



THÈSE

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ PARIS 8

Discipline : sciences du langage

présentée et soutenue publiquement par

Samantha Ruvoletto

le 29 juin 2016

Titre :

ÉTUDE LONGITUDINALE DES PHÉNOMÈNES DE RESYLLABATION CHEZ LES
ENFANTS FRANÇAIS ENTRE COMPÉTENCE PHONOLOGIQUE, LECTURE ET
ORTHOGRAPHE

Directrice de thèse :

Sophie Wauquier

JURY

M. Jean-Pierre Chevrot (Professeur Université Stendhal Grenoble 3), rapporteur

Mme Pascale Colé (Professeur Université d'Aix-Marseille)

Mme Saveria Colonna (Maître de Conférence Université Paris 8)

M. Giorgio Magri (Chargé de Recherche, UMR 7023 du CNRS)

Mme Annie Rialland (Directeur de Recherche Émérite, Université Paris 3), rapporteur

Résumé

Pendant l'acquisition, l'enfant francophone reçoit comme input une langue orale où les frontières entre les mots sont masquées à cause de trois phénomènes de resyllabation très communs : la liaison (ex. *les ours* [lezuʁs]), l'élision (ex. *l'avion* [la.vjɔ̃]) et l'enchaînement (ex. *une araignée* [y.na.ʁe.ɲe]). Ces phénomènes sont à l'origine de l'apparition de mauvaises segmentations dans les productions des enfants à partir de l'âge de 2 ans, classifiées sous les noms de « remplacements » ou « consonnes erronées » (ex. *de(s)[n]éléphants* [de.ne.le.fã] pour *des[z]éléphants* [de.ze.le.fã]) et de « omissions » ou « non-réalisations » (*u(n)[Ø]avion* [ɛ̃.a.vjɔ̃] pour *un[n]avion* [ɛ̃.na.vjɔ̃]). Les études développementales (longitudinales ou transversales) observent une diminution de ces erreurs entre 4 et 5 ans. Si dans la langue orale, les phénomènes de resyllabation causent un non-alignement entre les frontières syllabiques et les frontières lexicales, dans la langue écrite, les frontières lexicales des mots sont délimitées par des marques graphiques (<les ours>, <une ambulance>, <l'éléphant>). Au moyen d'une étude longitudinale conduite sur 43 enfants francophones à partir du CP (6;3) jusqu'au CE1 (7;6), nous montrons que l'apprentissage de l'écriture et de la lecture, dénommés *literacy*, aide les enfants à fixer les formes lexicales et à mieux produire à l'oral et que cette variable influence aussi le traitement cognitif des séquences resyllabifiées. À partir de ces résultats, nous proposons également un modèle phonologique pour l'acquisition de la liaison après 5 ans qui tient compte des résultats phonétiques enregistrés et de l'apprentissage de la langue écrite.

English Abstract

A longitudinal study on children language acquisition of French re-syllabification phenomena considering phonological and literacy skills

Early word segmentation in French is complicated by three phenomena of re-syllabification: liaison (word1 *les* [le], 'the' + word2 *ours* [uʁs], 'bears' → *les[z]ours* [le.zuʁs], 'the bears'), elision (word1 *le* [le], 'the' + *ours* [uʁs], word2 'bear' → *l'ours* [luʁs], 'the bear') and enchaînement (word1 *une* [yn], 'a' + word2 *autruche* [ot.ʁyʃ], 'ostrich' → *une austruche* [y.no.t.ʁyʃ], 'an ostrich'). These processes of re-syllabification entail children's errors in oral productions such as replacements (*le(s)[n]éléphants* [le.ne.le.fã] for *les[z]éléphants* [le.ze.le.fã], 'the elephants') and non-realizations (*u(n)[Ø]ours* [ɛ̃.a.vjɔ̃] for *un[n]ours* [ɛ̃.nuʁs], 'a bear'). Even though the re-syllabified sequences of word1+word2 are perceived blended, in writing forms they appear separated by graphic markers, as blanks (*les ours*, *une austruche*) or apostrophes (*l'ours*). In our longitudinal study we evaluate the role of literacy in French word segmentation of re-syllabified sequences comparing the results of oral tasks and writing-reading tasks in the same group of children in three points according to literacy skills. Data analysis shows that oral productions of re-syllabified sequences are influenced by literacy: the complete acquisition of written lexical forms and the strengthen reading skills help to correctly perform oral productions and phenomena of re-syllabification. We also notice that re-syllabified sequences entail a processing cost in word recognition. The thesis ends with the proposal of a phonological model that explains word boundaries acquisition after 5 years of age considering literacy as an important factor.

Remerciements

Après cette nouvelle (et inattendue) expérience de doctorat en France, je tiens à remercier les personnes qui ont marqué mon parcours professionnel et personnel et qui m'ont également permis d'achever cette thèse.

Mes remerciements les plus vifs vont d'abord à ma directrice, Sophie Wauquier. À partir de notre premier rencontre (quand j'étais encore une petite italienne qui ne parlait que l'anglais) tu as toujours cru en moi et en mon potentiel. Merci pour ton soutien par rapport à la thèse et aussi par rapport à tout le reste. Tu as été pour moi une figure maternelle au sein du monde académique et dans ma nouvelle vie parisienne. Grâce à toi, j'ai eu la possibilité de découvrir la France, sa langue et d'apprendre beaucoup de choses que je n'aurais jamais pensé apprendre il y a quatre ans.

J'adresse aussi des remerciements très chaleureux à tous les membres de mon jury de thèse.

Je remercie Jean Pierre Chevrot pour les échanges scientifiques très enrichissants et professionnels que nous avons eus pendant les conférences et les séminaires durant ces années de thèse. Je vous remercie pour la disponibilité et l'ouverture scientifique que vous avez toujours manifesté envers mon travail.

Je tiens à adresser mes remerciements à Saveria Colonna, non seulement pour avoir accepté de faire partie de mon jury de thèse, mais aussi pour m'avoir guidée sur l'analyse des données et m'avoir aidée sur le plan méthodologique.

J'adresse également mes remerciements à Giorgio Magri. Merci pour le soutien pendant ces années de thèse et pour avoir ouvert mon regard aux théories phonologiques et aux parcours académiques. Tu as été la première personne que j'ai rencontrée au CNRS à Paris (en juillet 2012) et je suis ravie que tu sois aussi présent à la fin de mon parcours. Grazie infinite, Giorgio !

J'adresse mes remerciements à Pascale Colé, dont les travaux m'inspirent énormément, pour avoir accepté de faire partie de mon jury de thèse. Je suis honorée d'avoir la possibilité d'écouter vos retours et commentaires sur ma recherche.

Je remercie vivement Annie Rialland pour avoir accepté de faire partie de mon jury de thèse. J'avais échangé de manière très fructueuse avec vous lors de la conférence RFP à Lille. Je suis ravie d'avoir de nouveaux échanges avec vous et que vous puissiez assister à l'évolution de mon travail de thèse.

J'adresse également mes remerciements à tous mes collègues de Paris 8 et du CNRS : Jean-François Bourdin et Corinne Marius pour m'avoir accueillie, Asaf Bachrach et Isabelle Roy pour avoir écouté mes présentations et m'avoir donné des *feed-backs* indispensables pendant le séminaire des doctorants, le groupe de « l'atelier de phonologie » pour avoir toujours stimulé mes connaissances phonologiques.

Pendant ces années de thèse, j'ai eu la chance d'être entourée par des gens formidables. La plupart de mes collègues de doctorat sont devenus des amis. Nous avons partagé des moments inoubliables au sein du labo mais aussi dans la vie quotidienne (des apéros, des « repas linguistiques », des crêpes, des anniversaires, des fondues au chocolat et des mariages). Merci Nori-chan (loin mais proche de mon cœur), Heglyn, Silvia, Mathilde, Adèle, Annie, Matilde et Rosaria. Merci pour avoir été là et pour avoir partagé de magnifiques moments avec moi, entre des discussions sur la langue et des moments *gossip* autour d'un café ou d'un repas à la cantine. Des remerciements spéciaux vont à Mathilde et à Annie pour avoir relu la thèse. Merci pour ce gros travail de correction très soigné.

Je veux également remercier toutes les personnes qui ont permis la réussite des expériences présentées dans cette thèse.

J'adresse mes remerciements au personnel de l'école De La Salle à Dunkerque, d'abord à la directrice Mme Raffin-Peyloz pour avoir permis de conduire une étude longitudinale dans son établissement. Un grand merci va aux maîtresses Sylvie et Nathalie pour leur disponibilité et les informations fournies qui ont permis le développement de ma recherche. Je remercie également les parents qui ont autorisé la passation des tâches et les 43 enfants qui ont rendu cette expérience spéciale : votre « Madame Italienne » ne va jamais vous oublier. Je tiens à remercier la ville de Dunkerque pour l'hospitalité en la personne de Janine, Zélie et Flavien Felice qui m'ont mis en contact avec l'école.

Je veux également remercier tout l'institut Marie Curie (Orsay Ville) et, en particulier, Francesca Rocchetti, Sara Maria Cigna et Charlène Lesgi pour m'avoir fourni de nombreux « cobayes » adultes francophones. Pour la même raison je remercie aussi Mathilde et ses amis, les étudiants en acquisition du langage à Paris 3, les doctorantes francophones de Paris 8, Giulia (et son copain), Sarra, Ariel, Paul, Thomas, Cyril, Stéphane et Rémi.

Je remercie Coralie Vincent pour l'aide matérielle aux expériences et pour tous les conseils techniques, Marie-Anne Sallandre et Stéphanie Roy pour m'avoir permis de tester leurs petits enfants, Sébastien Pacton pour ses conseils dans l'interprétation des données psycholinguistiques.

Viennent maintenant les remerciements plus personnels, à mes amis non-linguistes, pour leur soutien constant pendant ma thèse.

J'adresse d'abord mes remerciements les plus profonds à mon compagnon, Anthony Felice. Pendant ces années de thèse tu as été, pour utiliser les paroles de Coldplay, « *the rock upon which I stand* ». Merci pour avoir toujours cru en moi, et pour m'avoir toujours encouragée. Tes surprises folles, tes attentions et tes câlins m'ont permis d'oublier les difficultés quotidiennes et d'achever cette thèse. Merci également pour m'avoir écoutée quand je répétais les présentations pour les conférences (et ce, même si tu n'es pas dans le domaine) et surtout, pour avoir relu mes pages pendant ton temps libre, au moins 10 milliards de fois et pour avoir eu la patience de corriger plusieurs fois les mêmes fautes d'orthographe (sans TROP t'énerver). Après trois ans d'immersion totale dans la linguistique, tu mérites une partie de mon titre.

Un grand merci va aussi aux Parigini (*vecchi e nuovi*) et à tous les amis italiens que j'ai rencontrés à Paris pendant ces quatre ans. Vous étiez source de déstress, de confort dans les moments les plus difficiles, à l'écoute dans les moments de folie (et je ne parle pas seulement des colorations des cheveux). Vous avez remplie ma vie de thésarde de nouveaux stimuli et vous avez également enrichi mon bagage d'expérience.

Je remercie et j'embrasse fort mes amis de longue date, qui ont toujours été là avec moi (et ma thèse). La distance n'a pas fait changer nos rapports et à chaque fois que je retourne en Italie, vous êtes capables de m'accueillir les bras ouverts, comme si je n'étais jamais partie, comme si rien n'avait changé. Tantissime grazie per tutto quello che ero, sono e sarò, miei cari amici di sempre (Bruno, Erica, Federica, Samanta, Linda, Schiavon, Miriam, Francesca, Luca, Vera, Ila, Serena, Brusa, Ludo, Federico Piva). J'adresse également de vifs remerciements à Andrea Penso et Federico Donatiello pour avoir partagé avec moi tous les moments positifs et négatifs de la vie académique. Je remercie aussi toutes les personnes qui m'ont rendue visite à Paris pendant ces années, qui m'ont permis de redécouvrir la ville et de prendre des pauses constructives à mon travail de thèse.

Enfin, mes plus profonds remerciements vont à ma famille, ma maman, mon papa et mon petit frère adoré Gabriele. Vous m'avez toujours encouragée à partir, à faire ce que je sentais être le mieux pour moi, sans juger aucun de mes choix, même les plus fous. C'est grâce à vous que je suis là où je suis maintenant. Merci pour votre soutien qui m'a permis de réussir à faire une thèse, pour être toujours là pour moi et pour m'attendre pleins d'amour toutes les fois que je rentre en Italie.

Je termine cette thèse le cœur plein d'émotions, la tête pleine de souvenirs et de connaissances. Merci à tous d'y être associé d'une manière ou d'une autre.

À Anthony

Table des matières

Liste des abréviations.....	25
-----------------------------	----

Chapitre 1. Introduction générale.....	27
--	----

Partie 1. Le cadre théorique et les questions de recherche

Chapitre 2. Phonologie et orthographe : les influences réciproques.....	35
---	----

2.1 Introduction.....	35
2.2 L'influence de l'orthographe en diachronie : l'effet Buben.....	35
2.3 Comment les compétences de la langue écrite influencent les compétences phonologiques des locuteurs.....	36
2.3.1 L'influence de l'orthographe sur la phonologie chez l'enfant.....	37
2.3.2 L'activation de l'orthographe en présence d'un input acoustique chez l'adulte.....	38
2.3.2.1 Les résultats des expériences en psycholinguistique.....	39
2.3.2.1.1 Voisinage orthographique et voisinage phonologique.....	39
2.3.2.2 L'activation par l'orthographe de zones cérébrales dédiées à l'oral.....	41
2.3.2.3 Différences de performances entre illettrés et alphabétisés.....	42
2.3.2.3.1 Différences neurologiques entre illettrés et alphabétisés.....	43
2.3.3 L'influence de l'orthographe sur la représentation phonologique.....	44
2.4 L'oral et l'écrit s'influencent réciproquement.....	47

Chapitre 3. Les mots enchaînés : les phénomènes de resyllabation en français.....	49
---	----

3.1 Les phénomènes de resyllabation en français.....	49
3.1.1 L'élision.....	49
3.1.2 L'enchaînement.....	50
3.1.2.1 L'enchaînement : un phénomène prosodique peu étudié.....	51
3.1.2.2 La modélisation de l'enchaînement.....	52
3.1.3 La liaison.....	53
3.1.3.1 Les contextes de la liaison et leur classification.....	54
3.1.3.1.1 Obligatoire ou catégorique ? Facultative ou variable ? Interdite par qui ?.....	56
3.1.3.2 L'analyse linguistique du corpus PFC.....	58
3.1.3.3 Le statut de la consonne de liaison (CL).....	61
3.2 Le <i>h</i> muet et le <i>h</i> aspiré.....	68
3.2.1 Le schwa dans le contexte <i>h</i> aspiré.....	70
3.2.2 Le coup de glotte dans le contexte <i>h</i> aspiré.....	71
3.2.3 Un contexte particulier à formaliser.....	72
3.3 Les phénomènes de resyllabation entre la phonologie et l'orthographe.....	74

Chapitre 4. Segmenter l'oral.....	77
-----------------------------------	----

4.1 Introduction.....	77
4.2 Segmenter à l'oral.....	77
4.2.1 La segmentation de la chaîne parlée chez les enfants.....	77
4.2.2 Le contexte déterminant-nom.....	80
4.2.3 L'acquisition de la liaison.....	81
4.2.3.1 Acquisition de la liaison : deux explications possibles.....	87

4.3 La segmentation chez les adultes : reconnaître les mots dans la chaîne orale...	94
4.3.1 L'accès au lexique chez l'adulte et l'input non-aligné.....	95
4.3.2 La reconnaissance des mots chez l'adulte francophone en présence de phénomènes de resyllabation.....	96
4.3.2.1 Perception de la CL et accès au lexique.....	98
4.4 De l'enfance à l'âge adulte : une question d'orthographe.....	100
Chapitre 5. Représenter et reconnaître les mots écrits.....	101
5.1 Introduction.....	101
5.2 <i>Emergent literacy</i> et <i>literacy</i>	101
5.2.1 La sensibilité à l'écrit, ou <i>print awareness</i> , chez les enfants entre <i>emergent literacy</i> et <i>literacy</i>	102
5.2.2 Le rôle de la lecture partagée et la sensibilité à l'écrit.....	104
5.3 Les compétences phonologiques qui caractérisent <i>literacy</i> et <i>emergent literacy</i>	105
5.4 <i>Literacy</i> : représenter à l'écrit.....	107
5.5 Reconnaître les mots écrits.....	110
5.5.1 Les modèles de lecture.....	111
5.5.1.1 Le modèle à double voie.....	111
5.5.1.1.1 Les modèles à étapes.....	113
5.5.1.1.2 Le modèle par analogie.....	113
5.5.1.1.2.1 L'acquisition de la lecture par analogie.....	114
5.5.1.3 Le modèle connexionniste.....	114
5.5.1.3.1 Le modèle connexionniste et l'acquisition.....	115
5.6 Lire et écrire en français.....	116
5.6.1 La connaissance des régularités orthographiques en français.....	116
5.6.2 Écrire en français.....	117
5.6.2.1 La morphologie et l'écriture du français.....	119
5.6.2.1.1 Le marquage du pluriel et la liaison.....	119
5.6.2.2 Orthographier la liaison.....	120
5.6.3 La lecture en français.....	122
5.6.3.1 La lecture de séquences en liaison.....	124
5.6.4 La syllabe comme unité de base dans les processus de lecture et d'écriture chez les enfants français.....	125
5.6.4.1 Le modèle IAS : les unités syllabiques dans le modèle de lecture....	127
5.6.4.2 Le rôle de la syllabe change selon le niveau de lecture des enfants. .	129
5.6.5 Apprentissage de la <i>literacy</i> à l'école en France : la méthode globale et la méthode alphabétique.....	130
5.7 Le rapport entre oral et écrit.....	131
Chapitre 6. Questions de recherche et hypothèses.....	133
6.1 Introduction.....	133
6.2 Apport empirique et nouvelles données.....	133
6.2.1 Des nouvelles données sur la langue orale.....	133
6.2.2 Des nouvelles données sur la langue écrite.....	134
6.2.3 La variable « lecture partagée ».....	134
6.3 Répondre aux questions de recherche et vérifier les hypothèses.....	135

Partie 2. La méthodologie utilisée et les modalités de traitement des données

Chapitre 7. Méthodologie et outils expérimentaux.....	143
7.1 Les participants.....	143
7.1.1 Les enfants.....	143
7.1.1.1 Les compétences scolaires.....	145
7.1.2 Les parents.....	150
7.2 La construction du matériel.....	154
7.2.1 La base de données pour les tâches orales : MANULEX.....	154
7.2.1.1 La construction du répertoire de mots pour les tâches orales : le choix du Mot1 et du Mot2.....	156
7.2.2 La base de données pour les tâches liées à la langue écrite.....	161
7.2.3 Le traitement des données.....	168

Partie 3. L'étude longitudinale

Chapitre 8. Première étape de l'étude longitudinale : le début de l'école primaire	171
.....
8.1 La première étape de l'étude longitudinale.....	171
8.1.2 Les hypothèses pour le début du CP.....	171
8.1.3 La procédure des premières tâches : octobre 2013.....	173
8.2 PERC ^{CP1} : la perception d'erreurs et leur correction.....	173
8.2.1 Procédure de la tâche PERC ^{CP1}	173
8.2.2 Matériel de la tâche PERC ^{CP1}	174
8.2.3 Le traitement de données pour la tâche PERC ^{CP1}	177
8.2.4 Résultats de la tâche PERC ^{CP1}	178
8.2.4.1 Résultats pour CONDcons.err (<i>le(s)[n]agneaux</i>).....	179
8.2.4.2 Résultats pour CONDnon-réal (<i>le(s)[Ø]anneaux</i>).....	181
8.2.4.3 Résultats pour CONDréal (<i>des[z]ananas</i>).....	183
8.2.4.4 Comparaison entre les conditions : CONDnon-réal vs. CONDcons.err.....	184
8.2.5 Les limites de la tâche PERC ^{CP1}	185
8.3 DENa ^{CP1} : la tâche de dénomination guidée au début du CP.....	187
8.3.1 Procédure de la tâche DENa ^{CP1}	187
8.3.2 Matériel de la tâche DENa ^{CP1}	187
8.3.3 Résultats de la tâche DENa ^{CP1}	189
8.3.3.1 L'analyse phonétique des non-réalisations au début du CP.....	193
8.3.3.2 La production problématique d'élision.....	197
8.3.3.2.1 L'analyse phonétique : [z].....	198
8.3.3.2.2 L'analyse phonétique : [Ø].....	199
8.3.3.2.3 L'analyse phonétique de [ə].....	200
8.3.3.2.4 La production d'élision comme une conséquence de l'analyse de la liaison.....	201
8.3.3.3 Le rôle de la fréquence.....	202
8.4 Discussion des données de production et de perception au début du CP.....	204
8.5 La variante sociolinguistique.....	208
8.6 Perception d'erreurs, dénomination guidée et lecture partagée : y a-t-il une corrélation ?.....	212

Chapitre 9. Deuxième étape de l'étude longitudinale : la fin du CP.....	215
9.1 La deuxième étape de l'étude.....	215
9.1.1 Les hypothèses pour la fin du CP.....	216
9.1.2 La procédure des tâches orales : mai 2014.....	217
9.2. Tâche PERC ^{CP2} : perception d'erreurs et reformulation.....	218
9.2.1 Procédure de la tâche PERC ^{CP2}	218
9.2.2 Résultats généraux de la tâche PERC ^{CP2}	219
9.3 Tâche DENa ^{CP2} : dénomination guidée.....	220
9.3.1 Procédure de la tâche DENa ^{CP2}	220
9.3.2 Résultats généraux de la tâche DENa ^{CP2}	220
9.3.3 L'analyse phonétique des non-réalisations à la fin du CP.....	223
9.3.4 Le rôle de la fréquence.....	224
9.3.5 La variante sociolinguistique.....	225
9.4 Tâche DENb ^{CP2} : la tâche de dénomination guidée, une autre façon de produire	226
9.4.1 Procédure de la tâche DENb ^{CP2}	227
9.4.2 Résultats de la tâche DENb ^{CP2}	228
9.4.3 Une comparaison entre les deux tâches de dénomination en mai : DENa ^{CP2} vs. DENb ^{CP2}	230
9.5 Les tâches liées à la langue écrite : juin 2014.....	234
9.5.1 La tâche PROD.ECR ^{CP2}	234
9.5.1.1 Procédure de la tâche PROD.ECR ^{CP2}	234
9.5.1.2 Résultats de la tâche PROD.ECR ^{CP2}	235
9.5.1.2.1 La production écrite de noms.....	236
9.5.1.2.2 Liaison, élision, enchaînement dans PROD.ECR ^{CP2} : frontières lexicales vs. frontières syllabiques.....	236
9.5.1.2.3 Les productions écrites qui respectent les frontières lexicales..	238
9.5.1.2.4 Des questions morphologiques.....	239
9.5.1.2.5 Les productions écrites qui respectent les frontières syllabiques	240
9.5.1.2.6 <i>H</i> comme initiale à l'écrit.....	245
9.5.1.2.7 Le rôle de la fréquence et de la familiarité dans PROD.ECR ^{CP2}	248
9.5.2 La tâche LECT ^{CP2} de lecture à voix haute.....	250
9.5.2.1 Procédure et matériel de la tâche LECT ^{CP2}	250
9.5.2.2 Résultats généraux de la tâche LECT ^{CP2}	251
9.5.2.2.1 Les productions des phénomènes de resyllabation dans LECT ^{CP2}	253
9.5.2.2.2 Les séquences à <i>h</i> muet et <i>h</i> aspiré au début du nom.....	256
9.5.2.2.3 Analyses phonétiques de la tâche LECT ^{CP2}	259
9.5.2.2.3.1 Analyse phonétique des non-réalisations (NON- REAL.LECT ^{CP2}).....	259
9.5.2.2.3.2 Les non-réalisations en présence de <i>h</i> aspiré et de <i>h</i> muet.	262
9.5.2.2.3.3 Analyse phonétique de PRA.LECT ^{CP2}	264
9.5.2.2.4 Le rôle de la fréquence et de la familiarité dans la tâche LECT ^{CP2}	265
9.5.3 Le rapport entre les productions écrites et les productions en lecture à la fin du CP.....	267
9.5.3.1 Le rapport entre productions écrites et lecture pour <i>h</i> muet et <i>h</i> aspiré	

.....	270
9.6 Le rapport entre la <i>literacy</i> et les productions orales à la fin du CP.....	271
9.7 Le rapport entre la fréquence des activités de lecture partagée et les résultats des tâches à la fin du CP.....	276
9.8 Discussions des hypothèses à la fin du CP.....	277
9.8.1 Les productions orales des phénomènes de resyllabation à la fin du CP. .	277
9.8.2 Les productions écrites des phénomènes de resyllabation à la fin du CP. .	280
9.8.3 Les relations entre productions écrites et productions orales.....	284
Chapitre 10. Troisième étape de l'étude longitudinale : l'année au CE1.....	287
10.1 Troisième étape de l'étude longitudinale : janvier 2015.....	287
10.1.1 Les hypothèses au CE1.....	288
10.1.2 Participants.....	289
10.2 DENa ^{CE1} : la tâche de dénomination guidée au CE1.....	290
10.2.1 Procédure de la tâche DENa ^{CE1}	290
10.2.2 Résultats de la tâche DENa ^{CE1}	290
10.2.3 Analyse phonétique des non-réalisations au CE1.....	293
10.2.4 Le rôle de la fréquence.....	294
10.2.5 La variante sociolinguistique.....	295
10.3 Les tâches liées à la langue écrite : janvier 2015.....	296
10.3.1 La tâche d'écriture PROD.ECR ^{CE1}	297
10.3.1.1 Procédure de la tâche PROD.ECR ^{CE1}	297
10.3.1.2 Résultats de la tâche PROD.ECR ^{CE1}	298
10.3.1.2.1 Les productions écrites qui respectent les frontières lexicales.....	299
10.3.1.2.2 Des questions morphologiques.....	300
10.3.1.2.3 Les productions écrites qui respectent les frontières syllabiques.....	301
10.3.1.2.4 Les productions écrites des <i>fillers</i>	302
10.3.1.2.5 Orthographier les séquences où Mot2 commence par <i>h</i>	303
10.3.1.2.6 Le rôle de la fréquence et de la familiarité dans la PROD.ECR ^{CE1}	304
10.3.2 La tâche LECT ^{CE1} de lecture à voix haute.....	306
10.3.2.1 Procédure et matériel pour la tâche LECT ^{CE1}	306
10.3.2.2 Résultats généraux de la tâche LECT ^{CE1}	306
10.3.2.2.1 Les productions de phénomènes de resyllabation en LECT ^{CE1}	307
10.3.2.2.2 Analyse phonétique des non-réalisations (NON-REAL.ECT ^{CE1}).....	311
10.3.2.2.3 Les non-réalisations en présence de <i>h</i> aspiré et de <i>h</i> muet.....	312
10.3.2.2.4 L'analyse phonétique de PRA.LECT ^{CE1}	314
10.3.2.2.5 Le rôle de la fréquence et de la familiarité de la tâche LECT ^{CE1}	316
10.3.3 Le rapport entre la lecture et l'écriture au CE1.....	318
10.4 Le rapport entre la <i>literacy</i> et les productions orales au CE1.....	319
10.5 Les tâches d'amorçage (AMOA et AMOB).....	322
10.5.1 Le groupe d'adultes.....	323
10.5.2 Procédure.....	324
10.5.3 La tâche d'amorçage AMOA.....	327
10.5.3.1 Répertoire pour AMOA.....	327

10.5.3.2 Les hypothèses pour AMOa.....	329
10.5.3.3 Les résultats de la tâche AMOa.....	331
10.5.3.4 Discussion sur les résultats de AMOa.....	333
10.5.4 La tâche d'AMOb.....	334
10.5.4.1 Le répertoire pour la tâche AMOb.....	335
10.5.4.2 Les hypothèses pour AMOb.....	336
10.5.4.3 Les résultats de la tâche AMOb.....	337
10.5.4.4 Discussion des résultats de AMOb.....	339
10.6 Le rapport entre les résultats des tâches au CE1 et la fréquence des activités de lecture partagée.....	340
10.7 Discussion et conclusions.....	341
10.7.1 Acquisition de la langue orale et <i>literacy</i>	341
10.7.2 Après le CE1 : une comparaison entre nos résultats et les données de Dugua et Baclesse (2014).....	343
Chapitre 11. L'évolution entre CP et CE1 : des données aux conclusions.....	347
11.1 Introduction.....	347
11.2 Des problématiques méthodologiques dans les tâches proposées.....	348
11.2.1 La tâche de perception d'erreurs.....	348
11.2.2 La tâche de dénomination guidée.....	348
11.2.3 Les tâches liées à la langue écrite.....	349
11.2.4 La tâche d'amorçage.....	349
11.3 Les trois étapes de l'étude longitudinale : la comparaison des données.....	350
11.3.1 La tâche de perception d'erreurs : PERC ^{CP1} et PERC ^{CP2}	350
11.3.2 La tâche orale de dénomination guidée (DENa) entre CP et CE1.....	351
11.3.2.1 L'analyse phonétique du traitement des non-réalisations en production entre CP et CE1.....	356
11.3.3 L'analyse de la fréquence dans les productions orales pour DENa ^{CP1} , DENa ^{CP2} et DENa ^{CE1}	359
11.3.4 Les tâches liées à la langue écrite entre CP et CE1.....	359
11.3.4.1 La tâche d'écriture : PROD.ECR ^{CP2} et PROD.ECR ^{CE1}	360
11.3.4.1.1 Les erreurs dans les productions écrites.....	360
11.3.4.1.2 Les productions écrites qui respectent les frontières lexicales ou les frontières syllabiques.....	361
11.3.4.1.3 Le marquage du pluriel à l'écrit.....	362
11.3.4.1.4 <i>H</i> aspiré et <i>h</i> muet à l'écrit.....	363
11.3.4.1.5 La familiarité des contextes et la fréquence des items à écrire.....	364
11.3.4.2 La tâche de lecture : LECT ^{CP2} et LECT ^{CE1}	364
11.3.4.2.1 Les lectures correctes et les lectures syllabiques.....	365
11.3.4.2.2 Les erreurs de segmentation en lecture.....	366
11.3.4.2.3 L'analyse phonétique des NON-REAL.LECT ^{CP2} et NON-REAL.LECT ^{CE1}	367
.....	367
11.3.4.2.4 <i>H</i> aspiré et <i>h</i> muet en lecture.....	368
11.3.4.2.4.1 L'analyse phonétique de NON-REAL.LECT ^{CP2} et NON-REAL.LECT ^{CE1} dans les contextes à <i>h</i> initial.....	370
11.3.4.2.5 La familiarité des contextes et la fréquence des items à lire.....	371
11.3.5 Comparaison des corrélations.....	373

11.3.5.1 Rapport entre la <i>literacy</i> et les productions orales.....	373
11.3.5.1.1 Les productions écrites et les productions orales.....	374
11.3.5.1.2 Le marquage du pluriel à l'écrit et la réalisation correcte de la liaison en [z].....	375
11.3.5.1.3 La lecture et la production orale.....	375
11.3.5.1.4 Le rapport entre la lecture et l'écriture au CP et au CE1.....	376
11.3.6 Les autres variables analysées.....	377
11.3.6.1 La variable sociolinguistique.....	377
11.3.6.2 La variable « lecture partagée ».....	378

Partie 4. Discussion et conclusions

Chapitre 12. Discussion des résultats et vérification des hypothèses.....	381
12.1 Les données sur la langue orale après 6 ans.....	381
12.2 La variable sociolinguistique.....	383
12.3 Les données sur la langue écrite.....	384
12.4 L'analyse phonétique des non-réalisations en lecture.....	385
12.5 La lecture et l'écriture : quel rapport ?.....	386
12.6 Les noms à h initial.....	387
12.7 Le rôle de la lecture partagée.....	387
12.8 Discussion des questions de recherche et vérification des hypothèses.....	389
12.9 Un modèle pour l'acquisition de la liaison après 6 ans.....	393
12.10 L'élision comme un phénomène facile à maîtriser.....	397
12.11 L'enchaînement : des résultats vagues.....	398
12.12 Perspectives futures.....	401
12.12.1 Amélioration de la base de données.....	401
12.12.2 Les données phonétiques chez les plus petits.....	402
12.12.3 La tâche d'amorçage : tester d'autres groupes de participants et formaliser un modèle de la lecture.....	402
12.12.4 D'autres recherches qui mettent en rapport la langue orale et la <i>literacy</i>	403
Références bibliographiques.....	407
Annexes.....	425

Figures

Figure 1. Spectrogramme d'une phrase produite par un locuteur natif du français.....	27
Figure 2. Modèle autosegmental pour le mot italien <i>frutto</i> (Ruvoletto, 2011).....	46
Figure 3. L'étape précédente du modèle de représentation autosegmental pour le mot <i>frutto</i> (Ruvoletto, 2011).....	47
Figure 4. La représentation d'après Clements et Keyser (1981) dans Encrevé (1983 : 56).....	52
Figure 5. Représentation par Encrevé (1983 : 57).....	52
Figure 6. Représentation de la séquence <i>dont un ami</i> par Clements et Keyser (1983 : 103).....	62
Figure 7. Représentation de la CL comme consonne flottante dans Encrevé (1983 : 57).....	62
Figure 8. Représentation de la liaison enchaînée.....	63
Figure 9. Représentation de la liaison non-enchaînée.....	63

Figure 10. Représentation des séquences sans liaison.....	63
Figure 11. Représentation de Pagliano dans Côté (2005 : 67).....	73
Figure 12. Représentation dans Côté (2005 : 68).....	73
Figure 13. Schéma d'acquisition de la liaison : le premier stade (Wauquier & Braud, 2005).....	91
Figure 14. Schéma d'acquisition de la liaison : le deuxième stade (Wauquier & Braud, 2005).....	92
Figure 15. Schéma d'acquisition de la liaison : le stade final (Wauquier & Braud, 2005).....	93
Figure 16. Schéma qui met en rapport oral et écrit dans l'acquisition de la <i>literacy</i> par Ziegler et Goswami (2005 : 4).....	108
Figure 17. Modèle cognitif à double voie de la lecture des mots isolés (Morton & Patterson, 1980).....	111
Figure 18. Le modèle connexionniste par Seidenberg et McClelland (1989).....	115
Figure 19. Lecture des mots et pseudo-mots dans plusieurs langues (Seymour <i>et al.</i> , 2003 par Ziegler & Goswami, 2005 : 11).....	123
Figure 20. Le modèle IAS (Mathey <i>et al.</i> , 2006).....	128
Figure 21. Outils pour les connections graphèmes-phonèmes sur les tables des élèves.....	144
Figure 22. Tableaux dans les classes pour l'apprentissage de l'écriture.....	144
Figure 23. Enregistrement des données pour les exercices du livret.....	148
Figure 24. Classification des mots écrits pour créer la base de données.....	163
Figure 25: Liste de mots traités dans la base de données.....	164
Figure 26. Exemple d'une feuille <i>Excel</i> utilisée pour classer les données.....	168
Figure 27. Spectrogramme pour la production de continuum.....	194
Figure 28. Spectrogramme pour l'insertion d'un coup de glotte.....	194
Figure 29. Spectrogramme pour une production ambiguë.....	195
Figure 30. Spectrogramme pour la production d'une pause.....	195
Figure 31. Spectrogramme pour l'insertion d'un coup de glotte en élision.....	200
Figure 32. Représentation de la liaison par Encrevé (1988).....	206
Figure 33. Graphique des résultats chez les enfants de cadres et d'ouvriers (Nardy & Dugua, 2011 : 114).....	210
Figure 34. Spectrogramme pour la production en lecture de [le/a.fiʃ].....	260
Figure 35. Spectrogramme pour la production en lecture [le.ʔa.fiʃ].....	260
Figure 36. Spectrogramme pour la production en lecture [ɛ.wa.zo].....	261
Figure 37. Représentation des résultats pour la production d'enchaînement.....	278
Figure 38. Représentation de l'insertion d'un coup de glotte en liaison.....	279
Figure 39. Structures de AMOa et AMOb.....	325
Figure 40. Structure d'un trial dans AMOa et AMOb.....	326
Figure 41. Stade 2 dans le modèle de l'acquisition de la liaison proposé par Wauquier et Braud (2005).....	393
Figure 42. Étape 3 bis.....	394
Figure 43. Dernière étape, acquisition du double flottement.....	395
Figure 44. Nouvelle version du modèle phonologique (Scheer <i>et al.</i> , 2015).....	395

Tableaux

Tableau 1. Classification de la liaison et des exemplifications de Wauquier (2010 : 5) après Delatre (1947).....	55
--	----

