sur les performances orthographiques des apprenants chinois en français L2. Pour ce faire, les données résultant d'une tâche écrite semi-spontanée accomplie par des apprenants chinois suivant des cours de français dans l'enseignement supérieur chinois, ont été examinées. Dans ce cadre, nous cherchions à présenter une étude à la fois descriptive et comparative. D'une part, l'analyse empirique a porté essentiellement sur la réalisation orthographique de deux domaines : l'orthographe lexicale et l'orthographe grammaticale. D'autre part, en insistant sur le développement orthographique en acquisition L2, nous avons comparé les performances des apprenants à trois niveaux linguistiques différents.

La langue écrite est « une source de communication puissante en classe de langue étrangère » (Ågren, 2005). Ainsi, il nous semble indispensable de mener une recherche dans une perspective linguistique et didactique, afin de mieux connaître le développement de la langue écrite en L2, ainsi que les divers facteurs qui jouent un rôle dans la mise en œuvre de la production écrite. Dans la présente étude, les résultats concernant les performances orthographiques des apprenants chinois ont été commentés et nous avons également proposé les hypothèses possibles pour interpréter les formes erronées, en mettant la réalisation orthographique en relation avec le développement langagier des apprenants L2, ainsi qu'avec le processus de la production écrite. Notre but a été d'identifier les différents traitements orthographiques adoptés par les apprenants chinois lors de leur production écrite et de mettre en lumière le processus d'apprentissage de l'écrit en français L2.

Dans l'introduction de la thèse, nous avons souligné que le questionnement essentiel de cette recherche était de décrire le développement orthographique du français L2 écrit, ainsi que les tendances interlangagières des apprenants chinois du

français L2. Nous souhaitons ainsi dans cette conclusion revenir aux principaux résultats obtenus dans ce domaine.

### **5.1. Le développement orthographique des apprenants chinois**

Dans la partie empirique, nous avons examiné dans un premier temps l'acquisition de l'orthographe des mots du lexique, en analysant les erreurs lexicales de trois catégories différentes, à savoir les erreurs orthographiques à dominante phonétique, les erreurs phonogrammiques, ainsi que les erreurs lexicales sans incidence phonique. Les erreurs orthographiques à dominante phonétique sont nombreuses chez les apprenants chinois<sup>59</sup>. De plus, ce type d'erreurs ne diminue pas en nombre au cours de la scolarité. En outre, les erreurs phonogrammiques et les erreurs lexicales sans incidence phonique sont beaucoup moins lourdes en poids relatif. Néanmoins, les erreurs phonogrammiques diminuent en nombre en fonction du niveau scolaire des apprenants, alors que les erreurs lexicales sans incidence phonique persistent et même augmentent légèrement en nombre au cours de la scolarité. Ainsi, nous constatons que le profil développemental ne diffère pas nettement entre le stade initial et le stade assez avancé.

Sans contact permanent avec l'image auditive correcte du mot et avec les dissemblances existant entre les deux systèmes phonologiques, il est fort probable que les informations phonologiques stockées à l'entrée lexicale chez les étudiants chinois soient gravement discréditées. Lors de la production écrite, les apprenants récupèrent à priori directement les formes orthographiques stockées à l'entrée lexicale. Si la forme des mots n'est que partiellement connue par l'apprenant, ou si l'apprenant est mentalement surchargé pendant la tâche d'écriture, la mobilisation de la forme orthographique serait alors difficile et l'apprenant doit par conséquent recourir à d'autres stratégies possibles afin de compenser des risques, telles que l'emploi des correspondances entre phonèmes et graphèmes, de l'analogie, des règles apprises implicitement. Néanmoins, l'emploi de ces stratégies aboutit à des solutions qui ne

---

<sup>59</sup> A la différence des apprenants chinois du français L2, ce type d'erreur disparaît rapidement chez les élèves français (voir par exemple Manesse & Cogis, 2007)

rencontrent pas toujours la graphie normée, ce qui explique la persistance des erreurs orthographiques à dominante phonétique observée dans notre corpus.

Le même phénomène est également observé si l'on examine la réalisation de l'orthographe grammaticale. En ce qui concerne le problème de l'accord en nombre et en genre, les apprenants chinois présentent une maîtrise précoce dès le stade initial : le taux de réussite est de presque 90% pour l'accord du nombre et de plus de 90% pour l'accord en genre. Cependant le nombre des erreurs ne diminuent pas significativement au fil des trois ans d'étude du français. En d'autres termes, le profil développemental ne varie pas significativement de la Licence 2 à la Licence 4.

**Tableau 50.** Taux de réussite pour le problème de l'accord

Stade	Accord en nombre			Accord en genre	
	Nom	Verbe	Adjectif	Déterminant	Adjectif
L2	89.6%	90%	86%	95.92%	99.06%
L3	90.2%	76%	96%	93.63%	96.02%
L4	91%	83%	90%	93.60%	95.15%

Rappelons que l'opposition singulier/pluriel ou l'opposition féminin/masculin n'existe pas dans le système graphique du chinois<sup>60</sup>. A l'aide de l'enseignement explicite et d'exercices écrits intensifs, les apprenants chinois du français L2 sont en mesure de mettre en œuvre l'accord en nombre et en genre à l'écrit de manière productive. Cependant, les informations morphologiques concernant le nombre ou le genre ne sont pas intégrées dans les entrées lexicales. En revanche, elles sont apprises et stockées hors du système lexical des apprenants, en tant que connaissances explicites. Ainsi la sélection de la flexion pertinente devient-elle un processus conscient et strictement séquentiel : sélection de la racine liée à l'item lexical à produire, et sélection de la flexion sur la base des connaissances morphologiques du scripteur. Le coût de la sélection de la flexion congruente est par conséquent élevé chez les apprenants chinois

---

<sup>60</sup> La notion du pluriel peut s'exprimer par certains quantifieurs en chinois.

et les erreurs surgissent lorsque les apprenants ne disposent pas d'une capacité attentionnelle suffisante pour gérer l'application de règles. En outre, en ce qui concerne le problème du genre, il faut indiquer que les associations entre les noms et les informations de leur catégorie du genre ne seraient pas aussi fortes dans le lexique mental en L2 que dans celui de L1. Ainsi les apprenants chinois se trouvent-ils confrontés à la difficulté d'attribuer le genre congruent à une forme nominale. Et une grande partie des erreurs relevant du problème du genre observées dans notre corpus s'explique par cela. L'ensemble de ces éléments liés avec la représentation lexicale spécifique aux apprenants L2 favorise alors la persistance des erreurs relevant du problème de l'accord.

Nous avons identifié le scénario développemental similaire lors de l'analyse des erreurs relevant des terminaisons verbales. Bien que les apprenants démontrent une bonne maîtrise de la morphologie verbale après un an d'étude du français, il faut noter que les graphies erronées ne diminuent pas du stade initial au stade assez avancé. Il nous semble qu'après trois ans d'étude du français, les formes verbales sont encore loin d'être stables chez les apprenants chinois. En d'autres termes, les différentes flexions d'un verbe ne sont pas complètement intégrées dans son entrée lexicale et par conséquent, elles ne pourraient pas être automatiquement disponibles dans la production. Comme nous l'avons évoqué auparavant, la sélection de la flexion pertinente est coûteuse en attention chez les apprenants chinois. Avec l'introduction du temps du passé simple dans le texte de source à partir de la Licence 3, la sélection des terminaisons verbales devient encore plus compliquée et par conséquent, les formes erronées persistent au fil de la scolarité.