Tableau 2. Synthèse dans Côté (2014 : 13) : différentes options pour la représentation de la CL.....	68
Tableau 3. Description des erreurs de segmentation des enfants après Wauquier (2010 : 8).....	84
Tableau 4. Erreurs et contextes à partir de Basset (2000) dans Chevrot et Fayol (2001).....	86
Tableau 5. Description des classes des participants au CP et au CE1.....	144
Tableau 6. Les exercices du livret de passation, en détails.....	146
Tableau 7. Résultats des tâches qui évaluent les composantes sonores et visuelles dans CP A et CP B.....	149
Tableau 8. Pourcentage de réussite pour les 8 tâches qui évaluent les composantes sonores et visuelles.....	149
Tableau 9. Langues parlées par les parents des enfants.....	151
Tableau 10. Classification du niveau scolaire des parents.....	151
Tableau 11. Classification des catégories socio-professionnelles des parents des enfants au CP.....	152
Tableau 12. Fréquences des activités de lecture partagée déclarées par les parents....	153
Tableau 13. Résultats des tâches d'écriture du livret de passation selon la fréquence de la lecture partagée.....	154
Tableau 14. Description du répertoire utilisé dans les tâches orales.....	159
Tableau 15. Répertoire des adjectifs utilisés comme Mots1 dans les tâches orales....	160
Tableau 16. Répertoire des Mots1 utilisés dans les tâches orales.....	160
Tableau 17. Répertoire des Mots2 utilisés dans les tâches liées à la langue écrite et leurs fréquences.....	165
Tableau 18. Répertoire des Mots2 à initiale <i>h</i> utilisés dans les tâches liées à la langue écrite et leurs fréquences.....	165
Tableau 19. Séquences de Mots2 utilisées dans les tâches liées à la langue écrite et classification des contextes (connus vs. inconnus).....	167
Tableau 20. Alternance de phrases tests et <i>fillers</i> avec les réponses attendues dans les 6 blocs de PERC.....	175
Tableau 21. Répertoire de phrases tests dans les 3 conditions pour les contextes dans PERC.....	177
Tableau 22. Réponses possibles et évaluations données dans PERC.....	178
Tableau 23. Résultats de PERC ^{CP1} pour les trois conditions.....	179
Tableau 24. Classification des erreurs des enfants après l'analyse des corrections pour CONDcons.err dans PERC ^{CP1}	180
Tableau 25. Identifications de l'erreur-cible pour CONDcons.err en PERC ^{CP1}	181
Tableau 26. Classification des erreurs après l'analyse des reformulations pour CONDnon-réal dans PERC ^{CP1}	182
Tableau 27. Identifications de l'erreur-cible pour CONDnon.réal dans PERC ^{CP1}	183
Tableau 28. Répertoire des séquences tests pour DENa.....	188
Tableau 29. Ordre de présentation du répertoire pour DENa.....	189
Tableau 30. Résultats généraux pour DENa ^{CP1}	189
Tableau 31. Résultats pour chaque phénomène de resyllabation dans DENa ^{CP1}	191
Tableau 32. Productions phonétiques dans les non-réalisations dans DENa ^{CP1}	196
Tableau 33. Relation entre le phénomène à produire et la réalisation de coups de glotte dans DENa ^{CP1}	197
Tableau 34. Longueur de [z] dans les deux productions (DENa ^{CP1}).....	199

Tableau 35. Indice SFI et nombre d'erreurs dans DENa ^{CP1}	202
Tableau 36. Scores moyens d'erreurs dans DENa ^{CP1} selon le niveau scolaire des parents	209
Tableau 37. Scores de réponses correctes pour les trois conditions dans PERC ^{CP2}	219
Tableau 38. Résultats généraux dans DENa ^{CP2}	220
Tableau 39. Scores par type d'erreur et pour chaque phénomène dans DENa ^{CP2}	221
Tableau 40. Analyse phonétique pour les non-réalisations dans DENa ^{CP2}	223
Tableau 41. Scores moyens d'erreurs des enfants dans DENa ^{CP2} selon le niveau d'études des parents.....	225
Tableau 42. Répertoire pour DENb.....	227
Tableau 43. Résultats généraux dans DENb ^{CP2}	228
Tableau 44. Classification des erreurs dans DENb ^{CP2} selon le type d'erreur et le phénomène de resyllabation.....	229
Tableau 45. Comparaison des résultats généraux : DENa vs. DENb.....	231
Tableau 46. Résultats par type d'erreur et par phénomène : DENa vs. DENb.....	232
Tableau 47. Bloc de séquences tests utilisées pour PROD.ECR ^{CP2}	234
Tableau 48. Pourcentage des productions écrites qui respectent les frontières lexicales dans PROD.ECR ^{CP2}	238
Tableau 49. Pourcentages du marquage du pluriel à l'écrit dans les séquences en liaison en [z] dans PROD.ECR ^{CP2}	240
Tableau 50. Pourcentages de productions écrites à "frontières lexicales" dans PROD.ECR ^{CP2}	241
Tableau 51. Segmentations incorrectes à l'écrit pour la liaison en [z] dans PROD.ECR ^{CP2}	242
Tableau 52. Segmentations incorrectes à l'écrit pour la liaison en [n] dans PROD.ECR ^{CP2}	243
Tableau 53. Segmentations incorrectes à l'écrit pour l'élision dans PROD.ECR ^{CP2}	244
Tableau 54. Segmentations incorrectes à l'écrit pour l'enchaînement dans PROD.ECR ^{CP2}	245
Tableau 55. Pourcentage de <h> orthographié au début du Mot2 dans PROD.ECR ^{CP2}	246
Tableau 56. Segmentations incorrectes à l'écrit en présence d'un Mot2 à h initial dans PROD.ECR ^{CP2}	247
Tableau 57. Pourcentages des lectures selon les trois catégories dans LECT ^{CP2}	252
Tableau 58. ERR.LECT ^{CP2} pour les phénomènes de resyllabation en lecture.....	253
Tableau 59. Pourcentages de NON-REAL et PRA pour la liaison dans LECT ^{CP2}	254
Tableau 60. Pourcentages de NON-REAL et PRA pour l'élision dans LECT ^{CP2}	256
Tableau 61. Pourcentages de PRA, REAL et NON-REAL dans les contextes où Mot2 commence par h aspiré dans LECT ^{CP2}	257
Tableau 62. Pourcentages de PRA, REAL et NON-REAL dans les contextes où Mot2 commence par h muet dans LECT ^{CP2}	258
Tableau 63. Analyse des non-réalisations pour LECT ^{CP2}	261
Tableau 64. Analyse des non-réalisations pour le Mot2 à h aspiré initial en lecture dans LECT ^{CP2}	263
Tableau 65. Analyse de non-réalisations pour le Mot2 à h muet initial dans LECT ^{CP2}	263
Tableau 66. Résultats généraux de la tâche DENa ^{CE1}	290
Tableau 67. Scores par type d'erreur et pour chaque phénomène dans DENa ^{CE1}	291

Tableau 68. Analyse phonétique pour les non-réalisations dans DENa ^{CE1}	293
Tableau 69. Scores moyens des erreurs des enfants dans DENa ^{CE1} selon le niveau d'études des parents.....	295
Tableau 70. <i>Fillers</i> pour PROD.ECR ^{CE1}	297
Tableau 71. Insertion des <i>fillers</i> dans le bloc de séquences tests dans PROD.ECR ^{CE1}	297
Tableau 72. Pourcentages de productions écrites à « frontières lexicales » dans PROD.ECR ^{CE1}	299
Tableau 73. Pourcentages des productions écrites globalement correctes dans PROD.ECR ^{CE1}	300
Tableau 74. Pourcentages du marquage du pluriel à l'écrit dans la liaison en [z] dans PROD.ECR ^{CE1}	301
Tableau 75. Pourcentages de productions écrites à « frontières lexicales » dans PROD.ECR ^{CE1}	301
Tableau 76. Erreurs de segmentation à l'écrit pour les <i>fillers</i> au CE1 dans PROD.ECR ^{CE1}	302
Tableau 77. Pourcentage de <h> orthographié au début du Mot2 au CE1 dans PROD.ECR ^{CE1}	303
Tableau 78. Mauvaises segmentations en contexte <i>h</i> aspiré ou <i>h</i> muet dans PROD.ECR ^{CE1}	304
Tableau 79. Pourcentages pour les trois types de lecture chez les enfants dans LECT ^{CE1}	307
Tableau 80. ERR.LECT ^{CE1} pour les 4 phénomènes.....	307
Tableau 81. Pourcentages de NON-REAL ^{CE1} , PRA ^{CE1} et NONEnch ^{CE1}	309
Tableau 82. Réalisation et non-réalisation des phénomènes de resyllabation dans LECT ^{CE1} pour les séquences à <i>h</i> aspiré initial.....	310
Tableau 83. Réalisation et non-réalisation des phénomènes de resyllabation dans LECT ^{CE1} pour les séquences à <i>h</i> muet initial.....	311
Tableau 84. Analyse phonétique des non-réalisations dans LECT ^{CE1}	312
Tableau 85. Analyse phonétique de non-réalisations dans LECT ^{CE1} pour le contexte à <i>h</i> aspiré initial.....	313
Tableau 86. Analyse phonétique de non-réalisations dans LECTCE1 pour le contexte à <i>h</i> muet initial.....	313
Tableau 87. Classification des diplômes du groupe d'adultes testés.....	323
Tableau 88. Classification des métiers des adultes testés.....	324
Tableau 89. Répertoire pour AMOa.....	329
Tableau 90. Répertoire pour AMOb.....	336
Tableau 91. Comparaison des résultats pour la dénomination guidée entre Ruvoletto (2016) et Dugua & Baclesse (2014).....	344
Tableau 92. Comparaison des résultats pour la lecture entre Ruvoletto (2016) et Dugua & Baclesse (2014).....	345
Tableau 93. Analyses phonétiques des non-réalisations de liaison entre CP et CE1...	356
Tableau 94. Analyses phonétiques des non-réalisations d'élision entre CP et CE1....	358
Tableau 95. Analyses phonétiques des non-réalisations d'enchaînement entre CP et CE1.....	358
Tableau 96. Indice de corrélation entre les erreurs de production et la fréquence entre CP et CE1.....	359
Tableau 97. Marquage du pluriel entre CP et CE1.....	362

Tableau 98. Analyses phonétiques des lectures entre CP et CE1 pour la liaison et l'enchaînement.....	367
Tableau 99. Analyses phonétiques des lectures en contexte à <i>h</i> initial.....	370
Tableau 100. Corrélations entre la lecture et la fréquence au CP et au CE1.....	372
Tableau 101. Corrélations écrit-oral entre CP et CE1.....	374
Tableau 102. Rapport entre oral et écrit pour le marquage du pluriel.....	375
Tableau 103. Rapports productions orales et lectures au CP et au CE1.....	376
Tableau 104. Rapport entre l'écriture et la lecture au CP et au CE1.....	377

Graphiques

Graphique 1. Distribution logarithmique des fréquences SFI dans MANULEX.....	157
Graphique 2. Identification de l'erreur cible dans CONDcons.err et CONDnon-réal.	184
Graphique 3. Non-réalisations vs. Insertions d'une consonne erronée dans DENa ^{CP1} .	192
Graphique 4. Pourcentages de formes produites pour l'élosion dans DENa ^{CP1}	198
Graphique 5. Longueur de [ə] dans les trois productions analysées.....	201
Graphique 6. Scores des enfants dans DENa ^{CP1} selon la profession des parents.....	211
Graphique 7. Fréquence des lectures partagées selon la profession des parents.....	213
Graphique 8. Fréquence des lectures partagées selon le niveau scolaire des parents.	214
Graphique 9. Erreurs d'insertion d'une consonne erronée ou de non-réalisation dans DENa ^{CP2}	222
Graphique 10. Les productions des enfants dans DENa ^{CP2} selon le métier des parents.....	226
Graphique 11. Erreurs d'insertion d'une consonne erronée et de non-réalisations dans DENb ^{CP2}	230
Graphique 12. Scores moyens de productions écrites (frontières lexicales vs. frontières syllabiques) dans PROD.ECR ^{CP2}	237
Graphique 13. Scores d'erreurs à l'écrit : contexte connu vs. contexte inconnu.....	248
Graphique 14. Scores moyens de lectures syllabiques dans LECT ^{CP2} selon la longueur de Mot2.....	252
Graphique 15. Longueur de [z] et [e] dans les deux productions dans LECT ^{CP2}	264
Graphique 16. Relation entre les scores de productions écrites à "frontières syllabiques" et les scores de lectures syllabiques.....	269
Graphique 17. Erreurs de non-réalisations et d'insertion d'une consonne erronée dans DENa ^{CE1}	292
Graphique 18. Scores des enfants dans DENa ^{CE1} selon la profession des parents.....	296
Graphique 19. Scores de productions écrites au CE1 : « frontières syllabiques » vs. « frontières lexicales ».....	299
Graphique 20. Longueur de [z] et de [e] dans les trois types de lecture au CE1.....	315
Graphique 21. Longueur de [z] et de [e] en lecture et en dénomination guidée.....	316
Graphique 22. Réponses correctes dans AMOa pour les trois conditions chez les enfants et les adultes.....	332
Graphique 23. TR dans AMOa pour les trois conditions chez les enfants et les adultes.....	332
Graphique 24. Réponses correctes à AMOb dans les trois conditions chez les enfants et les adultes.....	338
Graphique 25. TR dans AMOb pour les trois conditions chez les enfants et chez les adultes.....	339
Graphique 26. Comparaison des réponses correctes pour PERC : CP vs. CE1.....	351

Graphique 27. Comparaison des réponses correctes DENa dans les trois moments...	352
Graphique 28. Comparaison des erreurs en liaison et élision dans les trois moments	353
Graphique 29. Non-réalisations et insertions d'une consonne erronée entre CP et CE1	354
Graphique 30. Non-réalisations et insertions d'une consonne erronée pour la liaison	355
Graphique 31. Erreurs en PROD.ECR du CP au CE1	360
Graphique 32. Productions écrites : frontières syllabiques et frontières lexicales entre CP et CE1	361
Graphique 33. Scores moyens de productions écrites de <h> au début du Mot2	363
Graphique 34. Niveaux de lecture des enfants entre CP et CE1	365
Graphique 35. PRA et NON-REAL en lecture au CP et au CE1	367
Graphique 36. Réalisations des phénomènes consonantiques en lecture entre CP et CE1	369
Graphique 37. Non-réalisations des phénomènes de resyllabation en lecture entre CP et CE1	370

Liste des abréviations

AMOa : première tâche d'amorçage.

AMOb : deuxième tâche d'amorçage.

CP¹ : la tâche a eu lieu au début du cours préparatoire.

CP² : la tâche a eu lieu à la fin du cours préparatoire.

CE¹ : la tâche a eu lieu au cours élémentaire première année.

CF : consonne fixe.

CGP : correspondance graphème-phonème.

CL : consonne de liaison.

CONDcons.err : condition d'insertion d'une consonne erroné dans PERC (ex. *le(s)* [n]agneaux [le.na.no]).

CONDnon-réal : condition de non-réalisation dans la tâche PERC (ex. *le(s)* [Ø]anneaux [leØa.no]).

CONDréal : condition de réalisation dans la tâche PERC (ex. *des[z]*ananas [de.za.na.na]).

DENa : première tâche de dénomination guidée.

DENb : deuxième tâche de dénomination guidée.

ERR.LECT : erreur en lecture dans la réalisation du phénomène de resyllabation.

FILL.LECT : lecture d'une séquence *fillers*.

Fréq. : fréquence des items dans la base de données que nous avons créée.

FS : forme de surface.

FSJ : forme sous-jacente.

LECT : tâche de lecture.

NONench.LECT : type de ERR.LECT où la CL est réalisée à la fin du Mot1 et, après une pause, le Mot2 est lu comme un mot à voyelle initiale (ex. *les[z]*affiches [lez/a.fi] pour [le.za.fi]).

NON-REAL.LECT : type de ERR.LECT qui implique la non-réalisation du phénomène de resyllabation dans la tâche de lecture (ex. *les[Ø]*affiches [leØa.fi]).

PERC : tâche de perception et de la reformulation.

PRA.LECT : type de ERR.LECT où l'enfant, après avoir lu l'article, propose une pause et ensuite un Mot2 qui commence par la CL (Pause + Remplissage de l'Attaque, ex. *les-pause-[z]affiches* [le/za.fiʃ] pour [le.za.fiʃ]).

PROD.ECR : tâche de production écrite.

REAL.LECT : réalisation du phénomène de resyllabation dans la tâche de lecture.

SFI : Indice de fréquence standard dans MANULEX.

SYMBOLES

Dans la représentation phonétique :

∅ : un segment qui est normalement réalisé n'est pas réalisé.

/ : réalisation d'une pause entre deux mots

Les productions écrites sont transcrites entre guillemets (ex. <avion>).

Chapitre 1

Introduction générale

Quand nous entendons une langue étrangère, nous percevons un continuum où nous n'avons pas conscience d'individualiser les unités qui le composent, ou de déterminer leurs frontières. De plus, à l'inverse de la langue écrite (où les espaces blancs sur papier signalent les frontières), il n'y a pas de pauses bien marquées dans la langue orale qui aident au découpage de la chaîne acoustique. Comme le soulignent Vroomen & de Gelder (1999 : 413) "the 'silence' that we hear between words is in the head of the speaker, and not in the signal".

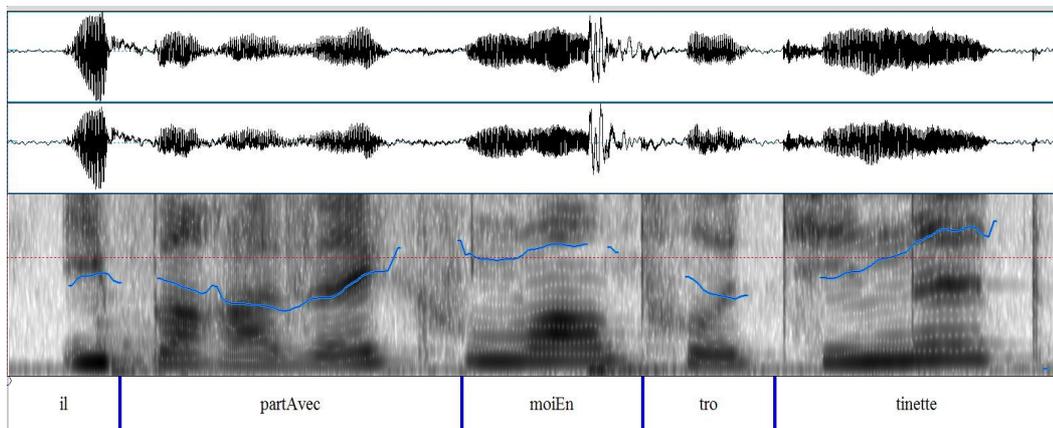


Figure 1. Spectrogramme d'une phrase produite par un locuteur natif du français

Le spectrogramme ci-dessus (Figure 1) montre que dans le parlé spontané, une phrase comme « il part avec moi en trottinette », composée de 6 mots séparés, ne présente aucun signe permettant un découpage du signal acoustique en 6 morceaux correspondants. Nous appelons « mot » un « son ou groupe de sons articulés ou figurés graphiquement, constituant une unité porteuse de signification à laquelle est liée, dans

une langue donnée, une représentation d'un être, d'un objet, d'un concept, etc » en utilisant la définition donnée dans le *Trésor de la langue française* (TLF, 1985 : 1108-1113). Dans la parole spontanée, il n'y a pas de signes spécifiques qui aident à isoler les mots que nous pouvons très facilement isoler à l'écrit grâce aux blancs. Reconnaître les mots dans une langue en situation de communication aide non seulement à décoder les informations que notre locuteur veut transmettre, mais nous permet aussi, en tant qu'interlocuteurs, de pouvoir les utiliser pour transmettre un message. À l'oral, savoir découper les mots est essentiel pour toutes les communications.

Par ailleurs, dans la langue parlée, les phénomènes de coarticulation agissent entre les frontières des mots. Par exemple, la cooccurrence d'un mot à consonne finale et d'un mot à voyelle initiale cause un non-alignement entre les frontières syllabiques et les frontières lexicales. En anglais, dans la séquence *cold ice* (fr. *glace froide*), la consonne finale [d] de *cold* (fr. *froide*), en position de coda, devient l'attaque du mot *ice* (fr. *glace*) : *cold* [kəʊld] + *ice* [aɪs] → *cold ice* [kəʊl.dais]. Cependant, si nous lisons la phrase à l'écrit, nous pouvons voir qu'elle est formée de deux mots séparés. Nous pouvons intuitivement en déduire que connaître l'orthographe aide à isoler les mots et établir les frontières entre deux formes lexicales.

Le non-alignement entre frontières syllabiques et lexicales à l'oral caractérise particulièrement la langue française. Comme l'a dit Delattre (1951 : 62), « le français enchaîne les mots, quitte à perdre de vue leurs limites. D'où l'impression unie, liée, que donne la chaîne parlée ». Cette impression d'unité est causée par les caractéristiques linguistiques du français. L'accent de mot qui en détermine la prosodie fait que les mots sont plus difficilement discernables à l'oral (Encrevé, 1988 ; Blanche-Benveniste & Jeanjan, 1987). De plus, l'organisation du français en structures syllabiques ouvertes (Delattre, 1965; Kaye & Lowenstamm, 1984) rend les frontières des mots perméables aux influences contextuelles. Ainsi nous pouvons enregistrer des phénomènes très communs à l'oral : la liaison, l'élision et l'enchaînement. À cause de ces phénomènes qui impliquent un processus de resyllabation, les frontières syllabiques masquent les frontières lexicales. Ce masquage et le non-alignement de ces deux éléments font émerger des problématiques de segmentation de l'oral chez les enfants qui commencent à parler français et chez les adultes qui apprennent le français comme langue étrangère.

Cependant, l'orthographe peut énormément aider à traiter les frontières des mots en français dans les cas de la liaison, de l'élision et de l'enchaînement. La connaissance de la représentation du mot dans sa forme visuelle et l'acquisition d'un lexique orthographique peuvent influencer sa reconnaissance à l'oral.

Nous avons décidé de conduire une étude longitudinale chez un groupe d'enfants francophones fréquentant l'école primaire. L'étude a eu une durée de 2 ans pendant lesquels nous avons testé les enfants à trois périodes : début du CP, fin du CP et CE1. Cette étude nous a permis de mieux comprendre leurs compétences de segmentation à l'écrit et à l'oral en présence de phénomènes entraînant une resyllabation. Les données ont également permis d'établir une relation entre ces deux types de compétences et d'analyser l'existence d'influences réciproques entre représentations phonologiques et représentations orthographiques.

Organisation de la thèse

Cette thèse est divisée en quatre parties. La première partie présente le cadre théorique et les questions de recherche. Elle inclut les chapitres 2, 3, 4, 5 et 6.

Dans le chapitre 2, nous présenterons les études mettant en rapport les compétences orales et écrites chez les enfants et les adultes. Nous établirons une relation entre phonologie et orthographe qui permettra de développer l'étude au centre de cette thèse.

Dans le chapitre 3, nous présenterons en détail les trois phénomènes de resyllabation qui caractérisent la langue française : la liaison, l'élision et l'enchaînement. Pour chaque phénomène, nous proposerons une description linguistique et un résumé des études existantes concernant la modélisation phonologique. Pour la liaison, qui est le plus étudié des trois, nous traiterons également des différentes classifications existantes et des hypothèses sur le statut et la position de la consonne de liaison. Nous consacrerons une section à part pour les questions de *h* muet et de *h* aspiré en rapport avec les phénomènes de resyllabation.

Dans le chapitre 4, nous décrirons les dynamiques de segmentation des séquences orales chez les adultes et chez les enfants francophones. Nous évoquerons d'abord le processus de segmentation chez les enfants, les erreurs en présence de phénomènes de resyllabation et les modèles qui cherchent d'expliquer l'acquisition de la liaison et des frontières des mots. Nous expliquerons également les difficultés

caractérisant la segmentation en mots chez l'adulte. Nous mettrons en évidence les stratégies permettant la segmentation correcte en mots en présence de la liaison et des autres phénomènes de resyllabation. La comparaison des différentes stratégies remettra en jeu l'acquisition de l'orthographe et de la lecture dans l'identification des frontières des mots.

Le chapitre 5 présentera un panorama sur la lecture et sur l'écriture. Nous traiterons d'abord des études d'*émergent literacy* et de *literacy* pour montrer comment l'enfant développe ses connaissances de la langue écrite. Nous présenterons les compétences principales qui interviennent dans l'acquisition de la lecture et de l'écriture en présentant aussi des modèles cognitifs. Ensuite, nous concentrerons notre attention sur les stratégies d'apprentissage de la langue française écrite et des problématiques liés à ces caractéristiques linguistiques, dont la présence de phénomènes de resyllabation. Enfin, nous présenterons les modèles d'apprentissage utilisés dans les écoles en France, pour comprendre les dynamiques d'apprentissage concernant la scolarisation.

Les descriptions de la segmentation à l'oral et à l'écrit en présence de phénomènes de resyllabation nous amèneront au chapitre 6, où nous développerons les questions de recherche et les hypothèses.

La deuxième partie de la thèse décrira la méthodologie utilisée et les modalités de traitement des données. Cette partie correspond au chapitre 7, où nous présenterons les participants à l'étude longitudinale, les bases de données utilisées et le répertoire choisi pour les différentes tâches. À la fin du chapitre, nous montrerons les analyses utilisées pour traiter la grosse quantité de données récoltées.

Les chapitres 8, 9, 10 et 11 sont inclus dans la troisième partie de la thèse. Nous exposerons les données récoltées dans l'étude longitudinale.

Dans le chapitre 8, nous présenterons les données de la première étape de notre étude longitudinale : le début du CP. Ce chapitre comprendra seulement les données sur la langue orale.

Dans le chapitre 9, nous décrirons les données récoltées dans la deuxième étape de notre étude, la fin du CP. Dans le chapitre 10, nous montrerons les résultats enregistrés pendant la troisième étape, le CE1. Dans ces deux chapitres, nous présenterons des données sur la langue orale et des données sur la langue écrite. Nous chercherons également à établir des relations entre la production orale, la lecture et l'écriture.

Dans le chapitre 11, nous résumerons les données récoltées aux trois moments pour comprendre l'évolution et le développement des compétences orales et écrites chez les enfants entre CP et CE1. À partir de cette comparaison, dans la quatrième partie de la thèse, composée par le chapitre 12, nous infirmerons ou confirmerons les hypothèses et nous discuterons les questions de recherche présentées au chapitre 6. À la fin du chapitre 12, nous proposerons des idées pouvant être le point de départ pour des recherches futures.

Partie 1.

Le cadre théorique et les questions de recherche

Chapitre 2

Phonologie et orthographe : les influences réciproques

2.1 Introduction

Dans cette thèse, nous cherchons à établir une relation entre la langue orale et la langue écrite. Nous présenterons donc dans ce premier chapitre de la partie théorique les études montrant l'existence de ce rapport qui constituent le point de départ de notre recherche.

2.2 L'influence de l'orthographe en diachronie : l'effet Buben

Dans son ouvrage de 1935, Vladimir Buben explique l'évolution phonétique du français par l'influence de l'orthographe. Cette influence de l'orthographe sur la prononciation des mots en français prend le nom de l'auteur et devient connue dans la littérature comme l'effet Buben (Blanche-Benveniste & Chervel, 1978 ; Chevrot & Malderez, 1999). Cet effet est visible au moyen de l'analyse de données diachroniques du français et il agit selon deux modalités :

- restitution de lettres muettes à la fin du mot, surtout dans les noms propres et dans les mots monosyllabiques (ex. dans *fil*s ou *but*) ou en position interne (ex. dans *obscure* ou *adversaire*). Mais également la réalisation d'un son plus long correspondant aux géminées graphiques (ex. le /m/ en *grammaire*) ;

- réinterprétation phonologique de certaines séquences graphiques : sur la base de la graphie <cisne> correspondant à l'évolution de la forme du bas latin *ciscismus* et dont l'orthographe restaurée <cygne> sur la base de la racine latine savante *cygnus* a entraîné ensuite qu'il soit prononcé aujourd'hui [sjɛ̃].

Pour expliquer le changement historique du français oral, Buben considère l'existence de deux variantes de la prononciation : une variante qui a des caractéristiques phonétiques connues pour favoriser son succès dans le changement (ex. la fréquence, Hay & Foulkes, 2016) et une deuxième variante qui montre les caractéristiques orthographiques, c'est-à-dire qui prend en compte à l'oral des correspondances graphème-phonème. Entre ces deux variantes, l'orthographe joue le rôle d'arbitre (Chevrot & Malderez, 1999) soit en permettant à la variante orthographique de s'imposer à l'oral, soit en maintenant la forme orale dans la langue d'usage.

L'effet Buben s'applique très facilement à la langue française car « [...]ce n'est pas la très grande régularité globale d'un système phonographique qui favorise la susceptibilité d'une langue à l'effet Buben, mais la présence simultanée de régularités phonographiques locales – par exemple, la correspondance régulière entre un graphème et un phonème particulier – et d'exceptions lexicales ponctuelles à ces régularités » (Chevrot & Malderez, 1999 : 121). La langue française est une langue inconsistante, où aux régularités s'ajoutent également de nombreuses exceptions.

De cette façon l'orthographe se présente comme une composante importante des évolutions historiques de la langue orale. Cependant, son impact est également visible en synchronie, comme nous allons le voir dans les prochains paragraphes.

2.3 Comment les compétences de la langue écrite influencent les compétences phonologiques des locuteurs

Dans ces paragraphes nous présentons les influences que la langue écrite a sur la représentation acoustique des mots. Comme Perfetti l'a écrit (1992 :160) : “A fully specific orthographic representation is bonded to the phonemic representation much as in chemical bonding”. Les résultats de différentes recherches qui utilisent différentes

méthodes d'analyse confirment cette idée. D'abord nous présentons les résultats chez les enfants pour ensuite prendre en considération les recherches conduites chez les adultes.

2.3.1 L'influence de l'orthographe sur la phonologie chez l'enfant

Ehri et Wilce (1982) décrivent très bien l'influence de la représentation graphique sur les compétences phonologiques. Les mots dans des langues opaques comme le français ou l'anglais n'ont pas de correspondance un à un entre graphème et phonème ; parfois ils présentent aussi des lettres n'étant pas prononcées à l'oral, dites lettres « muettes » comme par exemple le <t> dans *escargot* (/ɛs.kɑʁ.go/) ou le <k> dans *know* (/nəʊ/). Selon Ehri et Wilce (1985), quand les enfants trouvent des lettres muettes à l'écrit, ils focalisent leur attention sur elles car elles sont différentes des autres : elles ne correspondent pas à des sons : “the silent letters are salient in the memory representation of words spelling”. Quand on demande combien de sons a le mot anglais *witch* (/wɪtʃ/, fr. *sorcière*) à des enfants anglais de 12 ans qui maîtrisent lecture et écriture et qui connaissent bien le *spelling* (niveau normal dans le test standardisé pour le *spelling* anglais), ils vont dire qu'il a 5 phonèmes, parce qu'ils comptent comme phonème la lettre muette (Landerl *et al.*, 1996). Les enfants testés catégorisent un graphème sans réalisation phonétique, comme s'il s'agissait d'un phonème.

Chevrot (1998) montre que l'acquisition de l'orthographe peut avoir des effets durables sur la production orale. Par exemple, il montre que l'apprentissage de la graphie de *parce que* modifie sa prononciation. Les enfants pré-lecteurs au début de la première année de l'école primaire produisent d'abord la forme [pas.kə] sans /ʁ/ et ils jugent que cette forme est la seule à être correcte. Ce résultat est enregistré aussi chez les enfants entre 10 et 12 ans qui ont plus de problèmes en écriture et en lecture. Après 2 ans d'école primaire, les enfants qui jugent correcte la production de [paʁs.kə] avec un /ʁ/ sont aussi les enfants qui réussissent le mieux à orthographier ce mot. Nous pouvons dire que les résultats de cette recherche montrent que les enfants reconnaissent [paʁs.kə] comme une variante possible quand ils connaissent également la graphie de ce mot. Par conséquent, à nouveau, les compétences orales des enfants doivent être mises en rapport avec les compétences de l'écrit.

Ce rapport entre l'écriture et la production orale peut être directement vérifié au moyen d'une tâche de lecture. Morgan (1988-1999) propose une tâche de lecture à voix haute de mots et pseudo-mots à des enfants locuteurs du castillan. Il trouve que, en castillan, le <k> en fin de syllabe est plus souvent réalisé [θ] quand il correspond à la lettre <c> qui peut aussi transcrire /θ/ en attaque de syllabe, plutôt qu'à la lettre <x> ou <k>. Chevrot (1996) propose aussi une tâche de lecture chez des enfants français d'âge compris entre 10 et 11 ans. Il teste la lecture d'une série de nombres contenant le chiffre *quatre*, écrits soit en chiffres, soit en lettres (*4000* vs. *quatre mille*). Quand les enfants doivent lire les nombres en chiffres, dans 75 % des cas ils suppriment le /ʁ/ post-consonantique. En revanche, quand ils doivent lire les nombres en lettres, ce pourcentage de non-production de /ʁ/ chute à 42 %. Ces résultats montrent une possible influence de l'orthographe sur la production orale : les enfants qui doivent rendre à l'oral la représentation écrite sont influencés par la forme orthographique et le rapport existant entre les graphèmes et les phonèmes.

À partir de tous ces résultats, nous pouvons conclure que le niveau de connaissance des représentations orthographiques influence les performances dans les tâches méta-phonologiques, et de manière plus générale que la connaissance de la forme orthographique peut influencer l'idée que les enfants peuvent avoir de la structure phonologique d'un mot.

2.3.2 L'activation de l'orthographe en présence d'un input acoustique chez l'adulte

Quand un adulte alphabétisé écoute un mot dans sa propre langue, il active aussi la représentation orthographique du mot, qui va faciliter (ou compliquer), l'exécution d'une tâche orale, qui normalement va reposer exclusivement sur les compétences phonologiques.

Des résultats de tâches méta-phonologiques montrent qu'il est plus rapide de décider si un mot rime avec un autre quand les deux mots ont des représentations orthographiques plus proches. En anglais, la décision de la rime entre *toast* (fr. *toast*) et *roast* (fr. *rôti*) est plus rapide que pour *toast* et *ghost* (fr. *fantôme*), car *toast* et *roast* ont plus de graphèmes en commun que *toast* et *ghost* (Seidenberg & Tanenhaus, 1979).

De plus, l'orthographe influence les résultats dans les tâches de *phoneme monitoring* où les participants doivent trouver certains phonèmes-cibles dans des mots acoustiques (Dijkstra *et al.*, 1995 ; Hallé *et al.*, 2000). Les performances dans ce type de tâche sont influencées par le *spelling* du phonème à identifier dans le mot-cible (Cheung *et al.*, 2001 ; Morais *et al.*, 1979). Si le phonème à détecter peut être orthographié par plusieurs graphèmes, le temps pour le détecter s'allonge. En revanche, si le phonème n'a qu'une seule graphie, il a plus de chances d'être détecté plus rapidement.

Les études sur le français confirment cette tendance : il est plus facile de détecter dans un mot oral un phonème comme /p/ qu'un phonème comme /k/, parce que ce dernier segment a plusieurs réalisations orthographiques comme <c>, <cc>, <k>, <ck>, <que> qui interfèrent avec l'exécution de la tâche, la rendant moins rapide (Frauenfelder *et al.*, 1990).

L'influence de la représentation orthographique et de ses caractéristiques sur les tâches phonologiques a été bien étudiée par la psycholinguistique. Nous présentons les résultats ci-dessous.

2.3.2.1 Les résultats des expériences en psycholinguistique

2.3.2.1.1 Voisinage orthographique et voisinage phonologique

Pour présenter les résultats des expériences de psycholinguistique qui mettent en rapport phonologie et orthographe, il faut d'abord présenter deux des variables utilisées : le voisinage orthographique et le voisinage phonologique. Pour mieux comprendre ces formes de voisinage, nous prenons les exemples suivants. Si on considère le mot *aide*, en changeant un seul graphème nous pouvons créer ses voisins orthographiques : *vide* et *aidé*. Ces deux mots, existant dans le lexique du français, sont des voisins orthographiques du mot *aide*, créés au moyen d'un changement de graphème. En revanche, les voisins phonologiques d'un mot sont les mots qu'on peut créer en changeant un seul phonème, sans modifier les autres. Par exemple *contre* ([kɔ̃tʁ]) est un voisin phonologique de *montre* ([mɔ̃tʁ]), car entre l'un et l'autre il y a un seul phonème différent, le phonème initial (/k/ vs. /m/).

Ces deux types de voisinage présentent plusieurs tailles ou densités selon les indices statistiques caractéristiques de la langue à analyser. Si un mot a de nombreux voisins dans le lexique de la langue de référence, on définit son voisinage comme dense ; s'il a peu de voisins, on dit que son voisinage est faible. À l'aide de certaines bases de données, nous pouvons repérer les voisins et calculer la taille du voisinage. Pour le français, ces informations sont fournies par la base de données en ligne LEXIQUE.org (New *et al.*, 2001).

Au moyen d'une manipulation du voisinage orthographique et phonologique, Ziegler *et al.* (2003) et Ziegler et Muneaux (2007) révèlent une influence de l'orthographe dans le processus de reconnaissance des mots acoustiques chez les adultes lecteurs avancés (comparés à des lecteurs débutants ou à un groupe de dysphasiques). Les résultats montrent que les mots acoustiques avec un voisinage orthographique plus dense sont reconnus plus vite à l'oral. En revanche, les mots avec un voisinage phonologique plus dense sont reconnus plus lentement. Les résultats montrent un effet de facilitation du voisinage orthographique et un effet de compétition du voisinage phonologique qui rend la tâche plus difficile. Ziegler *et al.* (2004) analysent aussi les influences des correspondances graphèmes-phonèmes sur la reconnaissance des mots à l'oral. Ils enregistrent ce qu'ils ont appelé *consistency effect* : un effet de facilitation à l'oral pour les mots où la prononciation a peu de graphies possibles (ex. *espoir*, consistant) par rapport aux mots où la forme orale est caractérisée par plusieurs graphies possibles (ex.. *maintien*, inconsistent). L'effet de consistance orthographique a été enregistré également dans une tâche d'écoute passive (Perre *et al.*, 2011).

L'influence de l'orthographe sur la reconnaissance de mots acoustiques a été également montrée par des expériences de *priming paradigm* ou *amorçage*. Dans une tâche de décision lexicale acoustique, Perre *et al.* (2009) demandent à des adultes anglophones d'indiquer si le mot acoustique cible est un mot possible dans leur langue après la présentation d'une amorce acoustique. Le rapport entre amorce et cible est sujet à 2 conditions tests et à une condition de contrôle :

1. condition test 1 : rapport orthographique-phonologique (comme dans le couple *beef-leef*) ;
2. condition test 2 : rapport seulement phonologique (comme *leaf-leef*) ;

3. condition contrôle : aucun rapport entre amorce et cible (*sick-reef*).

Les chercheurs analysent le temps de réponse (TR) et le pourcentage de réponses erronées. Ils établissent une facilitation (soit des TR plus courts comparés à la condition contrôle) dans la condition 1 où le rapport amorce-cible est orthographique et phonologique (*beef-leef*) : une inhibition (soit des TR plus longs comparés à la condition contrôle) dans la condition 2 lorsque le rapport entre amorce et cible est seulement phonologique (*leaf-leef*).

Des résultats similaires ont été enregistrés dans une autre tâche de décision lexicale par Jakimik *et al.* (1985) qui utilisent le même paradigme d'amorçage. Ils observent un effet de facilitation si l'amorce a la même orthographe et la même phonologie que la cible (par exemple, *napkin-nap*). L'influence de l'orthographe sur les résultats dans des tâches orales a été vérifiée aussi par d'autres tâches de décision lexicale acoustique (Ziegler & Ferrand, 1998 ; Ventura *et al.*, 2004 ; Ziegler *et al.*, 2004) et pour des tâches qui prévoient la catégorisation sémantique des mots oraux ou la détermination du genre grammatical (Peereman *et al.*, 2009).

Les expériences présentées dans ces paragraphes montrent que, quand les adultes écoutent un input acoustique, non seulement la représentation phonologique est activée, mais aussi la représentation orthographique associée.

2.3.2.2 L'activation par l'orthographe de zones cérébrales dédiées à l'oral

Les rapports entre orthographe et phonologie ont été testés au moyen des potentiels évoqués et des électro-encéphalographies pour vérifier si les régions du cerveau activées par la langue orale et la langue écrite s'influencent réciproquement au niveau neurologique. Dans une tâche d'amorçage qui prévoit de mesurer l'effet de la consistance (*consistency effect*) de l'input orthographique, les chercheurs ont enregistré l'activation par l'orthographe de zones cérébrales destinées à la reconnaissance des mots acoustiques comme le gyrus fusiforme postérieur gauche (Perre & Ziegler, 2008 ; Pattamadilok *et al.*, 2009).

Les mesures neurologiques montrent une influence réciproque entre le traitement de mots à l'oral et des mots à l'écrit. La manière dont le cerveau procède dans l'analyse des mots écrits va activer des zones dédiées aux processus oraux. Ces résultats confirment le lien très étroit entre orthographe et phonologie.

Dans les paragraphes suivants nous présentons une comparaison des recherches qui tiennent compte des différences entre deux groupes de participants : illettrés et alphabétisés. Ces études permettent d'évaluer la contribution de la *literacy* aux performances dans certaines tâches et aussi au niveau cognitif et neurologique.

2.3.2.3 Différences de performances entre illettrés et alphabétisés

Connaître l'orthographe d'un mot aide les locuteurs à traiter le mot à l'oral. Nous pouvons penser qu'un locuteur qui connaît l'orthographe, capable de lire et d'écrire, aurait de meilleures performances sur des tâches orales qu'un locuteur qui ne connaît pas l'orthographe de sa langue. Pour vérifier cette idée, les chercheurs ont créé des expériences qui mettent en rapport deux groupes : un groupe de locuteurs illettrés et un groupe de locuteurs alphabétisés.

Ehri et Wilce (1980) montrent que connaître le *spelling* des mots influence la compétence phonologique des locuteurs américains. En anglais américain les phonèmes alvéolaires /d/ et /t/ dans des mots comme *metal* (fr. *métal*) et *ladder* (fr. *échelle*), sont représentés par deux graphèmes différents mais sont entendus et prononcés par la majorité des locuteurs comme la consonne [t]. Ehri et Wilce montrent qu'un américain adulte qui maîtrise mieux sa langue écrite (avec de bons résultats dans des épreuves de *spelling*) réussit à mieux distinguer et réaliser les deux prononciations [d] et [t] dans *metal* et *ladder* qu'un adulte illettré. Le locuteur alphabétisé a une prononciation des mots plus fine et arrive aussi à distinguer mieux les deux productions à l'oral grâce à l'étayage que l'orthographe va constituer pour sa compétence phonologique.

Morais *et al.* (1979) trouvent des résultats similaires pour le français. Les chercheurs proposent une tâche de suppression ou d'ajout d'un phonème dans des mots en français. Les performances sont meilleures chez le groupe d'alphabétisés. En revanche, un haut pourcentage d'erreurs a été enregistré dans le groupe d'illettrés.

Par ailleurs, l'existence de correspondances a été étudiée dans des travaux de neuroscience comparant alphabétisés et illettrés. Nous allons voir ces résultats dans les paragraphes ci-dessous.

2.3.2.3.1 Différences neurologiques entre illettrés et alphabétisés

Comme nous l'avons montré, plusieurs études ont soulevé une différence de performances pour les tâches phonologiques entre adultes alphabétisés et illettrés. Ces deux groupes ont été comparés pour établir des différences dans leurs configuration cérébrale. L'idée qui sous-tend cette considération est exprimée par Petersson *et al.* (2009 : 179) "literacy and formal education influence the development of the human brain and its capacity to interact with the environment".

En effet, des études de neuroscience montrent des différences dans les hémisphères cérébrales dues à l'apprentissage de l'écriture-lecture. Pendant les années dédiées à cet apprentissage, la partie centrale postérieure du corps calleux est soumise à une myélinisation majeure (Thomson *et al.*, 2000). Les données de Thomson *et al.* (2000) montrent que la partie cérébrale en question est plus épaisse chez les personnes alphabétisées et plus fine chez les non alphabétisées. Ainsi, selon eux, la densité de la substance blanche est modifiée par l'acquisition de la langue écrite qui cause une augmentation de cette substance et donc une augmentation de la connectivité neuronale. Ces différences neurologiques entre personnes illettrées et alphabétisées peuvent expliquer les moins bonnes performances des personnes illettrées sur des tâches linguistiques et non-linguistiques (Friederichi & Thierry, 2008 ; Petersson *et al.*, 2007).

Cependant la question n'est pas seulement anatomique mais aussi fonctionnelle. Castro-Caldas *et al.* (1998) comparent les résultats d'une tâche phonologique de répétition de pseudo-mots et mots chez un groupe d'illettrés et un groupe d'alphabétisés. Pour la lecture des mots, les chercheurs n'enregistrent pas une différence dans le taux d'erreur et dans les régions du cerveau activées. En revanche, pour les pseudo-mots, les illettrés et les alphabétisés activent des régions neuronales différentes et, chez les illettrés, ils remarquent plus d'erreurs en répétition. Ces résultats montrent que l'apprentissage de l'écriture et de la lecture influence l'organisation

fonctionnelle du cerveau. Cette idée est confirmée par l'étude de Dehaene *et al.* (2010) qui compare les résultats des différentes tâches chez un groupe d'illettrés et un groupe d'alphabétisés. Les chercheurs concluent "literacy refines spoken language processing by enhancing a phonological region, the planum temporale" (p.6).

En conclusion, selon ces résultats, les illettrés et les alphabétisés développeraient différemment les structures cérébrales et l'organisation des régions cérébrales. Chez les alphabétisés, le traitement de la langue orale est plus fine et la région qui s'occupe du traitement phonologique de l'oral subit une amélioration grâce à l'apprentissage de la lecture et de l'écriture.

2.3.3 L'influence de l'orthographe sur la représentation phonologique

Tous les résultats que nous avons montrés dans ce chapitre nous conduisent à établir un rapport très étroit - et parfois réciproque - entre la phonologie et l'orthographe. Nous pouvons penser que l'orthographe influence la phonologie et qu'elle peut modifier la représentation phonologique des mots chez les locuteurs. À la lumière des résultats de leur étude psycholinguistique, Peerman *et al.* (2009 : 367) confirment cette idée : "the orthographic properties of words may partially shape how phonological representation are organized". Chevrot *et al.* (2000) vérifient l'hypothèse selon laquelle apprendre la forme écrite de certains mots modifie les représentations phonologiques correspondantes. Les chercheurs ont testé des enfants de 3^{ème} année de l'école primaire au moyen d'une tâche d'apprentissage de pseudo-mots qui font référence à des images. Chacun de ces pseudo-mots se termine par un /ʁ/ final post-consonantique qui pouvait être supprimé dans la production orale (ex. *mallopre* ou *quaçontre*). L'apprentissage de ces pseudo-mots a eu lieu en quatre séances selon trois modalités :

1. Groupe « oral ». Les enfants de ce groupe écoutent les pseudo-mots à l'oral et le /ʁ/ était supprimé dans 75 % des cas.
2. Groupe « oral-écrit ». Dans ce groupe, les enfants écoutent les mêmes enregistrements que les enfants du groupe "oral" mais dans les dernières séances, les expérimentateurs introduisent aussi les formes orthographiques en support.

3. Groupe « écrit ». Les enfants apprennent les pseudo-mots seulement via la forme écrite.

Pour vérifier le pré-apprentissage des séries de pseudo-mots, les chercheurs proposent deux tâches de production orale – dénomination d'images et syllabation. Les résultats montrent que les productions orales sont différentes selon le groupe d'enfants. Dans le groupe « oral », les enfants produisent le /ʁ/ seulement dans 50 % des cas ; dans le groupe « écrit » ce pourcentage augmente jusqu'à être proche de 100 %. Enfin, dans le groupe « oral-écrit », le pourcentage de production de /ʁ/ augmente quand l'orthographe est introduite dans le processus d'apprentissage. Ces résultats montrent que l'apprentissage de la forme graphique modifie la production orale. Une expérience similaire faite par Bürki et Frauenfelder (2012) sur 23 adultes francophones et la production de schwa dans des pseudo-mots confirme que la variante avec schwa est plus facilement reconnue si elle est présentée aussi au niveau graphique par le <e>. Ces résultats à partir de ces tâches d'apprentissage de pseudo-mots sont ultérieurement vérifiés par des recherches concernant les sujets sourds de naissance qui doivent construire des représentations phonologiques des mots sur la base de la seule orthographe. Alegria et Leybaert (1991) montrent que même si la forme écrite n'est pas présente, les locuteurs sourds ont tendance à avoir une prononciation plutôt orthographique. Comme le résumant bien Metsala et Walley (1998 : 118) : « l'apprentissage des représentations graphiques des mots conforte un processus continu de réorganisation des représentations phonologiques, qui deviennent de plus en plus spécifiées, détaillées, stables, organisées en segments ».

À partir de l'idée que l'orthographe peut changer la représentation phonologique, nous avons développé nos études dans notre mémoire de Master sur les géminées italiennes.

À la différence du français, en italien, les géminées sont phonologiques (ex. *cane* [kane], fr. *chien* vs. *canne* [ken:e], fr. *canes*). Les enfants italiens qui ne sont pas encore capables d'écrire ne réussissent pas à distinguer à l'oral un mot avec géminées (*canne* [ken:e], fr. *canes*) et un mot sans géminées (ex. *cane* [kane], fr. *chien*). C'est seulement à partir de l'école primaire et au moyen de l'écriture que les enfants se rendent compte qu'il y a une différence entre les mots avec et sans géminées. La

création d'une correspondance entre les caractéristiques visuelles (une suite de deux lettres identiques) et les caractéristiques phonologiques et phonétiques permet d'établir la différence entre la présence et l'absence de gémées (Ruvoletto, 2011).

Pour expliquer ce phénomène, nous avons proposé l'hypothèse suivante : l'évolution de la représentation phonologique est influencée par l'acquisition de la forme orthographique matérialisant visuellement les deux consonnes. Si nous prenons en considération la représentation autosegmentale, les gémées en italien sont représentées par deux positions squelettales (les deux X associés au phonème /t/ en figure 2). Les compétences de tous les paliers de cette représentation sont acquis par tous les enfants avant l'apprentissage de la *literacy* alors que l'intégration de l'information représentée sur le palier squelettal semble se faire uniquement lors de l'enseignement de l'écriture. Avant de commencer l'apprentissage de l'écriture, les enfants traitent les gémées comme une seule unité et il n'arrivent pas à différencier les mots avec doubles consonnes des mots sans doubles consonnes (Figure 3). Grâce à l'apprentissage de l'écriture qui matérialise visuellement l'existence de deux consonnes, ils modifient la représentation phonologique de gémées. Les gémées s'encodent phonologiquement : les enfants intègrent les connaissances sur le palier squelettal en réussissant à associer à un son deux positions ce qui permet de distinguer un mot comme *cane* d'un mot comme *canne*.

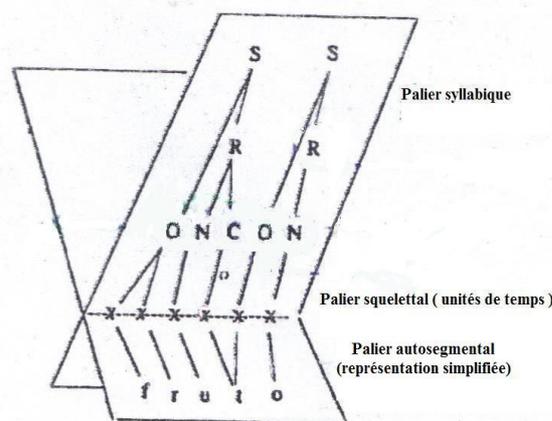


Figure 2. Modèle autosegmental pour le mot italien *frutto* (Ruvoletto, 2011)

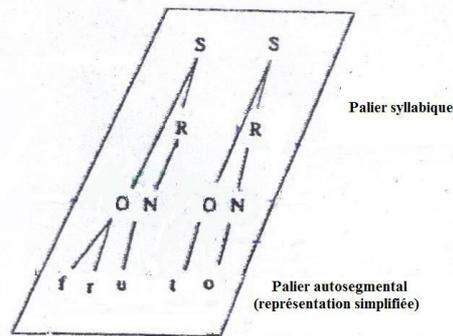


Figure 3. L'étape précédente du modèle de représentation autosegmental pour le mot *frutto* (Ruvoletto, 2011)

2.4 L'oral et l'écrit s'influencent réciproquement

Dans ce chapitre, nous avons d'abord montré comment la connaissance de l'orthographe influence les résultats des tâches phonologiques chez les enfants et chez les adultes. Ensuite, nous avons établi une interaction entre orthographe et phonologie au moyen des résultats des tâches psycholinguistiques. Nous avons également présenté les différences dans les performances et dans l'activation cérébrale entre illettrés et alphabétisés. Enfin nous avons montré comment la connaissance de la forme orthographique peut impliquer des changements dans la représentation phonologique.

À partir de ce lien très étroit entre écrit et oral, nous pouvons chercher à mettre en relation la segmentation à l'oral et à l'écrit en présence des phénomènes de resyllabation. Nous traiterons de ces phénomènes dans le prochain chapitre.

Chapitre 3

Les mots enchaînés : les phénomènes de resyllabation en français

3.1 Les phénomènes de resyllabation en français

La liaison, l'élision et l'enchaînement sont des phénomènes linguistiques qui agissent sur deux mots : Mot1, le mot de gauche et Mot2, le mot de droite. Dans les trois cas, le Mot2 est un mot à voyelle initiale, dont l'attaque syllabique est vide. Une consonne du Mot1 va remplir l'attaque vide du Mot2. La manière dont a lieu ce remplissage et la nature des segments impliqués permettent de distinguer ces trois phénomènes que nous allons développer dans les prochaines sections.

3.1.1 L'élision

En considérant la définition de Schane (1967 : 37), « l'élision est définie comme la suppression ou effacement de la voyelle finale d'un mot devant un autre mot commençant aussi par un son vocalique ». Pour montrer ce phénomène, Schane donne deux exemples : *le + ami* → [lami] où l'élision a lieu et *le + camarade* → [le.ka.ma.ʁad] où le processus d'élision n'est pas enregistré. Comme exemplifié en (1), la voyelle finale du mot de gauche (Mot1) s'amuit devant un deuxième mot (Mot2) débutant par une voyelle. Ensuite la consonne qui était l'attaque du même mot devient l'attaque du Mot2, en conséquence de la chute d'un noyau vocalique.

- (1) Mot 1 : *le* [lə] + Mot 2 : *ours* [uʁs] → *l'ours* [luʁs] où [ə] → ∅ / _V

La voyelle finale du Mot1 impliquée dans l'élision est le schwa [ə] : l'élision est possible quand cette voyelle occupe le noyau vocalique du Mot1. L'élision du schwa [ə] permet à la consonne [l] de remplir l'attaque vide au début du mot à droite. Le Mot1 *le* [lə] est donc resyllabifié dans le Mot2 *ours* [uʁs]. La conséquence est le non-alignement entre frontières syllabiques et lexicales (*l'ours* vs. [luʁs]).

L'élision est un phénomène très fréquent en français qui permet d'éviter le hiatus quand il n'y a pas l'insertion d'une consonne euphonique (ex. dans les formes interrogatives avec l'inversion sujet-verbe *il y a* → *y a-[t]-il*).

L'élision est obligatoire dans le cas d'un Mot1 qui termine par [ə] (ex. *le* [lə] + *ours* [uʁs] → *l'ours* [luʁs]), [a] ouvert (ex. *la* [la] + *abeille* [a.bɛj] → *l'abeille* [la.bɛj]) et [i] (ex. *si* [si]+*il* [il] → *s'il* [sil]) suivi par un Mot2 commençant par voyelle (*abeille*, *ours*, *hippopotame*¹). Si le Mot2 commence par *h* aspiré, l'élision n'a pas lieu (*le* [lə]+*hibou* [i.bu] → *le hibou* [lə.i.bu]).

3.1.2 L'enchaînement

L'enchaînement est un phénomène de resyllabation qui agit entre les frontières des mots qui constituent une seule unité prosodique. Dans le processus d'enchaînement, la consonne du Mot1 que nous pouvons appeler consonne fixe (CF), en utilisant la mention d'Encrevé (1983), va se lier par enchaînement à la voyelle initiale du Mot2 au moyen d'une resyllabation. Ainsi, la consonne qui est en position de coda dans la dernière syllabe du Mot1, se retrouve dans la position d'attaque de la syllabe suivante comme il est possible de voir en (2), où le deuxième [t] de *petite* [pə.tit] est la CF.

$$(2) \quad \text{Mot 1 : } \textit{petite} \text{ [pə.tit]} + \text{Mot 2 : } \textit{orange} \text{ [ɔ.ʁɑ̃ʒ]} \rightarrow \textit{petite[t]orange} \\ \text{[pə.ti.tɔ.ʁɑ̃ʒ]}$$

Dans cette opération, la CF ne change pas sa nature. Dans la quasi totalité des cas d'enchaînement, il n'y a pas de modification de la mélodie du segment à quelques exceptions près (le [f] dans le couple *neuf[v]ans* qui est réalisé [v] ; le [s] dans *six[z]enfants* qui se réalise en [z]). De plus dans l'enchaînement, aucun segment

1 Le mot commence orthographiquement par un *h* muet mais phonologiquement par une voyelle.

n'apparaît en surface : la consonne du Mot1 qui s'enchaîne est toujours prononcée, avec ou sans enchaînement (*la petite fille* [la.pə.tit.fij]). Cette CF appartient aussi à la forme orthographique du Mot1 (*petite*).

L'enchaînement peut impliquer soit une seule consonne (C), soit deux consonnes (CC). Si le Mot1 finit par une double consonne (CC), une des deux consonnes reste dans la position de coda de la première syllabe et la deuxième consonne devient l'attaque de la syllabe suivante comme en (3).

- (3) Parc à huîtres [par.ka.ɥitr]
 Concept archaïque [kõ.sɛp.tar.ka.ik]

3.1.2.1 L'enchaînement : un phénomène prosodique peu étudié

Comme le souligne Encrevé (1983 : 42), l'enchaînement est un phénomène phonétique qui n'a pas d'enjeu social et qui ne constitue pas une variable stylistique ailleurs que sur la liaison. C'est pour cette raison que, selon Encrevé, les générativistes se sont peu intéressés à ce phénomène.

« Considéré comme automatique de la part de tous les francophones (sans doute parce que la consonne y est, par définition, toujours entendue et que l'orthographe, de ce fait, ne pose pas ici de problème puisque, de toute façon, la resyllabation n'y est jamais notée), l'enchaînement des consonnes fixes n'a jamais fait l'objet d'un enseignement explicite dans l'école française, ni ne fut jamais un des domaines où se jouaient la légitimité et la distinction en matière de langage, raison, sans doute, pour laquelle les générativistes ne l'ont pas remarqué ».

En tout cas, nous avons des études d'Encrevé (1983) sur l'enchaînement d'une CF en français qui montrent que ce phénomène prosodique doit être traité comme un phénomène variable du français. Dans le parlé spontané les locuteurs peuvent en effet réaliser une CF non-enchaînée au Mot2. Cette CF peut être séparée de la voyelle du Mot2 par l'insertion d'une occlusive glottale. Son analyse du discours politique montre très bien la variabilité de ce phénomène en ajoutant aussi l'idée que le caractère occlusif de la CF favorise le non-enchaînement.

3.1.2.2 La modélisation de l'enchaînement

Si peu d'études ont été entreprises sur l'enchaînement, plusieurs phonologues ont essayé de modéliser ce phénomène et de traiter la CF.

Dans le cadre théorique de la phonologie autosegmentale, Clements et Keyser (1981) essaient de représenter la CF au moyen d'un modèle tridimensionnel qui comprend 3 couches (tiers) : une couche syllabique ; une couche centrale occupée par des positions ; le squelette (*skeleton*) et une couche segmentale où il est possible de représenter les segments phonologiques. Pour Clements et Keyser, la CF est liée à une position du squelette qui est aussi associée à un nœud syllabique, comme en (Figure 4).

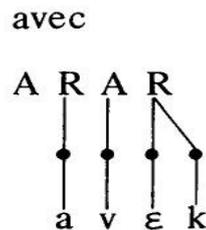


Figure 4. La représentation d'après Clements et Keyser (1981) dans Encrevé (1983 : 56)

Cette représentation ne tient pas compte de la nature de la CF et de sa potentialité de devenir attaque du Mot2 en cas d'enchaînement.

Puisque l'enchaînement peut affecter la CF, Encrevé propose une représentation où la CF est liée à la position squelettale sans être ancrée à la position syllabique (Figure 5).

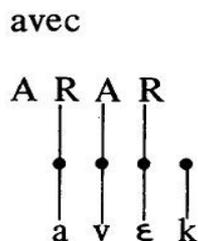


Figure 5. Représentation par Encrevé (1983 : 57)

Dans cette représentation, la CF n'est pas directement liée à la position syllabique, et elle peut donc se lier à l'attaque vide au début du Mot2 en cas d'enchaînement.

3.1.3 La liaison

Comme l'enchaînement et l'élision, la liaison implique un décalage entre frontières lexicales et frontières syllabiques de Mot1 et Mot2.

(5) Mot1 : *les* [le] + Mot2 : *ami* [ami] → *les***[z]***ami* [le.za.mi]

Dans ce processus exemplifié en (5), une consonne latente, dite consonne de liaison (CL) apparaît en surface (le [z] dans [le.za.mi] vs. *les* [le] en isolation où le [z] n'apparaît pas en surface). La CL est resyllabifiée dans le Mot2 si le mot commence par une voyelle (ex. *les***[z]***amis* [le.za.mi] ; *les garçons* [le.ɡaʁ.sɔ̃]). Le résultat est une séquence en liaison où les frontières de mots ne correspondent pas aux frontières de syllabes (*les amis* [le.za.mi]).

La liaison donne lieu à des changements au niveau segmental. Le premier changement phonétique est le dévoisement ou le voisement de la consonne en liaison. Par exemple, la CL fricative /s/ se voise et se réalise en [z] (ex. *gros homme* [ɡʁo.zɔ̃m]) ainsi que l'occlusive /d/ se dévoise et devient [t] (ex. *grand homme* [ɡʁɑ̃.tɔ̃m]). Le deuxième changement phonétique est lié à la réalisation des voyelles nasales. Nous avons deux réalisations possibles : une réalisation qui implique la dénasalisation de la voyelle nasale (ex. *bon* [bɔ̃] réalisé [bɔ̃n] dans *bon ami* [bɔ̃.na.mi]) ; une réalisation qui est le résultat d'une conservation de la nasalité sur la voyelle (ex. *mon enfant* [mɔ̃.nɑ̃.fɑ̃]).

Dans l'exemple en (5), la CL est resyllabifiée dans Mot2. C'est le cas de la liaison dite « enchaînée » car une resyllabation est prévue. Mais, en français, nous pouvons aussi trouver des exemples de liaison « non-enchaînée ». Dans ce type de liaison, la resyllabation n'apparaît pas et la CL reste en position de coda du Mot1 (ex. *vos amis* [voz.a.mi]). Dans ce dernier cas, il y a une correspondance entre les frontières syllabiques et les frontières lexicales des mots car la CL garde la place qu'elle occupe

dans la forme orthographique, soit consonne finale du Mot1. Pierre Encrevé (1983) est le premier linguiste à proposer une analyse de la liaison non enchaînée à partir d'un corpus de discours politiques. Il établit une connexion entre la variabilité de la liaison et ses caractéristiques prosodiques. Il montre que, dans le contexte où la liaison est obligatoire (toujours produite), la CL est invariablement enchaînée au Mot2. En revanche, pour la liaison facultative, l'enchaînement est variable et il est défavorisé par le caractère occlusif de la CL. Une étude phonétique de Malécot (1975) montre que la séquence *cet émoussement* peut être prononcée avec enchaînement ([se.te.mu.smã]) ou sans enchaînement ([set.ʔe.mu.smã]) avec l'insertion d'un coup de glotte entre la CL non enchaînée et la voyelle au début du Mot2.

La liaison fait aussi appel à la morphologie. Le phénomène de liaison est plutôt associé au marquage du pluriel et, en effet, d'un point de vue morphologique, il y a la tendance chez les locuteurs à marquer le pluriel par une liaison en [z] et le singulier par l'absence de liaison (Delattre, 1966) comme on peut le voir dans l'exemple en (6) :

(6) *des soldats [z] espagnols vs. un soldat [Ø]² espagnol*

Pour cette raison, parfois les locuteurs ajoutent entre Mot1 et Mot2 une consonne qui n'appartient ni à Mot1 ni à Mot2, correspondant à la marque du pluriel à l'oral, [z]. Ce type de phénomène est appelé « fausse liaison », et est exemplifié en (7).

(7) **vingt[z]oiseaux [vẽ.zwa.zo] au lieu de vingt[t]oiseaux [vẽ.twa.zo]*

3.1.3.1 Les contextes de la liaison et leur classification

Les éléments impliqués dans le processus de liaison sont deux mots (Mot1 et Mot2) et une consonne de liaison (CL). La cohésion syntaxique entre Mot1 et Mot2 est posée comme déterminante pour la réalisation de la liaison : sans cohésion syntaxique, la liaison ne se fera pas. Les études des catégories syntaxiques possibles des Mot1 et Mot2 ont permis une première classification de la liaison. La première a été faite par

2 Le symbole indique la non-réalisation de la liaison.

Delattre (1947). Il distingue trois catégories de liaison : obligatoire, facultative et interdite. Les différences entre ces catégories sont explicitées selon le rapport syntaxique entre Mot1 et Mot2 en liaison (résumé dans le tableau 1).

La liaison est obligatoire dans le contexte nominal entre déterminant + nom (*un[n]ami*), adjectif + nom (*grand[t]ami*), dans les contextes verbaux entre pronom personnel + verbe (*ils[z]ont*) et vice versa (*ont-[t]ils*). Les cas de liaison obligatoire sont aussi notables dans des couples de mots invariables comme un mot monosyllabique suivi par un mot commençant par voyelle (*très[z]en retard* ou *trop[p]important*) ou dans des constructions figées (*les États-[z]Unis*).

	Liaisons obligatoires	Liaisons facultatives	Liaisons interdites
Noms	<ul style="list-style-type: none"> • Det + N /Adj Un [n] ami Deux [z] amis • Adj + N Un grand [t] ami 	<ul style="list-style-type: none"> • N_{pl} + X Les jeunes filles [z]? en fleur Les filles [z]? émues Les jardins [z]? ont fleuri 	<ul style="list-style-type: none"> • N_{sg} + X Le soldat // en fête Le soldat // ému Le soldat // est revenu
Verbes	<ul style="list-style-type: none"> • Pro_{pers} + Vbe Ils [z] ont • Vbe + Pro_{pers} Ont-[t] ils 	<ul style="list-style-type: none"> • Aux + Vbe Ils ont [t]? été Je vais [z]? essayer • Vbe + X Il arrivait [t]? à l'heure 	
Mots invariables	<ul style="list-style-type: none"> • Monosyll + X En [n] avion Très [z] en retard Trop [p] important 	<ul style="list-style-type: none"> • Polysyll + X Pendant [t] ? une heure Toujours [z]? en retard Extrêmement [t]? important 	<ul style="list-style-type: none"> • Et + X Et // alors Et // avec lui
Constructions figées et cas particuliers	<ul style="list-style-type: none"> Tout-[t]-à-coup Les États-[z] Unis 		<ul style="list-style-type: none"> • H « aspiré » Des // héros En // haut Cent // huitième fois

Tableau 1. Classification de la liaison et des exemplifications de Wauquier (2010 : 5) après Delattre (1947)

Dans le cas de la liaison facultative, le locuteur peut prononcer la séquence de deux mots soit avec liaison, soit sans liaison. La production de la liaison n'est pas fautive, et en même temps la non-production n'est pas une erreur. Dans le contexte nominal, nous trouvons des liaisons facultatives entre nom pluriel et Mot2 à voyelle initiale (ex. *les filles[z]émues*). Pour les contextes verbaux, entre auxiliaire et verbe (*Ils ont[t]été*) ou entre verbe et Mot2 commençant par une voyelle (*Il arrivait[t]à l'heure*). Pour les mots invariables, il y a la possibilité de liaison entre un mot polysyllabique et un nom à voyelle initiale (*Pendant[t]une heure*).

Le dernier type de liaison est la liaison interdite. Produire ce type de liaison est considéré comme une erreur. Les contextes interdits sont : nom singulier suivi par Mot2 à voyelle initiale (*Le soldat//en fête*), après la conjonction *et* (*Et//alors*) ou quand Mot2 commence par *h* aspiré (*Les//Halles*).

Comme nous avons parlé de Mot1 et Mot2, il faudrait aussi parler de CL. Notamment, dans toutes les catégories, nous avons différentes CL possibles mais les plus fréquentes sont [t], [n], [z]. Il y a aussi la possibilité de repérer d'autres CL comme [p], [k] ou [ʁ] mais elles sont moins fréquentes, si non rares (comme [g]).

3.1.3.1.1 Obligatoire ou catégorique ? Facultative ou variable ? Interdite par qui ?

La classification de la liaison en trois catégories a été reconsidérée à plusieurs reprises depuis Delattre.

Dans son livre de 1988, Encrevé propose de traiter des contextes où Mot2 est un mot monosyllabique comme des contextes sujets à une liaison variable. Une révision de la classification classique a aussi été proposée par Boula de Mareüil *et al.* (2003). Depuis leur analyse du corpus du français BREF³, les auteurs montrent que le contexte où Mot1 est un adverbe monosyllabique est plutôt un contexte de liaison facultative. Ils concluent que “liaison rates very significantly depending on word and word bigram identities. This confirms that liaison is strongly linked to both syntactic and lexical levels” (p.1332).

Grâce à l'analyse du corpus du projet *Phonologie du Français Contemporain* (PFC) et à son protocole d'enquête⁴, les appellations de liaison comme « obligatoire » ou « facultative » ont été remises en question. À la lumière des données orales du corpus PFC, Durand *et al.* (2011) remarquent que dans la catégorie définie par Delattre comme « obligatoire », il y a une variation de réalisation dépendant des phénomènes morphologiques et catégoriels. Il propose donc une description de la liaison en ‘mot à

3 Le corpus BREF est un corpus basé sur 100 heures de lecture des magazines français.

4 Le protocole PFC règle l'encodage des données sur la liaison pour la lecture à voix haute et la conversation guidée et spontanée chez des locuteurs français en différentes zones géographiques. Les fichiers audio sont sujets à une transcription qui s'intègre à la base de données.

mot' qui va réviser la tripartition de la liaison en obligatoire, facultative et interdite et qui donne plus d'attention à la production des locuteurs et aux composantes sociolinguistiques.

Les analyses du corpus *PFC* (Mallet, 2008) montrent que la liaison est presque toujours réalisée dans les contextes suivants : déterminant + nom (*les[z]amis*) ; proclitique + verbe (*vous[z]allez*), verbe + enclitique (*allez-[z]y*) ; expressions figées (*de temps[z]en temps*). En revanche, la liaison est très variable dans le contexte monosyllabique + X. Dans ce contexte, les pourcentages de réalisation changent selon le type de mots monosyllabiques : la liaison est presque toujours réalisée avec *en* (13 non réalisations pour 2006 réalisations), beaucoup moins réalisée après *dans* (59 non réalisations pour 859 réalisations) et encore plus variable pour *chez* (18 non réalisations pour 74 réalisations). Les résultats sur le contexte monosyllabique + X mettent encore en danger l'analyse de Delattre et sa classification, car, pour lui, ce dernier contexte devait être traité comme obligatoire.

À la lumière des résultats du *PFC*, une nouvelle classification de la liaison basée sur l'usage oral a été proposée. Cette classification prévoit deux types de liaison : une liaison catégorique, toujours réalisée par les locuteurs ; une liaison variable qui est sujet à une variation de réalisations selon différentes composantes.

La liaison catégorique trouve sa réalisation dans des contextes catégoriques typiques :

- déterminant + X à voyelle initiale en syntagme nominale (avec Mot1 *un, les, des, ton, mes, tes, ses*) du type *les enfants, un enfant* ;
- proclitiques : *ils, elles, on, nous, vous, en*, du type *il en[n] a* ;
- enclitiques : *mange-t[t]-il ? , comment dit[t]on ?* ;
- composés et syntagmes figés : *du type tout-[t]à fait, pot[t]au feu*.

Nous pouvons remarquer qu'il y a un lien entre les CL, leurs fréquences et les possibilités de réalisation à l'oral. Dans le corpus *PFC*, la liaison en [n] est attestée plutôt dans des contextes catégoriques que variables, avec des Mots1 comme *on, mon* et *son*. Elle est très souvent réalisée pour *un* et *en* (98 %) et peu réalisée pour *bien* (47 %). En revanche, les liaisons en [z] et [t] sont attestées dans des contextes catégoriques ou variables, et [ʁ] seulement dans les contextes variables. Ce type de données permet de mieux comprendre la fréquence de ces liaisons à l'oral : une liaison

en [n] serait plus entendue qu'une liaison en [ʁ], car la première est plus souvent réalisée dans les contextes catégoriques. La liaison en [ʁ], liée à un contexte variable, serait donc moins fréquente.

3.1.3.2 L'analyse linguistique du corpus *PFC*

La plus récente analyse du corpus de langue orale *PFC* fait émerger des données très importantes sur la liaison qui révèlent différentes grandes tendances contemporaines (Eychenne *et al.*, 2014). Les données montrent que la liaison est un phénomène variable sujet à des facteurs linguistiques. Mais ceux-ci ne sont pas les seuls à régler la réalisation de la liaison. Elle souffre aussi de plusieurs contraintes para-linguistiques ou extra-linguistiques comme, par exemple, des facteurs sociolinguistiques et stylistiques. Nous allons présenter ces facteurs dans les prochains paragraphes.

Les facteurs linguistiques

Comme déjà vu par Durand et Lyche (2008), la plupart des liaisons attestées sont des liaisons enchaînées. Les liaisons non-enchaînées sont peu nombreuses et le pourcentage de réalisation de liaison est plus élevé en lecture où les formes orthographiques de Mot1 sont disponibles. Ces résultats sont confirmés par l'analyse de Mallet (2008). D'autres résultats, en accord avec sa thèse, concernent la CL : les seules consonnes observées en liaison sont [z, n, t, p, ʁ] mais les fréquences de réalisation sont très différentes : la liaison en [z] est la plus fréquente (46,26 %), suivie de la liaison en [n] (36,06%), [t] (17,25 %), et encore plus loin [ʁ] (0,18%) et [p] (0,06%).

Les facteurs stylistiques

Chaque locuteur du français se trouve confronté à différentes situations de communication qui l'obligent à utiliser différents registres ou styles. Par exemple, en situation de parole spontanée ou relâchée, il utilisera un registre familier qui pourra aller jusqu'à l'argotique. Dans d'autres situations, comme par exemple la récitation de vers, il se servira d'un registre soutenu ; lors d'une conférence, il utilisera un registre standard.

Delattre (1966) note une forte différence dans la fréquence des liaisons selon les registres utilisés par les locuteurs. Entre le registre familial et le registre plus soutenu, la fréquence de production de liaisons facultatives augmente. Ainsi, dans une conversation familière, il est plus fréquent de produire la liaison seulement dans les contextes obligatoires (*des[z]hommes/illustres ont/attendu*). Dans une conversation soignée, le locuteur ajoute aux liaisons obligatoires les liaisons facultatives verbe-participe (ex. *des[z]hommes/illustres/ont[t]attendu*). Pendant une conférence, plus de liaisons facultatives seront produites (*Des[z]hommes[z]illustres ont[t]attendu*) et enfin pour la récitation de vers, toutes les liaisons possibles seront réalisées (*Des[z]hommes[z]illustres ont[t]attendu*).

Les trois modalités prévues dans le protocole *PFC* permettent d'avoir des données sur des « styles » différents, plus ou moins surveillés. Les résultats de l'analyse du corpus (Eychenne *et al.*, 2014) montrent qu'un pourcentage plus élevé de liaison est produit en lecture (cela peut être influencé par la disponibilité de la forme orthographique, comme dans Laks, 2005). Le pourcentage de liaison en conversation guidée est moins élevé et le pourcentage de réalisation encore inférieur dans la conversation spontanée. Ces résultats confirment l'analyse de Delattre, soulignent que dans le parlé spontané, comme dans le registre familial, les liaisons sont moins réalisées.

Les facteurs individuels : l'âge, le sexe, la qualification et le niveau d'études

La production de liaisons peut être également influencée par l'âge, le sexe ou le niveau scolaire des locuteurs.

Des études sur le corpus *PFC* montrent que les locuteurs plus âgés réalisent plus de liaisons facultatives (Malécot, 1975; Ashby, 1981; Ranson, 2008), surtout en conversation spontanée.