La maîtrise précoce de la morphologie écrite constitue une autre caractéristique en ce qui concerne les performances orthographiques des apprenants chinois du français L2. Dès le stade initial, les apprenants sont en mesure de mettre en œuvre à l'écrit les marques flexionnelles de manière productive : le taux de réussite est à peu près de 90% tant pour l'accord en genre et en nombre que pour les formes verbales. Cela s'explique sans doute par le fait que les apprenants chinois sont plus âgés et disposent d'une capacité cognitive assez avancée. Contrairement aux enfants français, qui cherchent à

faire des tentatives de catégorisation lors de leur construction du système écrit, les apprenants chinois ont déjà construit le système écrit et peuvent transférer leurs connaissances linguistiques préalables en étudiant une langue étrangère. Avec un enseignement explicite et un entraînement intensif, ils sont en mesure de faire une analyse morphosyntaxique suffisante lors de la production écrite.

La performance des apprenants chinois est liée aussi, nous semble-t-il, à la tâche de production choisie par cette étude : il s'agit d'une production écrite. A l'écrit, l'apprenant a un temps suffisant pour réfléchir avant de mettre en œuvre la forme orthographique. Cela pourrait expliquer partiellement la bonne maîtrise de la morphologie française que les apprenants chinois présentent dans leur production. Ågren (2005) et Santens (2016) indiquent également dans leur recherche que la production écrite des apprenants du français L2 montre « une émergence relativement précoce de l'accord verbal morphologique » (Ågren, 2005 : 141).

Finalement, dans cette analyse des performances orthographiques, nous avons évoqué plusieurs différences et ressemblances en ce qui concerne la production et l'acquisition de l'orthographe en français L1 et L2. Rappelons que les enfants français disposent d' « un système basé sur un input majoritairement oral » (Ågren, 2005 : 140), tandis que les étudiants chinois abordent simultanément le français oral et écrit dès le début de leur étude. De surcroît, les apprenants chinois disposent des connaissances linguistiques des deux autres langues, à savoir le chinois et l'anglais. Au moment d'apprendre à orthographier en français, les apprenants chinois sont plus âgés et ont par conséquent une capacité cognitive plus élevée par rapport aux enfants français. Comme Ågren l'indique, « l'opération motrice d'écrire est depuis longtemps automatisée » dans leur langue maternelle (2005 : 140). L'analyse des performances orthographiques des apprenants chinois nous montre trois différences essentielles entre les apprenants en français L1 et L2. Tout d'abord, il semble que le côté silencieux de la morphologie française ne gêne pas les apprenants chinois, ce qui se voit à la fois sur l'accord en nombre et en genre, ainsi que sur les formes verbales homophones en /E/. Dans ces trois domaines, les apprenants chinois présentent une bonne maîtrise dès le stade initial. Deuxièmement, en ce qui concerne le genre et le déterminant, nous avons discuté du

statut différent du déterminant en acquisition L2 : en français L1, l'enfant apprend le genre des noms à l'oral, en s'appuyant essentiellement sur l'association très forte entre déterminant et nom. Néanmoins, ce lien fort n'existe pas en français L2 et nous avons identifié de nombreuses erreurs relevant de l'attribution du genre, qui sont en revanche assez rares dans la production du français L1. En dernier lieu, à partir de l'analyse de notre corpus, il est difficile d'identifier un ordre d'acquisition en ce qui concerne la morphologie écrite du nombre et du genre. La littérature sur la mise en œuvre de l'accord en français L1 met l'accent sur les étapes de l'acquisition. Par exemple, les apprenants apprennent la morphologie verbale du nombre plus tardivement que celle du nom. Cependant, cet ordre d'acquisition ne se présente pas dans notre corpus.

## **5.2. Perspectives**

L'objectif de cette étude exploratrice a porté sur la description d'un domaine linguistique jusqu'à maintenant assez peu connu dans les recherches du français L2 : les performances orthographiques des apprenants chinois du français L2 et leur développement orthographique dans la production écrite. Si tout scripteur a un long trajet à parcourir avant d'arriver à la maîtrise complète de l'orthographe du français, ceci est sans doute d'autant plus vrai lorsqu'il s'agit de s'approprier ce système français dans une langue étrangère, dont l'organisation graphique présente des bases complètement différentes de celles du français. A partir de ce travail de collecte de données, de constitution et d'analyse du corpus, nous avons souhaité apporter de nouvelles perspectives dans l'étude du français L2 écrit.

Étant donné les particularités du français écrit ainsi que le rôle majeur de la langue écrite dans l'enseignement du français L2, des études approfondies dans ce domaine sont nécessaires, en offrant « un intérêt aussi bien pour la conceptualisation des phénomènes linguistiques que pour une application didactique différenciée » (Gerolimich & Stabarin, 2017). Comme Ågren l'évoque, « de nombreux apprenants guidés de français L2 consacrent un temps considérable à apprendre à écrire en français ». Ainsi, il est important de mieux connaître les contraintes linguistiques et

cognitives qui se présentent à ces apprenants lors de leur apprentissage de l'orthographe.

Il n'est pas inutile d'indiquer ici que les analyses et les commentaires que nous avons présentés dans ce travail, bien que d'un grand intérêt, ne nous semblent pas suffisants pour répondre aux questions sur les procédures du traitement orthographique que les apprenants adoptent, leurs rôles respectifs et leur mode d'action. Les hypothèses proposées nécessitent une vérification pragmatique et les expériences opérationnelles devront être introduites dans la suite du travail.

En outre, rappelons que les apprenants du stade initial de notre corpus présentent déjà une bonne maîtrise de la morphologie écrite. Ainsi leurs productions ne nous permettent pas d'observer les tâtonnements orthographiques pratiqués par les apprenants chinois au début de leur apprentissage. Afin de mieux connaître le développement de la réalisation orthographique, un élargissement du corpus vers les apprenants les moins avancés semble nécessaire pour permettre des analyses affinées. De plus, il est souhaitable de réaliser d'autres recherches similaires auprès des échantillons plus larges d'apprenants chinois en français L2 et avec d'autres tâches d'écriture pour savoir si les résultats obtenus sont conformes à ceux de la présente étude.

Finalement, il nous semble très intéressant d'amener une comparaison systématique entre les productions orthographiques des apprenants chinois, celles des apprenants ayant une langue maternelle plus proche du français et celles des apprenants français. Notons que chaque population affronte des défis différents lors de l'apprentissage du français. Ainsi, les données issues de ce type de recherche ouvriront le champ vers des dimensions complémentaires dans le domaine de l'acquisition du français L2 et mettront en exergue les spécificités de la langue française.

# BIBLIOGRAPHIE

- Ågren, M. (2005). *Développement de la morphologie du nombre en français langue étrangère à l'écrit: Étude transversale* (Thèse de licentiat). Lunds universitet, Suède.
- Ågren, M. (2008). *À la recherche de la morphologie silencieuse : sur le développement du pluriel en français L2 écrit* (Thèse de Doctorat). Lund University.
- Ågren, M., & Hedbor, C. (2006). Acquisition et mise en oeuvre de la morphologie flexionnelle en français langue étrangère. *Rééducation Orthophonique*, (225), 159-189.
- Allal, L., Rouiller, Y., Saada-Robert, M., & Wegmuller, E. (1999). Gestion des connaissances orthographiques en situation de production textuelle. *Revue française de pédagogie*, 126.
- Alleton, V. (2010). *L'Écriture chinoise*. Paris: Presses Universitaires de France - PUF.
- Allmen, M. R.-V. (1982). *Apprentissage de l'orthographe française élémentaire par des adolescents non francophones enfants de travailleurs migrants: problèmes linguistiques et sociolinguistiques*. Peter Lang.
- Ameur-Amokrane, S. (2007). Pour améliorer la maîtrise de l'orthographe en FLE/FLS ... renforcer l'enseignement de la phonétique. *Annales de l'université d'Alger*, 17(1), 29-37.
- Ameur-Amokrane, S. (2011). Du lien entre la maîtrise du système phonologique et les compétences orthographiques. *TRANEL (Travaux neuchâtelois de linguistique)*, 54, 49-61.
- Anderson, J. R. (1983). *The Architecture of Cognition* (Reprint). Mahwah, NJ: Psychology Press.
- Anderson, J. R. (1993). *Rules of the Mind* (1 edition). Hillsdale, N.J.: Psychology Press.
- Anderson, J. R. (1995). *Learning and Memory: An Integrated Approach* (2nd Edition). Hoboken, NJ: John Wiley & Sons.