En ce qui concerne le sexe des locuteurs, les résultats sont controversés : les analyses de Malécot (1975) ou Booij et De Jong (1987) montrent un usage plus fréquent de la liaison chez les femmes ; en revanche Ashby (1981) trouve une réalisation plus fréquente chez les hommes. Les résultats de recherches plus récentes montrent cependant qu'entre les productions des deux sexes, il n'y a pas une différence statistiquement significative (Durand *et al.*, 2011).

Pour la variation de la liaison selon le niveau d'études, l'analyse d'échantillonnage *PFC* par Mallet (2008) ne montre aucune différence d'usage de la liaison.

Les facteurs géographiques

Le programme *PFC* a conduit une recherche pour comprendre la variation de l'usage de la liaison selon les différentes zones géographiques francophones. Les sujets ont été regroupés en trois groupes : l'espace « France hexagonale », l'espace francophone « français langue première » (Belgique, Suisse, Canada) et l'espace francophone « français langue seconde » (Afrique).

D'abord, nous concentrons notre attention sur la « France hexagonale » pour voir si la réalisation de la liaison change selon les zones géographiques de France. Coquillon *et al.* (2010) ont considéré 67 enquêtes de la zone nord et 67 enquêtes de la zone sud de la France. Ils ont d'abord examiné le contexte Nom (pl.) + X dans la lecture à voix haute. Les résultats montrent qu'il n'y a pas une différence importante entre la réalisation de la liaison au nord et au sud de la France dans ce contexte.

D'autres données ont été récoltées par Bordal et Lyche (2008) sur le français parlé en Afrique sur 4 points d'enquêtes : Abidjan, Bamako, Ouagadougou et Bangui. Les données africaines ont été comparées aux données de Paris (en considérant la variété parisienne comme variété standard). Les résultats montrent que si, à Paris, 35 % des liaisons sont réalisées, nous avons seulement 2 % de réalisation à Abidjan, 7 % à Ouagadougou, 10 % à Bamako et 11 % à Bangui. Le pourcentage le plus bas est enregistré à Abidjan qui est, parmi ces quatre villes africaines, considérée comme la plus francophone. De plus, dans les quatre points d'enquête, la liaison en contexte adjectif + nom est traitée comme variable. Selon les auteurs, ces comportements constatés dans les territoires africains doivent être associés aux caractéristiques prosodiques des langues maternelles des locuteurs.

Pour l'espace francophone « français langue première », les données *PFC* ne montrent pas de véritables différences dans l'emploi de la liaison. Cependant, il faut souligner l'exception du Québec, où le pronom *ils* suivi d'un verbe n'est jamais sujet d'une liaison, alors que la liaison dans ce contexte est considérée comme catégorique en français standard. De plus, *ils* est prononcé [i] (comme c'est souvent le cas dans la

plupart des variétés de français), et le *l* n'enchaîne pas avec le Mot2 qui suit. Il est remplacé par une glide avec une préservation de CV sur frontière (ex. *ils ont* [ijɔ̃] *acheté* pour [il.zɔ̃]).

3.1.3.3 Le statut de la consonne de liaison (CL)

La liaison cause l'apparition d'un segment latent (la CL). Différents chercheurs ont cherché à comprendre son statut et la nécessité de la formaliser comme telle dans la représentation lexicale. Plusieurs options ont été proposées pour le définir.

- ***La CL est rattachée au Mot1***

Dans cette option, la CL est traitée comme un segment attaché au Mot1. Nous présentons les descriptions de troncation de Schane (1968) et Selkirk (1974), d'allomorphie de Gaatone (1978a) et Plénat (2008) et enfin de la phonologie autosegmentale par Clements et Keyser (1983) et Encrevé (1988).

Pour Schane (1968) et Selkirk (1974), la forme sous-jacente (FSJ) porte une consonne sous-jacente qui est détruite par troncation dans les cas où Mot2 commence par une consonne, comme dans (8) ; de même, elle apparaît dans la forme de surface (FS) en présence de liaison, comme dans (9).

(8) FSJ /ɛ̃n#/tapi/ → Troncation → FS [ɛ̃.ta.pi]

(9) FSJ /ɛ̃n#/elefã/ → FS [ɛ̃.ne.le.fã]

L'idée centrale de cette explication est que « la non-liaison est le résultat d'un processus actif alors que la liaison elle-même ne résulterait d'aucun processus particulier et serait transparente » (Durand *et al.*, 2011 : 107). Selon cette idée, la CL est une consonne lexicale toujours présente dans la FSJ.

Pour Gaatone (1978a) et Plénat (2008), la CL est une consonne stable et elle appartient à un allomorphe distinct. Nous avons deux FSJ pour *petit* dans (10) et (11).

(10) /pə.tit/ quand le mot qui suit commence par une voyelle.

(11) /pe.ti/ quand le mot qui suit commence par une consonne.

Clements et Keyser (1983) traitent la liaison comme un segment flottant qui est associé, par une règle, au sommet syllabique de la syllabe suivante, donc à l'attaque du Mot2. Ils sont les premiers à proposer cette interprétation, représentée en figure 6 pour la séquence *dont un ami*.

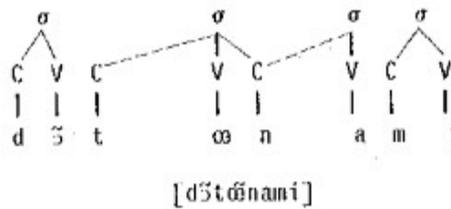


Figure 6. Représentation de la séquence *dont un ami* par Clements et Keyser (1983 : 103)

Pour Vergnaud (1982) « [...] toute règle est inutile et la liaison se déduit directement des propriétés analytiques du système : la condition de liage s'appliquant aux autosegments flottants associe directement la consonne flottante et l'attaque flottante » (Encrevé, 1983 : 57). La CL doit donc être interprétée comme l'attaque du Mot2 (Figure 7). Le problème de cette représentation est que la CL peut seulement se lier à l'attaque du Mot2. Cependant, nous savons qu'il y a des cas où la CL n'est pas liée au Mot2, comme les cas de liaison non-enchaînée.

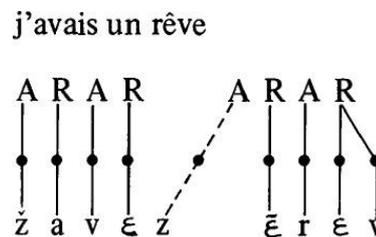


Figure 7. Représentation de la CL comme consonne flottante dans Encrevé (1983 : 57)

Encrevé (1983), en partant de l'idée que la CL doit être représentée différemment d'une CF toujours réalisée, propose donc une nouvelle interprétation autosegmentale qui a aussi ses racines dans les travaux de Goldsmith, dans les

propositions de John McCarthy et dans les élaborations successives données par Halle et Vergnaud (1980), Clements et Keyser (1981), Kaye et Lowenstamm (1982). Dans sa représentation, la CL est toujours un segment flottant mais elle flotte sur la ligne syllabique et segmentale. Dans le cas où l'attaque du Mot2 est vide (c'est-à-dire quand le Mot2 commence par une voyelle), le matériel segmental de la CL flottante peut se lier à la position vide (attaque du Mot2) et réaliser une liaison enchaînée (Figure 8). Si la liaison n'est pas enchaînée, la consonne flottante est liée à la coda du Mot1 (Figure 9).

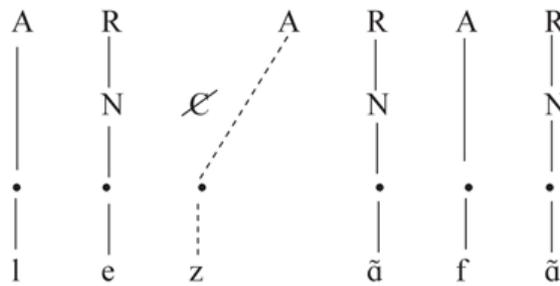


Figure 8. Représentation de la liaison enchaînée

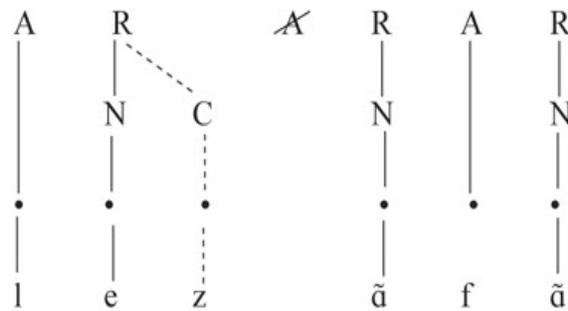


Figure 9. Représentation de la liaison non-enchaînée

En revanche si la liaison n'est pas réalisée, le contenu segmental de la CL n'est pas lié au squelette ou à la position prosodique, et il reste non réalisé (Figure 10).



Figure 10. Représentation des séquences sans liaison

Au moyen d'un segment flottant, il est possible d'expliquer deux FS avec une seule FSJ. En effet, dans le cas où le Mot2 commence par une voyelle, l'attaque est vide et donc la CL va remplir cette position. Dans le cas où le Mot2 commence par une consonne, la CL ne s'attache pas à l'attaque tant que la position n'est pas vide et la CL reste plutôt liée au Mot1.

Cette première option, où la CL est attachée au Mot1, comporte des avantages parce que la plupart des consonnes de liaison sont historiquement et orthographiquement des consonnes finales du Mot1. De plus, elle montre la prévisibilité de la CL, qui est notée dans la représentation sous-jacente, et elle permet d'expliquer le contenu mélodique de la consonne.

- ***La CL comme épenthétique***

Dans une deuxième option, la CL est épenthétique, c'est-à-dire qu'elle est un segment ajouté entre Mot1 et Mot2 (Klausenburger, 1974 ; Côté, 2005, 2008a). En tant qu'épenthétique, elle n'appartient ni à Mot1 ni à Mot2 et la représentation sous-jacente ne porte pas de consonnes. La CL est insérée entre FSJ et FS par épenthèse dans le cas de réalisation de la liaison comme en (12). Dans cette conception, la liaison est un processus actif, car dans le cas de non-liaison aucun changement n'est prévu entre FSJ et FS (comme en (13)).

(12) FSJ /*ẽ*/ # /e.le.fã/ → Épenthèse → FS [*ẽ*.ne.le.fã]

(13) FSJ /*ẽ*/ #/ta.pi/ → FS [*ẽ*-ta.pi]

Côté (2014) propose que la CL ne puisse pas être traitée comme initiale de Mot1 ou comme finale de Mot2, en présentant l'option épenthétique. Ces données sont basées sur l'assibilation en français laurentien⁵. En confrontant les consonnes [t] et [d] en position initiale et de liaison, elle trouve que ces consonnes sont affriquées de façon catégorique si elles sont des consonnes initiales. Par contre, il n'y a pas une assibilation catégorique si les consonnes en question sont des CL. Pour cette raison la CL ne doit pas être traitée comme consonne initiale du Mot2. De plus, les données sur le laurentien montrent que « les consonnes de liaison sont assibilées plus fréquemment que les consonnes finales. Cela suggère que les consonnes de liaison ne correspondent

⁵ Variété de français parlée au Québec, Canada.

pas à des consonnes finales stables » (Côté, 2014 : 23). Ainsi la CL ne peut pas être traitée comme finale du Mot1. Les résultats sur le laurentien soutiennent l'hypothèse d'une CL épenthétique.

L'explication épenthétique que nous avons présentée nous paraît satisfaisante pour la fausse liaison. Dans celle-ci, en effet, une consonne inattendue apparaît à l'attaque du Mot2 (ex. *chemins de fer*[z]anglais, *quatre*[z]enfants ou *Il va*[t]à la foire) dans des contextes où les autres CL devraient être prononcées (ex. [z] à la place de [p] dans la séquence *trop*[z]important). L'apparition de cette consonne inattendue est vue comme le résultat d'une CL insérée par épenthèse entre les deux mots.

Cependant, l'inconvénient d'une telle solution est de ne donner aucune représentation de la CL : « Le *t* de *petit ami* n'est inscrit en rien dans la représentation lexicale du mot *petit* (/pətit/), c'est-à-dire qu'il a le même statut que n'importe quelle autre consonne n'appartenant pas à ce mot, et que n'importe quel son n'appartenant pas à la langue française... » (Encrevé, 1988 : 105). De plus, il est impossible d'expliquer pourquoi c'est un [t] ou [z] ou toute autre consonne qui fait surface. Dans ces cas, parfois la CL a l'une ou l'autre mélodie ([t], [z]) ou toute autre encore. Enfin si c'était une consonne épenthétique, nous ne pourrions pas expliquer pourquoi il n'y aurait pas une seule consonne épenthétique (la même partout quelle que soit la finale de mot) ou pour quelle raison il y aurait [t] dans *petit* et pas [z] qui est la plus fréquente.

- ***La CL comme affixe***

Morin (1992, 2003, 2005a, 2005b) ne voit pas la CL comme une consonne lexicale mais plutôt comme un morphème affixal. Dans son analyse de la liaison avec des adjectifs pré-nominaux, la CL est partiellement autonome par rapport aux Mot1 et Mot2 et elle est vue comme un affixe, au même titre que le <s> du pluriel dans le nom anglais *boys*. Cette explication ne considère pas la liaison comme un phénomène simplement phonologique, mais comme l'expression d'un morphème. La présence et la nature du préfixe sont déterminés par l'adjectif qui précède le nom. De cette façon, nous pouvons interpréter ce phénomène comme dans le schéma en (14), où la liaison adjectivale devient une sorte de déclinaison du nom.

- (14) *ami* /ami/ *t-ami* [tami] (ex: *grand ami*)
 n-ami [nami] (ex: *mon ami*)
 z-ami [zami] (ex: *les amis*)

Selon le schéma, le locuteur français possède trois allomorphes pour le mot *ami*. Il réalisera celui qui conviendra à la situation.

- ***La CL attachée au Mot2***

L'idée est que la CL est une consonne lexicale rattachée plutôt au Mot2. Dans l'explication donnée par Ternes (1977), la CL est totalement attachée au Mot2 et elle en est la consonne initiale. La FSJ du Mot2 contient donc la CL (/le/ /zami/ pour représenter la séquence *les[z]amis*). Dans cette option, la CL n'est pas un préfixe, comme dans l'explication donnée par Morin, mais elle est intégrée au Mot2 et elle fait partie du morphème suivant.

- ***La CL intégrée dans les constructions***

Pour Bybee (2001, 2005), la CL est une consonne lexicale qui fait partie d'une construction plus large incluant Mot1 et Mot2 (pour *un ami*, la construction est [un ami] et pas [un] + [ami]). La CL est comme une consonne interne à cette construction et elle est réalisée par rapport aux caractéristiques de la construction en elle-même et aux fréquences des cooccurrences de Mot1 et Mot2 stockés dans la mémoire. Nous allons présenter plus en détail, dans le prochain chapitre (chapitre 5), les questions de « construction » et de « fréquence » pour mieux comprendre les données d'acquisition. En ce qui concerne le modèle de Bybee (2001a), la CL est partiellement lexicalisée car « même si elle apparaît entre ce qu'il est convenu d'appeler des mots et non à l'intérieur même de ces mots, elle est très proche des alternances conditionnées morphologiquement et lexicalement qui ont lieu à l'intérieur du mot » (Bybee, 2005 : 30). La CL est insérée dans la structure abstraite en (15), qui représente la liaison en [z] entre déterminant et nom (du type *les[z]amis*). Cette séquence tient compte des questions morphologiques et syntaxiques.

(15) [[DETERMINANT-z-[voyelle]-NOM] pluriel]

Pour mieux comprendre la position de la CL, Bybee considère les recherches de Ågren (1973), de Morin et Kaye (1982) et de Green et Hintze (1988) qui signalent que la CL peut être réalisée même si les locuteurs produisent une pause ou une hésitation entre Mot1 et Mot2. Elle prend l'exemple de Green et Hintze (1988 : 159) que nous présentons en (16) :

(16) *quatre euh... [z]obligations*

En (16), après une hésitation entre Mot1 et Mot2, la CL [z] est produite comme consonne initiale du Mot2. « La position de la pause dans cet exemple et dans d'autres tirés de ces travaux suggère que la consonne de liaison est davantage associée avec le second mot, dont elle commence la première syllabe, qu'avec le premier, qui en est la source historique » conclut Bybee (2005 : 29).

Cette dernière hypothèse donne une explication forte pour la fausse liaison et pour l'analyse de liaison chez les enfants de 2-3 ans ou chez les locuteurs ne sachant pas l'orthographe, comme nous allons voir dans le chapitre suivant. De plus, elle est en accord avec les données sur les langues créoles qui montrent que la CL a été parfois traitée par les locuteurs comme consonne lexicale du Mot2 (Rhodes, 2008). Cependant cette explication n'est pas économique parce que tous les contextes de liaison doivent être mémorisés dans des constructions afin que l'allomorphe approprié soit réalisé selon le Mot1 présent dans la séquence.

Dans ces paragraphes, nous avons synthétisé les différentes options pour la représentation de la CL. Le tableau 2, dans Côté (2014 : 13), nous en donne une synthèse.

Option pour le statut lexical des consonnes de liaison		Représentation lexicale
(1) C finales : les CL sont des consonnes finales du Mot1	a. Analyse par troncation : les CL sont des consonnes stables qui chutent dans des contextes de non liaison (Schane, 1968 ; Selkirk, 1974)	/døz/ /ami/
	b. Analyse par supplétion : les CL sont des consonnes stables appartenant à des allomorphes distincts (Gaatone, 1978 ; Plénat, 2008)	/dø, døz/ /ami/
	c. Analyse autosegmentale : les CL sont des consonnes stables par rapport au squelette ou à la syllabe, avec une représentation distincte de celle des consonnes stables (Clements & Keyser, 1983 ; Encrevé, 1988 ; Tranel, 1995)	/dø(z)/ /ami/
(2) C épenthétiques : les CL sont insérées par épenthèse (Klausenburger, 1974 ; Coté, 2005, 2008)		/dø/ /ami/
(3) C initiales : les CL sont des consonnes initiales du Mot2 (Ternes, 1977)		/dø/ /zami/
(4) C morphémiques : les CL correspondent à des morphèmes affixaux	a. Analyse préfixale : les CL sont des préfixes du Mot2 (Gougenheim, 1938 ; Morin & Kayes, 1982 ; Morin 2003)	/dø/ /z+ami/
	b. Analyse suffixale : les CL sont des suffixes du Mot1 (Morin, 1992)	/dø+z/ /ami/
(5) C intégrées à des constructions : les CL font partie de constructions (partialement) lexicalisées plus larges que le mot (Bybee, 2001a, 2005)		/dø z NOM/ /ami/

Tableau 2. Synthèse dans Côté (2014 : 13) : différentes options pour la représentation de la CL

3.2 Le *h* muet et le *h* aspiré

Nous développons dans cette section la question du *h* aspiré et du *h* muet, car ces deux contextes ont une influence sur les phénomènes de resyllabation évoqués dans les paragraphes précédents.

Dans la forme graphique de certains mots, nous pouvons trouver un *h* graphique qui est orthographiquement l'initiale du mot (ex. *hippopotame*, *hibou*, *hasard*). Cependant, ce *h* orthographique n'a pas de réalisation phonétique correspondante (ex. *hippopotame* [i.pø.pø.tam]). Même si dans toutes les réalisations à l'oral de mots isolés commençant par *h* nous n'avons aucune trace de cet élément, tous les mots qui commencent par *h*, ne sont pas identiques. Il y a des mots où le *h* appartient seulement à la forme orthographique. Ces mots sont appelés « à *h* muet

initial » : le *h* muet graphique est un indice de l'étymologie latine. Par exemple, dans le mot à *h* muet initial *honneur* [ɔ.nœʁ], le *h* orthographique fait référence à la forme étymologique du mot latin *honor*. En revanche, il y d'autres mots où le *h* initial n'est pas seulement orthographique mais implique des conséquences phonologiques. C'est le cas des mots dit à « *h* aspiré initial ».

Pour définir le phénomène du *h* aspiré initial, nous nous référons à l'article de Côté (2008a : 65).

"The so-called *h* aspiré corresponds to the property of some morpheme-initial vowels to block process that apply with normal initial vowels, and trigger processes that do not apply with regular vowels".

En effet, la présence d'un Mot2 qui commence par *h* aspiré bloque la réalisation de la liaison ou de l'élision. Dans le cas de la liaison, la CL n'est pas réalisée (17) ; dans le cas d'élision, la voyelle ne chute pas avant une autre voyelle (18). En plus, le [ə] est phonétiquement réalisé avant un *h* aspiré, ce qui n'est pas le cas avant une voyelle (19).

(17) *h* aspiré : *le(s) hérissons* → [lɛ.e.ʁi.sɔ̃]

voyelle : *les[z]abeilles* → [lɛ.za.bɛj]

(18) *h* aspiré : *la hauteur* → [la.o.tœʁ]

voyelle : *l'abeille* → [la.bɛj]

(19) *h* aspiré : *le hérisson* → [lə.e.ʁi.sɔ̃]

voyelle : *l'état* → [lɛ.ta]

En ce qui concerne le troisième phénomène traité dans cette thèse, l'enchaînement, la situation du contexte *h* aspiré est plus compliquée. Boersma (2005 : 3) affirme que "*h* aspiré words again act like consonant-initial words". En revanche, Encrevé (1988) considère que l'enchaînement est tout à fait possible dans le cas de Mot2 qui commence par *h* aspiré (20a). Cependant, cette réalisation alterne avec l'insertion d'un schwa (20b) ou d'une occlusive glottale (20c) qui bloque la resyllabation.

(20) *mille hannetons*

a. [mi.lan.tɔ̃]

b. [mi.lə.an.tɔ̃]

c. [mil.ʔan.tɔ̃]

Côté (2005) ajoute que l'enchaînement est plus systématique avec *h* aspiré en français québécois car "schwa insertion [...] is not an option, since schwas do not normally surface between (nonclitic) words in this variety, even after consonant clusters (bourse bleue 'blue purse' can be pronounced [bursəblø] or [bursblø] in Standard French but only [bursblø] in Québec French). As for glottal stop insertion, it seems somewhat marginal in non-emphatic contexts" (p. 66).

En plus, le *h* aspiré bloque la supplétion de certains adjectifs pronominaux (ex. *beau/bel*), de pronoms possessifs (ex. *sa/son*) et de séquences préposition-déterminant (ex. *du/de/l'*). Nous présentons les exemples de Côté (2008a : 65) en (21) et en (22) qui permettent de mieux expliciter la question.

- (21) *h* aspiré : *sa hauteur* [sa otœr] *[sɔ̃ n otœr]
beau héros [bo ero] *[bɛl ero]
du héros [dy ero] *[dø l ero]

- (22) voyelle : *son odeur* [sɔ̃ n odœr]
bel étau [bɛl eto]
de l'étau [dø l eto]

3.2.1 Le schwa dans le contexte *h* aspiré

Comme Encrevé, plusieurs auteurs ont remarqué l'apparition d'un schwa phonétiquement réalisé entre l'adjectif et le nom à *h* aspiré initial. D'abord, Dell (1973a, 1973b) remarque qu'en français nous pouvons entendre *quelle housse* [kɛləus] mais aussi *quel hêtre* [kɛ.lɛtʁ], où dans le même contexte il n'y a pas la réalisation du schwa avant le mot qui commence par *h* aspiré. À partir de ces deux exemples, Dell établit un lien entre phonologie et morphologie, en proposant deux formes sous-jacentes pour les adjectifs *quel* et *quelle* : une forme /kɛl+ə/ pour le féminin ; une forme /kɛl/ pour le masculin. Cependant, Tranel (1981) remarque qu'en français la séquence *quel hasard* [kɛ.lə.a.zaʁ], où nous pouvons noter l'apparition d'un schwa avec un adjectif non-invariable (*quel*) et un mot masculin à *h* aspiré initial (*hasard*), est

possible. Il propose donc que le processus d'insertion du schwa dans ce contexte soit optionnel et "constrained by orthography" (Tranel, 1981 : 288). Dans le cas où nous avons un adjectif qui présente une voyelle finale à la forme écrite, il est plus probable d'avoir une insertion de schwa à l'oral avant un *h* aspiré.

Ces interprétations ne prennent pas en considération l'emphase qui, pour Pagliano (2003), est responsable de l'épenthèse du schwa entre Mot1 et Mot2.

3.2.2 Le coup de glotte dans le contexte *h* aspiré

Dans le cas de l'enchaînement, Encrevé remarque l'insertion d'un coup de glotte avant le mot qui commence par *h* aspiré, pour emphatiser la syllabe initiale de ce mot. Cette insertion empêche la resyllabation de l'enchaînement. Les données expérimentales confirment cette insertion dans la non-réalisation des phénomènes de resyllabation. Gabriel et Meisenburg (2004) enregistrent les lectures de 12 locuteurs natifs du français qui produisent des liaisons lorsque le Mot2 commence par *h* aspiré (ex. *tout hongrois* [tu.tõg.ɤwa]). Ils observent que presque la moitié des locuteurs produisent la liaison, même si dans ce contexte elle ne doit pas être réalisée. L'autre moitié produit un coup de glotte entre Mot1 et Mot2, qui permet de séparer les deux mots et empêcher la liaison ([tu.ʔõg.ɤwa]).

Mais concernant cette occlusive glottale, quelle est sa nature et quel est son statut ? D'abord il faut dire que l'occlusive glottale n'est pas dans le répertoire phonologique du français et qu'elle n'apparaît, dans la plupart des cas, que lorsque le locuteur veut emphatiser sa production orale. Scheer & Ségéral (2001 : 117) nous disent que « une prononciation [gɤosʔus] pour *grosse housse* - qui est objectivement possible - est aussi nécessairement insistante ou emphatique, bref, relève seulement de l'expressivité ». Dans sa thèse, Pagliano (2003 : 640) est en accord avec cette interprétation mais défend aussi l'hypothèse selon laquelle « l'emphase se traduit dans la représentation sous-jacente » et que les mots à *h* aspiré initial comportent des coups de glotte lexicaux initiaux. À la différence d'Encrevé, pour Pagliano, cette occlusive glottale est réalisée phonétiquement, sous emphase ou non, car elle est la manifestation d'une position forte. De plus, comme pour Dell (1973a), pour Pagliano elle est présente dans la représentation lexicale de noms à *h* aspiré, afin d'empêcher les phénomènes de resyllabation.

3.2.3 Un contexte particulier à formaliser

Comme le contexte *h* aspiré bloque les phénomènes de resyllabation, les mots qui commencent par *h* aspiré sont notamment comparés aux mots qui commencent avec une consonne. Ce traitement n'est pas entièrement correct, car nous pouvons remarquer une différence entre les contextes *h* aspiré et les contextes à consonne initiale en ce qui concerne l'insertion d'un schwa clitique. Comme l'affirme Côté (2008a), le schwa clitique est toujours obligatoire avant un mot à *h* aspiré, mais il est optionnel avec un mot qui commence par une consonne (23).

(23)	<i>h</i> aspiré	<i>ce héros</i>	[sə ero] *[s ero]
	consonne	<i>ce bateau</i>	[s(ə) bato]

De plus, le schwa peut apparaître avant un Mot2 qui commence par *h* aspiré et après un Mot1 qui termine par une consonne. Cela n'est pas le cas si le Mot2 commence par une consonne (24).

(24)	<i>h</i> aspiré	<i>mille héros</i>	[mil (ə) ero]
	consonne	<i>mille bateau</i>	[mil bato] *[mil ə bato]

Comme on l'a vu, Pagliano (2003) considère que les mots à *h* aspiré initial intègrent une occlusive lexicale. La présence d'une occlusive glottale permet de distinguer les mots à *h* aspiré des mots à initiale vocalique ou consonantique. Étant donné que cette occlusive glottale n'apparaît pas systématiquement dans toutes les productions en contexte *h* aspiré, elle affirme qu'il n'est pas lié au constituant syllabique. Elle propose qu'un mot à initiale *h* aspiré présente un coup de glotte flottant dans une représentation phonologique (Figure 11) élaborée dans le cadre de la théorie CVCV (Lowenstamm, 1996 ; Scheer, 1998, 1999, 2000).

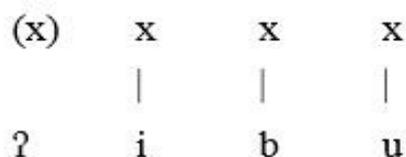


Figure 11. Représentation de Pagliano dans Côté
(2005 : 67)

Comme nous l'avons déjà dit, la liaison avec un mot qui commence par *h* aspiré n'a pas lieu. Cependant le *h* aspiré n'est pas réalisé au niveau segmental et un mot à *h* aspiré en isolation est identique à n'importe quel mot qui commence par une voyelle. Dans sa représentation autosegmentale, Encrevé présente l'élément *h* aspiré comme une attaque : il est « lié à la position initiale du squelette dans la couche du mot, mais sans correspondance dans la couche segmentale » (1983 : 57). Ensuite, il propose que cette position puisse être remplie, sous certaines conditions, par une occlusive glottale (Encrevé, 1988). Puisque le remplissage par l'occlusive glottale n'est pas systématique, l'attaque peut parfois rester vide (Sung, 1989 ; Girard & Lyche, 1997) et dans ce cas, la position est associée à une position prosodique vide à l'attaque du Mot2, indiquée avec \emptyset (Figure 12).

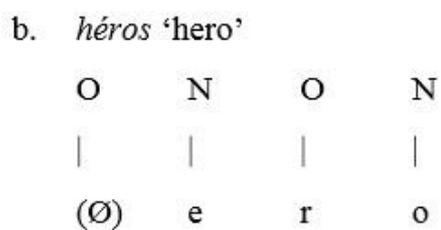


Figure 12. Représentation dans Côté (2005 : 68)

L'hypothèse d'une position vide et d'un possible remplissage de cette position par une occlusive glottale permet d'établir une distinction lexicale entre les mots à voyelle initiale (ex. *étai* [eto]) et les mots à *h* aspiré initial (ex. *héros* [∅ero]/[ʔero]).

L'existence d'une position sous-jacente pour le *h* aspiré qui bloque la liaison est très intéressante pour notre étude, car cette position squelettale peut être signalée par le graphème *h* au début du Mot2 dans la représentation orthographique. Le schwa ou le coup de glotte peuvent être une manière de réaliser ce *h* graphique. Dans cette thèse, où nous cherchons à voir les relations entre l'écrit et l'oral, la phonologie et l'orthographe, la question du *h* aspiré constitue un champ d'observation intéressant à explorer.

3.3 Les phénomènes de resyllabation entre la phonologie et l'orthographe

Nous avons montré dans ce chapitre l'existence de certains phénomènes qui causent une non-correspondance entre les frontières syllabiques et les frontières lexicales. Cette non-correspondance émerge aussi si nous comparons la représentation phonologique et la représentation orthographique en français. Comme Côté (2005) l'a bien décrit dans son article, le français est caractérisé à l'oral par des phénomènes qui, au niveau phonologique, font référence à des *defective segments*, c'est-à-dire des positions prosodiques vides. Il s'agit des cas de deux phénomènes que nous avons présentés : la liaison et le *h* aspiré. L'analyse du *h* aspiré invite à théoriser une position syllabique vide qui n'est pas associée à un contenu segmental au niveau phonologique ; la position vide au début du nom empêche la mise en place des phénomènes consonantiques comme l'élision, la liaison et l'enchaînement. Pour la liaison, dans la représentation autosegmentale l'attaque est vide et remplie par la CL flottante.

Si, du point de vue phonologique, la représentation se sert d'une position vide, dans l'orthographe, des symboles prennent la place de ces positions vides en leur donnant un contenu orthographique. Pour le *h* aspiré, la position phonologique vide est représentée par un *h* graphique au début du nom. Pour la liaison, si le segment flottant est latent du point de vue de la représentation phonologique, il est toujours présent à l'écrit. La CL associée à un graphème du Mot1 va remplir l'attaque du Mot2 phonologiquement vide. L'écrit donnerait donc une correspondance matérielle et visuelle à un segment latent ou à une position vide. Cette représentation orthographique peut interférer avec la représentation phonologique. Quand les locuteurs apprennent à lire ou à écrire, ces représentations phonologiques peuvent être modifiées par l'information orthographique. En ce qui concerne les phénomènes de resyllabation en français, nous pouvons penser que l'acquisition des frontières des mots

à l'oral et leurs représentations phonologiques peuvent être influencées par l'orthographe et par les formes visuelles du lexique orthographique, grâce à l'apprentissage de la lecture et de l'écriture.

Dans les prochains chapitres, nous présenterons l'acquisition de la langue orale et de la langue écrite chez les enfants afin d'établir un rapport entre les deux types de représentation, surtout en ce qui concerne les phénomènes de resyllabation, au centre de notre étude. De plus, pour montrer comment la reconnaissance des mots dans la chaîne orale peut être modifiée par l'acquisition de l'orthographe, nous présenterons aussi la situation des adultes alphabétisés.

Chapitre 4

Segmenter l'oral

4.1 Introduction

Les caractéristiques linguistiques de la langue française et la présence de phénomènes de resyllabation à l'oral compliquent la tâche de segmentation en mots de l'input acoustique, surtout chez l'enfant qui a comme seule référence la chaîne parlée. Dans ce chapitre, nous présenterons d'abord une synthèse des travaux sur l'acquisition des frontières de mots à l'oral chez les plus petits. Nous traitons en détail la segmentation du français en présence de la liaison, phénomène au centre de notre étude.

Ensuite, nous présenterons les pratiques de segmentation chez l'adulte. Ceci nous permettra de comparer le traitement des données des enfants qui n'ont pas de représentations des mots écrits à celles des adultes, dont le lexique orthographique est stocké de manière stable dans leur représentation des mots.

4.2 Segmenter à l'oral

4.2.1 La segmentation de la chaîne parlée chez les enfants

Trouver les frontières des mots dans la parole continue se révèle être l'une des activités les plus importantes que les enfants doivent accomplir dans l'acquisition du langage. Or, atteindre ce but n'est pas aussi facile que nous pourrions le penser de prime abord. Le langage adressé aux enfants n'est pas plus simple ou référentiel que celui adressé aux adultes. Dans les interactions parents-enfants, le langage est, en fait,

assez spontané et les mots présentés en isolation sont limités en nombre et fréquence (Jacob & D'Odorico, 2001). Les mots entendus par l'enfant ne sont pas des événements stables et distincts. En revanche, comme dans le langage adulte, ils sont sujets aux phénomènes de coarticulation ou aux phénomènes de resyllabation qui agissent entre les frontières des mots. De plus, les enfants avant l'entrée à l'école ne peuvent pas se servir de l'orthographe pour segmenter à l'oral. Comment les enfants français reconnaissent-ils les mots dans la chaîne orale ?

Dès lors, pour répondre à cette question, il faut d'abord évoquer certaines caractéristiques de la langue française. À partir des années 1940, grâce aux travaux de Pike (1945) et d'Abercrombie (1967), les langues naturelles ont été classées en deux catégories : les langues à stress ou *stress-timed*, qui comprennent les langues caractérisées par une succession régulière d'unités accentuelles (comme l'anglais), et les langues syllabiques ou *syllable-timed*, qui tiennent compte de la cooccurrence régulière des unités syllabiques, comme le français.

Cette classification, basée au départ sur l'existence d'une isochronie de leurs unités rythmiques (syllabes ou pieds) dans les langues, a été confirmée depuis par des mesures acoustiques. Ramus *et al.* (1999) ont mesuré la proportion de la durée des intervalles vocaliques et la déviation standard d'intervalles consonantiques par phrase dans huit langues différentes. Au moyen de ces deux mesures, ils ont remarqué que ces huit langues se répartissent en trois classes rythmiques (à stress, syllabiques et moraïques). Le français a été classifié comme une langue syllabique, où l'unité sous-jacente à son organisation rythmique est la syllabe. En effet le français est une langue « cursus » qui s'organise en structures à syllabe ouverte du type CV (Delattre, 1965 ; Pulgram, 1970 ; Kaye & Lowenstamm, 1984 ; Béchade, 1992 ; Di Cristo, 1998).

Cette caractéristique du français doit être prise en considération pour parler de segmentation chez les enfants. Selon la "rhythmic segmentation hypothesis" (Nazzi *et al.*, 2006), pour bien segmenter la chaîne orale, l'enfant s'appuie sur les caractéristiques rythmiques de sa langue maternelle, la langue cible : "segmentation differs according to the global rhythmic properties of one's native language" (Nazzi *et al.*, 2006 : 285). Si les enfants anglophones utilisent les indices accentuels offerts par leur langue, comme l'alternance entre syllabe accentuée et syllabe non-accentuée (Cutler & Carter, 1987, Cutler & Norris, 1988) et les unités trochaïques (Jusczyk *et al.*, 1999 ; Curtin *et al.*, 2005), les enfants francophones qui acquièrent une langue syllabique s'appuient au contraire plutôt sur la syllabe. Les résultats des études de Nazzi *et al.* (2006) et de

Nazzi (2008) confirment cette idée : à partir de 6 mois, les enfants francophones commencent à utiliser les unités rythmiques de leur langue, c'est-à-dire les syllabes, pour segmenter en mots. Ils réussissent donc à isoler des mots monosyllabiques. En revanche, ils ne réussissent pas à isoler les mots dissyllabiques, car ils font encore référence aux frontières des syllabes pour établir les frontières des mots. Petit à petit, ils se concentrent sur des unités plus larges, comme les mots lexicaux constitués de plusieurs unités syllabiques. Ils commencent aussi à se servir d'autres indices pour segmenter l'oral (Nishibayashi *et al.*, 2015).

Dans la parole continue, l'enfant dispose des indices linguistiques qui aident à segmenter la chaîne orale et qui lui permettent d'identifier les frontières des mots. Les indices sont de quatre types :

1. Les indices prosodiques.

Ces indices prennent en considération l'accent et l'intonation qui caractérisent les unités à l'intérieur des mots (ex. la syllabe). En français, les syllabes en fin d'une unité prosodique sont sujettes à un allongement (Delattre, 1966 ; Fletcher, 1991). Grâce à l'identification de cet allongement, les enfants francophones peuvent traiter cette unité prosodique comme une unité finale qui détermine la frontière d'un mot (Christophe, 1993).