- Anne-Flore, B. (2012). *De l'origine de l'accent à l'enseignement de la phonétique en FLE* (Mémoire). Université Stendhal - Grenoble 3.
- Arabski, J. (1979). *Errors as indications of the development of interlanguage*. Katowice: Uniwersytet Slaski.
- Bailey, N., Madden, C., & Krashen, S. D. (1974). Is There a "Natural Sequence" in Adult Second Language Learning? *Language Learning*, 24(2), 235-243.
- Barbé, G. (2005). *Apprentissage d'une langue étrangère/secode: Parcours et stratégies de formation*. Bruxelles: De Boeck.
- Barbier, M.-L. (1997). *Rédaction de texte en langue première et en langue seconde* (Thèse de Doctorat). Université de Provence, Aix-en-Provence.
- Bartning, I. (2000). Gender agreement in L2 French: Pre-advanced vs advanced learners. *Studia Linguistica*, 54(2), 225-237.
- Bartning, I., & Schlyter, S. (2004). Itinéraires acquisitionnels et stades de développement en français L2. *Journal of French Language Studies*, 14(3), 281-299.
- Bassano, D., Maillochon, I., Klampfer, S., & Dressler, W. U. (2001). L'acquisition de la morphologie verbale à travers les langues, SUMMARY. *Enfance*, 53(1), 81-99.
- Beacco, J.-C. (2000). *Les dimensions culturelles des enseignements de langue* (01 éd.). Paris: Hachette.
- Béatrice, B., Christel, L., & Bourbon, C. (2004). Acquisition de l'accord en genre de l'adjectif en Français écrit. *Le Langage et l'Homme*, XXXIX(2), 35-43.
- Beck, M.-L. (1998). L2 Acquisition and obligatory head movement: English-Speaking Learners of German and the Local Impairment Hypothesis. *Studies in Second Language Acquisition*, 20(3), 311-348.
- Benazzo, S. (2005). Le développement des lectures d'apprenants et l'acquisition de la portée à distance en L2. *Acquisition et interaction en langue étrangère*, (23), 65-93.
- Bereiter, C., & Scardamalia, M. (Éd.). (1987). *The Psychology of Written Composition* (1 edition). New York London: Routledge.

- Betrix Koehler, D. (1993). Du bon usage de l'erreur dans une didactique de l'orthographe. In D. Bain, P. Perrenoud, & L. Allal (Éd.), *Evaluation formative et didactique du français* (p. 161-170). Neuchâtel: Delachaux et Niestlé.
- Blanche-Benveniste, C. (1990). *Le Français parlé: études grammaticales*. Éditions du Centre national de la Recherche scientifique.
- Bonin, P. (2002). *Production verbale de mots. Approche cognitive*. Bruxelles: De Boeck.
- Bosse, M.-L., Valdois, S., & Tainturier, M.-J. (2003). Analogy without priming in early spelling development. *Reading and Writing, 16*, 693-716.
- Bourdin, B., Leuwers, C., & Bourbon, C. (2011). Impact des contraintes linguistiques et cognitives sur l'acquisition de l'accord en genre de l'adjectif en français écrit. *Psychologie Française, 56*(3), 133-143.
- Brissaud, C., & Chevrot, J.-P. (2000). Acquisition de la morphographie entre 10 et 15 ans : le cas du pluriel des formes verbales en /E/. *Verbum, 22*(4), 425-439.
- 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. *Langue française, 151*(3), 74-93.
- Brissaud, C., & Cogis, D. (2002). La morphologie verbale écrite, ou ce qu'ils en savent au CM2. *Lidil, 25*, 31-42.
- Brissaud, C., & Jaffré, J.-P. (2003). Présentation : Regards nouveaux sur la lecture et l'écriture. *Revue française de linguistique appliquée, VIII*(1), 5-14.
- Brissaud, C., & Sandon, J.-M. (1999). L'acquisition des formes verbales en /E/ à l'école élémentaire et au collège, entre phonographie et morphographie. *Langue française, 124*(1), 40-57.
- Capelle, G., & Gidon, N. (2006). *Reflets*. Pékin: Foreign language teaching and research press.
- Carroll, S. (2006). Second-Language Acquisition and the Computational Paradigm\*. *Language Learning, 39*(4), 535-594.

- Catach, N. (1968). La bataille de l'orthographe aux alentours de 1900. In G. Antoine & R. Martin (Éd.), *Histoire de la langue française 1880-1914* (p. 237-251). Paris: CNRS Éditions.
- Catach, N. (1980). *L'Orthographe du français*. Paris: Nathan.
- Catach, N. (1982). Norme et orthographe dans la perspective du français langue étrangère. *Français Dans Le Monde*, 169, 32-37.
- Catach, N. (1991). *L'ORTHOGRAPHE EN DEBAT. Dossier pour un changement*. Paris: Nathan Université.
- Catach, N. (1995). Présentation. *Langue française*, 108(1), 3-9.
- Catach, N. (2001). *Histoire de l'orthographe française*. Paris: Honoré Champion.
- Catach, N. (2003a). *L'Orthographe* (9e éd.). Paris: Presses Universitaires de France - PUF.
- Catach, N. (2003b). *L'Orthographe française*. Paris: Nathan Université.
- Chamot, A. U., & Hatch, E. M. (1978). Grammatical problems in learning English as a third language. In *Second language acquisition: a book of readings*. Rowley, MA: Newbury House.
- Chervel, A., & Manesse, D. (1989). *Comparaison de deux ensembles de dictées 1873-1987. Méthodologie et résultats*. Paris: INRP.
- Chevrot, J.-P., Brissaud, C., & Lefrançois, P. (2003). Normes et variations dans l'acquisition de la morphographie verbale en /E/ : tendances, conflits de tendance, résolution. *Faits de langues*, 22, 57-66.
- Chini, M. (1995). Un aspect du syntagme nominal en italien L2 : le genre. *Acquisition et interaction en langue étrangère*, (5), 115-142.
- Cibois, P. (1991). Enquêtes d'opinion sur l'orthographe. *Mots. Les langages du politique*, 28(1), 86-98.
- Clahsen, H. (1999). Lexical Entries and Rules of Language: A Multidisciplinary Study of German Inflection. *Behavioral and Brain Sciences*, 22(6), 991-1013.
- Cogis, D. (2005). *Pour enseigner et apprendre l'orthographe. Nouveaux enjeux – Pratiques nouvelles. École/Collège*. Paris: Delagrave.

- Cogis, D. (2007). L'orthographe grammaticale : une difficulté majeure. In D. Manesse & D. Cogis (Éd.), *Orthographe à qui la faute?* (p. 97-136). Issy-les-Moulineaux: ESF éditeur.
- Comrie, B. (1999). Grammatical Gender Systems: A Linguist's Assessment. *Journal of Psycholinguistic Research*, 28(5), 457-466.
- Corder, S. P. (1980). La sollicitation de données d'interlangue. *Langages*, 14(57), 29-38.
- Cuet, C., & Audrey, M. (2007). Enseigner le français en Chine, quelques pistes de recherche sur les plans interculturel et interlinguistique. Présenté à Actes du colloque de Nantes.
- Daigle, D., Mathieu, I., & Montésinos-Gelet, I. (2008). Évaluation du traitement des correspondances phonèmes/phonogrammes chez des lecteurs/scripteurs adultes en début d'apprentissage du français langue seconde. *Canadian Modern Language Review*, 64(4), 637-661.
- David, J., Guyon, O., & Brissaud, C. (2006). Apprendre à orthographier les verbes : le cas de l'homophonie des finales e /E/. *Langue française*, 151(3), 109-126.
- de Weck, G., & Michel, F. (2009). L'orthographe en production de textes chez les enfants avec et sans dysorthographe. *Langage & Pratiques*, 43, 46-58.
- Debyser, F. (1970). La linguistique contrastive et les interférences. *Langue française*, 8(1), 31-61.
- Delen Karragaç, N. (2012). Problème de la terminologie dans la grammaire traditionnelle et/ou moderne du français. *Anadiss, revue du Centre de Recherche Analyse de discours*, 13, 100-120.
- Desbordes, F. (1990). *Idées romaines sur l'écriture*. Presses Universitaires de Lille.
- Dewaele, J.-M., & Véronique, D. (2000). Relating gender errors to morphosyntax and lexicon in advanced French interlanguage. *Studia Linguistica*, 54(2), 212-224.
- Dewaele, J.-M., & Véronique, D. (2001). Gender assignment and gender agreement in advanced French interlanguage: a cross-sectional study. *Bilingualism: Language and Cognition*, 4(3), 275-297.