2. Les indices allophoniques.

La réalisation de certains phonèmes change selon la position qu'ils ont dans les mots. Un phonème peut avoir une certaine réalisation en position de frontière et une réalisation différente à l'intérieur du mot. Par exemple, en anglais, les phonèmes /t/ et /ɾ/ ont des réalisations différentes à l'intérieur du mot comme *nitrate* et entre frontières des mots dans la séquence *night rate* (Hohne & Jusczyk, 1994). L'analyse des différentes réalisations des phonèmes aide les enfants à identifier les frontières.

3. Les indices phonotactiques.

Les enfants se servent des propriétés distributionnelles des patterns acoustiques pour découvrir des règles phonotactiques qui font référence aux combinaisons des sons possibles dans une langue (Saffran *et al.*, 1996b). De cette façon, l'indice sonore aide l'enfant à installer la frontière du mot, là où il y a violation d'une règle phonotactique. Par exemple, en entendant la séquence de l'italien *midaiunbacio* [midaiunba.tʃo] (fr. *fais moi un bisous*), l'enfant italoophone utilise les règles phonotactiques pour segmenter la phrase : il sépare *dai* [dai] de *un*

[un] en sachant que les deux voyelles *i* [i] et *u* [u] sont très rarement internes aux mots ; il segmente ensuite *un* [un] de *bacio* [ba.tʃo] en considérant que la séquence *nb* [nb] n'est pas possible dans la langue italienne. La même chose s'applique au français en présence de séquences comme [kf] et [vg], qui ne se retrouvent jamais à l'intérieur d'un mot et donc aident à poser ses frontières.

4. Les indices statistiques et distributionnels.

Comme l'écrit Nazzi (2008 : 312), « l'agencement des phonèmes et des syllabes consécutives au sein des mots peut fournir des indications sur la probabilité que deux sons appartiennent ou non au même mot ». Certaines transitions phonémiques sont plus probables à l'intérieur du mot qu'entre les frontières de deux mots. Les enfants analysent ces probabilités pour déterminer les frontières des mots.

Cependant, même si les enfants peuvent se servir de ces indices présents dans la langue française, il reste des contextes difficiles à segmenter.

4.2.2 Le contexte déterminant-nom

Comme nous l'avons déjà exposé dans la section précédente, le français est une langue à rythme syllabique qui s'organise autour des syllabes ouvertes. De plus, 53 % des noms sont monosyllabiques (vs. par exemple une autre langue romane comme l'italien où il y a seulement 4 % de monosyllabes). À l'oral, la prosodie induit une perception des suites déterminant + nom comme une seule unité prosodique.

Non seulement l'article fait partie de la même unité prosodique que le nom en perception mais encore il aide à distinguer, en tant que clitique, le genre et le nombre dans la plupart des cas où ils ne sont pas morphologiquement et phonologiquement exprimés sur le nom (ex. *le vase* [lə.vaz] vs. *la vase* [la.vaz] ; *les phoques* [lə.fɔk] vs. *le phoque* [lə.fɔk]) et peut être considéré quasiment comme un préfixe détaché.

À cause de ces caractéristiques prosodiques et syntaxiques, le déterminant est produit très fréquemment dans l'input adulte (Guasti *et al.* 2008 ; Bassano 2010) et par conséquent, il est produit plus tôt en français, à 2 ans et demi, comparé aux langues germaniques comme l'allemand ou le hollandais (Bassano, 2010).

Ces caractéristiques morpho-syntaxique du déterminant en français et le fait que dans la suite déterminant + nom les deux éléments soient perçus comme une seule unité prosodique ont des conséquences sur la segmentation. L'analyse des premières

productions des enfants révèle ces difficultés. Dans ses travaux, Bassano (2010) montre qu'avant la production des articles dans la forme standard reconnue à l'âge adulte, il y a la production de *fillers* constitués de structures monosyllabiques placés systématiquement avant les noms. Plusieurs données transversales (Braud, 2003) et longitudinales (Wauquier-Gravelines, 2004) ont confirmé cette tendance en montrant que les enfants entre 12 et 26 mois produisent le nom commun avec la trace du déterminant défini interprétable comme un « proto-déterminant », sous-spécifié morphologiquement mais spécifié phonologiquement et prosodiquement (ex. production comme [lə.tɛ] pour *le chien* [lə.ʃjɛ̃]). Dans une étude de 2006, Dugua remarque dans la production des séquences déterminant-nom du type *un latrompe* ou *un lalampe* qui montrent que l'article n'est pas détaché du nom et qu'ils ne forment qu'une seule unité. La prosodie du français permet, en effet, d'expliquer l'apparition de ce proto-déterminant ou *filler* : les enfants adaptent le nom en produisant une structure dissyllabique accentuée sur la deuxième syllabe. L'idée que met en valeur le rôle de la prosodie dans l'apparition du proto-déterminant a bien été développée par les travaux de Demuth et Fee (1995) et Demuth (2001). Ils proposent un modèle prosodique de l'acquisition de la langue où les premières unités produites par les enfants sont influencés par leurs niveau d'accès aux unités prosodiques.

Ces types de productions confirment la difficulté pour les enfants francophones L1 de traiter le déterminant et le nom comme deux unités séparées. La structure de la langue française, qui est codifiée dans l'input, cause aux jeunes locuteurs des problèmes de segmentation et de reconnaissance du mot lexical au tout début de l'acquisition. Pour compliquer encore les choses, il y a la présence de phénomènes de resyllabation qui agissent sur ce contexte, parmi lesquels la liaison.

4.2.3 L'acquisition de la liaison

Comme les enfants francophones utilisent la syllabe pour la segmentation (§ 4.2.1) et que l'input qu'ils entendent est sujet aux phénomènes de resyllabation qui entraînent le non-alignement entre frontières lexicales et syllabiques (chapitre 3), nous pouvons penser que ces phénomènes peuvent être responsables des difficultés de segmentation.

Pour l'anglais, le non-alignement entre frontières syllabiques et lexicales implique un retard dans la segmentation de séquences resyllabifiées. Mattys et Jusczyk (2001) montrent que les mots à voyelle initiale, sujets à resyllabation, sont bien segmentés seulement à environ 16 mois. Avant, les enfants ne sont pas capables, par exemple, d'isoler le mot *ice* (fr. *glace*) dans la phrase *weird ice [wiər.daɪs] no longer surprises anyone* (fr. *la glace bizarre ne surprend plus personne*), où le mot cible est précédé par un adjectif à consonne finale qui est resyllabifié et responsable d'un non-alignement.

Comme cela a été le cas pour l'anglais, les enfants francophones peuvent avoir des difficultés à segmenter les séquences des mots sujets à la liaison, à l'élision ou à l'enchaînement. En utilisant la syllabe comme unité fondamentale pour la segmentation, ils peuvent faire correspondre le début de la syllabe avec le début du mot et produire des formes comme *zours [zuʁs]* pour *ours [uʁs]*. La segmentation devient plus difficile et problématique.

Dans ce panorama, nous pouvons nous poser la question suivante : comment les enfants segmentent-ils en français dans le cas d'un input non-aligné ? Plus précisément, comment arrivent-ils à segmenter les séquences sujettes aux phénomènes de resyllabation ?

Nous présentons d'abord ci-après les données expérimentales en perception pour analyser ensuite les données en production.

L'étude de Babineau et Shi (2011) analyse la segmentation en mots, en présence de liaison, chez les enfants de 20 et 24 mois. Au moyen de plusieurs tâches de perception où les propriétés distributionnelles des sons et les unités syllabiques sont en conflit, ils montrent que les enfants francophones de 20 mois préfèrent les mots où les frontières de syllabes sont préservées. Ils utilisent alors dans la segmentation l'unité syllabique plutôt que les propriétés distributionnelles des sons. Ces chercheurs concluent qu'à 20 mois, les enfants francophones “relied on the syllabic edges to find word boundaries, i.e. a strong bias for syllabic integrity at this age” (p. 6). Donc s'il n'y a pas d'alignement entre les frontières de mots et les frontières de syllabes, les enfants n'utilisent pas les propriétés distributionnelles des sons mais préfèrent se servir de la structure syllabique pour opérer une segmentation. C'est seulement à partir de 24 mois que les enfants utilisent les propriétés distributionnelles pour segmenter en mots.

Comme Babineau et Shi (2011 : 9) le disent : “the older infants appeared to have learned to overcome the syllable-alignment constraints for parsing liaison cases”.

Or, c'est exactement autour de 24 mois que des erreurs de segmentation apparaissent dans la production des enfants francophones. Ces erreurs de liaison sont décrites et analysées par plusieurs auteurs, surtout dans le contexte spécifique déterminant-nom.

Chevrot et Fayol (2001) ont conduit une étude longitudinale au moyen de l'enregistrement du parlé spontané d'une enfant francophone (Sophie) dans un contexte familial. La production de Sophie a été analysée entre 25 et 42 mois. Les résultats montrent 276 cas appelés de « remplacement » ou de « consonne erronée » où elle utilise la consonne de liaison [n] au lieu de [z] (ex. [trwa.nuʁs] pour [trwa.zuʁs] dans *trois ours*). De plus, certaines erreurs ressemblent plutôt à des épenthèses : elle ajoute [n], [z], [t] ou [l] (qui sont toutes des consonnes possibles de liaison ou d'élision) entre des mots où il n'y a pas la liaison (ex. [papa.nuʁs] dans *papa ours*).

Plusieurs études (Wauquier & Braud, 2005 ; Dugua, 2006 ; Nardy & Dugua, 2011 ; Wauquier & Shoemaker, 2013 ; Chevrot *et al.*, 2013) ont confirmé des productions caractérisées par l'intégration de la CL au début du nom (ex. *de(s) [n]éléphants* [de.ne.le.fã] pour *des[z]éléphants* [de.ze.le.fã] ou [pa.pa.nuʁs] dans *papa ours*). Cette « lexicalisation » peut se produire en présence ou non d'une allomorphie du déterminant et d'une perte de nasalité (*je vais à ma [n]école* [ʒə.vɛ.a.ma.ne.lɔl] pour *je vais à mon école* [ʒə.vɛ.a.mɔ̃.ne.kɔl]). Dans une moindre mesure, les chercheurs enregistrent la complète omission de la CL (*u(n)[Ø]avion* [ɛ̃.a.vjɔ̃] sans réalisation de la CL [n] pour *un[n]avion* [ɛ̃.na.vjɔ̃]) substituée parfois par un coup de glotte.

La CL peut être traitée comme une consonne lexicale et le contraire est aussi tout-à-fait possible. Parfois les enfants traitent la consonne lexicale comme CL et leurs productions montrent la chute de la consonne lexicale qui est omise comme si c'était une CL (*Blanche-Neige et les 7 ains* [blãf.nɛ.ʒe.le.se.tɛ̃] pour *Blanche-Neige et les 7 [n]ains* [blãf.nɛ.ʒe.le.set.nɛ̃]). La consonne lexicale est parfois aussi remplacée par une autre consonne lexicale erronée. Cela est le cas des mots qui commencent par une consonne lexicale qui est aussi dans le répertoire de consonnes de liaison (ex. [z]). Les

erreurs que nous avons remarquées sont des séquences comme *un-nèbre* [ɛ̃.nɛbr] pour *un zébre* [ɛ̃.zɛbr]. Dans le tableau 3 nous résumons tous les types d'erreurs possibles en production observés chez les enfants.

Types d'erreurs	Exemples
Consonne erronée ou remplacement	[le.nan] pour [le.zan] Les [n]ânes pour les [z]ânes
Allomorphes du déterminant + nasalité	[ʒə.vɛ.a.ma.ne.kəl] pour [ʒə.vɛ.a.mɔ̃.ne.kəl] Je vais à ma [n]école pour je vais à mon école
Omission : absence de consonne ou coup de glotte en attaque	[le.e.le.fã] pour [le.ze.le.fã] le(s)[Ø]éléphants pour les[z]éléphants
Chute de consonne lexicale avec liaison	[blãʃ.ne.ʒe.le.se.tɛ̃] pour [blãʃ.ne.ʒe.le.set.nɛ̃] Blanche-Neige et les 7 ains pour Blanche-Neige et les 7 [n]ains
Consonne lexicale erronée	[ɛ̃.nɛbr] pour [ɛ̃.zɛbr] un nèbre pour un zébre
Épenthèse	[pa.pa.turs, pa.pa.nuʁs] pour [pa.pa.uʁs] Papa [t]ours ou Papa [n]ours pour Papa Ours

Tableau 3. Description des erreurs de segmentation des enfants après Wauquier (2010 : 8)

Toutes les erreurs que nous avons présentées dans le tableau ne sont pas enregistrées au même âge ou avec la même fréquence. Par exemple, l'erreur de consonne erronée, l'allomorphie du déterminant et l'épenthèse sont les erreurs les plus fréquentes (Wauquier, 2009). De plus, la fréquence dépend de la CL impliquée dans le processus. Les consonnes occasionnant le plus d'erreurs sont les consonnes obligatoires fréquentes, d'abord [n], puis [z] (Chevrot *et al.*, 2009). Les erreurs où la consonne lexicale est interprétée comme une CL (chute de consonne lexicale avec liaison, consonne lexicale erronée) apparaissent plus tardivement que les autres (Wauquier, 2009).

Les données en production montrent que les enfants ont des difficultés à comprendre la différence entre CL et consonne lexicale. Cette difficulté est compréhensible car la CL prend la place de la consonne initiale du nom et constitue de plus l'attaque de syllabe qui, comme nous l'avons dit, est utilisée comme indice pour isoler le début des mots dans la chaîne parlée. L'étude transversale de Dugua (2006) cherche à établir l'âge à partir duquel les enfants commencent à différencier le traitement de la CL de la consonne lexicale. Elle propose une étude transversale qui permet de recueillir des données sur deux cent enfants de 2 ans et 4 mois à 6 ans et 1 mois répartis dans quatre tranches d'âge. Ils sont testés dans la production guidée d'un répertoire de pseudo-mots représentant des animaux imaginaires. Dans une des tâches, l'expérimentateur montrait à l'enfant une image représentant un de ces animaux et disait : *c'est un[n]ivak*. Lorsque l'animal était double ? Si l'enfant disait *deux nivaks*, en gardant le [n] au début du nom, on admettait qu'il avait segmenté la séquence entendue en privilégiant la syllabe CV, où la CL [n] est l'initiale : *un nivak*. S'il produisait *deux-z-ivak* en substituant [z] au [n] entendu, on en déduisait qu'il avait segmenté la séquence entendue en interprétant le [n] de *un[n]ivak* comme une CL. Dugua voit qu'entre 2-3 ans et 3-4 ans les enfants privilégient une segmentation qui préserve la séquence CV, où la CL [n] est intégrée au début du nom. Entre 4 et 5 ans, les enfants admettent deux types de segmentations : dans une première segmentation, le [n] est traité comme la consonne initiale du nom ; dans une deuxième segmentation, le [n] est une véritable CL. En revanche, dans la dernière tranche d'âge, entre 5 et 6 ans, le [n] est exclusivement traité comme CL. Le bon traitement de [n] dans la tâche de Dugua est en accord avec la diminution d'erreurs de liaison en production à l'âge de 5-6 ans. Ces résultats sont également en accord avec une autre expérience (Dugua & Chabanal, 2006) dans laquelle les enfants sont amenés à manipuler des figurines (*âne, écureuil, éléphant, ours*) pour produire des mots à initiale vocalique en isolation dans la situation d'apostrophe (*ours, viens ici !*). Les enfants testés sont divisés en quatre tranches d'âge de 2 ans à 6 ans (2;4-3;1 / 3;2-4;1 / 4;2-5;0 / 5;2-6;1). Les résultats montrent que les enfants les plus jeunes réalisent 35 % de productions avec une consonne [n] ou [z] au début du mot (*zours, nours*). Plus l'âge augmente, plus le pourcentage baisse. Autour de 4 ans, ces productions erronées tombent à 19 %, à 8 % vers 5 ans et 2,8 % à l'âge de 6 ans. Par ailleurs les seules erreurs qui persistent après 5

ans sont les omissions, où la liaison n'est pas réalisée (Chevrot *et al.*, 2013). Ces données confirment l'idée que, vers 6 ans, les frontières des mots commençant par une voyelle sont maîtrisées.

A partir de 7 ans, les enfants commencent à bien maîtriser cette particularité de la langue, avec une production correcte de liaisons obligatoires-catégoriques entre 97 % et 100 % dans les trois contextes : « après déterminant », « pré-verbale » et « expression figée » (Basset, 2000). Le tableau 4 résume le taux de réussite dans différents types de liaison selon les contextes. Les résultats d'une étude récente de Dugua et Baclesse (2014) montrent qu'au CE1 (sept ans), les enfants réalisent bien la liaison dans le contexte déterminant-nom en dénomination. Les erreurs de consonnes erronées (*un[z]ours* pour *un[n]ours*) arrivent seulement dans 0,15 % des cas et les omissions subissent une chute : 1,85 % de cas où la liaison n'est pas réalisée. Nous pouvons affirmer, à partir de ces données, que les enfants à cet âge maîtrisent la liaison obligatoire-catégorique.

Contextes obligatoires	3 ans	7 ans	11 ans
Après déterminant « les ours »	77 %	97 %	100 %
Clitique pré-verbal « j'en ai, ils ont »	91 %	100 %	99 %
Expression figée « tout à coup »	100 %	100 %	100 %
Total liaisons obligatoires	84 %	99 %	99 %
Contextes facultatifs			
Après adjectif épithète « un petit ours »	48 %	-	-
Après un nom pluriel « des enfants idiots »	2 %	0 %	0 %
Après auxiliaire avoir « ils ont un »	-	0 %	0 %
Après auxiliaire être « ils sont ici »	0 %	12 %	20 %
Après verbe « il vient aussi »	1 %	0 %	5 %
Après mot invariable « quand un, en Espagne »	3 %	24 %	12 %
Total liaisons facultatives	5 %	12 %	12 %

Tableau 4. Erreurs et contextes à partir de Basset (2000) dans Chevrot et Fayol (2001)

Les données que nous avons présentées font référence à la liaison qui est toujours réalisée dans l'input adulte et qui, par la certitude de sa réalisation, est appelée obligatoire-catégorique. La situation est différente pour la liaison facultative qui n'est

pas réalisée de façon catégorique dans l'input. Dans le tableau 4, nous pouvons voir que le pourcentage de réalisation correcte de la liaison facultative entre 7 et 11 ans reste autour de 12 % (Basset, 2000). Ce type de liaison est rarement réalisé avant 6 ans (Dugua & Spinelli, 2009) et sa non-réalisation reste la production majoritaire entre 3 ans et 11 ans (Chevrot *et al.* 2007 ; Dugua, 2006). Des études basées sur la méthode d'élicitation montrent que le taux de réalisation des liaisons facultatives est lié aux milieux sociaux des parents. Selon l'activité professionnelle des parents, favorisant ou pas un langage formel, les enfants réalisent plus ou moins de liaisons facultatives (Nardy, 2008 ; Chevrot *et al.*, 2011 ; Martin, 2005). De plus, ce taux est influencé par le sociolecte des parents et par la conscience stylistique des jeunes locuteurs (Chabanal & Embarki, 2002 ; Nardy *et al.*, 2014).

L'acquisition de la liaison obligatoire semble faire référence à un processus différent de celui qui caractérise la liaison facultative. Si l'acquisition de la liaison obligatoire relève, à partir des premières productions des enfants, de l'acquisition au niveau phonologique, l'acquisition de la liaison facultative n'apparaît qu'après, quand la contrainte phonologique est acquise. Elle peut aussi être conditionnée par d'autres composantes que la compétence linguistique, comme par exemple les compétences pragmatiques ou la connaissance de l'usage lexical. Nous avons décidé, pour cette raison, de concentrer notre étude sur la liaison obligatoire-catégorique.

4.2.3.1 Acquisition de la liaison : deux explications possibles

Dans une analyse développementale de la liaison, nous remarquons une proportionnalité inverse entre le pourcentage d'erreurs et l'âge : à une augmentation de l'âge correspond une diminution du pourcentage d'erreurs. Deux explications ont été données pour rendre compte de ce pattern du développement : une explication lexicale (Chevrot & Fayol, 2001) reposant sur la fréquence des mots, et une explication phonologique autosegmentale (Wauquier & Braud, 2005 ; Wauquier, 2009).

L'explication lexicale de Chevrot et Fayol (2001) s'appuie sur une modélisation constructionniste. Dans ce modèle, les représentations internes du locuteur sont constituées de « constructions » créées selon le rapport entre formes sonores et sens que l'usage a rendu conventionnel. Ces constructions ont des niveaux d'abstraction variés. Elles peuvent être des séquences lexicales mémorisées globalement à partir de

l'input phonétique, être associées aux catégories syntaxiques plus abstraites, comme la séquence sujet-verbe-objet (S-V-O), ou avoir des caractéristiques mixtes alliant des catégories abstraites et des éléments lexicaux comme par exemple « c'est + X ». Pour mieux comprendre l'idée de constructions, nous nous référons au travail de Bybee (2005 : 26) qui parle de constructions comme « des séquences récurrentes de morphèmes ou des mots qui sont reliés entre eux par des relations sémantiques, fonctionnelles ou statistiques particulières lorsqu'ils sont utilisés ensemble, relations qu'ils n'entretiennent pas forcément dans d'autres contextes ». Les constructions ne sont pas isolées les unes des autres mais vont progressivement s'organiser en réseau. Grâce à cette mise en réseau, des catégories plus abstraites peuvent émerger. À partir de ce cadre théorique, Tomasello (2003) propose une modélisation de l'acquisition de la syntaxe où les enfants mémorisent des « morceaux concrets de langage », ou *chunks*, qui peuvent être des séquences phonologiques auxquelles ils ajoutent l'intention communicative et les informations contextuelles ou des séquences de mots stockés globalement. Des exemples de *chunks* sont les séquences « regardé (le) livre », « regardé maman », « regardé (la) télé » où le verbe « regarder » est stocké chez les enfants en différentes séquences selon son usage. Les enfants organisent ces éléments par leur mise en relation au travers de schémas locaux. Cela permet d'établir des relations entre *chunks* du type « regardé + X », « mangé + X » et « fini + X ». L'étape suivante est l'abstraction par la création d'un réseau de schémas et l'émergence d'une séquence comme « base verbale + i/é + nom ».

Le même modèle est appliqué par Chevrot *et al.* (2013) à l'acquisition de la liaison : entre 2 et 4 ans, l'enfant français mémorise des *chunks* fréquents constitués par déterminant + nom (ex. *les-[z]-avions*, *un-[n]-avion* ou *les canards*). Puis ces *chunks* sont mis en relation les uns avec les autres. Du point de vue fonctionnel, l'enfant construit des schémas généraux déterminant-nom (ex. *les + X* ou *un + X*) ; du point de vue formel, les noms ayant une consonne initiale qui est une consonne de liaison sont organisés entre eux et comparés avec les autres noms à consonne initiale lexicale ([*na.vjɔ̃*], [*za.vjɔ̃*] vs. [*naviʁ*]). La présence des différentes formes lexicales concurrentes explique les insertions d'une consonne erronée dans la production infantile. Ensuite, entre 4 et 5 ans, les schémas deviennent plus spécifiques et grâce à cela, l'enfant arrive à comprendre quelle variante du nom, dans sa mémoire, est associée à quel déterminant particulier (ex. *les* est suivi par des noms de type /zX/ vs. *un* /nX/, avant voyelle en cas de liaison). La disparition progressive des erreurs dans la

production de liaison chez les enfants est interprétée comme la preuve de la réorganisation lexicale permettant une association pertinente entre déterminant et nom. Selon les auteurs, après cette phase, l'enfant va rendre ses schémas plus abstraits et il va réussir à s'approprier d'une « construction » [[DETERMINANT-z-[voyelle]-NOM] pluriel] (Bybee, 2001) pour s'en servir dans la production.

Ce cadre théorique prédit que les généralisations sur les noms les plus fréquents doivent être maîtrisées avant les autres puisque plusieurs exemplaires du mot en différents contextes seront disponibles dans le lexique de l'enfant. Par conséquent, nous supposons que les liens et l'organisation en schémas seront plus fixes et stables. Comme Bybee le résume bien (2005 : 25) « des mots entiers et des suites de mots sont mémorisés dans le lexique et le critère de mémorisation est la fréquence d'usage, les mots et suites de mots employés fréquemment étant mémorisés globalement » et aussi « à force d'être utilisés, les items fréquents voient leur force lexicale augmenter et les items qui ont une grande force lexicale sont plus facilement accessibles que les autres et résistent mieux au changement analogique » (Bybee, 2005 : 25). Les données récoltées par Chabanal et Liégeois (2014) montrent l'influence de la fréquence des items du langage adulte sur la production infantile. Au moyen d'une étude qui met en rapport le langage de trois enfants et de leurs parents, Chabanal et Liégeois montrent que pour la liaison catégorique-obligatoire « les choix des exemplaires opérés par Prune⁶ sont effectués à partir d'une sensibilité à la fréquence des mots² présents dans l'input parental et non pas en ayant recours à des connaissances phonologiques plus générales, sans lien avec le lexique mémorisé » (p. 280). L'effet de la fréquence des mots est vérifié dans des tâches expérimentales (Dugua *et al.*, 2009). Les chercheurs proposent une tâche de dénomination guidée chez 120 enfants entre 3 et 6 ans et une tâche de perception sur 20 enfants de 5 à 6 ans. Pour créer le matériel, les chercheurs ont demandé à 70 adultes francophones de juger l'orientation (singulier ou pluriel) d'un certain nombre de noms en répondant à une question du type « Entend-on plus souvent *un avion* ou *des avions* ? ». Selon la réponse à la question, les chercheurs ont pu établir une variante rare (pour *avion*, la variante plurielle) et une variante fréquente (pour *avion*, la variante singulière). Il était ainsi possible d'établir un degré d'orientation pluriel et singulier des noms. Le matériel a été utilisé pour créer les deux tâches. Dans la tâche de dénomination guidée, les 120 enfants doivent nommer des cartes qui

6 Prune est une des 3 enfants participant aux enregistrements de l'étude de Chabanal et Liégeois (2014)

représentent des objets. Les résultats de la tâche sont différents selon la tranche d'âge testée. Entre 3 et 4 ans, l'orientation au pluriel des noms corrèle avec la production de liaisons correctes en [z] (*les[z]éléphant*) et aussi avec la production de liaisons erronées en [z] (*un[z]éléphant*). Par conséquent, plus un nom est orienté pluriel, plus les enfants produisent des liaisons en [z] ou des erreurs en [z] devant ce nom. Cet effet d'orientation semble disparaître après 4 ans. Aucune corrélation n'est enregistrée pour les enfants de 4-5 ans et 5-6 ans. La tâche de dénomination montre un possible effet de la fréquence avant 4 ans : les variantes les plus fréquentes dans l'input sont les plus mémorisées et les plus disponibles à la production.

La tâche de perception vérifie ensuite ces résultats entre 5 et 6 ans. 120 enfants sont testés dans cette tâche utilisant le même matériel que la tâche de dénomination guidée, jugé par les adultes (donc l'orientation des noms). Les enfants doivent détecter des noms orientés singulier ou pluriel après un déterminant singulier ou pluriel. Dans la procédure, les enfants voient d'abord une image (ex. un avion) et doivent appuyer sur une touche dès qu'ils entendent le nom correspondant à l'image dans une phrase acoustique. Les chercheurs proposent différentes phrases pour introduire différentes conditions. Dans certaines phrases, l'objet représenté dans l'image peut apparaître dans sa variante en /n/ fréquente (ex. *elle a rêvé d'un AVION géant*) ou dans sa variante en /z/ rare (ex. *il a vu des AVIONS à l'aéroport*), mais aussi ne pas apparaître du tout (phrases *fillers*, ex. *les acrobates étaient impressionnants*). Quand les enfants entendent le nom de l'objet qui était représenté dans l'image (dans notre exemple, *avion*), ils doivent appuyer sur « oui » ; en revanche quand il n'entendent pas ce nom, ils doivent appuyer sur « non ». Les chercheurs enregistrent les temps de réponses pour les différentes conditions et ils remarquent une interaction significative entre orientation du nom et liaison : les noms orientés au pluriel étaient détectés plus vite après liaison en /z/ et les noms orientés au singulier étaient détectés plus vite après liaison en /n/. La conclusion de cette tâche est que les noms sont reconnus plus vite s'ils sont entendus sous la forme de la variante la plus fréquente dans l'environnement. Cet effet de la fréquence semble caractériser seulement les enfants. L'analyse de données de la même tâche chez 36 adultes (Chevrot *et al.*, à paraître) montre qu'il n'y a pas d'effet de la fréquence dans la tâche de perception. L'explication donnée par les chercheurs à ces résultats est que la variante des noms à initiale consonantique (ex. *[n]avion*) a disparu à cause de l'acquisition de la forme orthographique du mot qui fait que la seule variante

possible et disponible soit la variante à voyelle initiale (*avion*), comme nous l'avons vu dans le chapitre 2.

Nous remarquons d'emblée que, si cette hypothèse permet de rendre compte de l'acquisition de la liaison dans les contextes du syntagme nominal [un + X] et [les + X], elle ne permet pas d'expliquer les autres contextes où il y a des rapports syntaxiques différents de [déterminant + nom] d'un contexte adjectival de type *les gentils éléphants*. De plus, elle ne pose pas la question de la liaison en termes syllabiques et positionnels ou de l'acquisition d'une forme sous-jacente associée. Nous discuterons les avantages et les inconvénients de ce modèle vers la fin de la section.

Si la première hypothèse prévoit une construction du lexique et une fixation des schémas selon la fréquence du mot, la deuxième hypothèse, dite morpho-phonologique, est formulée dans le cadre lexical de la *whole-word hypothesis* (Francescato, 1970 ; Ferguson & Farwell, 1975 ; Menn & Matthei, 1992). Dans cette hypothèse, l'acquisition d'un mot passe par des généralisations phonologiques précoces sans aucune médiation des autres unités phonologiques (à la différence des *chunks*) mais au travers d'un gabarit (Vihman, 1996). Le gabarit, ou *template*, est la base spécifique par laquelle l'enfant peut « généraliser ».

C'est dans ce cadre que l'acquisition de la liaison a été formalisée par Wauquier et Braud (2005), Wauquier (2009). Grâce à ce concept de gabarit, l'acquisition de la liaison est décrite en quatre stades. Le premier stade concerne les enfants de 12 et 13 mois. Les associations dans les gabarits, créées par les segments et les syllabes, sont en relation de manière strictement alignée avec la position squelettale. Toutes les unités sont associées les unes aux autres dans un rapport de bi-univocité où ne sont présentes ni lignes autonomes, ni attaque branchante, ni coda (Figure 13).



Figure 13. Schéma d'acquisition de la liaison : le premier stade (Wauquier & Braud, 2005)

Ceci explique l'absence d'erreurs de liaison et soutient la conception selon laquelle la plupart des substantifs sont produits avec un déterminant non segmenté (voir § 4.2.2, Bassano, 2010). Dans la deuxième étape, à 2 ans, l'input est mieux analysé et les unités phonologiques sont connues plus finement. Dans le gabarit, le contenu segmental est désassocié et diversifié. Nous retrouvons la première manifestation d'autonomie entre information syllabique et information segmentale. Les enfants s'appuient sur les unités phonologiques les moins marquées (Peters, 1985) et remplissent l'attaque vide pour garder des schémas CV à l'attaque des mots. En accord avec cela, l'enfant va rattacher la position prosodique de la consonne de frontière sur l'attaque du nom, pour maximaliser l'attaque au détriment de la coda précédente. La position syllabique reste associée à son point squelettal mais est rattachée à l'attaque du nom. De plus, le contenu segmental est détaché et devient « flottant » sur la ligne segmentale. L'attaque vide est remplie par l'enfant avec une consonne qu'il a rencontré dans ce contexte, une CL ([n] puis [l], [j] et [w] par ordre de fréquence), avec pour conséquence la CL qui devient la consonne initiale du mot (Figure 14).

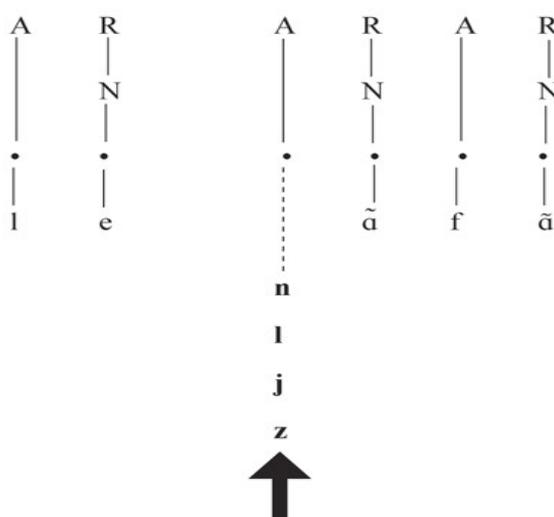


Figure 14. Schéma d'acquisition de la liaison : le deuxième stade (Wauquier & Braud, 2005)

C'est à ce stade que les erreurs systématiques de liaison sont enregistrées, dont en particulier l'utilisation de la CL erronée et l'omission. À l'âge de 3 ans a lieu la troisième étape avec l'acquisition du double flottement sur le palier segmental et syllabique. Encore une fois, le gabarit change et l'enfant encode la liaison à la fin du déterminant mais sous la forme d'un segment flottant (Figure 15).

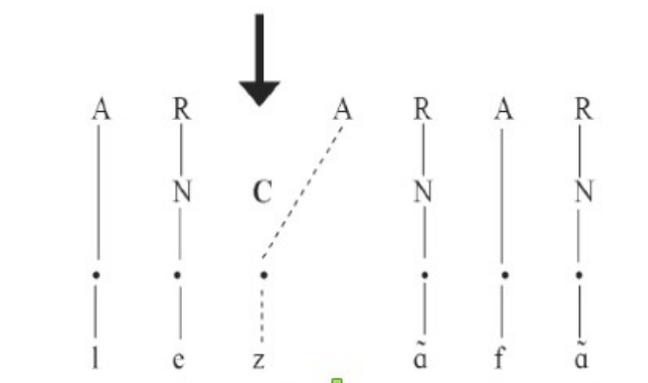


Figure 15. Schéma d'acquisition de la liaison : le stade final
(Wauquier & Braud, 2005)

A ce stade, un rôle central est joué par le *bootstrapping* morpho-phonologique : l'enfant utilise des informations morphologiques pour améliorer sa représentation phonologique. Par exemple, dans le cas de *les* dans la séquence *les enfants*, la CL [z] est encodée dans le déterminant comme morphème grammatical du pluriel qui porte un contenu segmental flottant. Ce contenu flottant peut remplir l'attaque de la première syllabe du nom qui est vide s'il commence par une voyelle. De cette façon, la consonne flottante est encodée dans la représentation lexicale sous-jacente et, par conséquent, les erreurs de liaison obligatoire vont peu à peu disparaître.

Les deux explications ne sont pas très différentes en ce qui concerne la description du début de l'acquisition. Dans les deux cas, dans le premier stade de l'acquisition, le traitement de suites en liaison est basé sur l'usage et Mot1 et Mot2 sont traités comme une seule unité. L'analyse de séquences devient ensuite plus fine. Cependant, les enfants ne réussissent pas encore à traiter la CL différemment d'une consonne lexicale. Qu'elle soit représentée de façon autosegmentale ou qu'elle soit vue comme une forme de surface, elle est dans les deux cas une consonne lexicale attachée au Mot2. Comme les enfants ne tiennent pas encore compte du statut particulier de la CL, ils produisent très fréquemment les erreurs de remplacement (ex. [le.nuʁs]) que nous avons signalées dans la section 4.2.3.

La grande différence entre les deux modèles apparaît dans le traitement du dernier stade de l'acquisition de la liaison. Chevrot propose l'existence d'une structure abstraite du type /un + nX/, /des + zX/, /un petit + tX/ où la CL (*n*, *z* ou *t*) est stockée à la fin du Mot1 et également au début du Mot2. La CL n'est pas représentée de façon

différente d'une consonne lexicale car ce modèle ne prévoit pas une différence des représentations sous-jacente entre les deux. En revanche, dans l'explication autosegmentale-phonologique proposée par Wauquier, à la base de l'acquisition de la liaison, il y a l'acquisition d'une forme sous-jacente particulière où la CL diffère d'une consonne lexicale commune car elle flotte entre le palier mélodique et le palier prosodique : la CL fait surface au début du Mot2 mais le contenu segmental est porté par le Mot1. Les enfants acquièrent la liaison car ils prennent conscience que la CL est associée au double-flottement qui ne caractérise pas la consonne lexicale où les contenus mélodique et prosodique sont seulement portés par le Mot2.

Comme il y a une forte différence entre la représentation de la CL et celle d'une consonne lexicale commune et que cette différence est liée à la représentation des contenus prosodiques et segmentaux, il est très important de faire référence à un gabarit pour décrire les différentes étapes de l'acquisition de la liaison. Pour cette raison, Wauquier propose que la suite déterminant-nom soit stockée et représentée avec ses aspects syllabiques et segmentaux au début de l'acquisition.

Puisque la présence des informations segmentales et prosodiques n'est pas nécessaire pour distinguer CL et consonne lexicale dans le modèle constructionniste, ces aspects ne sont pas pris en compte dans son explication.