- Doquet, C. (2003). Écriture et traitement de texte à l'école élémentaire: méthodes d'analyse et traitement de texte. *Language et Société*, 103, 11-29.
- Dubois, J. (1965). *Grammaire structurale du français : nom et pronom* (1ere édition). Librairie Larousse.
- Ducard, D., Honvault, R., & Jaffré, J.-P. (1995). *L'orthographe en trois dimensions*. Nathan.
- Dulay, H. C., & Burt, M. K. (1973). Should We Teach Children Syntax? *Language Learning*, 23(2), 245-258.
- Dulay, H. C., & Burt, M. K. (1974a). Natural sequences in child second language acquisition. *Language Learning*, 24(1), 37-53.
- Dulay, H. C., & Burt, M. K. (1974b). Natural Sequences in Child Second Language Acquisition1. *Language Learning*, 24(1), 37-53.
- Duteil, F. P. (1997). *Le verbe français en conjugaison orale*. P. Lang.
- Drocourt-Yang, Z. (2007). *Parlons chinois*. Paris: Editions L'Harmattan.
- Drocourt-Yang, Z. (2015). L'écriture chinoise : entre universaux et spécificités — Quelle approche dans son enseignement à l'université ? *Langues Modernes*, « Enseigner et apprendre les langues orientales à l'université : quelles approches? », n° 4-2015. ISSN 0023-8376.
- Ellis, N., & Collins, L. (2009). Input and Second Language Acquisition: The Roles of Frequency, Form, and Function Introduction to the Special Issue. *The Modern Language Journal*, 93(3), 329-335.
- Ervin-Tripp, S. M. (1974). Is Second Language Learning like the First. *TESOL Quarterly*, 8(2), 111-127.
- Escoubas-Benveniste, M.-P., & Di Domenico, S. (2017). La dictée de français et les étudiants italophones. Analyse exploratoire des erreurs de morphologie verbale. *Bulletin Suisse de Linguistique Appliquée*, 105, 77-94.
- Escudero, P. (2005). *Linguistic Perception and Second Language Acquisition : Explaining the Attainment of Optimal Phonological Categorization* (Thèse de Doctorat). Utrecht University & LOT.

- Eubank, L. (1993). On the Transfer of Parametric Values in L2 Development. *Language Acquisition*, 3(3), 183-208.
- Fathman, A. (1975). The Relationship Between Age and Second Language Productive Ability. *Language Learning*, 25(2), 245-253.
- Fayol, M. (2003a). L'acquisition/apprentissage de la morphologie du nombre. Bilan et perspectives. *Rééducation orthophonique*, 213, 151-166.
- Fayol, M. (2003b). L'apprentissage de l'accord en genre et en nombre en français écrit. In *Faits de Langue : Dynamiques de l'écriture, approches pluridisciplinaires*. (p. 47-58). Paris: Orphrys.
- Fayol, M. (2002). *Production du langage*. Paris: Hermes Science Publications.
- Fayol, M., & Jaffré, J.-P. (2008). *Orthographier*. Paris: Presses Universitaires de France - PUF.
- Fayol, M., & Jaffré, J.-P. (2016). L'orthographe : des systèmes aux usages. *Pratiques. Linguistique, littérature, didactique*, (169-170).
- Fayol, M., & Largy, P. (1992). Une approche fonctionnelle de l'orthographe grammaticale. *Langue française*, 95(1), 80-98.
- Fayol, M., & Pacton, S. (2006). L'accord du participe passé : entre compétition de procédures et récupération en mémoire. *Langue française*, 151(3), 59-73.
- Fayol, M., Thevenin, M.-G., & Lemaire, P. (1994). Subject-verb agreement errors in French. *Quarterly Journal of Experimental Psychology*, 47A, 437-464.
- Febvre, L., & Martin, H.-J. (1973). *The Coming of the Book: The Impact of Printing, 1450-1800*. London: Verso.
- Fenoglio, I., & Boucheron-Pétillon, S. (Éd.). (2002). Processus d'écriture et marques linguistiques, sous la direction de Irène Fenoglio et Sabine Boucheron-Pétillon. *Langages*, 147.
- Flege, J. (1992). Speech learning in a second language. In L. Menn & C. Stoel-Gammon (Éd.), *Phonological Development: Models, Research, and Application* (p. 565-604). Timonium, MD: York Press.
- Flege, J., Olaf Schiller, N., Meyer, A., & Gruyter, W. (2003). Assessing constraints on second-language segmental production and perception. In *Phonetics and*

- Phonology in Language Comprehension and Production* (p. 319-355). Berlin: Mouton de Gruyter.
- Flower, L., Hayes, J. R., Carey, L., Schriver, K., & Stratman, J. (1986). Detection, Diagnosis, and the Strategies of Revision. *College Composition and Communication*, 37(1), 16-55.
- Foulin, J.-N. (2003). *La connaissance orthographique. Des lettres aux mots* (Document de synthèse pour l'Habilitation à Diriger des Recherches). Université de Clermont-Ferrand.
- Fries, C. C. (1945). *Teaching and Learning English as a Foreign Language*. Ann Arbor.
- Fu, R. (2006). Politique et stratégies linguistiques dans l'enseignement supérieur des langues étrangères en Chine nouvelle. *Synergies Chine*, 1, 27-39.
- Gak, V. G. (1976). *L'orthographe du français: essai de description théorique et pratique*. Paris: SELAF.
- Galligani, S. (2003). Réflexion autour du concept d'interlangue pour décrire des variétés non natives avancées en français. *Linx. Revue des linguistes de l'université Paris X Nanterre*, (49), 141-152.
- Ganoac'h, D. (2004). *Théories d'apprentissage et acquisition d'une langue étrangère*. Paris: Didier.
- George, H. V. (1972). *Common Errors in Language Learning*. Massachusetts: Newbury House Publishers.
- Gerolimich, S., & Stabarin, I. (2017). Introduction. *Bulletin Suisse de Linguistique Appliquée*, 1-6.
- Ghellai, M. (1997). *Analyse des erreurs et des représentations orthographiques du français par des intellectuels arabophones* (Thèse de Doctorat). Université Stendhal - Grenoble 3.
- Girolami-Boulinier, A. (1997). *Les Niveaux actuels dans la pratique du langage oral et écrit*. Paris: Editions Masson.
- Gonac'h, J., & Mortamet, C. (2011). Pratiques orthographiques en français d'étudiants étrangers: le cas d'étudiants hispano-américains et afghans. *TRANEL (Travaux neuchâtelois de linguistique)*, 54, 113-127.

- Granget, C. (2005). Développement de l'accord verbal avec un sujet pluriel dans les récits écrits d'apprenants germanophones scolarisés en français. In J. Granfeldt & S. Schlyter (Éd.), *Acquisition et production de la morphologie flexionnelle*. Lund: Institut d'études romanes de l'université de Lund.
- Grüter, T., Lew-Williams, C., & Fernald, A. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research*, 28(2), 191-215.
- Guyon, O. (2003). Évolution des procédures d'accord nominal et verbal en français : perspective psycholinguistique. *Les Dossiers des Sciences de l'Éducation*, 9(1), 55-66.
- Hakuta, K., & Cancino, H. (1977). Trends in Second-Language-Acquisition Research. *Harvard Educational Review*, 47(3), 294-316.
- Hale, K. (1988). Linguistic Theory: Generative Grammar. In *Linguistic Theory in Second Language Acquisition* (p. 26-33). Dordrecht: Kluwer Academic.
- Hamon, P. (2009). *Le Signe et la Consigne*. Genève: Droz. Consulté à l'adresse
- Hayes, J., Flower, L., Gregg, L. W., & Steinberg, E. R. (1980). Identifying the organization of writing processes. In *Cognitive Processes in Writing* (p. 3-30).
- Housen, A., Kemps, N., & Pierrard, M. (2006). Stades développementaux et contexte instructionnel : l'acquisition de la morphologie verbale chez les apprenants avancés de FLE › Research Explorer. *Cahiers de l'Institut de Linguistique de Louvain*, 32, 273-293.
- Housen, A., Kemps, N., & Pierrard, M. (s. d.). The use of verb morphology of advanced L2 learners and native speakers of French › Research Explorer. In E. Labeau & F. Myles (Éd.), *The Advanced Learner Variety - The Case of French* (p. 41-62). Bern, London, New York: Peter Lang.
- Hu, Y. (2010). Synopsis d'adaptation de manuel français en version chinoise. *Synergies Chine*, 5, 39-46.
- Humeau, G. (2012). Le CECR, un outil pour réformer l'enseignement/ apprentissage du français en milieu universitaire chinois. *Synergies Chine*, 7, 129-139.
- Jaffré, J.-P. (1992). *Didactiques de l'orthographe*. Paris: Hachette.



- Jaffré, J.-P. (1998). L'écriture du français : une exception ? *Le français aujourd'hui*, 122, 43-53.
- Jaffré, J.-P. (2005). L'orthographe du français, une exception ? *Le français aujourd'hui*, (148), 23-31.
- Jaffré, J.-P. (2006). Pourquoi distinguer les homophones ? *Langue française*, 151(3), 25-40.
- Jaffré, J.-P. (2009). L'orthographe du français: une histoire, des problèmes. *Langage & Pratiques*, 43, 10-18.
- Jaffré, J.-P., & Brissaud, C. (2006). Présentation. *Langue française*, 151(3), 3-6.
- Jaffré, J.-P., & David, J. (1999). Le nombre : essai d'analyse génétique. *Langue française*, 124(1), 7-22.
- Jakobson, R. (1963). Linguistique et poétique. In *Essais de linguistique générale* (p. 209-248). Paris: Minuit.
- Jiang, N. (2000). Lexical representation and development in a second language. *Applied Linguistics*, 21(1), 47-77.
- Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics*, 25(4), 603-634.
- Jiang, N. (2007). Selective Integration of Linguistic Knowledge in Adult Second Language Learning. *Language Learning*, 57(1), 1-33.
- Katoozian, K. (2013). *Analyse des erreurs orthographiques des étudiants iraniens en langue française* (Thèse de Doctorat). Université Paris 3, Paris.
- Katoozian, K. (2014). Une revue d'articles sur les erreurs orthographiques à dominante phonétique: réflexions pour le FLE. *Lingvarvm Arena*, 5, 31-45.
- Klein, W., & Perdue, C. (1997). The Basic Variety (or: Couldn't natural languages be much simpler?). *Second Language Research*, 13(4), 301-347.
- Krashen, S. D. (1988). *Second Language Acquisition and Second Language Learning*. New York: Prentice-Hall.
- Krashen, S. D., & Pon, P. (1975). An Error Analysis of an Advanced Learner of ESL: The Importance of the Monitor. Working Papers on Bilingualism, No. 7. *Working Papers on Bilingualism*, 7, 125-129.

- Krashen, S. D., Sferlazza, V., Feldman, L., & Fathman, A. K. (1976). Adult Performance on the Slope Test: More Evidence for a Natural Sequence in Adult Second Language Acquisition. *Language Learning*, 26(1), 145-151.
- Kuhl, P., & Iverson, P. (1995). Linguistic experience and the 'perceptual magnet effect'. In W. Strange (Éd.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research*. Baltimore, MD: York Press.
- Lalande, J.-P. (1988). *Élaboration d'un modèle théorique de l'apprentissage de l'orthographe lexicale au primaire* (Thèse de Doctorat). Université de Montréal, Montréal.
- Lam, S. (2006). *Temps, aspect et nombre en chinois et en français une étude contrastive*. Université de Sherbrooke.
- Landron, S., Gao, J., Chang, Y., & Tian, Y. (2017). Les sinophones. In S. Detey, I. Racine, Y. Kawaguchi, & J. Eychenne (Éd.), *La prononciation du français dans le monde : du natif à l'apprenant* (p. 193-199). Paris: Clé International.
- Lardiere, D. (1998). Case and Tense in the 'fossilized' steady state. *Second Language Research*, 14(1), 1-26.
- Largy, P., & Cousin, M.-P. (2004). Apprendre implicitement le -s du pluriel. In A. Piolat (Éd.), *Écriture : Approches en sciences cognitives*. Aix-en-Provence: Publications de l'Université de Provence.
- Largy, P., Fayol, M., & Cousin, M.-P. (2004). Acquérir le pluriel des noms: existe-t-il un effet de fréquence du nom? *LIDIL: Revue de linguistique et de didactique des langues*, (30), 39-54.
- Largy, P., Fayol, M., & Lemaire, P. (1996). The Homophone Effect in Written French: The Case of Verb-Noun Inflection Errors. *Language and Cognitive Processes*, 11(3), 217-256.
- Larsen-Freeman, D., & Long, M. H. (1991). *An Introduction to Second Language Acquisition Research*. London: Routledge.
- Lassaad, K. (2011, septembre 28). *L'erreur orthographique dans les productions écrites des élèves tunisiens : origine du dysfonctionnement et contribution à la*

- maîtrise de la compétence orthographique* (Thèse de Doctorat). Université de la Sorbonne nouvelle - Paris 3.
- Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. Cambridge: MIT Press.
- Levelt, W. J. M. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition*, 42(1), 1-22.
- Lindström, E. (2013). *L'acquisition du genre en français L2 développement et variation* Eva Lindström Forskningsrapporter / Cahiers de la Recherche 50 - PDF (Thèse de Doctorat). Stockholms universitet.
- Long, M. H., & Sato, C. J. (1984). Methodological issues in interlanguage studies: an interactionist perspective. In *Interlanguage* (p. 253-279). Edinburgh: Edinburgh University Press.
- Lumbroso, O. (2004). *Zola. La Plume et le Compas*. Paris: Honoré Champion.
- Ma, X. (2010). Pour un manuel de français de qualité. *Synergies Chine*, 5, 61-69.
- Ma, X., & Liu, L. (1992). *Le français*. Pékin: Foreign language teaching and research press.
- Makassikis, M., & Pellat, J.-C. (2011). Les étudiants natifs et allophones face à l'orthographe française: le cas des homonymes. *Revue Tranel (Travaux neuchâtelois de linguistique)*, 54, 21-48.
- Manesse, D. (2007). L'orthographe lexicale: statu quo. In D. Manesse & D. Cogis (Éd.), *Orthographe à qui la faute?* Issy-les-Moulineaux: ESF éditeur.
- Manesse, D., & Cogis, D. (2007). *Orthographe, à qui la faute?* Issy-les-Moulineaux: Editeur ESF.
- Mariol, M., Schelstraete, M. A., & Van Rompaey, M. (2004). Sensibilité aux régularités orthographiques et acquisition des marques morphologiques du genre en français. *Le Langage et l'Homme*, XXXIX(2), 19-35.
- Martinet, A. (1974). Homonymes et polysèmes. *La Linguistique*, 10(2), 37-45.
- Martinet, C., Valdois, S., & Fayol, M. (2004). Lexical orthographic knowledge develops from the beginning of literacy acquisition. *Cognition*, 91(2), 11-22.