4.3 La segmentation chez les adultes : reconnaître les mots dans la chaîne orale

Comme nous l'avons dit, le français a une structure linguistique qui permet l'usage de la syllabe pour segmenter l'oral. Si les enfants sont aidés par la syllabe pour segmenter l'oral, les adultes s'appuient aussi sur cette unité pour reconnaître les mots. Mehler *et al.* (1981) ont remarqué que des locuteurs francophones adultes détectent plus vite le segment *PA* dans un mot comme *palace* (*pa.la.ce*), où la séquence cible est alignée avec la syllabe, plutôt que dans un mot comme *palmier* (*pal.mier*) où le segment cible ne correspond pas à l'unité syllabique. Des résultats analogues dans plusieurs langues confirment le rôle central de la syllabe dans le processus de reconnaissance des mots (pour l'espagnol Bradley *et al.*, 1993, pour l'italien Tabossi *et al.*, 2000, pour l'anglais Dumay *et al.*, 2002). Ces données montrent, selon les chercheurs, une sensibilité pour la syllabe et un traitement sous-lexical de cette unité en français. L'idée de base est que la représentation syllabique stockée en mémoire au

début de nos vies est utilisée ensuite pour capter les mots dans la langue parlée et les organiser au niveau pré-lexical.

Cependant la syllabe n'est pas la seule unité qui intervient dans le processus de segmentation de l'oral. Les locuteurs peuvent également réussir à percevoir des unités mineures, comme les phonèmes, et s'en servir pour construire un lexique. Cet accès aux unités mineures dépend de la complexité du système phonologique propre à la langue. Par exemple, le système vocalique de l'italien comporte 8 voyelles alors que celui du français en prévoit de 12 à 16, selon le dialecte. Il serait donc plus facile à un italoophone d'utiliser le phonème vocalique pour identifier les mots parce qu'il lui suffirait d'entendre la première consonne et de choisir entre un plus petit nombre de combinaisons. En revanche pour un francophone, il serait plus rapide de considérer les syllabes plutôt que d'essayer toutes les combinaisons dans le répertoire de ses voyelles. Savin et Bever (1970) confirment cette intuition. Dans leur étude, ces chercheurs montrent que les adultes francophones détectent plus facilement les syllabes que les phonèmes, confirmant que seule une analyse plus profonde rendrait actives ces unités mineures pour la reconnaissance des mots.

4.3.1 L'accès au lexique chez l'adulte et l'input non-aligné

Nous avons vu que la syllabe est une unité fondamentale pour trouver les frontières lexicales dès l'enfance. Les résultats obtenus concernent un input où le début du mot correspond au début de la syllabe, comme *palace*, où la première syllabe *pa* coïncide avec le début du nom. Mais dans certaines langues, y compris le français, cette correspondance peut ne pas avoir lieu (pour l'anglais *cold* [kəʊld] (fr. *froid*) + *ice* [aɪs] (fr. *glace*) → *cold ice* [kəʊl.dɑɪs] fr. *glace froide* ; pour le danois *de boot is gezonken*, fr. *le bateau a coulé* a une structure syllabique du type *de boo.tis.ge.zon.ken* où il n'y a pas d'alignement ; pour le français nous avons le cas de la liaison, de l'élision et de l'enchaînement).

Le non-alignement entre frontières lexicales et syllabiques qui en découle est responsable d'un problème potentiel de segmentation chez l'adulte. En effet, les mots resyllabifiés sont plus difficiles à reconnaître que les mots où la structure syllabique canonique est préservée. Cutler et Norris (1998) montrent que les locuteurs

anglophones ont certaines difficultés à détecter le mot *mint* (fr. *menthe*) dans la séquence *min.tayf*. Les mêmes résultats sont trouvés pour le danois (Vroomen *et al.* 1996).

La reconnaissance des mots chez l'adulte implique un accès au lexique qui a lieu de façon active (*top-down*) ou rétroactive (*bottom-up*). Accéder au lexique à partir d'un certain input oral peut impliquer un coût de traitement cognitif. Ce coût est mesuré au moyen de la longueur du temps de réaction (TR). Un coût de traitement cognitif plus important consécutif à une activité cognitive plus complexe, implique des TR plus longs. Les expériences psycholinguistiques ont montré qu'il existe un coût dans le traitement des mots où les frontières syllabiques ne sont pas alignées avec les frontières lexicales : l'accès aux représentations des mots est plus difficile. Dans une autre expérience sur le danois, Vroomen et De Gelder (1999) utilisent une tâche de détection du phonème /t/ (*generalized phoneme monitoring*) dans des séquences où le phonème est resyllabifié (*de poort is open*, de.poor.ti.so.pen, fr. *la porte est ouverte*) en comparaison avec des séquences où le phonème n'est pas resyllabifié (*de poort bleef open*, de.poort.bleef o.pen, fr. *la porte reste ouverte*). Dans la tâche, les sujets doivent détecter le phonème cible, sans tenir compte de la position occupée dans la chaîne. Les résultats montrent que “phonemes in non-resyllabified words were detected more quickly than those in resyllabified words” (p.419). Les résultats suggèrent aussi que les phonèmes sont détectés plus vite si la position qu'ils occupent correspond à la position de la représentation lexicale. Détecter un phonème resyllabifié demande un temps plus long car “[...] in resyllabified words, the phonetic information may need to be reassembled across a syllabic boundary (the *t* belongs to the previous syllable *boo*), but not in nonresyllabified words”.(Vroomen et De Gelder, 1999 : 420).

4.3.2 La reconnaissance des mots chez l'adulte francophone en présence de phénomènes de resyllabation

En français, la liaison et l'enchaînement constituent des phénomènes linguistiques qui entraînent une resyllabation sur les frontières lexicales. Ainsi, les frontières lexicales deviennent ambiguës à cause d'un processus de masquage. L'accès au lexique du nom à initiale vocalique peut être plus difficile et impliquer un coût de

traitement cognitif (et un TR plus long). Le traitement de séquences en liaison ou en enchaînement, a-t-il un coût de traitement dans l'accès au lexique chez les adultes ?

Wauquier-Gravelines (1996) observe que le temps nécessaire à la détection du mot *éléphant* dans la séquence acoustique *petit éléphant* est le même que dans la séquence *joli éléphant*. Les mêmes résultats sont trouvés par Gaskell *et al.* (2002) : les auteurs montrent qu'il n'y a pas un coût dans le traitement de séquences non-alignées pour la liaison et pour l'enchaînement. Au moyen du paradigme de *mixing priming*⁷ dans une tâche de reconnaissance de mots écrits (*visual word recognition*), les chercheurs proposent quatre conditions d'amorçage acoustique :

- condition de liaison (*un généreux[z]italien*),
- condition d'enchaînement (*un virtuose[z]italien*),
- condition d'alignement⁸ (*un chapeau italien*),
- condition contrôle (*un mystérieux organisme*).

L'amorçage acoustique (l'écoute d'une de ces quatre conditions, ex. *un généreux italien*), est suivi par une cible visuelle <ITALIEN>. Les participants doivent reconnaître la cible visuelle <ITALIEN> après l'écoute de la séquence acoustique. Pour l'hypothèse selon laquelle le non-alignement entre syllabe et mot est responsable d'un coût dans le processus d'accès au lexique, les chercheurs s'attendent à des TR plus élevés après l'amorce dans les conditions de liaison et d'enchaînement plutôt que dans la condition d'alignement (TR : *un généreux italien* > *un virtuose italien* > *un chapeau italien*). Cependant, les résultats montrent qu'il n'y a pas de différence entre les TR pour les 3 conditions d'amorçage : liaison, enchaînement et alignement (TR : *un généreux italien* = *un virtuose italien* = *un chapeau italien*). Le mot cible <ITALIEN> est reconnu à la même vitesse dans toutes les conditions. Les résultats de cette étude peuvent être discutés car les conditions testées ne sont pas tout-à-fait les mêmes du

7 La base du *priming paradigm* ou "amorçage" est la présentation d'un stimulus (amorce) qui doit influencer la cible et provoquer des résultats anormaux. La proposition d'amorces de types différents permet de voir les influences de plusieurs composantes sur la tâche et donc de confirmer des hypothèses sur la dynamique de certains processus cognitifs. Les variables toujours considérées dans le cas d'amorçage sont au nombre de deux dans la plupart des expériences : temps de réaction (TR), c'est-à-dire le temps que les participants mettent à donner une réponse, et le pourcentage de réponses correctes.

Le paradigme du *mixing priming* met ensemble une amorce et une cible dans deux modalités différentes. Dans ce cas, les expérimentateurs proposent une amorce acoustique (phrase prononcée) et une cible visuelle (mot écrit affiché sur l'écran).

Pour plus d'informations sur le *priming*, nous invitons à la lecture du chapitre 10.

8 Quand nous utilisons le mot « alignement » nous voulons dire que les frontières syllabiques sont alignées avec les frontières lexicales et il n'y a pas une resyllabation entre les mots.

point de vue syntaxique. Dans les conditions de liaison et d'enchaînement, l'amorçage proposé est une séquence adjectif-nom. En revanche, dans la condition d'alignement, la séquence est une succession nom-adjectif. Les résultats ne tiennent pas compte du rôle que la syntaxe peut jouer dans ces phénomènes. De plus, la présentation d'une cible visuelle implique l'activation du lexique orthographique. L'orthographe peut intervenir dans le traitement des frontières lexicales et annuler la différence dans les TR. Comme l'orthographe peut jouer un rôle dans la reconnaissance des mots en français, nous avons décidé de dédier le chapitre à cette composante (chapitre 5) et de tester aussi son influence dans notre protocole expérimental (chapitres 8, 9, 10).

Des résultats différents de ceux que nous avons décrits ont été enregistrés par Dejean de la Bâtie et Bradley (1995). Les auteurs proposent une tâche de détection du phonème initial du mot (*word-initial phoneme monitoring task*) pour vérifier si les phonèmes de liaison (CL) sont détectés plus difficilement que les autres phonèmes. Ils proposent des séquences où le phonème /t/ a quatre statuts différents :

- liaison : *c'est un petit[t]éléphant* ([pə.ti.te.le.fã]),
- alignement : *c'est un vrai éléphant* ([vʁε.e.le.fã]),
- lexical (liaison potentielle) : *c'est un petit talent* ([pə.ti.ta.lã]),
- lexical (aligné) : *c'est un vrai talent* ([vʁε.ta.lã]).

Les auteurs trouvent que le phonème initial /t/ est plus difficile à détecter quand il est précédé par le mot *petit* que quand il est précédé par le mot *vrai* (TR : *petit talent* > *vrai talent*). La différence entre *petit* et *vrai* est que dans le cas de *petit*, le /t/ est une consonne latente qui ne peut être réalisée que si le mot qui suit commence par une voyelle. Ces résultats montrent que la liaison potentielle dérange le traitement des frontières lexicales et cause un coût dans la détection du phonème entre Mot1 et Mot2.

Mais derrière ce type de traitement cognitif, y a-t-il une différence au niveau acoustique entre la CL et la consonne lexicale au début du Mot2 ? Les caractéristiques phonétiques de la CL et de la consonne lexicale influencent-elles l'accès au lexique ? Nous répondons à ces questions dans la section suivante.

4.3.2.1 Perception de la CL et accès au lexique

Plusieurs auteurs ont cherché à comprendre si la CL a des propriétés spécifiques au niveau acoustique qui peuvent aider les enfants et les adultes à la différencier de la

consonne lexicale initiale du mot. Si les études plus anciennes n'établissent aucune spécificité acoustique à la CL (Grammont, 1914 ; Passy, 1917 ; Nyrop, 1925), celle de Delattre (1947) propose que la position d'une consonne dans une structure syllabique et à l'intérieur d'un cluster implique une réalisation différente. Il suggère une réalisation plus relâchée pour la consonne d'enchaînement que pour la consonne initiale. Ces données sont confirmées par les résultats empiriques de Durand (1936) et Fougeron & Delais-Roussarie (2004). De plus, il semble que les locuteurs utilisent ces différences acoustiques pour trier les deux types de consonnes (Fougeron *et al.*, 2001. Fougeron & Delais-Roussarie, 2004). Et pour la CL ? Les analyses phonétiques de Dejean de la Bâtie (1993) nous présentent des différences de durée d'occlusion et de durée de VOT entre CL et consonne initiale. La CL [t] a des durées de VOT et d'occlusion inférieures à [t] consonne initiale. Des données similaires ont été enregistrées pour la même consonne par Wauquier-Gravelines (1996) et par Gaskell *et al.* (2002) et Spinelli *et al.* (2003) qui nous confirment une réalisation plus courte pour la CL que pour la consonne initiale. En ce qui concerne le contexte de réalisation de la liaison, les résultats chez les adultes montrent que la voyelle avant une CL est plus longue qu'une voyelle avant une autre consonne en position de coda (Nguyen *et al.* 2007). Sur la base de ces données de longueurs, Schoemaker (2010) essaie de vérifier la durée acoustique dans la segmentation du français *online*. En effet, la longueur phonétique peut moduler l'interprétation lexicale des séquences qui sont rendues ambiguës à cause de la liaison. Les résultats montrent que nous ne pouvons pas encore dire avec certitude si les caractéristiques acoustiques de la CL sont utilisées par les locuteurs en perception pour distinguer la CL de la consonne initiale et accéder au lexique. Wauquier-Gravelines (1996) ne trouve aucune corrélation statistique entre les différences de réalisation acoustique et l'accès au lexique chez les locuteurs francophones natifs. En revanche, Gaskell *et al.* (2002) et Spinelli *et al.* (2003) se montrent partants pour une sensibilité aux variantes acoustiques qui intervient dans l'accès au lexique. Wauquier dans son article (2009 : 112) invite à une « certaine prudence à l'égard d'une hypothèse selon laquelle l'information acoustique permettrait prioritairement aux locuteurs natifs de résoudre les ambiguïtés de frontière suscitées par la liaison ». Elle ajoute en utilisant l'idée que Nguyen *et al.* (2007) ont déjà formulée :

« On suggérera plutôt des modèles où les sujets vont explorer simultanément les informations acoustiques, phonologiques, lexicales et post-lexicales et les utiliser alternativement et selon leur disponibilité respective sur la base

de stratégies dont la logique dynamique et non strictement modulaire nous échappe encore en partie ».

4.4 De l'enfance à l'âge adulte : une question d'orthographe

Dans ce chapitre, nous avons présenté les données d'acquisition de la liaison chez les enfants (§ 4.2.3). Les productions des enfants entre 2 et 6 ans montrent une difficulté dans la production de séquences en liaison qui peut être la conséquence d'un problème d'identification des frontières des mots dans la chaîne orale. Nous avons vu que chez les adultes, la situation est différente (§ 4.3). La reconnaissance d'un mot à voyelle initiale (*italien*) par les locuteurs du français comporte les mêmes difficultés, qu'il soit présenté dans une séquence resyllabifiée (*un genereux italien*), ou dans une séquence alignée (*un chapeau italien*).

Or, les enfants ont comme seul moyen pour identifier les frontières des mots la langue parlée. En revanche, les adultes ayant acquis l'orthographe ont aussi une représentation écrite. Comme un mot appris à l'écrit est une forme visuelle séparée des autres par des espaces blancs, ses frontières lexicales, dans sa représentation orthographique, ne sont pas perturbées par les phénomènes de resyllabation de l'oral. Un adulte alphabétisé qui connaît la forme orthographique du mot *ours* sait que ce mot commence par une voyelle et que, même si à l'oral il est sujet aux phénomènes de resyllabation, il reste toujours un mot qui commence par une voyelle. Les adultes peuvent donc utiliser ces représentations orthographiques dans la tâche de perception et la connaissance de l'orthographe peut être responsable des résultats que nous avons décrit dans la section 4.3.

Par conséquent, nous pouvons penser que l'orthographe joue un rôle dans la segmentation en mots et dans l'identification des frontières lexicales. Cette hypothèse est confortée par les données d'acquisition. Comme nous l'avons vu (§ 4.2.3), les erreurs de liaison obligatoire-catégorique disparaissent vers 6 ans. Les seules erreurs qui restent stables après cet âge sont les omissions. À partir de 6 ans, les enfants francophones entrent à l'école primaire et commencent l'apprentissage spécifique de la langue écrite au moyen de la lecture et de l'écriture. Cet apprentissage influence-t-il l'acquisition des frontières des mots et vient-il à l'aide des productions orales ?

Nous apporterons des réponses à ces questions au cours de cette thèse, après avoir traité plus en détail l'acquisition de la lecture et de l'écriture.

Chapitre 5.

Représenter et reconnaître les mots écrits

5.1 Introduction

Dans ce chapitre, nous décrirons les processus que les enfants et les adultes mettent en place quand ils doivent donner une représentation écrite à la langue orale, pour ensuite associer une production orale à ces représentations, et ainsi lire.

Nous nous concentrons d'abord sur la description de deux concepts fondamentaux : *emergent literacy* et *literacy*. Ensuite, nous nous focalisons sur la problématique de représentation et de reconnaissance des mots écrits qui font référence à l'activité d'écriture et de lecture. Enfin, nous traitons dans des paragraphes spécifiques des difficultés d'écriture et de lecture de la langue française et des phénomènes de resyllabation au centre de cette recherche : le français. Nous évaluons aussi le rôle de la syllabe en lecture et en écriture. Étant donné que dans la langue orale, elle est au centre des processus de resyllabation, nous cherchons à montrer si elle est aussi impliquée dans l'écriture et dans la lecture du français.

5.2 *Emergent literacy et literacy*

Avant de commencer à traiter les activités de l'écriture et de la lecture, nous présentons deux concepts clés, très utilisés dans la littérature qui s'occupe de l'acquisition de la langue écrite. Ces deux concepts sont connus dans la littérature en anglais comme *emergent literacy* et *literacy*.

“Emergent literacy consists of the skills, knowledges, and attitudes that are developmental precursors to reading and writing” (Whitehurst & Lonigan, 1998 : 848). L'*emergent literacy*, qu'on pourrait traduire par « littéracie émergente », décrit donc les activités concernant la langue écrite que les enfants pratiquent avant l'apprentissage explicite de l'écriture. Quand nous parlons d'*emergent literacy*, nous considérons toutes les connaissances et les compétences de la langue écrite, caractérisant l'enfant dans sa phase pré-scolaire. Ces connaissances et les activités plutôt spontanées associées sont à la base du futur apprentissage scolaire de la lecture et de l'écriture.

Dès que les enfants commencent l'apprentissage explicite de la langue écrite au moyen des pratiques scolaires, nous commençons à parler de *literacy*. La *literacy*, ou « littéracie », comprend les compétences de lecture et d'écriture qui commencent à se développer à partir de la scolarisation et qui caractérisent les connaissances de l'enfant et de l'adulte alphabétisés.

Dans cette recherche nous analysons la lecture et l'écriture au début de l'école primaire, qui est le tournant entre *emergent literacy* et *literacy*. À ce moment, la frontière entre ces deux composantes n'est pas encore bien délimitée, et certains principes d'*emergent literacy* sont utilisés pour la construction de la *literacy*. Nous avons décidé de traiter en détail les compétences impliquées dans l'*emergent literacy* et dans la phase de *literacy* au début de l'école primaire dans les prochains paragraphes.

5.2.1 La sensibilité à l'écrit, ou *print awareness*, chez les enfants entre *emergent literacy* et *literacy*

Les enfants qui commencent à découvrir le monde ne sont pas seulement exposés aux sons des mots prononcés. Ils sont aussi plongés dans un environnement caractérisé par des mots écrits qui sont imprimés sur des objets physiques, par exemple les titres des journaux ou des livres qu'ils ont chez eux, les panneaux publicitaires, les signalisations qu'ils peuvent rencontrer dans la rue. Les enfants sont donc exposés dès leur naissance à un input linguistique écrit, et l'analyse de cet input permet la création d'une idée du code écrit avant l'entrée à l'école. Cette exploration de l'écrit constitue la base pour le développement de la lecture et de l'écriture scolaires.

Au moyen de l'analyse de la langue écrite, les enfants comprennent les régularités orthographiques et deviennent sensibles à ces régularités. Cette sensibilité orthographique (*print awareness*) est une aptitude précoce qui permet de comprendre

les caractéristiques et l'organisation de la langue écrite. Par exemple, elle permet de reconnaître que la langue écrite est « linéaire », qu'elle se développe de gauche à droite (Ganapole, 1987 ; Lavine, 1977), et qu'entre deux mots écrits il y a un espace blanc. À partir de 5 ans, les enfants commencent à comprendre que les mots écrits ne sont pas faits de lettres isolées, mais sont une série d'unités qui doivent répondre à deux conditions formelles appelées par Ferreiro (2000) « principes ». Le premier principe a été appelé « principe de quantité minimale » : un mot avec une seule lettre n'est pas considéré par les enfants comme une séquence lisible. Pour avoir un « vrai » mot il est nécessaire d'avoir un minimum de graphèmes à l'écrit, sinon « il n'y en a pas assez » (Ferreiro, 2000 : 250). Les enfants refusent en effet de considérer une séquence comme pouvant constituer un mot si celle-ci a moins de 3 lettres ; une carte avec la représentation d'un < A > ne présente pas assez de symboles pour construire une séquence légale dans la langue écrite. La deuxième condition formelle qui a été observée chez les enfants est appelée « principe de variété interne ». Un mot, pour être considéré comme tel, ne doit pas seulement avoir une quantité minimale de lettres, celles-ci doivent aussi être différentes. Une séquence lisible est constituée par des symboles différents les uns des autres. Par conséquent, une carte avec une séquence de lettres comme <MMMM> ou <AAAA> ne peut pas être une séquence licite à l'écrit car « c'est faux si on a deux fois la même chose » (Ferreiro, 2000 : 160). Il est intéressant de noter que ces principes ont été observés dans des langues différentes : pour l'espagnol (Ferreiro & Teberosky, 1979), ensuite pour le français (Ferreiro, 2000) et pour l'italien (Ferreiro *et al.*, 1996 ; Ruvoletto, 2011).

La sensibilité orthographique ne comprend pas seulement des jugements sur les régularités orthographiques, mais aussi des compétences qui associent aux lettres une signification. Les enfants, avant l'entrée à l'école, peuvent comprendre qu'un mot est associé à une signification (Justice & Ezell, 2001) et que le rapport avec ce signifiant est très différent du rapport qu'on peut avoir avec un dessin, car il manque d'iconicité (Tolchinsky, 2003). Snow *et al.* (1998) ont vu que des enfants très petits reconnaissent le < M > du symbole de *McDonald* comme une représentation pour le fast-food. Ils arrivent donc à comprendre qu'une lettre peut représenter un objet. Ce type de connaissance est un élément préalable à la construction des associations entre son et symbole qui sont à la base de l'alphabet. Plusieurs études ont en effet montré une

corrélation entre la sensibilité orthographique et les niveaux de lecture accomplis pendant la scolarité (Adams, 1990 ; Badian, 2001 ; Stuart, 1995 ; Weiss & Hagen, 1988).

La sensibilité orthographique n'est pas seulement une sensibilité à la langue écrite en général, elle est aussi en rapport avec les caractéristiques de la langue maternelle de l'enfant. Les données de Treiman *et al.* (2007) montrent l'existence d'une sensibilité orthographique de l'anglais chez les enfants anglophone de 4 ans. Avant l'entrée à l'école, ces enfants sont déjà capables d'identifier ce qui est écrit avec leur propre alphabet et ce qui est écrit avec un autre alphabet. Ils reconnaissent, en plus, des éléments de la langue écrite très spécifiques, comme par exemple la forme de la première lettre de leur prénom.

Cette sensibilité aux caractéristiques de la langue écrite devient une véritable connaissance des ses régularités quand les enfants commencent à apprendre explicitement la langue écrite. Dans une autre étude concernant la pratique de la *literacy*, Treiman (1993) montre qu'il y a une correspondance entre les caractéristiques du système écrit de l'anglais et les connaissances sur l'orthographe à la fin du CP : les élèves ont tendance à doubler plus fréquemment les lettres qui se doublent effectivement le plus souvent dans l'anglais écrit plutôt que celles qui ne le sont jamais.

Cette sensibilité orthographique est liée aux activités de lecture partagée que nous présentons dans les paragraphes suivants.

5.2.2 Le rôle de la lecture partagée et la sensibilité à l'écrit

Le rapport entre enfants et adultes est très important pour le développement psycho-cognitif de l'enfant, car l'adulte peut influencer le comportement de l'enfant, son bien-être mental et physique et ses compétences langagières (Dawson *et al.*, 1992 ; DiPietro, 2000).

La présence de l'adulte et la création de situations de lecture partagée (*storybook reading*) peut avoir une incidence sur le développement de la *literacy* chez l'enfant. La fréquence des épisodes de lecture partagée à la maison (Crain-Thoreson & Dale, 1992 ; Purcell-Gares, 1996) influence le développement de la langue écrite. À partir des renseignements donnés par les parents sur la langue écrite, les enfants peuvent renforcer certains principes et certaines conceptions qui concernent *l'emergent literacy*, et créer des bases pour la *literacy* à apprendre à l'école (Landry & Smith,

2006). Une des compétences qui va se renforcer grâce aux moments de lecture partagée entre enfant et adulte est la sensibilité à l'orthographe (Lundberg, 2006). La lecture partagée permet aux enfants de se familiariser avec les mots écrits et les régularités de la langue et de renforcer certaines hypothèses sur l'orthographe dans la phase pré-scolaire.

5.3 Les compétences phonologiques qui caractérisent *literacy* et *emergent literacy*

Les compétences phonologiques sont à la base de l'apprentissage de la lecture et de l'écriture. Plusieurs études ont montré l'existence d'une relation très étroite entre les compétences phonologiques et les niveaux de lecture et d'écriture au début de l'école primaire (Ecalte & Magnan, 2008 ; Goswami, 2001 ; Bertelson *et al.*, 1985 ; Casalis & Louis-Alexandre, 2000 ; Ecalte, 2000). D'autres études, plus concentrées sur les troubles du langage, ont mis en lumière une correspondance entre les difficultés d'apprentissage du code écrit et les problèmes phonologiques (Bradley & Bryant, 1978 ; Goswami & Bryant, 1990 ; Lecocq, 1986 ; Lundberg, 1982 ; Lundberg *et al.*, 1980 ; Morais *et al.*, 1979 ; Perfetti, 1985 ; Shankweiler, 1999 ; Stanovich, 1988 ; Stanovich & Siegel, 1994). De plus, dans certaines tâches phonologiques, les dyslexiques ont des scores inférieurs à ceux de normo-lecteurs (Colé & Sprenger-Charolles, 1999 ; Lecocq, 1991 ; Sprenger-Charolles *et al.*, 2000) .

Plusieurs types de composantes phonologiques aident l'acquisition de la *literacy* et influencent la construction de la connaissance orthographique dans ses premières phases. Nous en présentons trois : la *phonological awareness*, la mémoire phonologique et le *phonological naming*. La première compétence fait référence à l'analyse métalinguistique de la langue. Les deux autres compétences soulignent le rôle des mémoires à court et à long terme dans l'acquisition de la *literacy*.

La phonological awareness

Tous les travaux portant sur la *literacy* ont souligné d'emblée le rôle fondamental de la *phonological awareness*, ou compétence phonologique. Il s'agit de l'ensemble des capacités à reconnaître, identifier et manipuler les unités phonologiques

comme les phonèmes, les rimes et les syllabes (Morais *et al.*, 1979). La compétence phonologique chez l'enfant agit comme une loupe. Les premières unités identifiées par l'enfant sont les unités les plus grandes, par exemple les syllabes. Ensuite il peut se rendre compte de l'existence des rimes, et de la séparation entre attaque et rime. Il peut arriver à un autre niveau d'analyse, plus profond, avec la conscience de la structure syllabique noyau-coda. Enfin, en concomitance avec l'apprentissage de l'écriture, il a accès aux unités plus petites et abstraites, les phonèmes. Cette séquence d'analyse, construite sur la base de la taille des unités phonologiques (à partir des unités plus grandes jusqu'aux unités plus petites), est commune à toutes les langues, mais la taille des unités maîtrisées au début de l'apprentissage change selon les caractéristiques du langage parlé par l'enfant. Par exemple, les enfants turcs montrent une sensibilité aux phonèmes déjà avant le début de l'école primaire parce que, pour comprendre leur langue à l'oral, ils doivent être attentifs aux changements phonémiques (par exemple l'harmonie vocalique pour la formation du pluriel). Pour cette raison, à 5 mois d'école primaire, ils ont déjà une bonne association entre lettres et sons même si la méthode d'enseignement de la langue écrite est centrée sur les unités phrasales (Durgunoğlu & Öney, 2002). Les enfants locuteurs du turc, du grec et de l'italien ont une bonne connaissance des syllabes avant le début de l'école parce que la structure syllabique de leur langue est très facile à analyser (dans la plupart des cas nous sommes en présence d'une structure syllabique du type CV) et les unités syllabiques sont d'accès facile. Dans d'autres langues où la structure syllabique est plus complexe, comme l'anglais ou le français, l'analyse des unités linguistiques est plus difficile et la prise de conscience des unités plus petites demande plus de temps. Même si les unités linguistiques maîtrisées au début de l'acquisition de l'écrit changent selon la langue, il a été montré que le processus de départ a toujours lieu à partir des unités les plus grandes et que la connaissance des phonèmes s'améliore au cours de l'apprentissage de la *literacy* (Calfée *et al.*, 1972 ; Lewkowicz, 1980 ; Liberman *et al.*, 1974).

La mémoire phonologique

Quand nous parlons de mémoire phonologique, nous faisons référence à la mémoire à court terme des informations sonores (Baddeley, 1986) qui est mesurée au moyen de tâches de mémoire verbale, où les enfants doivent répéter des pseudo-mots de longueur croissante.

La mémoire phonologique aide les enfants dans l'acquisition de la lecture, car elle permet de garder la représentation des phonèmes associés aux lettres en phase de décodage (passage de la lettre au son) de la forme écrite. De cette façon, les enfants peuvent mobiliser plus d'attention et de ressources cognitives à d'autres activités que le décodage en lui-même, comme par exemple la compréhension des séquences écrites.

La dénomination phonologique ou phonological naming

La dénomination phonologique (ou *phonological naming*) fait référence à la possibilité de repérer des informations phonologiques stockées dans la mémoire à long terme. Plus les enfants sont capables de retrouver les informations phonologiques, plus ils peuvent s'en servir pour les associer aux lettres, aux syllabes ou aux morceaux des mots.

5.4 *Literacy : représenter à l'écrit*

Dans les premières périodes qui caractérisent *l'emergent literacy*, vers 4 ans, les productions écrites de l'enfant sont encore indifférenciées, car ni l'enfant ni l'adulte ne peuvent distinguer les différents mots écrits (Tolchinsky, 2006). Ensuite, l'enfant établit un lien entre le monde réel et la graphie : ses premières formes écrites représentent donc des entités réelles et elles sont associées à des significations. Vers 6 ans, l'enfant commence à établir des associations entre phonème et graphème. L'écriture devient une écriture du type « alphabétique » ou « phonétique », où l'enfant utilise les liens entre les lettres et les sons pour écrire/lire. Au départ, les associations sont établies entre les sons et les lettres, ensuite l'enfant arrive à associer un son à un graphème (donc à un groupe de lettres). Un mot comme *table* /tabl/, où chaque lettre est associée à un son, serait donc écrit correctement avant un mot comme *soute* /sut/ où le son [u] est représenté par deux lettres <ou>. On passe alors de *l'emergent literacy* à la *literacy* proprement dite. En conséquence, les productions écrites des enfants commencent à être sujets à l'effet de la fréquence (les mots écrits plus souvent sont mieux écrits) et de la lexicalisation (par ailleurs les mots sont mieux écrits que les non-mots). Toutes ces évolutions dans les compétences des enfants sont liées aussi à l'entrée à l'école (Sprengrer-Charolle *et al.*, 1997). À partir de la première année de l'école primaire (CP),

l'enfant intègre la médiation phonologique à la médiation orthographique pour arriver à la construction d'un lexique orthographique (Share, 1995; Frith, 1985; 1986 ; Morton, 1989 ; Harris & Coltheart, 1986). Dans le cadre de l'acquisition de la *literacy* à l'école primaire, pour apprendre à bien orthographier, il est donc nécessaire de passer à un niveau plus abstrait. Pour devenir d'habiles écrivains, il faut construire un *mapping* bien structuré entre les phonèmes et les graphèmes où les liens entre ces deux unités sont bien stables. Nous utilisons le schéma (Figure 16) que nous retrouvons dans l'article de Ziegler et Goswami (2005) afin de mieux comprendre le déroulement de l'apprentissage de l'écriture chez l'enfant après 6 ans.

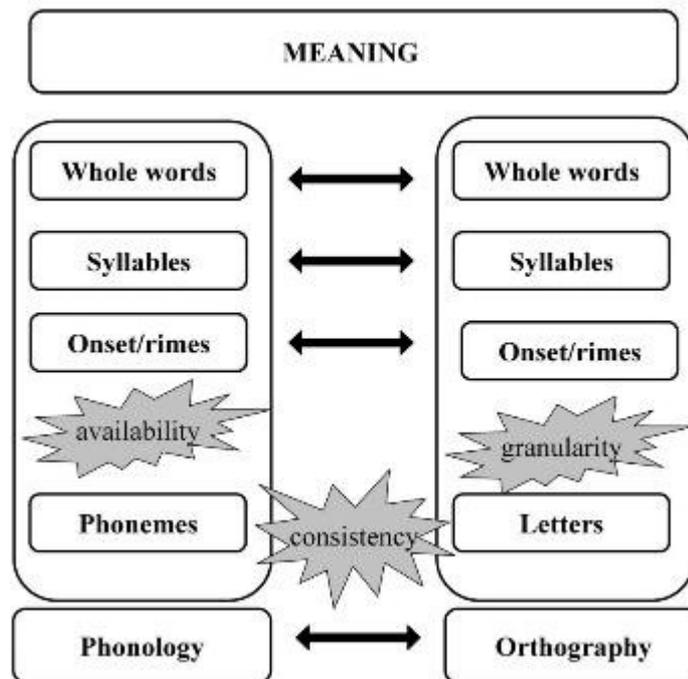


Figure 16. Schéma qui met en rapport oral et écrit dans l'acquisition de la *literacy* par Ziegler et Goswami (2005 : 4)

Pour apprendre à lire et à écrire, il est nécessaire de réussir d'abord à maîtriser les différents niveaux orthographiques sous-lexicaux : mots, syllabes, attaques/rimes et graphèmes/lettres (la colonne droite du schéma). À cette période, il est aussi nécessaire de reconnaître les unités phonologiques sous-lexicales (colonne gauche du schéma). Enfin, les enfants doivent réussir à installer des connexions (flèches) entre les deux types d'unités (les deux colonnes). Ces connexions seront ensuite stockées en mémoire dans ce qu'on appelle le « lexique mental » où les mots sont représentés comme des séquences de caractères et de phonèmes. Établir les connexions entre les sous-éléments

phonologiques ou orthographiques et entre unités phonologiques et lexicales fait émerger trois problématiques (en gris dans le schéma) : la disponibilité (*availability*), la consistance (*consistency*) et la granularité du *mapping* graphème-phonème (*granularity of the mapping spelling to sound*).

- **Disponibilité (*availability* dans le schéma)**

Toutes les unités phonologiques ne sont pas accessibles explicitement à partir de la forme de surface qu'est le mot dans son ensemble (*whole word* dans le schéma). Accéder aux sous-unités phonologiques nécessite un niveau de développement cognitif suffisant pour construire une compétence métalinguistique au service de cet accès. Cette compétence, que nous avons décrite dans la section 5.3 comme *phonological awareness*, permet d'analyser les unités phonologiques plus en profondeur, en passant des mots aux syllabes, aux attaques-rimes et enfin aux phonèmes. La disponibilité des unités mineures dans une langue dépend de ces caractéristiques. Dans certaines langues comme l'italien ou l'espagnol, les phonèmes sont plus disponibles parce que le lien entre les unités de niveau supérieur (syllabes, mots) et les unités de niveau inférieur (phonèmes) est plus linéaire et transparent. Pour l'anglais et le français, accéder aux unités de niveau inférieur est plus difficile.

- **Consistance (*consistency* dans le schéma)**

Comme nous pouvons le voir dans le schéma, la consistance règle le lien entre les unités orthographiques et les unités phonologiques. Les unités orthographiques n'ont pas de rapport une à une avec les unités phonologiques dans toutes les langues. Dans certaines langues, les unités orthographiques correspondent à plusieurs représentations phonologiques et les unités phonologiques correspondent à plusieurs graphies. Les différents degrés de consistance dans les rapports entre ces deux unités influencent l'acquisition de la lecture et de l'écriture. Les langues où le rapport entre graphèmes et phonèmes est plus biunivoque, dites consistantes, sont apprises mieux et plus vite. Les langues où la plupart des phonèmes sont représentés par différents graphèmes, dites inconsistantes, demandent un temps d'acquisition plus long et des efforts plus importants.

- **Granularité (*granularity* dans le schéma)**

Entre les unités phonologiques majeures (mots, syllabes, attaques/rimes ; dans le schéma, à droite) et le niveau des graphèmes, Ziegler et Goswami posent la granularité. La granularité se réfère à la grandeur des unités sous-lexicales (*grain size*) utilisées par les enfants dans les activités de *literacy*. Selon les chercheurs, la taille de ces unités change selon la langue et elle implique les difficultés d'apprentissage. Les unités utilisées ont en effet une taille différente et à l'augmentation de cette taille correspond aussi l'augmentation du nombre des exemplaires possibles pour chaque unité : nous pouvons compter plus de mots que de syllabes, plus de syllabes que de rimes, plus de rimes que de graphèmes, plus de graphèmes que de lettres. Une langue qui pose sa base orthographique dans des *grains* (unités sous-lexicales) plus grands (par exemple sur une granularité syllabique plutôt que graphémique) serait plus difficile à apprendre à l'écrit car elle a besoin de faire référence à des unités plus nombreuses qui seront stockées et récupérées plus difficilement et plus lentement dans la mémoire (Ziegler & Goswami, 2005). En revanche, une langue dans laquelle l'orthographe permet d'accéder directement aux unités plus petites (comme les graphèmes) serait plus facile à apprendre à l'écrit parce qu'elle implique un nombre d'unités moindres stockées dans la mémoire qui seront plus facilement et plus rapidement récupérées.

Nous avons présenté les trois concepts de disponibilité, de consistance et de granularité en général pour voir quelles variables interviennent dans l'acquisition de la *literacy*. Nous montrerons dans la section 5.6.2 quel rôle jouent ces composantes pour la langue française.