- Masseron, C., & Luste-Chaa, O. (2008). Typologie d'erreurs lexicales : difficultés et enjeux (p. 519-531). Présenté à Actes du 1er Congrès mondial de linguistique française, Paris.
- Mayard, S. (2007). *La maîtrise de l'orthographe lexicale du français et de l'espagnol* (Mémoire). Université du Québec à Montréal, Montréal.
- Meisel, J. M. (2016). The acquisition of the syntax of negation in French and German: contrasting first and second language development: *Second Language Research*, 227-263.
- Michel, F. (2007). La production de textes et son apprentissage. Présenté à Les journées de l'ONL, Écrire des textes, l'apprentissage et le plaisir.
- Michot, M.-E. (2013). Stades de progression de la finitude verbale en FL2. *Círculo de Lingüística Aplicada a la Comunicación*, 54(0), 44-64.
- Michot, M.-E. (2015a). *L'acquisition de la morphologie verbale en français langue seconde* (Thèse de Doctorat). Université de Bruxelles.
- Michot, M.-E. (2015b). *L'acquisition de la morphologie verbale en français langue seconde* (Thèse de Doctorat). Université de Bruxelles.
- Mirande, J., Alliet, O., & Roca, F. (2010). *Contes et légendes : Les chevaliers de la Table Ronde*. Paris: Nathan.
- Millet, A., Lucci, V. & Billiez, J. (1990). *Orthographe mon amour*. Grenoble : Presses universitaires de Grenoble.
- Montonen, O. (2001). *Le système graphique dans La deffence et illustration de la langue francoyse de Joachim Du Bellay*.
- Murray, D. M. (1978). Write before Writing. *College Composition and Communication*, 29(4), 375-381.
- Nouveau, D. (2017). Morphologie verbale et FLE: les groupes verbaux en -ir. *Bulletin Suisse de Linguistique Appliquee*, 105, 43-58.
- Olson, L. L., & Samuels, S. J. (1973). The Relationship between Age and Accuracy of Foreign Language Pronunciation. *The Journal of Educational Research*, 66(6), 263-268.

- Pacton, S., Foulin, J.-N., & Fayol, M. (2005). L'apprentissage de l'orthographe lexicale. *Rééducation Orthophonique*, 222, 47-68.
- Paloheimo, M. (2014). L'acquisition de la morphologie verbale française dans les productions écrites des apprenants finnophones avancés. *Cahiers de l'AFLS*, 19(1).
- Peereman, R., & Content, A. (1999). LEXOP: A lexical database providing orthography-phonology statistics for French monosyllabic words. *Behavior Research Methods, Instruments, & Computers*, 31(2), 376-379.
- Pellat, J.-C., & Andrieux-Reix, N. (2006). Histoire d'É ou de la variation des usages graphiques à la différenciation réglée, Abstract. *Langue française*, (151), 7-24.
- Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009a). Statistical learning in a natural language by 8-month-old infants. *Child Development*, 80(3), 685.
- Pelucchi, B., Hay, J., & Saffran, J. (2009b). Learning in reverse: Eight-month-old infants track backward transitional probabilities. *Cognition*, 113(2), 244-247.
- Perche, V. (s. d.). Vers une didactique complexe des langues-cultures : d'un éclectisme empirique à un éclectisme cohérent. *Synergies Chine*, 6, 177-188.
- Perdue, C., Benazzo, R., & Giuliano, P. (2002). When finiteness gets marked : the relation between morphosyntactic development and use of scopal items in adult language acquisition. *Linguistics*, 40(4), 849-890.
- Pérez, M., & Garcia-Debanc, C. (2016). Les zones de fragilité orthographique dans les productions sous dictée contrainte d'un texte classique par des adultes diplômé·e·s. *SHS Web of Conferences*, 27, 7013.
- Pernet-Liu, A. (2012). *Approche du passage à l'écrit en français (langue étrangère) chez des étudiants chinois en Chine: le mémoire universitaire entre références chinoises et occidentales* (Thèse de Doctorat). Université de Bourgogne, Dijon.
- Pienemann, M. (1998). *Language Processing and Second Language Development: Processability theory*. Amsterdam ; Philadelphia: John Benjamins Publishing Company.

- Plane, S. (2011). L'épreuve « Agir en fonctionnaire de l'État de façon éthique et responsable » au CAPES de lettres. In O. Narjoux & Bertrand (Éd.), *Préparer et réussir le CAPES externe de lettres modernes* (p. 161-182). Bruxelles: De Boeck.
- Plane, S. (2015). A la recherche du temps de l'écriture. Présenté à Actes du colloque Stratégies d'écriture, stratégies d'apprentissage de la maternelle à l'Université, Colmar.
- Plane, S., Alamargot, D., & Lebrave, J.-L. (2010). Temporalité de l'écriture et rôle du texte produit dans l'activité rédactionnelle. *Langages*, (177), 11-34.
- Plane, S., Rondelli, F., & Vénérin-Guénez, C. (2013). Variations, fidélité, infidélité : l'écriture et la réécriture de discours rapportés par de jeunes scripteurs. In C. Desoutter & C. Facq-Mellet (Éd.), *Le discours rapporté : approches linguistiques et perspectives didactiques* (p. 215-232). Peter Lang.
- Polivanov, E. (1931). La perception des sons d'une langue étrangère. *Travaux du Cercle Linguistique de Prague*, 4, 76-96.
- Pouradier Duteil, F. (1997). *Le verbe français en conjugaison orale*. Frankfurt am Main: Peter Lang.
- Prévost, P., & White, L. (1999). Finiteness and variability in SLA: More evidence of missing surface inflection. (p. 575-586). Présenté à Proceedings of the 23rd Annual Boston University Conference on Language Development, Somerville MA: Cascadilla Press.
- Prévost, P., & White, L. (2000). Missing Surface Inflection or Impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research*, 16(2), 103-133.
- Pu, Z. (2011). La perspective actionnelle par tâches et la culture d'apprentissage chinoise. *Synergies Chine*, 6, 37-45.
- Richards, J. C. (1974). *Error analysis: perspectives on second language acquisition*. Longman.
- Riegel, M., Pellat, J.-C., & Rioul, R. (1994). *Grammaire méthodique du français*. Paris: Presses Universitaires de France - PUF.

- Ringbom, H., Asher, R. E., & Simpson, J. M. Y. (1994). Contrastive analysis. *Encyclopedia of linguistics*, 2, 737-742.
- Rittle-Johnson, B., & Siegler, R. S. (1999). Learning to spell: variability, choice, and change in children's strategy use. *Child Development*, 70(2), 332-348.
- Robert, J.-M. (2002). Sensibilisation au public asiatique. *Ela. Études de linguistique appliquée*, 126(2), 135-143.
- Ros-Dupont, M. (2000). Acquisition de l'écrit: comment les enfants de 7 à 9 ans traitent-ils l'homophonie? *Linx*, (43), 63-76.
- Rosen, É., & Porquier, R. (2003). Présentation. L'actualité des notions d'Interlangue et d'interaction exolingue. *Linx. Revue des linguistes de l'université Paris X Nanterre*, (49), 7-17.
- Roulet-Amiot, L., & Jakubowicz, C. (2006). Production and perception of gender agreement in French SLI. *Advances in Speech Language Pathology*, 8(4), 335-346.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical Learning by 8-Month-Old Infants. *Science*, 274(5294), 1926-1928.
- Sagart, L. (2006). L'emploi de phonétiques dans l'écriture chinoise. In Bottero, F. et Djmouri, R. (ed). *Écriture chinoise, données, usages et représentations*. Paris : EHESS/CRLAO.
- Sagaz, M. (2011). Contextualisation du CECR et pratiques méthodologiques locales : le cas du Japon. *Synergies Europe*, 6, 75-84.
- Sagaz, M., & Dcatel, N. (s. d.). Pourquoi adopter le CECR au Japon ? *Synergies Europe*, 6, 153-163.
- Saillard, C. (2001). Le chinois mandarin entre véhicularisation et vernacularisation. In *Les langues de Communication, Mémoires de la Société de Linguistique de Paris*, Nouvelle série, Tome XI, Peeters, 121-134.
- Saillard, C., & Boutet, J. (2008). Construction des répertoires langagiers dans la migration Wenzhou (Chine) à Paris. *Cahiers de l'Observatoire des pratiques linguistiques*, 2, 6.

- Santens, S. (2016). *La maîtrise de la morphologie verbale en FL2 l'écrit et l'oral* (Mémoire). Université Gent.
- Schriefers, H., & Jescheniak, J. D. (1999). Representation and Processing of Grammatical Gender in Language Production: A Review. *Journal of Psycholinguistic Research*, 28(6), 575-600.
- Segui, J., & Ferrand, L. (2000). *Leçons de parole*. Paris: Editions Odile Jacob.
- Selinker, L. (1972). Interlangue. *IRAL - International Review of Applied Linguistics in Language Teaching*, 10(1-4), 209–232.
- Selinker, L., Swain, M., & Dumas, G. (1975). The Interlanguage Hypothesis Extended to Children1. *Language Learning*, 25(1), 139-152.
- Sharwood Smith, M. (10-20). The competence/control model, crosslinguistic influence and the creation of new grammars. In E. Kellerman & M. Sharwood Smith (Éd.), *Cross-linguistic influence in second language acquisition*. New York: Pergamon Press.
- Simon, T. (1924). *Pédagogie expérimentale. Écriture. Lecture. Orthographe*. Paris: Armand Colin.
- Stenson, N. (1974). Induced errors. In J. H. Schumann & N. Stenson (Éd.), *New frontiers in second language learning*. Massachusetts: Newbury House.
- Stockwell, R. P., Bowen, J. D., & Martin, J. W. (1965). *The Grammatical Structures of English and Spanish*. University of Chicago Press.
- Taylor, B. P. (1974). Toward a Theory of Language Acquisition1. *Language Learning*, 24(1), 23-35.
- Thevenin, M.-G., Totereau, C., Jarousse, J.-P., & Fayol, M. (1999). L'apprentissage/enseignement de la morphologie écrite du nombre en français. *Revue française de pédagogie*, 126(1), 39-52.
- Thomas, A. (2017). Le développement des formes dans les contextes de l'infinitif par des apprenants suédophones du français L2. *Bulletin Suisse de Linguistique Appliquée*, 105, 113-127.



- Totereau, C., Barrouillet, P., & Fayol, M. (1998). Overgeneralizations of number inflections in the learning of written French: The case of noun and verb. *British Journal of Developmental Psychology*, 16(4), 447-464.
- Totereau, C., & Brissaud, C. (2006). Acquisition des flexions verbales homophones en /E/ : une étude longitudinale du CE2 au CM2. *Rééducation Orthophonique*, (225), 59-74.
- Totereau, C., Thevenin, M. G., & Fayol, M. (1997). The development of the understanding of number morphology in written French. In C. A. Perfetti, L. Rieben, & M. Fayol (Éd.), *Learning to Spell*. New York: Routledge.
- Trévisiol, P. (2003). *Problèmes de référence dans la construction du discours par des apprenants japonais du français, langue 3* (Thèse de Doctorat). Université de Paris 8, Saint-Denis, France.
- Trubetzkoy, N.-S. (1967). *Principes de phonologie*. Paris: Klincksieck.
- Trudeau, N., Poulin-Dubois, D., Frank, I., Courcy, A., & Sutton, A. (2008). *Normalisation et validation de la version québécoise des MacArthur Communicative Development Inventories (MCDI)*. Université de Montréal.
- Tucker, G. R., Lambert, W. E., & Rigault, A. A. (1977). *The French Speakers Skill With Grammatical Gender: An Example of Rule-Governed Behavior*. The Hague: Mouton.
- Valois, D., Royle, P., Sutton, A., & Bourdua-Roy, E. (2009). L'ellipse du nom en français : le rôle des données de l'acquisition pour la théorie linguistique. *Canadian Journal of Linguistics/Revue Canadienne de Linguistique*, 54(2), 339-366.
- Véronique, D. (2009). *L'acquisition de la grammaire du français, langue étrangère*. Paris: Didier.
- Véronis, J. (1988). From sound to spelling in French: Simulation on a computer. *Cahiers de Psychologie Cognitive*, 8, 315-334.
- Wang, W. (2015). Potentiel d'adaptation du CECR dans les facultés de français en Chine. *Synergies Europes*, 6, 193-204.

- Wardhaugh, R. (1970). The Contrastive Analysis Hypothesis. *TESOL Quarterly*, 4(2), 123-130.
- Weinreich, U. (1953). *Languages in Contact: Findings and Problems*. The Hague: Mouton de Gruyter.
- Weinrich, H. (1989). *Grammaire textuelle du français*. Paris: Didier.
- Willemin, S. (2009). Présentation. *Langage & Pratiques*, 43, 2-9.
- Wise, H. (1997). *The Vocabulary of Modern French: Origins, Structure and Function* (1 edition). London: Routledge.
- Witte, S. P., & Freedman, S. W. (1985). Revising, composing theory, and research design. In *The acquisition of written language: response and revision* (p. 250-284). Norwood: Ablex Publishing Corporation.
- Xu, Y. (2010). Différences méthodologiques entre les manuels français et chinois de FLE sous l'angle de l'organisation structurale du contenu. *Synergies Chine*, 5, 89-98.
- Ziegler, J. C., Jacobs, A. M., & Stone, G. O. (1996). Statistical analysis of the bidirectional inconsistency of spelling and sound in French. *Behavior Research Methods, Instruments, & Computers*, 28(4), 504-515.

## Liste des figures

<b>Fig. 1</b> : Quatre couches d'annotation pour l'item.....	111
<b>Fig. 2</b> : Correction des annotations. ....	113
<b>Fig. 3</b> : Lancement de l'enquête sur les noms au pluriel .....	121
<b>Fig. 4</b> : Occurrences de noms au pluriel .....	122
<b>Fig. 5</b> : Contextes gauche et droit de noms au pluriel .....	123
<b>Fig. 6</b> : Distribution de noms au pluriel en fonction de niveau linguistique .....	123
<b>Fig. 7</b> : Lancement de la recherche de formes non-normatives.....	124
<b>Fig. 8</b> : Création de la cinquième couche d'annotation .....	126
<b>Fig. 9</b> : Recherche des formes normatives.....	126
<b>Fig. 10</b> : Création de la nouvelle annotation avec la valeur 0 .....	127
<b>Fig. 11</b> : Création de la sixième couche d'annotation .....	127
<b>Fig. 12</b> : Liste des formes non-normatives .....	128
<b>Fig. 13</b> : Recherche des formes verbales erronées dont l'infinitif se termine par <i>-er</i> .....	128
<b>Fig. 14</b> : Recherche des formes verbales erronées de l'imparfait.....	129
<b>Fig. 15</b> : Comparaison de la répartition des erreurs.....	155
<b>Fig. 16</b> : Pourcentage de réussite en fonction des temps différents.....	225
<b>Fig. 17</b> : Pourcentage de réussite de l'accord du nombre.....	231
<b>Fig. 18</b> : Réalisation de la morphologie du genre selon les différents types de déterminant .....	243
<b>Fig. 19</b> : Évolution du pourcentage de réussite pour les différentes catégories des formes verbales en fonction des niveaux linguistique différents .....	329