5.5 Reconnaître les mots écrits

Reconnaître et accéder aux mots écrits est à la base de l'activité de lecture. Pour mieux comprendre les façons dont les enfants reconnaissent les mots écrits en français, il faut d'abord présenter les modèles cognitifs de lecture.

5.5.1 Les modèles de lecture

Pour rendre compte de la reconnaissance des mots écrits chez l'adulte et de l'acquisition de la lecture chez l'enfant, les chercheurs ont proposé trois modèles cognitifs :

- le modèle à double voie,
- le modèle par analogie,
- le modèle connexionniste.

Ces modèles ont pour but de simuler la lecture à voix haute et la lecture silencieuse. Les deux premiers modèles affirment l'existence d'un lexique mental, une récolte de formes visuelles et sonores stockées dans la mémoire à long terme. Ces formes rendent compte des caractéristiques orthographiques et phonologiques des mots.

Ces modèles de lecture présentent des adaptations pour expliquer l'acquisition de la lecture chez les enfants. Nous présentons en détail les trois modèles de lecture et ses adaptations développementales dans les prochains paragraphes.

5.5.1.1 Le modèle à double voie

Le modèle de lecture à double voie (Coltheart *et al.*, 1993 ; Paap & Noel, 1991 ; Humphreys & Evett, 1985) fait référence à deux procédures qui agissent au moyen de deux différentes voies à partir d'un input écrit (Figure 17) :

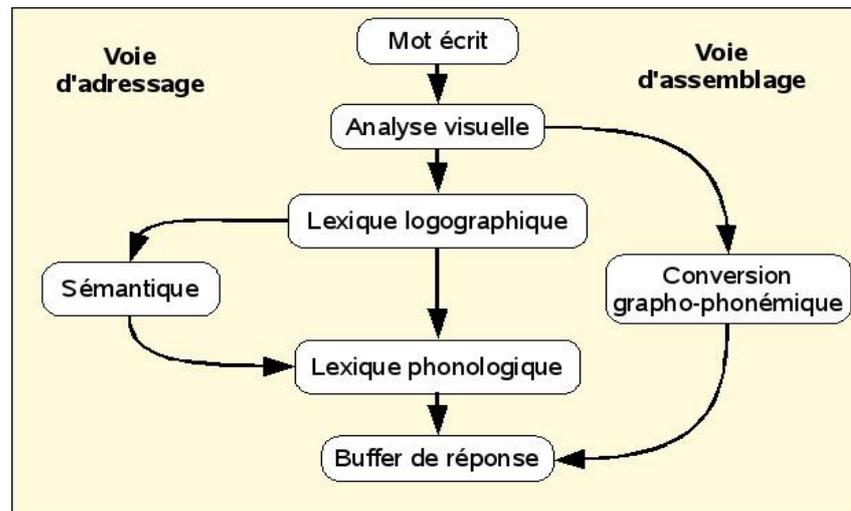


Figure 17. Modèle cognitif à double voie de la lecture des mots isolés (Morton & Patterson, 1980)

1. Adressage par la voie lexicale (« voie d'adressage » dans le graphique)

Dans l'adressage, la représentation orthographique du mot écrit est activée dans le lexique orthographique et donne accès directement à ses significations et à la représentation phonologique. Le processus d'adressage utilise la voie directe lexicale (ou phonologie adressée).

Au moyen de cette voie, le mot écrit est toute de suite associé à l'entrée lexicale qui lui correspond dans le lexique orthographique et active le sens qui lui est associé et la forme phonologique correspondante.

2. Assemblage par la voie pré-lexicale (« voie d'assemblage » dans le graphique)

À la base de l'assemblage, il y a la segmentation de la représentation orthographique qui permet la construction de la représentation phonologique. Ce processus est lié à l'utilisation de la voie indirecte pré-lexicale (ou phonologie assemblée). Le mot écrit est découpé dans les graphèmes qui le constituent. Ces unités sont mises en relation avec les phonèmes en activant le système de règles de correspondance graphèmes-phonèmes (CGP) qui permet de créer un lien entre chaque graphème et chaque phonème. Les phonèmes sont enfin assemblés et le mot peut être lu.

Les deux procédures de lecture aboutissent à un *buffer* de réponse. Le *buffer* est une sorte de mémoire de travail spécifique pour la langue orale où des séquences de phonèmes sont enregistrées pour un temps déterminé (Caramazza *et al.*, 1986). Il est utilisé pour la lecture à voix haute.

Normalement, si le lecteur est expert et qu'il connaît le mot, il va utiliser la voie lexicale pour le lire. Cependant, il peut se servir de la voie pré-lexicale si le mot n'est pas connu ou s'il n'est pas stocké dans sa mémoire. La voie pré-lexicale est utilisée surtout chez les enfants. Quand les enfants n'ont pas encore un vocabulaire de la langue écrite très développé (des formes stockées dans le lexique mental), les mots écrits ne sont pas reconnus tout de suite et ils sont obligés d'utiliser pour la lecture la voie pré-lexicale. En revanche, quand ils commencent à devenir de bons lecteurs, ils peuvent tranquillement utiliser la voie lexicale pour lire.

5.5.1.1.1 Les modèles à étapes

Les modèles à étapes (Frith, 1985, 1986; Harris & Coltheart, 1986 ; Marsh *et al.*, 1981 ; Morton, 1989 ; Seymour, 1986) sont l'adaptation développementale du modèle à double voie. Ces modèles d'acquisition de la lecture comportent trois différentes étapes, ou stades :

- Stade logographique : les enfants reconnaissent les mots écrits présents dans leur environnement, grâce à la maîtrise de lettres et de certaines configurations de lettres (comme par exemple les mots <coca-cola> en blanc sur un fond rouge).
- Stade phonologique : les enfants utilisent les unités phonologiques de différentes tailles et commencent à les associer aux unités graphiques. Les enfants utilisent ce que nous avons appelé la voie pré-lexicale pour lire les mots et la procédure mise en œuvre est l'assemblage.
- Stade orthographique : la lecture n'est pas médiée par les unités phonologiques. Les enfants accèdent directement au lexique au moyen de la procédure d'adressage. Les unités orthographiques sont traitées comme des unités de sens.

La progression développementale est linéaire. Pour accéder au stade ultérieur, les enfants doivent avoir acquis les compétences du stade antérieur.

5.5.1.2 Le modèle par analogie

À l'inverse de l'idée d'une procédure non-lexicale, le modèle par analogie (Glushko, 1981) propose que les analogies entre les mots à lire et les mots connus interviennent dans la procédure d'assemblage. Le modèle suppose qu'entre les graphèmes et les phonèmes, d'autres unités interviennent dans l'identification des mots, comme par exemple les successions attaque-rime. La structure attaque-rime est très accessible dans la langue orale et il a été montré qu'elle est également fondamentale dans le processus de découpage en mots de la chaîne parlée (Treiman, 1985). Dans les modèles par analogies, cette structure intervient dans la médiation entre phonèmes et graphèmes quand les lecteurs doivent lire les mots.

5.5.1.2.1 L'acquisition de la lecture par analogie

Comme son nom l'indique, le modèle analogique suppose que le processus de lecture repose sur le phénomène d'analogie et plus particulièrement sur des analogies de rime. En effet, dans son adaptation développementale par Goswami et Bryant (1990), avant d'apprendre à lire, les enfants établissent des relations entre les mots qui ont la même rime à l'oral. La reconnaissance des rimes ainsi que les rapports établis entre les mots qui ont la même rime à l'oral permettent la construction d'un lexique orthographique de mots écrits. Ensuite, pour enrichir leur lexique écrit et lire de nouveaux mots, les enfants utilisent les analogies avec des mots qu'ils connaissent déjà.

Ce modèle prédit également qu'ils n'utilisent pas seulement la procédure analogique pour lire. Il met aussi en relation les graphèmes avec les phonèmes. L'utilisation des analogies et la mise en relation des correspondances graphèmes-phonèmes sont deux procédures différentes qui impliquent des capacités phonologiques différentes mais qui, ensemble, conduisent à l'acquisition de la lecture.

5.5.1.3 Le modèle connexionniste

À la différence des modèles précédemment vus, où lire suppose que les enfants retrouvent les mots stockés dans la mémoire à long terme, les modèles connexionnistes n'envisagent pas l'existence d'un lexique mental. La lecture est plutôt la conséquence

de l'activation de certaines unités spécialisées (Seidenberg & McClelland, 1989). Les unités spécialisées qui sont activées encodent trois types d'information aux niveaux orthographique, phonologique et sémantique. Ces unités sont interconnectées dans un réseau et elles reçoivent et émettent des informations. L'activation des unités interconnectées permet la lecture des mots (Figure 18). Toutes les unités sont activées automatiquement et il n'y a pas la médiation de règles de correspondance graphème-phonème (CGP). Reconnaître un mot signifie atteindre un certain état d'activation de sa représentation. Chaque configuration d'activation spécifique aboutit à la reconnaissance d'un mot différent.

Le modèle connexionniste permet de bien simuler la lecture silencieuse chez l'adulte et permet d'unifier la modélisation de la lecture de mots et de pseudo-mots.

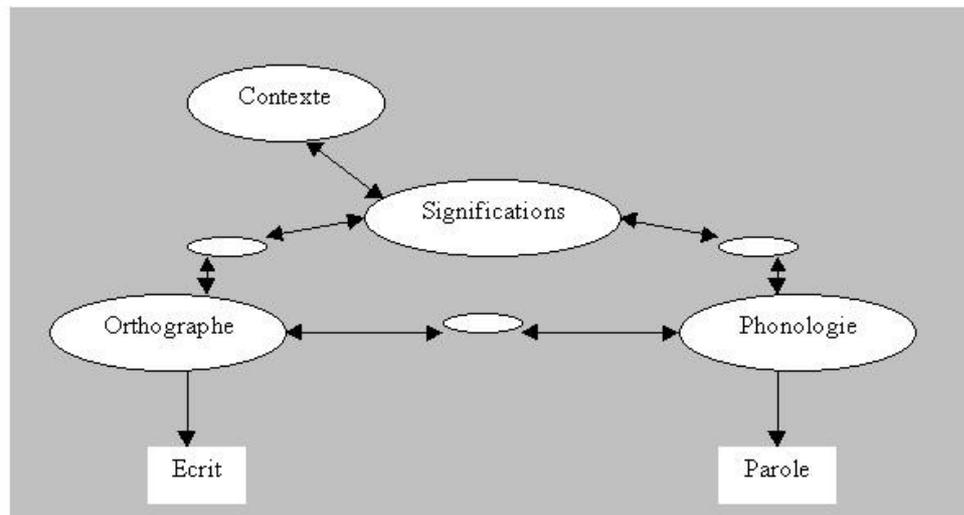


Figure 18. Le modèle connexionniste par Seidenberg et McClelland (1989)

5.5.1.3.1 Le modèle connexionniste et l'acquisition

Comme nous l'avons vu, le modèle connexionniste permet de simuler très bien la lecture de mots et de pseudo-mots chez les adultes. Cependant, il ne permet pas de rendre compte des résultats obtenus dans les études sur l'acquisition de la lecture. D'abord, il ne tient pas compte du rapport entre écriture et lecture mais il se concentre seulement sur la lecture. Ensuite, il n'intègre pas les connaissances métalinguistiques que les enfants utilisent pour apprendre à lire et qui sont des compétences fondamentales qu'il faut considérer dans un modèle qui simule la lecture. Enfin, ce

modèle ne tient pas compte de certaines informations segmentales et de leur organisation hiérarchique (comme la syllabe ou la rime) que les enfants utilisent pour la lecture au début de l'acquisition.

Pour conclure, le modèle connexionniste permet de simuler très bien la procédure de lecture silencieuse chez l'adulte mais beaucoup de progrès doivent encore être faits pour bien intégrer les composantes qui jouent un rôle dans l'acquisition de la *literacy* chez les enfants.

5.6 Lire et écrire en français

Lire et écrire sont deux activités complémentaires qui composent la *literacy* mais qui doivent néanmoins être distinguées car elles font référence à des processus cognitifs distincts (Cossu & Marshall, 1985). En effet, il existe des lecteurs qui peuvent lire correctement des mots sans être capables de les orthographier. De plus, en français, l'activité de lire est plus facile que l'activité d'écrire car, à l'écrit, il y a une redondance des marques silencieuses qui ne sont pas prononcées à l'oral et qu'il faut apprendre par cœur (ex. le <-s> dans <oranges>).

Dans les prochains paragraphes dédiés au français, nous avons décidé de traiter d'abord de la connaissance de régularité orthographique qui se place dans des questions communes entre l'écriture et la lecture. Ensuite, nous décrirons les processus d'écriture et de lecture séparément car comme nous l'avons dit, ils doivent être traités comme deux activités distinctes.

5.6.1 La connaissance des régularités orthographiques en français

Nous avons vu qu'au début de l'acquisition de l'orthographe, les enfants comprennent les régularités orthographiques de leur langue et la sensibilité orthographique devient une véritable connaissance de ces régularités présentes dans la langue (§ 5.2.1).

Une étude de Pacton *et al.* (1996) montre que les élèves francophones sont sensibles aux positions légales des doubles consonnes : dans une tâche de décision lexicale, un non-mot comme <bummor> est traité comme un mot possible alors que <bumorr> est plutôt vu comme impossible car il présente un doublement final illégal en français. Pour juger si un mot est possible ou non, les enfants utilisent leur sensibilité orthographique.

Les mêmes résultats ont été obtenus dans une autre étude conduite chez les enfants d'une école primaire française (Pacton *et al.*, 1998). Pour comprendre les résultats, il faut préciser qu'en français, le phonème /o/ est massivement transcrit comme <o> entre et <l>, comme <au> entre <f> et <t> et comme <eau> jamais au début d'un mot (hors du cas du mot *eau*). Si les enfants ont une connaissance des régularités orthographiques, ils doivent considérer dans une tâche de décision lexicale comme plus probable un mot où <f> et <t> sont suivis de <au>, et <l> de <o>. De plus, ils doivent refuser un mot qui commence par <eau>. Les résultats montrent que les enfants du CP refusent tous les non-mots qui commencent par <eau>, les considérant comme une séquence impossible dans leur langue. En analysant les régularités orthographiques de la langue française, les enfants ont compris qu'un mot qui commence par une séquence <eau> n'est pas un mot possible dans leur langue écrite.

5.6.2 Écrire en français

Écrire en français est une tâche très difficile. Pour un enfant français, écrire un mot régulier est aussi difficile qu'écrire un mot qui n'existe pas dans sa langue. Ce résultat peut « [...] s'expliquer par le fait que, pour écrire un mot, même régulier, il ne suffit pas de connaître les CGP (correspondances graphème-phonème), il faut en plus maîtriser l'orthographe canonique : alors que les mots *poteau* et *taureau* ne peuvent se lire que d'une seule façon, ils peuvent s'écrire de plusieurs manières, une seule étant correcte » (Sprenger-Charolles, 2012 : 119). Mais comment se déroule l'acquisition de l'écriture en français ? Pour mieux comprendre la spécificité du processus d'apprentissage pour le français, nous pouvons considérer le schéma de Ziegler et Goswami (§ 5.4). Nous présentons en détail les questions précédemment abordées de disponibilité, consistance et granularité pour le français.

- **Disponibilité dans la langue française**

La disponibilité est liée à la possibilité d'accéder aux unités pré-lexicales qui constituent le mot à partir de sa surface orale (du mot aux syllabes aux phonèmes) au moyen d'un travail métalinguistique.

Dans une langue comme le français, les liens entre phonèmes-syllabes-mots ne sont pas transparents et il est plus difficile d'accéder aux phonèmes. Il y a donc dans cette langue une faible disponibilité des liens entre les différents niveaux des unités phonologiques. Accéder aux unités de niveau inférieur (phonèmes) est donc plus long et plus difficile. Cette moindre disponibilité des unités plus petites a des conséquences sur l'apprentissage de la langue écrite : devenir alphabétisé en français demande plus de temps et comporte plus d'obstacles que dans des langues comme l'italien, par exemple, où la transparence est totale et où, par conséquent, les liens entre phonèmes, syllabes et mots sont directs.

- **Consistance du français**

La consistance fait référence au rapport entre graphèmes et phonèmes. Si, dans la langue, il y a une correspondance d'un à un entre le graphème et le phonème (un seul graphème peut être représenté par un seul phonème et vice-versa), la langue est considérée comme consistante. Si la correspondance entre graphèmes et phonèmes est plus complexe, la langue est plutôt considérée comme inconsistante.

Le français est une langue inconsistante. Nous pouvons dire que les correspondances graphèmes-phonèmes sont très régulières et faciles à prédire car plusieurs graphèmes sont associés à un seul phonème. Il s'agit par exemple des trois graphèmes <eau>, <au> et <o> qui représentent le seul phonème /o/ et donc une seule réalisation. En revanche, les correspondances phonème-graphème qui jouent un rôle dans l'écriture, sont moins intuitives. Par exemple, pour bien orthographier les mots *domino*, *bateau* et *landau*, il est nécessaire de connaître les normes orthographiques et de bien sélectionner la bonne représentation écrite de /o/ entre les trois orthographes. Pour cette raison, le passage entre l'oral et l'écrit est très compliqué.

- **Granularité en français**

Les enfants de langues maternelles différentes considèrent des unités de différentes tailles pour écrire (*grain*). Pour maîtriser la langue écrite, les enfants français doivent avoir comme référence la régularité de leur propre langue, ils utilisent donc comme unités dans ces langues les rimes. Malheureusement, le français, comme

l'anglais, l'allemand et le hollandais, est une langue où le voisinage entre les rimes est très dense, et où il y a donc beaucoup de mots qui se terminent par la même rime. Pour cette raison, les enfants doivent utiliser d'autres unités de différente taille pour pouvoir distinguer les mots, comme par exemple les syllabes (Goswami, 1993, 1999 ; Goswami & Bryant, 1990 ; Goswami & East, 2000). Comme nous l'avons déjà vu dans la section 5.4, l'utilisation des unités de grande taille implique le traitement d'un plus grand nombre d'unités et donc d'un processus d'acquisition de l'écrit plus lent. Nous traiterons dans la section 5.6.4 de la question du rôle de la syllabe dans le processus d'écriture.

5.6.2.1 La morphologie et l'écriture du français

La morphologie du français joue un rôle très important dans l'acquisition de l'écrit, car dans la plupart des cas les marques que nous trouvons à l'écrit n'ont pas de correspondance dans la langue orale (Catach, 1986 ; Dubois, 1965). La morphologie flexionnelle du pluriel en français est silencieuse : dans la flexion du nom (<s>), de l'adjectif (<s>) et du verbe (<nt>), plusieurs consonnes qui apparaissent à l'écrit ne sont pas prononcées à l'oral. Seulement 50 mots sur 25000 présentent une différence entre pluriel et singulier à l'oral (Toterou *et al.*, 1997), et ce sont des verbes dans la plupart de cas (ex. *a/ont, est/sont*). En revanche, à l'écrit, un haut pourcentage de mots présente une distinction de marque. Rares sont les mots qui ne présentent pas de marque à l'écrit pour indiquer la différence de nombre (ex. *nez*).

L'assymétrie entre oral et écrit, causée par la redondance du marquage à l'écrit, implique que les enfants qui doivent apprendre le marquage morphologique à l'écrit en français ne peuvent pas faire référence à l'oral (Chilounaki & Bryant, 2007) alors qu'ils doivent pourtant développer aussi une conscience morphologique (Colé *et al.*, 2004).

Plusieurs études ont cherché à comprendre le développement de cette conscience morphologique. Les résultats des recherches montrent que la morphologie du nombre est une des premières contraintes orthographiques que l'enfant arrive à maîtriser (Keeney & Wolfe, 1972 ; Nicolaci-da Costa & Harris, 1983). Thévenin *et al.* (1999) ont présenté une progression graduelle de l'apprentissage de l'écriture entre CP et CE1. Le premier morphème acquis à l'écrit est le <s> nominal ; le deuxième

morphème est le <s> adjectival et enfin il y a le <ent> des verbes. Les deux morphèmes en <s> sont tout de suite plus stables, même si des difficultés résiduelles subsistent avec l'adjectif post-nominalisé.

5.6.2.1.1 Le marquage du pluriel et la liaison

Comme nous l'avons dit, le français écrit a une forme silencieuse de marquage. Le seul cas pour lequel une consonne flexionnelle qui marque le pluriel est prononcée est dans le contexte de liaison. Dès lors, la production à l'oral d'une consonne qui marque le pluriel aide-t-elle son marquage à l'écrit ?

Une étude transversale de Mesquida *et al.* (2010), entre deux classes de l'école primaire, essaie de répondre à cette question. L'hypothèse de recherche s'appuie sur les travaux de Largy *et al.* (2004). Ceux-ci considèrent que si l'acquisition morphologique s'appuie sur la phonologie, la production de la CL qui matérialise la marque du pluriel va aider à comprendre la morphologie du pluriel du français écrit. Les tâches proposées étaient les suivantes : 19 enfants du CE1 et 26 enfants du CM2 devaient écrire dans un carnet une série de mots dictés dans deux contextes :

- déterminant + adjectif + nom, avec liaison et sans liaison,
- déterminant + nom, avec et sans liaison,

Les résultats montrent qu'au CE1, le pourcentage de <s> finaux est plus élevé en contexte de non liaison (32,6 %) que dans le contexte de liaison (25,6 %). Au CM2, c'est l'inverse qui émerge : la plupart des flexions au pluriel sont exprimées (77 % vs 70 % sans liaison). L'expérience de Mesquida *et al.* (2010) montre que la CL n'est pas considérée comme la marque flexionnelle du pluriel dans les premières productions écrites. La réalisation phonétique du <s> final du déterminant ou de l'adjectif ne constitue pas, au début de l'acquisition de l'orthographe, un soutien à l'acquisition du marquage du pluriel qu'ils maîtriseront complètement à l'issue de leur apprentissage.

5.6.2.2 Orthographier la liaison

Les enfants qui doivent écrire des séquences de deux mots en français, sont confrontés à une langue caractérisée à l'oral par la présence de phénomènes consonantiques que nous avons présentés dans le chapitre 3. Ces phénomènes qui

impliquent une resyllabation de Mot1 sur Mot2, masquent les frontières lexicales des mots et peuvent causer des difficultés quand les enfants apprennent l'orthographe. Par exemple, orthographier la CL <s> dans une séquence comme *les[z]ours* peut être une tâche difficile car il faut comprendre, qu'au niveau orthographique, elle correspond à la fin du Mot1 et qu'en revanche le Mot2, même si la CL est l'attaque de la première syllabe à l'oral, doit être traité comme un mot qui commence par une voyelle. La tâche est compliquée par la tendance, que nous avons signalée plus haut, qui consiste pour les enfants à appuyer leur acquisition orthographique sur une unité de type syllabique. On s'attend donc à ce qu'éventuellement, ils utilisent le constituant syllabique pour orthographier et donnent lieu à des orthographe fautive comme <les sours> ou <les zours> où la CL est orthographiée dans sa position phonologique.

La thèse de Soum (1997) analyse exactement les productions écrites des enfants pour vérifier le traitement des phénomènes de resyllabation. Soum se focalise sur les phénomènes d'enchaînement et de liaison en proposant une analyse des productions écrites des enfants de 7 ans, au CE1. En général, elle remarque des erreurs de groupement et de découpage phonographique dans les syntagmes qui ne présentent ni enchaînement, ni liaison (<un moteur da vion> pour *un moteur d'avion* ou <parcœur> pour *par cœur*). Pour ce qui concerne la liaison, les enfants présentent un nombre non négligeable d'erreurs d'omission de la CL qui se resyllabe sur le mot à droite (ex. <le amis> pour *les[z]amis*). La plupart des erreurs à l'écrit sont caractérisées par la présence d'une consonne qui va remplir l'attaque vide à droite (ex. <un lécureuil> pour *un[n]écureuil*) ou par la conservation de la CL simultanément dans le mot de gauche et le mot de droite (ex. <un gros sorage> pour *un gros[z]orage* ou <un nange> pour *un[n]ange*). La même problématique se manifeste pour l'enchaînement. Dans 50,6 % des cas, les productions écrites des enfants du CE1 montrent un problème avec la frontière gauche (ex. <leu enfant> pour *leur enfant*). Si les erreurs entre déterminant et nom semblent disparaître avec la pratique de l'écriture, Sabio dans son étude (2000) observe une persistance d'erreurs sur les frontières des mots, en contextes différents, dans le deuxième cycle de l'école primaire (<chez s'eu> pour *chez[z]eux* ou <ce qui se passent entre trelle> pour *entre elles*).

Soum-Favaro *et al.* (2014) ont proposé récemment les résultats d'une étude transversale qui tient compte de l'évolution des connaissances à l'écrit des enfants pendant l'école primaire, du CE1 (sept ans) au CM2 (dix ans). Il s'agit d'une comparaison entre les productions écrites de syntagmes nominaux en liaison avec [n],

[z] et [t] (<un avion>, <un gros aigle>, <un petit agneau>) et des syntagmes contenant des consonnes de même contenu segmental mais initiales de mot <n>, <z> et <t> (<un navire>, <un gros zèbre>, <un petit tableau>). Au CE1, les enfants montrent encore des difficultés à orthographier la CL. Dans beaucoup de productions écrites, les enfants gardent la CL au début du mot à droite (<un navion>). Les auteurs concluent qu'à ce stade-là, le contexte de liaison est encore le contexte qui pose le plus de problèmes. En revanche au CE2 (huit ans) et au CM1 (neuf ans), le pourcentage d'erreurs diminue sur le contexte de liaison pour augmenter sur le contexte à consonne initiale. Enfin, au CM2 (10 ans), les pourcentages vont s'inverser à nouveau : la plupart des erreurs sont encore enregistrées sur le contexte de liaison au lieu de se manifester sur le contexte de la consonne initiale. Les auteurs ont bien résumé ce pattern développemental : « Au moment où le nombre d'erreurs de segmentation diminue (dès le CE2), on observe une progression d'erreurs, non attendues, sur les mots à initiale consonantique consistant systématiquement à omettre la consonne initiale (<un navire> → <un avire>). Le pourcentage le plus important d'erreurs d'omission de la consonne initiale apparaît au CE2, moment où le pourcentage d'erreurs de segmentation baisse de façon considérable. » (p. 158-159). Ils expliquent ce comportement par une sur-généralisation : le contexte à consonne initiale est traité comme un contexte de liaison au CE2.

5.6.3 La lecture en français

La langue française est très difficile à lire comparée à d'autres langues plus consistantes comme l'italien ou l'espagnol, où les correspondances graphèmes-phonèmes sont plus immédiates. Des études récentes sur la dyslexie montrent que dans les premières années de l'école primaire, les performances en lecture des enfants francophones à développement typique sont comparables aux performances de lecture des enfants dyslexiques italiens (Sprenger-Charolles, 2013). À la fin du CP, les enfants italophones peuvent lire correctement une moyenne de 94 % des mots et 82 % de pseudo-mots (Cossu *et al.*, 1995). Ces résultats sont une conséquence du haut niveau de consistance de la langue. En revanche, les enfants francophones au même niveau scolaire peuvent seulement décoder 50 % des mots familiers (Goswami *et al.*, 2005).

De plus, après les premières années d'école, entre 6 et 8 ans, le *mapping* graphème-phonème reste toujours moins bon en français comparé à l'espagnol, langue très consistante (Goswami *et al.*, 1998). Le tableau (Figure 19) reflète ce qu'on a déjà expliqué : les meilleurs résultats en lecture sont enregistrés dans les langues les plus consistantes (grecs, finlandais, allemand, italien et espagnol), plutôt que dans les langues moins consistantes, en bas du tableau (anglais, danois, portugais et français).

Data (% Correct) From Seymour, Aro, and Erskine's (2003) Large-Scale Study of Reading Skills at the End of Grade 1 in 14 European Languages

Language	Familiar real words	Pseudowords
Greek	98	92
Finnish	98	95
German	98	94
Austrian German	97	92
Italian	95	89
Spanish	95	89
Swedish	95	88
Dutch	95	82
Icelandic	94	86
Norwegian	92	91
French	79	85
Portuguese	73	77
Danish	71	54
Scottish English	34	29

Note. From "Foundation Literacy Acquisition in European Orthographies," by P. H. K. Seymour, M. Aro, and J. M. Erskine, 2003, *British Journal of Psychology*, 94, pp. 153, 156. Copyright 2003 by the British Psychological Society. Reprinted with permission.

Figure 19. Lecture des mots et pseudo-mots dans plusieurs langues (Seymour *et al.*, 2003 par Ziegler & Goswami, 2005 : 11)

Cependant, comme pour les autres langues et pour le français, plusieurs études ont souligné l'importance des compétences phonologiques dans l'apprentissage de la lecture. Dans une étude longitudinale, Sprenger-Charolles (2012) montre que les enfants francophones pendant le CP se servent de la phonologie pour lire. L'utilisation de la phonologie est encore plus forte à la fin du CP qu'au début. Elle conclut que « les études longitudinales avec des enfants français indiquent aussi que l'apprentissage de la lecture est facilité par un recours précoce à la procédure phonologique de lecture, ainsi que par de bonnes capacités de segmentation phonémique. » (Sprenger-Charolles, 2012 : 20). L'importance de la phonologie a été confirmée par d'autres études

longitudinales qui montrent qu'en français, “phonological procedure seems to be the bootstrapping mechanism for reading acquisition.” (Sprenger-Charolles *et al.*, 2003 : 194).

Les enfants francophones commencent donc à lire et écrire en utilisant prioritairement la voie phonologique. Aucun effet de fréquence (et donc de la procédure lexicale) n'est enregistré au début de l'année scolaire dans ces études. Si on considère comme modèle de lecture à adopter le modèle de lecture à double voie (§ 5.5.1.1), l'absence d'un effet de fréquence implique la non-utilisation de la voie lexicale. Or, c'est seulement à la fin du CP qu'on constate un possible effet de fréquence dans la tâche de lecture (Sprenger-Charolles *et al.*, 1997). Cela ne veut pas dire que les enfants au CP utilisent seulement la voie phonologique mais plutôt qu'ils l'utilisent plus précocement et afin de construire leur lexique orthographique. La médiation phonologique aide à la constitution du lexique orthographique et la voie lexicale commence à être utilisée en même temps que la voie phonologique. À partir de la fin du CE1, la voie lexicale devient de plus en plus utilisée en lecture (Sprenger-Charolles *et al.*, 2003)

Dans les prochains paragraphes, nous présentons des études qui se concentrent sur la lecture d'un des phénomènes de resyllabation au centre de notre recherche : la liaison.

5.6.3.1 La lecture de séquences en liaison

Nous n'avons pas beaucoup d'études sur la lecture de séquences resyllabifiées. La seule étude à notre connaissance est l'étude très récente de Dugua et Baclesse (2014). Elles envisagent de donner un panorama sur la lecture de séquences en contexte de liaison chez les enfants de CE1 et CM1 (âge moyen 9;2). Au moyen de trois différentes tâches, les deux chercheurs vérifient les compétences de liaisons obligatoires et facultatives chez 110 enfants francophones. Les trois tâches proposées sont : une tâche de dénomination guidée, une tâche de jugement normatif et une tâche de lecture à voix haute. Nous présentons seulement les résultats pour la liaison obligatoire-catégorique et pour la tâche de lecture à voix haute, qui sont plus relatifs à notre recherche.

Dans la tâche de lecture à voix haute, 96,5% des liaisons obligatoires sont bien réalisées. En revanche, Dugua et Baclesse enregistrent encore des erreurs d'omissions (3,1%) et de CL erronée (0,4 %). L'explication qu'ils donnent est que « du fait de sa situation à la frontière de deux mots, même si la liaison est systématique à l'oral, elle nécessite un traitement particulier en lecture chez des enfants » (p.136).

5.6.4 La syllabe comme unité de base dans les processus de lecture et d'écriture chez les enfants français

Plusieurs auteurs ont souligné l'importance de la syllabe dans l'acquisition de la langue orale, comme on l'a vu dans le chapitre 4. Cette unité est également d'importance fondamentale dans l'acquisition de la lecture et de l'écriture.

Avant d'apprendre à écrire à l'école primaire, les enfants à l'école maternelle produisent des séquences de lettres qui sont appelées « productions écrites inventées » car elles ne sont pas des graphies conventionnelles comme nous pouvons le penser. Une analyse des productions écrites inventées chez les enfants a été faite par Creuzet et Pasa (2009) au moyen d'une tâche d'écriture. Plus de cent enfants francophones de grande section de maternelle ont été testés. Les résultats de la tâche montrent que les enfants représentent chaque syllabe du mot avec une lettre à l'écrit, et confirment que l'unité syllabique joue un rôle majeur dans la production écrite pré-scolaire.

Plusieurs travaux concernant les productions écrites chez les adultes (Kandel *et al.*, 2006) et chez les enfants (Kandel & Valdois, 2006) ont mis en évidence que la syllabe peut être considérée comme une des unités de base de la *literacy*. Les études qui ont abouti à cette conclusion montrent que le temps de traitement en perception (lecture) et en production (écriture) est conditionné par l'appartenance ou non des unités produites à une même syllabe. En effet, il a été montré, par exemple, que passer d'une lettre à la suivante (ex. de <a> à <l> dans le digramme <al>) demande plus de temps quand les deux lettres appartiennent à deux syllabes différentes (ex. <al> dans *ba.leine*). De même, le temps de production à l'écrit d'un digramme est plus court si les deux lettres appartiennent à la même syllabe (ex. <al> dans *bal.con*).

Les erreurs en lecture montrent également le respect de cette unité, surtout dans sa forme plus basique, CV. Bastien-Toniazzo *et al.* (1999) montrent que la plupart des erreurs de lecture au CP sont des erreurs qui respectent la structure syllabique CV. Par

exemple, des erreurs d'inversion (*nu.tu.po* pour *an.tu.po*) ou rajout (*ba.ra.mi.de* pour *bar.mi.de*) ou suppression d'une lettre (*ba.mi.de* pour *bar.mi.de*). Ces erreurs montrent la tendance à simplifier toutes les syllabes pour les rendre plus proches de la structure CV et reflètent les erreurs phonologiques des enfants français dans l'acquisition de la langue orale (Wauquier & Yamaguchi, 2013).

Ensuite, les chercheurs ont montré que la syllabe joue un rôle dans la lecture au moyen des expériences qui mettent en rapport un stimulus et une cible.

La première expérience que nous présentons a été faite par Colé *et al.* (1999). Un groupe d'enfants du CP et un groupe d'adultes doivent détecter des segments CV (ex. *pa*) ou CVC (ex. *par*) dans des mots écrits en français qui ont comme première syllabe CV (ex. *parole*) ou CVC (ex. *pardon*). Les mots utilisés dans la tâche ont des niveaux de familiarité différents. Pour les enfants et pour tous les mots, il est plus facile de détecter la syllabe CV si les mots commencent par CV plutôt que par CVC. Ils ont donc enregistré une facilitation conséquente à ce qui est appelé « congruence syllabique » ou « compatibilité syllabique » entre la séquence à détecter et le stimulus. En revanche, les adultes ne présentent cette facilitation que pour les mots peu familiers. Ces résultats montrent la mise en place de stratégies différentes chez l'adulte et chez l'enfant au CP. Les adultes utilisent une stratégie phonologique (basée sur le traitement de la syllabe) seulement pour les mots peu fréquents, tandis que les enfants l'utilisent pour lire tous les mots. Chez les enfants, la procédure d'identification des mots n'est pas encore automatique et ils ont encore besoin de mettre en place un traitement syllabique pour tous les mots qu'ils rencontrent. C'est seulement au cours du CM2, lorsque le niveau de lecture s'améliore, que les enfants montrent des résultats plus proches de ceux des adultes (Colé, 1996 ; Colé & Sprenger-Charolles, 1999).

Pour confirmer l'effet de « congruence syllabique » chez les enfants face à l'écrit, Doignon et Zagar (2006) utilisent le paradigme de conjonction illusoire. Dans leur tâche, les mots écrits sont présentés en deux couleurs qui correspondent ou non aux frontières des syllabes (ex. <BAr.bu> avec <BA> en rouge et <r.bu> en bleu vs. <BAR.bu> où <BAR> est rouge et <bu> est bleu). La couleur rend saillante une certaine partie du mot qui peut influencer sa lecture. Après la présentation du mot écrit, une lettre est présentée brièvement sur l'écran (par exemple le <r> après la présentation de l'input <BAr.bu>). Les enfants doivent dire la couleur de cette lettre. Les résultats montrent qu'il y a plus d'erreurs dans l'identification de la couleur de la

lettre quand la lettre est colorée différemment de la syllabe d'appartenance (comme dans le cas de <BAr.bu>). Il est plus facile de deviner la couleur de la lettre si elle est colorée de la même couleur que la syllabe d'appartenance (comme dans <BAR.bu>).

Chetail & Mathey (2012), au moyen d'une tâche d'amorçage avec pseudo-mots, ont confirmé que l'accès au lexique à partir de l'écrit a lieu grâce à la syllabe. Ils enregistrent un effet d'amorçage dans le cas de congruence syllabique entre amorce et cible (quand l'amorce est congruente avec la cible par exemple dans le couple **vo.liar-VO.LUME**). En revanche, il n'y a pas l'effet en présence d'une non-congruence syllabique (ex. **vo.lode-VOL.CAN**).

La syllabe semble bien constituer une unité qui intervient dans le processus d'écriture et dans le processus de reconnaissance des mots écrits mais elle apparaît aussi comme une unité au moyen de laquelle a lieu l'accès au lexique. Les chercheurs ont donc tenté de l'intégrer dans le modèle cognitif. Nous présentons une intégration de l'unité syllabique dans le modèle IAS.

5.6.4.1 Le modèle IAS : les unités syllabiques dans le modèle de lecture

Étant donné que les unités syllabiques sont activées dans la reconnaissance des mots écrits et vu qu'elles jouent un rôle en lecture, Mathey *et al.* (2006) proposent l'intégration de ces unités dans les modèles connexionnistes à activation en parallèle et à activation interactive⁹ AI (McClelland & Rumelhart, 1981). C'est de là qu'a été élaboré le modèle IAS (Figure 20).