# Liste de tableaux

<b>Tableau 1.</b> La conjugaison du verbe « prendre » .....	68
<b>Tableau 2:</b> phonogrammes avec la phonétique de base : zhǔ 主 (Drocourt, 2015)...	82
<b>Tableau 3.</b> Hiérarchie de difficulté (Larsen-Freeman & H.Long, 1990 : 54) .....	98
<b>Tableau 4:</b> Convention de transcription .....	109
<b>Tableau 5.</b> Typologie des erreurs d'orthographe (Manesse & Cogis, 2007 : 88) .....	114
<b>Tableau 6.</b> Classement d'erreurs d'orthographe .....	116
<b>Tableau 7.</b> La répartition des erreurs orthographiques.....	129
<b>Tableau 8.</b> Typologie des erreurs lexicales .....	138
<b>Tableau 9.</b> Modèle de transition de Simon.....	140
<b>Tableau 10.</b> Répartition des erreurs orthographiques par type (Girolami- Boulinier, 1984) .....	141
<b>Tableau 11.</b> Répartition des erreurs orthographiques par type (Chervel & Manesse, 1989) .....	142
<b>Tableau 12.</b> Répartition des erreurs orthographiques chez les étudiants libyens du français L2 (Ghellai, 1997) .....	143
<b>Tableau 13.</b> Répartition des erreurs orthographiques chez les élèves tunisiens du français L2 (Lassaad, 2011) .....	144
<b>Tableau 14.</b> Répartition des erreurs orthographiques chez les étudiants iraniens en français L2 (Kattozian, 2013) .....	145
<b>Tableau 15.</b> Rapport entre syllabe, morphème et mot en français (Drocourt, 2007 : 155) .....	147
<b>Tableau 16.</b> Rapport entre syllabe, morphème et mot en chinois (Drocourt, 2007 : 156) .....	147
<b>Tableau 17.</b> Nombre des erreurs orthographiques phonétiques par niveau .....	155
<b>Tableau 18.</b> Nombre moyen d'erreurs phonétiques par classe (Manesse & Cogis, 2007) .....	156
<b>Tableau 19.</b> Nombre des erreurs orthographiques phonogrammiques par niveau ....	156

<b>Tableau 20.</b> Nombre moyen d'erreurs phonétiques par classe (Manesse & Cogis, 2007) .....	157
<b>Tableau 21.</b> Nombre des erreurs lexicales sans incidence phonique par niveau .....	157
<b>Tableau 22.</b> Nombre moyen d'erreurs phonétiques par classe (Manesse & Cogis, 2007) .....	158
<b>Tableau 23.</b> Situations comparées d'un enfant francophone et d'un enfant étranger face à l'apprentissage de l'oral et de l'écrit ( <i>Légende : Langue 1, langue maternelle ; langue 2, langue cible</i> ).....	173
<b>Tableau 24.</b> Exemples de morphologie silencieuse et audible dans le syntagme nominal et le syntagme verbal ( <i>Légende : souligné = morphologie audible ; en caractères gras = morphologie silencieuse</i> ) .....	191
<b>Tableau 25.</b> L'accord nominal en nombre dans la totalité des contextes au pluriel .	206
<b>Tableau 26.</b> L'accord nominal en nombre dans la totalité des contextes au pluriel chez les apprenants suédois (Ågren, 2005) .....	206
<b>Tableau 27.</b> Différences individuelles des omissions du pluriel nominal.....	208
<b>Tableau 28.</b> Erreurs relevant de l'accord du nombre nominal commises par les apprenants de L2 et leur contexte .....	209
<b>Tableau 29.</b> Erreurs relevant de l'accord du nombre nominal commises par les apprenants de L3 et leur contexte .....	212
<b>Tableau 30.</b> Erreurs relevant de l'accord du nombre nominal commises par les apprenants de L4 et leur contexte .....	214
<b>Tableau 31.</b> Taux de réussite de l'accord nominal du nombre selon les différents contextes .....	216
<b>Tableau 32.</b> Accord nominal du nombre dans un groupe nominal contenant un ou plusieurs adjectifs.....	217
<b>Tableau 33.</b> Nombre des adjectifs produits selon les différents niveaux .....	219
<b>Tableau 34.</b> Taux de réussite de l'accord adjectival du nombre selon les différents contextes .....	221
<b>Tableau 35.</b> Performance de l'accord du nombre dans le syntagme verbal.....	223
<b>Tableau 36.</b> Articles simples .....	236

<b>Tableau 37.</b> Articles spécifiques .....	236
<b>Tableau 38.</b> Taux de réussite de l'accord du genre pour les déterminants .....	240
<b>Tableau 39.</b> Taux de réussite de l'accord du genre pour les adjectifs .....	244
<b>Tableau 40.</b> Théorie de la processabilité de Pienemann appliquée à l'acquisition de l'accord adjectival en français (cité par Bartning 2001) .....	252
<b>Tableau 41.</b> Accord en genre pour les déterminants et les adjectifs .....	254
<b>Tableau 42.</b> Évolution du pourcentage de réussite en fonction du niveau des apprenants .....	282
<b>Tableau 43.</b> Formes erronées produites par les apprenants de trois niveaux quand la cible est <i>-er</i> .....	284
<b>Tableau 44.</b> Formes erronées produites par les apprenants de trois niveaux quand la cible est <i>-é</i> .....	284
<b>Tableau 45.</b> Formes erronées quand la forme de l'infinitif est attendue .....	322
<b>Tableau 46.</b> Formes erronées quand la forme du participe passé est attendue .....	323
<b>Tableau 47.</b> Formes erronées quand la forme de l'indicatif présent est attendue .....	324
<b>Tableau 48.</b> Formes erronées quand la forme du passé simple est attendue .....	325
<b>Tableau 49.</b> Classification verbale de Gerolimich et Stabarin (cité dans Nouveau, 2017) .....	340
<b>Tableau 50.</b> Taux de réussite pour le problème de l'accord .....	346

*L'apprentissage de l'orthographe chez les apprenants chinois du français L2 : une analyse portant sur un corpus écrit*

**Résumé.** Dans l'apprentissage du français écrit, la norme orthographique représente un lieu de difficulté majeure. Confronté à des complexités multiples, tout scripteur aura un long trajet à parcourir avant d'arriver à la maîtrise complète des traitements orthographiques, et ceci est sans doute d'autant plus vrai lorsqu'il s'agit de s'approprier le système orthographique du français dans une langue étrangère. Cette étude vise donc à analyser la réalisation orthographique, chez les apprenants chinois du français L2. Pour ce faire, nous examinons les données issues d'une tâche écrite semi-spontanée, en nous attachant plus particulièrement au profil développemental des apprenants chinois, et leurs tendances interlangagières face aux problèmes de l'orthographe lexicale, de l'accord en nombre et genre et de la morphologie verbale. En outre, une analyse des erreurs commises par les apprenants, ainsi que les explications possibles pour interpréter ces erreurs sont également présentées, en faisant référence aux caractéristiques spécifiques de l'apprentissage de la langue étrangère.

**Mots-clés :** orthographe, apprenants chinois, français L2, orthographe lexicale, morphologie écrite, erreur

*Orthographic learning by Chinese-speaking university learners of French L2: an analysis based on a written corpus*

**Abstract:** In the area of written French, the orthographic learning represents a major difficulty. Faced with multiple complexities, any writer will have a long way to go before arriving at the complete control of spelling, and this is probably more true when it comes to appropriating this French system in a foreign language which is typologically distinct from the French language. This study aims to analyze the orthographic realization in French L2, for Chinese learners. To do this, we examine data from a semi-spontaneous writing task, focusing in particular on the developmental profile of learners, and their interlanguage tendencies in relation to the problems of lexical orthography, number/gender agreement, and verbal morphology. In addition, an errors analysis as well as possible explanations for interpreting these errors are also presented, with reference to the specific characteristics of L2 learning.

**Keywords:** spelling, Chinese learners, French L2, lexical orthography, written morphology, error

UNIVERSITE SORBONNE NOUVELLE - PARIS 3

ED 268 – Langage et langue : description, théorisation, transmission EA 7345- CLESTHIA : langage, système, discours

Maison de la recherche, Bureau A006, 4 rue des Irlandais 75005 Paris.