Dans ce modèle, qui utilise les représentations locales¹⁰, est intégré un autre niveau appelé sous-lexical qui met en place la représentation des syllabes (*syllable units* dans le schéma). Celui-ci est connecté au niveau d'activation des lettres et des mots. L'activation des syllabes est sensible à leur fréquence dans la langue : cela veut dire que des syllabes plus fréquentes sont activées plus vite que les moins fréquentes. La lecture prévoit deux voies possibles : une voie orthographique (*orthographic route*) caractérisée par des connexions directes lettres-mots et une deuxième voie

9 Le traitement du mot est influencé par des informations de bas niveau et par des informations de haut niveau. Il existe deux types de connexions : excitatrices et inhibitrices. Les connexions excitatrices ont lieu globalement entre les différents niveaux alors que les connexions inhibitrices sont activées à l'intérieur des niveaux.

10 Chaque trait visuel, acoustique, chaque lettre et phonème et aussi chaque mot sont représentés par des unités distinctes

phonologique (*phonological route*) où interviennent les unités syllabiques. Les connexions ont lieu entre trois éléments : lettres, syllabes et mots. Dans le niveau « lettres » il y a une connexion entre les unités « lettres » en positions adjacentes. De cette façon, plus le digramme est fréquent, plus l'activation entre les deux lettres est forte. Ainsi, ce modèle tient compte de l'effet en lecture du voisinage orthographique et syllabique.

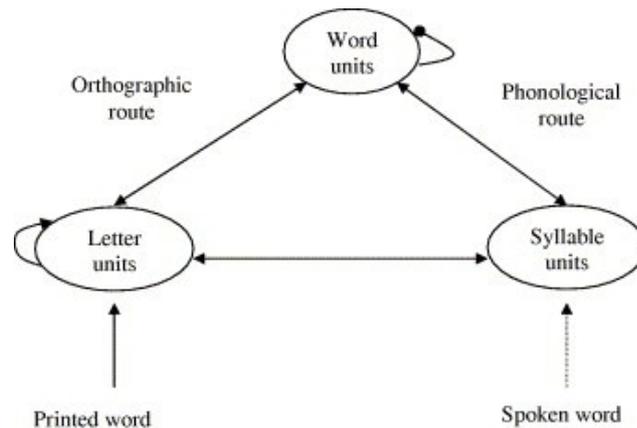


Figure 20. Le modèle IAS (Mathey *et al.*, 2006)

Dans cette configuration, la syllabe est à la base de deux processus : un processus de facilitation et un processus d'inhibition.

- Le processus de facilitation est sous-lexical et il résulte des congruences de structures syllabiques. L'effet de facilitation de la syllabe explique les résultats décrits précédemment (§ 5.6.4).
- Le processus d'inhibition a lieu au niveau lexical pour les connexions inhibitrices entre le mot qui doit être lu et les voisins syllabiques. Ce processus d'inhibition est lié à la fréquence syllabique (Mathey *et al.*, 2006) : plus la première syllabe d'un mot est fréquente, plus elle active ses voisines syllabiques et plus les temps de réponse dans des tâches de décision lexicale est long (Hutzler *et al.*, 2004).

Le modèle met aussi en rapport les effets de facilitation et d'inhibition : si l'effet de facilitation est plus important que l'effet d'inhibition, on s'attend à ce que la lecture soit plus rapide. En revanche, si l'effet d'inhibition est supérieur à l'effet de facilitation, la lecture sera plus lente.

Comme nous l'avons vu dans la section 5.5.1.3, les modèles connexionnistes simulent très bien la lecture de mots et de pseudo-mots chez les adultes mais ils avaient des limites en ce qui concerne la représentation des processus d'acquisition de la lecture. Le modèle IAS est un progrès vers l'intégration des composantes qui caractérisent le développement de la *literacy* dans les modèles cognitifs connexionnistes.

5.6.4.2 Le rôle de la syllabe change selon le niveau de lecture des enfants

L'étude de Chetail et Mathey (2009) conduite sur des enfants de CE1, montre qu'il y a une activation syllabique dans la reconnaissance des mots écrits en français mais que cette activation dépend du niveau de lecture de ces enfants.

Dans leur étude, 45 enfants sont testés au moyen d'une tâche de décision lexicale colorée. La procédure prévoit la présentation d'une série de mots sur l'écran. Ces mots sont colorés selon deux modalités différentes :

1. **congruence syllabique** : la première syllabe est colorée d'une couleur différente (ex. Carotte) ;
2. **non congruence syllabique** : la première partie du mot est colorée avec une couleur différente mais cette partie ne respecte pas la frontière syllabique (ex. CARotte).

On doit remarquer que quand une partie du mot est colorée avec une autre couleur, elle devient saillante, et cette importance peut produire des effets sur la lecture du mot et donc sur la décision lexicale.

Les syllabes testées sont CV et CVC. Les enfants ont été classés selon leurs différents niveaux de lecture et d'écriture.

Les résultats montrent que l'effet de congruence syllabique change selon le niveau de lecture. Les moins bons lecteurs reconnaissent le mot plus vite quand la couleur de sa première partie respecte les frontières de la syllabe (situation de congruence, ex. CARotte) plutôt que quand elle ne la respecte pas (situation de non congruence, ex. CARotte). Ainsi, les moins bons lecteurs sont aidés dans la lecture par la syllabe, devenue saillante à cause de la couleur différente. Ce comportement est expliqué facilement dans le modèle de lecture IAS : la syllabe colorée et saillante

permet la forte activation des unités sous-lexicales syllabiques correspondantes, qui d'une fois à l'autre activent le mot. La syllabe a, chez les moins bons lecteur, un effet de facilitation sous-lexicale dans le processus cognitif de la lecture.

En revanche chez les meilleurs lecteurs, quand la couleur rend la syllabe plus saillante, la reconnaissance des mots est plus lente. Ceci est expliqué par la compétition entre les voisines syllabiques activées au niveau lexical. La syllabe colorée active les unités syllabiques sous-lexicales correspondantes au niveau phonologique, mais simultanément, elle active au niveau lexical les voisines qui partagent la même syllabe, causant l'inhibition et allongeant les temps de réponses. Chez les meilleurs lecteurs, la syllabe a un effet d'inhibition lexicale dans le processus cognitif de la lecture.

5.6.5 Apprentissage de la *literacy* à l'école en France : la méthode globale et la méthode alphabétique

Comme nous l'avons vu, pour apprendre à lire dans les différentes langues, les enfants utilisent différentes stratégies qui s'appuient sur des unités de différentes tailles (§ 5.4). Dans l'enseignement de l'écriture-lecture de différentes langues, il est donc nécessaire d'adopter différentes modalités selon les unités que les enfants semblent considérer comme focus dans la procédure d'apprentissage. Le français est une langue inconsistante, il y a donc la nécessité de se focaliser sur des unités plus grandes que les phonèmes.

En France, deux méthodes d'apprentissage sont utilisées à l'école :

- **La méthode alphabétique :**

Elle prévoit l'enseignement progressif de correspondances entre lettres et sons du langage. Si nous considérons encore le modèle à double voie (§ 5.5.1.1), nous pouvons dire que cette méthode prévoit le développement de la procédure d'assemblage. Cette méthode marche très bien pour les écritures alphabétiques (comme l'italien ou le français).

- **La méthode globale :**

Dans cette méthode, l'apprentissage est basé sur la mémorisation de formes écrites du mot. L'accent est mis sur la compréhension et sur la communication, mais il n'y a pas un enseignement systématique sons-graphèmes. Si nous considérons encore une fois le modèle à double voie (§ 5.5.1.1), nous pouvons

dire que cette méthode prévoit le développement de la procédure d'adressage et du lexique orthographique. Cette méthode donne les meilleurs résultats dans les langues logographiques (comme le chinois).

Mais quelles sont les influences spécifiques de ces méthodes sur l'apprentissage de la lecture-écriture du français ? L'étude de Content et Leybaert (1992) nous donne des réponses. Les participants sont des enfants francophones fréquentant 2ème, 4ème et 6ème années de l'école divisés en deux groupes correspondant respectivement à la méthode utilisée par les enseignants (globale vs. alphabétique). Les tâches jugent la reconnaissance de mots, la segmentation et la compréhension de phrases. L'analyse transversale des données montre que la méthode globale donne de meilleurs résultats à long terme. Elle est susceptible de consolider les connexions dans la procédure d'adressage. Cependant les enfants qui ont appris la langue écrite par la méthode globale ont plus de problèmes à comprendre les différences entre les mots rares et les pseudo-mots homophones. Puisque cette méthode n'aide pas l'assemblage, comme peut le faire la méthode phonologique, les enfants de 4ème et 6ème années ont toujours des difficultés dans les tâches phonologiques, comme la suppression de consonnes. Les enfants de la méthode alphabétique montrent de meilleures performances dans la 2ème année surtout sur l'écriture des mots réguliers mais également la création de néologismes qui reflètent un assemblage parfait.

À la lumière de ces résultats, nous ne pouvons pas dire qu'il y a une meilleure méthode qui doit être privilégiée dans l'enseignement du français écrit. Utiliser les deux méthodes simultanément est la meilleure solution car elle permet aux enfants de développer en parallèle deux compétences fondamentales (adressage et assemblage).

5.7 Le rapport entre oral et écrit

Nous avons présenté dans les chapitres précédents les caractéristiques de la langue française, à l'écrit et à l'oral. Il ressort de l'étude de la littérature existante que le français peut être considéré comme :

- une langue où, à l'oral, définir les frontières des mots est une tâche difficile à cause de ses caractéristiques prosodiques (chapitres 3 et 4) ;

- une langue où, à l'oral, les frontières de syllabes masquent les frontières de mots à cause des phénomènes de resyllabation très fréquents de liaison, d'élision et d'enchaînement (chapitre 3) ;
- une langue inconsistante où les rapports entre graphèmes et phonèmes ne sont pas biunivoques. Par exemple, le phonème /o/ peut être représenté par trois graphies différentes <eau>, <au> ou <o> (chapitre 5).
- une langue où, à l'écrit, accéder aux unités de niveau inférieur (phonèmes) à partir d'unités supérieures est très difficile, car les unités de niveau inférieur sont d'accès difficile à partir de la surface. L'accès à l'écrit a donc lieu au moyen d'unités assez grandes comme la rime ou la syllabe (chapitre 5) ;
- une langue très difficile à acquérir à l'écrit à cause des caractéristiques que nous venons de présenter et également à cause de la présence des phénomènes consonantiques de liaison, d'élision et d'enchaînement qui créent des problèmes dans l'acquisition de l'orthographe (chapitre 5) ;

À l'oral comme à l'écrit, les phénomènes consonantiques causent beaucoup de problèmes chez les enfants et la question que l'on peut alors se poser est de savoir si apprendre à écrire les formes lexicales aide les performances orales dans les cas de liaison, d'élision et d'enchaînement. Cette question en soulève une autre bien plus fondamentale : un lien entre la connaissance de l'orthographe et les compétences phonologiques est-il possible ? L'orthographe peut-elle influencer la représentation phonologique et la changer ? Le rapport entre la phonologie et la *literacy* commence très tôt (chapitre 2) et guide les enfants dans l'acquisition de l'orthographe à l'école primaire. Nous cherchons à vérifier les relations entre l'apprentissage de l'écriture et de la lecture et l'identification des frontières de mots dans la production orale pour la liaison, l'élision et l'enchaînement. De plus, nous pouvons essayer de comprendre comment la connaissance des frontières orthographiques modifie et développe la représentation phonologique des mots. Ces relations sont au centre des hypothèses de cette thèse, développées dans le prochain chapitre.

Chapitre 6

Questions de recherche et hypothèses

6.1 Introduction

Dans cette thèse, nous voulons :

- récolter de nouvelles données sur la langue orale et sur la langue écrite
- évaluer l'évolution de ces données en accord avec le niveau de scolarisation des enfants, en proposant une étude longitudinale.

Au moyen de certaines tâches orales et écrites, nous testerons les compétences de 43 enfants français à partir du cours préparatoire (CP) jusqu'au cours élémentaire de 1^{re} année (CE1).

6.2 Apport empirique et nouvelles données

6.2.1 Des nouvelles données sur la langue orale

Comme nous l'avons vu dans le chapitre 4, nous disposons de nombreuses données sur l'acquisition des phénomènes de resyllabation à l'oral. Cependant la plupart de ces données prennent en considération les productions des enfants jusqu'à 5 ou 6 ans (§ 4.2.3). Nous avons donc décidé de nous occuper de la tranche d'âge qui suit, en analysant les productions orales après 6 ans. Nous cherchons à voir si, comme

proposé par certains chercheurs, les phénomènes de resyllabation sont acquis après 6 ans ou si, comme évoqué par certains autres, ils posent encore problème.

Nous avons vu dans le chapitre 3 que les productions de liaison peuvent être influencées par le niveau socioculturel des parents (§ 4.2.3). Dans l'analyse descriptive des données récoltées, nous prenons en considération la variable sociolinguistique. Nous cherchons à voir si les productions de liaison des enfants testés sont liées à cette variable.

6.2.2 Des nouvelles données sur la langue écrite

Jusqu'à présent, les études qui traitent des phénomènes de resyllabation à l'écrit prennent en considération les compétences des enfants à partir du CE1 (§ 5.6.2.2). Cette étude cherche à décrire les processus d'écriture et de lecture des phénomènes consonantiques dans une étape d'apprentissage de la *literacy* encore inexplorée, le CP. Elle cherche également à définir quelles méthodes pédagogiques doivent être mises en place pour réduire les erreurs et les difficultés dans la lecture et dans l'écriture des séquences en liaison, en élision et en enchaînement, en analysant les stratégies adoptées par les enfants au CP et au CE1.

6.2.3 La variable « lecture partagée »

Dans la description des données orales et écrites, nous avons aussi considéré la variable « lecture partagée ». Comme nous l'avons vu dans le chapitre 5, les enfants qui partagent plus de moments de lecture avec leurs parents sont des enfants qui ont de meilleures compétences de la langue écrite. Nous évaluerons l'influence de la fréquence des moments de lecture partagée sur les performances de lecture, d'écriture et de langue orale de nos participants pour comprendre son rôle au début du parcours scolaire.

6.3 Répondre aux questions de recherche et vérifier les hypothèses

Dans cette thèse nous examinerons les questions de recherche suivantes.

◆ **Question 1 (Q1) : quel modèle de l'acquisition de la liaison décrit le mieux le découpage en mots après 5 ans ?**

Nous avons vu que, pour décrire l'acquisition des frontières des mots en français, deux modèles ont été proposés (§ 4.2.3.1) :

- 1) le modèle lexical-constructionniste qui souligne l'importance de la fréquence des formes de surface dans le lexique pour l'acquisition de leurs frontières ;
- 2) le modèle morpho-phonologique qui met en rapport l'acquisition des frontières de mot avec des compétences phonologiques et des représentations sous-jacentes.

Nous essayons dans cette thèse de bien comprendre lequel de ces deux modèles de l'acquisition de la liaison et, de façon plus générale, des frontières de mot, est le plus adapté pour décrire ce stade.

Pour vérifier les influences des deux modèles, nous mettons en rapport les productions orales des enfants de 6 à 7 ans avec la fréquence des mots dans leur lexique. De plus, nous analyserons les productions orales des enfants pour évaluer l'existence d'une forme sous-jacente et d'une possible représentation phonologique, mais également pour comprendre le traitement des frontières des Mots1 et Mots2 à l'oral.

◆ **Question 2 (Q2) : quelle est l'influence de l'apprentissage de l'orthographe sur la production orale et sur le découpage en mots ?**

Quand les enfants apprennent à écrire, les séquences entendues à l'oral prennent forme sur papier. À l'écrit, les frontières des mots sont bien délimitées par des symboles graphiques, comme les espaces blancs (ex *les ours*). De plus, dans les séquences en liaison, la CL est pleinement représentée en tant que graphème attaché au mot à gauche (Mot1). Nous avons vu que l'apprentissage de l'écriture d'une langue influence les compétences orales de cette langue (chapitre 2). L'apprentissage de l'orthographe chez les enfants français au début de l'école

primaire peut donc être vu comme un facteur qui facilite la segmentation en mots à l'oral, car elle permet de fixer sur papier les frontières lexicales de ces mots.

Pour vérifier si connaître les formes écrites a un effet sur la langue orale, nous proposons une comparaison entre les résultats des tâches orales et des tâches liées à la langue écrite (lecture et écriture) pendant les premières années de l'école primaire. Nous nous attendons à trouver une influence des compétences de la langue écrite sur la production orale. Selon nos hypothèses, cette influence devrait se manifester sous 3 formes :

- **Hypothèse 1 (H1) : la disparition des erreurs de remplacement ou des consonnes erronées**

En ce qui concerne la liaison, nous avons vu que les enfants semblent avoir acquis les frontières des mots à partir de 5 ans, car une diminution des erreurs de remplacement ou de consonne erronée est enregistrée (§ 4.2.3).

Nous nous attendons à une nouvelle diminution de ces erreurs de consonne erronée pendant les deux premières années de l'école primaire, grâce à l'apprentissage de l'écriture.

- **Hypothèse 2 (H2) : la stabilisation des non-réalisations**

La littérature sur l'acquisition de la liaison montre une stabilité des non-réalisations à l'oral après 5 ans (§ 4.2.3). La stabilité des non-réalisations peut être vue comme une influence défavorable de l'écriture sur la production orale. Apprendre les mots en tant que formes visuelles séparées peut causer leur traitement à l'oral puisqu'il faut prononcer deux formes séparément et donc sans la réalisation d'un phénomène de resyllabation. Cela a pour conséquence une diminution des enchaînements entre Mot1 et Mot2 en faveur des non-réalisations. Pour vérifier l'influence de l'écriture sur la stabilité des non-réalisations, nous comparerons les résultats des tâches liées à la langue écrite et à la langue orale afin de vérifier s'il y a une corrélation.

Nous nous attendons à une stabilisation des non-réalisations pendant les deux premières années de l'école primaire à cause de l'apprentissage des formes écrites.

○ **Hypothèse 3 (H3) : une relation entre les connaissances morphologiques à l'écrit et les productions orales**

Nous avons montré dans les chapitres précédents que le français a une forme de marquage à l'écrit pour les éléments morphologiques pourtant silencieux à l'oral (§ 5.6.2.1). Cependant dans le cas de liaison en [z], la consonne flexionnelle qui marque le pluriel, est réalisée à l'oral. L'acquisition du marquage du pluriel à l'écrit permet-elle de mieux produire les séquences en liaison ? Nous vérifierons le rapport entre le marquage du pluriel à l'écrit et la production de liaison au pluriel en mettant en relation les résultats des tâches orales et des tâches liées à la langue écrite chez les enfants au début de l'école primaire.

Contrairement à ce que les chercheurs ont montré (§ 5.6.2.1), nous nous attendons à une amélioration du marquage du pluriel en parallèle avec les progrès dus à l'apprentissage de l'orthographe avec une influence positive de l'apprentissage du marquage du pluriel à l'écrit sur les réalisations de liaison en [z].

Le traitement de ces questions de recherche et la vérification de ces hypothèses, nous conduira à la création d'un modèle de l'acquisition des frontières de mot qui tient compte de l'introduction de la *literacy* dans les compétences des enfants.

◆ **Question 3 (Q3) : quel processus cognitif derrière le découpage en mots ?**

Nous cherchons à voir quels processus cognitifs sous-tendent le découpage en mots chez les enfants. Nous considérons que ce processus est différent de celui qui caractérise le découpage en mots chez l'adulte, car les enfants qui apprennent la langue écrite n'ont pas la même compétence de *literacy* qu'un adulte francophone scolarisé et diplômé.

Pour faire émerger cette différence, nous avons testé un groupe d'adultes francophones scolarisés en France qui utilisent la langue écrite tous les jours dans leur activité professionnelle. Nous nous attendons à remarquer une utilisation de l'unité syllabique dans le découpage en mots chez les enfants qui n'est pas présent chez les adultes.

○ **Hypothèse 4 (H4) : un coût pour le traitement cognitif de l'input en liaison chez les enfants**

Comme nous l'avons vu dans le chapitre 4, les résultats des recherches chez les adultes montrent que segmenter un input en liaison n'apporte pas un coût dans le traitement cognitif (§ 4.3.2). Nous pensons que cette absence de coût est une conséquence du bon niveau en segmentation des adultes et du fait qu'ils ont une connaissance de la langue écrite qui permet un stockage des mots à segmenter dans leur lexique mental. Comme nous l'avons vu dans le chapitre 2, quand les adultes entendent une séquence à l'oral, la représentation écrite est activée. Quand les adultes entendent une séquence en liaison à l'oral, ils activent les représentations orthographiques de Mot1 et Mot2, où Mot2 commence par une voyelle et où les frontières ne sont pas perturbées par la resyllabation. Grâce à cette activation, repérer un mot dans un input aligné (*un joli éléphant*) et dans un input en liaison (*un petit[t]éléphant*) ne fait pas de différence car dans les deux cas c'est la forme orthographique <éléphant> qui influence la segmentation.

Nous pouvons penser qu'au CE1, chez les enfants qui n'ont pas encore une connaissance complète de la langue écrite et qui ont encore des problèmes de segmentation en mots, le découpage d'un input oral en liaison peut avoir des coûts cognitifs.

Pour vérifier notre hypothèse, nous proposerons deux tâches d'amorçage.

À la différence des adultes, nous nous attendons au CE1 à des coûts dans le traitement cognitif des séquences resyllabifiées.

○ **Hypothèse 5 (H5) : le rôle de la syllabe entre écrit et oral chez les enfants**

Comme nous l'avons vu dans le chapitre 4, la syllabe est une unité importante qui guide la segmentation quelle que soit la langue. Ceci est particulièrement vrai en français. En effet, nous avons vu dans le chapitre 5 que la syllabe est aussi utilisée par les enfants français comme unité de reconnaissance des mots écrits au début de l'apprentissage de la *literacy* (§ 5.6.4). Pour mettre en rapport oral et écrit et vérifier le rôle de la syllabe dans le découpage en mots en français au début de l'école primaire, nous utilisons les données d'une des deux tâches d'amorçage. Puisqu'à l'oral et à

l'écrit la syllabe est une des unités les plus importantes pour le découpage et que, dans la liaison, les frontières syllabiques masquent les frontières lexicales, nous nous attendons à remarquer que le coût dans le traitement cognitif enregistré (H4) soit causé par l'utilisation de la syllabe dans le découpage. Contrairement aux adultes, les enfants à l'école primaire qui doivent repérer les Mot2 dans la séquence *le petit[t]éléphant* (en liaison) ou dans la séquence *le gentil éléphant* (alignée) n'activent pas tout de suite l'orthographe de <éléphant>. Ils utilisent encore l'unité syllabique pour le découpage et cette utilisation est responsable d'un coût cognitif pour l'input resyllabifié (le petit[t]éléphant).

Nous chercherons à répondre aux questions de recherche et à confirmer ou infirmer les hypothèses au cours de cette thèse. Avant de passer aux résultats de l'étude longitudinale que nous avons conduite, nous présenterons la méthodologie utilisée dans le chapitre 7.

Partie 2.
La méthodologie utilisée et les modalités de traitement des données

Chapitre 7

Méthodologie et outils expérimentaux

7.1 Les participants

7.1.1 Les enfants

Nous avons contacté une école privée de Dunkerque, dans la région du Nord, où la directrice a accepté que nous testions des enfants dans deux classes de CP pendant deux années scolaires (2013-2014 et 2014-2015). Les cours dans les deux classes sont assurés par deux maîtresses différentes. Cependant, les objectifs pédagogiques (les programmes, les méthodes, les tâches et le suivi des acquis attendus chez les enfants) sont communs, grâce à la communication quotidienne entre les deux maîtresses. De plus, toutes deux utilisent une méthode mixte en y intégrant les composantes des deux méthodes traditionnelles. Elles utilisent simultanément la méthode globale, où les enfants sont conduits à apprendre les mots dans la forme lexicale précédée par leur article (§ 5.6.5) et une méthode alphabétique-phonologique (§ 5.6.5), où les enfants sont incités à faire des liaisons entre les graphèmes et les phonèmes et à travailler sur la conscience méta-phonologique, par exemple à reconnaître le même son dans différents mots. Les enfants sont encouragés à connecter graphème-phonème et mots par des tableaux accrochés sur leur table (Figure 21) et sur les murs de la classe (Figure 22). De plus, ils se servent d'un outil appelé « cahier phonologique » où ils regroupent les mots écrits qui ont le même son initial. Par exemple, toutes les fois qu'ils retrouvent un mot qui commence avec le son [ʃ], ils doivent l'écrire dans la page dédiée au son [ʃ] où, à la fin, nous retrouvons une liste de formes orthographiques (ex. *château, chapeau, choix*, etc...).



Figure 21. Outils pour les connexions graphèmes-phonèmes sur les tables des élèves



Figure 22. Tableaux dans les classes pour l'apprentissage de l'écriture

Nous avons testé 43 enfants appartenant à deux classes du CP, que l'on appellera CP A et CP B, de respectivement 21 et 22 élèves. Sur les 43 enfants, 26 sont des filles et 17 sont des garçons. Les âges des enfants vont de 5 ans et 8 mois à 7 ans et 4 mois. L'âge moyen du groupe est de 73,4 mois, soit 6 ans et 2 mois. Les enfants ont aussi été testés au CE1. Au CE1, nous comptons seulement 37 enfants pour 43 au départ. Nous donnons dans le tableau 5 les détails par classe. Tous les enfants parlent français, leur langue maternelle, et leurs parents ont autorisé le traitement et l'utilisation scientifique des données. Les informations concernant le contrôle éthique de la prise de données et de leur utilisation sont en annexe 1.

	CPA	CP B	CE1 A	CE1 B
Nombre d'élèves	21	22	20	17
Filles	13	12	12	11
Garçons	7	10	8	6
Âge moyen (ans;mois)	6;2	6;1	7;5	7;7

Tableau 5. Description des classes des participants au CP et au CE1

7.1.1.1 Les compétences scolaires

À l'entrée à l'école primaire, les maîtresses proposent des tâches préliminaires standard qui permettent une évaluation globale. Cette évaluation globale est identique dans toutes les classes de CP car elle est prévue par les directives du ministère de l'Éducation Nationale. Les épreuves sont décrites dans deux livrets de passation : un premier livret qui teste la maîtrise des mathématiques et de la logique ; un deuxième livret qui teste la maîtrise de la langue. Nous détaillons ci-dessous le second qui intéresse nos objets d'étude.

Le livret de passation « maîtrise de la langue » est composé de 23 exercices. Ces exercices sont présentés aux enfants sur 4 jours, trois semaines après la rentrée scolaire, en septembre 2013.

Les exercices proposés sont destinés à évaluer plusieurs compétences des enfants au début de l'école primaire : la connaissance des composantes sonores et visuelles des mots, du vocabulaire technique de l'apprentissage (comme par exemple *mot, chiffres, couverture, titre*), des compétences sémantiques et de l'écriture. Nous présentons dans le tableau 6 les exercices avec les sigles du livret de passation et leurs buts.

Évaluation des compétences des composantes sonores et visuelles		
Exercice	Sigle ¹¹	But
1	NB SYL	dénombrer les syllabes d'un mot (avec un répertoire de 5 noms <i>mouton, éléphant, chat, caméléon, hérisson</i>)
2	LOC SYL	localiser une syllabe dans un mot
3	ISOL PHON	isoler un phonème dans un mot
4	DIS MEM MOT	discriminer et mémoriser un mot
5	ORAL EC	établir des relations entre la quantité d'oral et la quantité d'écrit
6	DV	discrimination visuelle (trouver le même groupe graphique de lettres à partir d'un exemple donné)
7	DIS MEM MOT 2	discriminer et mémoriser un mot ¹²
8	REC MOT	reconnaître des mots courants

¹¹ Nous donnons les sigles des exercices selon la notation donnée par l'école en phase de correction

¹² Cette tâche est égale à la tâche 4 (DIS MEM MOT), les items à mémoriser sont seulement différents.

Évaluation du « vocabulaire technique »		
Exercice	Sigle	But
9	REC EC	reconnaître les types d'écrits rencontrés dans la vie quotidienne
10	REP LIV	se repérer dans un livre (connaissance du vocabulaire technique)
11	SPA	s'orienter dans l'espace de la page
12	LGE TECH	connaître le langage technique (ex. <i>chiffres, lettres, mot, majuscule ...</i>)
Évaluation de l'« aspect sémantique »¹³		
Exercice	Sigle	But
13	COMP CONS	comprendre des consignes données de manière collective
14	TPS V	comprendre le temps des verbes pour exprimer le passé, le présent, le futur
15	CHRO	comprendre une histoire lue par l'enseignant, la raconter en restituant les enchaînements chronologiques
16	CHRO 2	comprendre un vocabulaire pertinent (entourer des images)
17	COMP VOC	comprendre un vocabulaire pertinent (dessiner)
18	COMP VOC 2	comprendre le système alphabétique de codage de l'écriture (correspondance graphème-phonème)
Évaluation de l'écriture		
Exercice	Sigle	But
19	CODA PHON	apprendre le geste de l'écriture
20	COMP PHRA	comprendre le système alphabétique de codage de l'écriture (segmenter une phrase prononcée à l'oral)
21	CODA MOT	comprendre le système alphabétique de codage de l'écriture (permanence de la forme écrite dans la copie des mots)
22	EC PRE	écrire son propre prénom
23	EC DATE	écrire la date du jour

Tableau 6. Les exercices du livret de passation, en détails

Chaque exercice présente une consigne bien détaillée dans le livret de passation, présenté dans les annexes (Annexe 2). Avec ce livret, les maîtresses disposent du matériel à donner aux enfants pour compléter les exercices (Annexe 3). Le livret donne aussi le codage de correction qui est constitué de 4 degrés numériques : 0, 1, 4 et 9. Le degré zéro correspond à l'absence de réponse ; le degré 9 aux performances insuffisantes et le degré 1 aux meilleures. Le degré 4 est associé à un

13 Nous utilisons ici la dénomination du livret. La terminologie adoptée est ambiguë car le mot « aspect » peut faire référence à la définition linguistique d'aspect. Les tâches de cette partie ne vérifient pas l'aspect verbal mais l'ensemble des compétences sémantiques.

niveau de performance moyen. Après la correction des tests, les résultats sont présentés sous la forme d'un code couleur qui matérialise visuellement les performances des enfants. Si le degré 1 est associé au vert, le degré 4 est associé au jaune et le degré 9 au rouge. Le marquage en rouge signale une situation de difficulté de l'enfant visant à alerter la maîtresse. Dans le cas du jaune, l'enfant présente des problèmes à la rentrée qui vont se résoudre au bout de 6 mois. Pour le vert, il n'y a pas de gros problème et le niveau au début de l'année scolaire est considéré comme assez élevé. Comme nous pouvons le voir dans cette capture d'écran (Figure 23), les différentes couleurs permettent d'illustrer les résultats de chaque élève pour chaque exercice. Quatre colonnes à droite résument les pourcentages des 4 codes. Cela permet d'isoler les élèves avec des situations plus problématiques dans la classe. De plus, en bas, nous observons les pourcentages de réussite par exercice et par classe, la moyenne et la médiane de réussite selon la catégorie testée par le groupe d'exercices. Ces types de classification et d'évaluation permettent de voir la situation initiale par élève et d'avoir également à disposition une vision globale de la classe.

Ce tableau a été réalisé pour les deux classes de façon séparée et les maîtresses ont bien voulu nous permettre de les consulter pour les items 1 à 23, où l'évaluation des compétences correspond à nos questions de recherche. Ils constituent pour nous un socle de travail sur lequel appuyer nos propres analyses.

EVE	MÂITRISE DE LA LANGUE FRANÇAISE																								
	COMPOSANTES SONORES ET VISUELLES								VOCABULAIRE TECHNIQUE					ASPECT SEMANTIQUE					ECRIURE					TOTALX PAR ELEVE, en %	
1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.	23.				
NS STYL	LOC	ISOL PHON	DIS MEMI EC	ORAL EC	DV	DIS MEMI MOTZ	REC REPR LVI MOT	REC REPR LVI MOT	REC REPR LVI SPA	LGE TECH	COMP LGE	COMP CON	TPE V	CHRO	CHROZ	COMP	COMP	PHON	CODA PHRA	CODA MOT	EC PREN	DATE	00000.0	00000.0	
12/12/2008	1	4	4	1	4	4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	48%	26%	26%
	1	4	4	1	4	4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	30%	26%	48%
	1	1	4	1	4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	52%	30%	17%
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	65%	13%	22%
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	30%	13%	43%
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	39%	13%	48%
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	30%	43%	26%
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	74%	22%	4%
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	70%	30%	0%
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	75%	22%	0%
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	57%	35%	9%
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	4%	13%	70%
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	61%	22%	17%
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	52%	17%	30%
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	57%	26%	17%
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	30%	30%	39%
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	57%	17%	26%
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	61%	26%	13%
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	39%	39%	9%
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	48%	22%	22%
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	65%	22%	13%
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	43%	30%	26%
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	57%	13%	26%
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	52%	30%	17%
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	43%	8%	43%

SAUVEZ	48%	21%	25%	2%
SAUVEZ	52%	22%	24%	0%
OUVERTEZ				

Figure 23. Enregistrement des données pour les exercices du livret

Les exercices de 1 à 8 sont très importants pour notre recherche. Ils constituent la catégorie appelée « composantes sonores et visuelles des mots » qui nous donnent une idée des compétences phonologiques de base des enfants. Nous présentons dans le tableau 7 le pourcentage de réussite générale et dans le tableau 8, les pourcentages de réussite aux 8 exercices pour le CP A et le CP B.

COMPOSANTES SONORES ET VISUELLES	
% de réussite - CP A	54,8
% de réussite - CP B	52,8

Tableau 7. Résultats des tâches qui évaluent les composantes sonores et visuelles dans CP A et CP B

	1 NB SYL	2 LOC SYL	3 ISOL PHON	4 DIS MEM MOT	5 ORAL EC	6 DV	7 DIS MEM MOT2	8 REC MOT
% de réussite CP A	88,5	42,3	15,4	76,9	11,5	73,2	69,2	61,5
% de réussite CP B	59,3	51,9	29,6	70,4	22,2	63,0	59,3	66,7

Tableau 8. Pourcentage de réussite pour les 8 tâches qui évaluent les composantes sonores et visuelles

En général, les performances des enfants ne subissent pas une grande variation ; en revanche la réussite aux exercices particuliers montre un autre panorama. Si, par exemple, pour l'exercice 8 (REC MOT), qui prévoit la reconnaissance de mots écrits, il n'y a pas de grande différence entre les deux classes (5,2 %), on peut distinguer pour les autres exercices deux compétences prédominantes différentes. Dans la classe du CP A, les enfants ont des résultats meilleurs dans les tâches qui contrôlent la mémorisation et les capacités visuelles de discrimination (exercices 4, 6, 7). Dans la classe de CP B, les enfants sont en général plus performants dans les tâches sonores mettant en rapport l'oral et l'écrit (exercices 5 et 8) comme, par exemple, isoler les phonèmes et localiser les syllabes. Le premier exercice, qui demande de compter les syllabes dans des mots oraux, est la seule exception. Ici les enfants du CP A nettement ont de meilleurs résultats que les enfants du CP B.

La situation et les données présentées à ce stade n'étaient pas influencées par l'environnement de la classe ou par l'enseignement de la maîtresse car les exercices avaient été donnés au début de l'année scolaire et les enfants venaient d'arriver à l'école primaire.

Ces résultats nous ont été très utiles, en particulier pour voir si ces compétences initiales des enfants, plutôt visuelles dans une classe et plutôt sonores dans l'autre, allaient influencer l'acquisition de l'écriture et de la lecture. Comme nous l'avons vu dans le chapitre 5, la conscience phonologique et la capacité à isoler les phonèmes dans la parole continue sont corrélées à de meilleures performances dans la première période de l'acquisition de la lecture et de l'écriture. À l'inverse, la mémorisation et la reconnaissance visuelle sont des compétences qui deviennent importantes dans une deuxième phase d'acquisition de la *literacy*, quand l'acquisition des correspondances graphème-phonème est déjà en place. Dans les classes du CP au début de l'année scolaire, il n'y a pas une grande différence de réussite pour les groupes d'exercices qui évaluent les niveaux d'écriture (CP A 47,7 % vs. CP B 39,3 % avec une différence de 8,4 %). Ni compétences sonores ni compétences visuelles ne semble donc être corrélées avec le niveau de production écrite spontanée avant l'apprentissage de la *literacy* à la rentrée.

7.1.2 Les parents

Lors de notre travail, nous avons pris en compte les facteurs sociolinguistiques et établi une enquête permettant de mieux connaître la situation linguistique des parents dont la production est l'input principal chez les enfants. Ceci nous a paru d'autant plus nécessaire que notre travail porte en partie sur la liaison, qui est un phénomène variable soumis à des contraintes sociolinguistiques telles que le milieu social et le niveau scolaire. L'input des parents, issus de milieux sociaux et de niveaux scolaires variés, peut plus ou moins influencer les compétences linguistiques des enfants et par conséquent leurs performances dans les tâches que nous proposons.

Pour ces raisons, nous avons demandé aux familles de remplir des questionnaires parentaux qui nous ont permis de recueillir des informations personnelles et sociolinguistiques (Annexe 1). La plupart des parents sont de langue maternelle française (Tableau 9). Quelques parents utilisent à l'oral la variété régionale du nord de la France. Trois parents parlent des langues étrangères, soit le néerlandais, soit le flamand. Les influences de ces langues sont la conséquence du fait que le territoire où nous avons conduit nos recherches est géographiquement très proche de la

frontière avec la Belgique et est originellement flamand. Malgré la compétence des langues étrangères ou des variétés différentes, 100 % des parents déclarent communiquer en français standard à la maison, avec leurs enfants.

Langue parlée	Nombre d'enfants avec un père locuteur de cette langue	Nombre d'enfants avec une mère locutrice de cette langue
Français standard	24	30
Français du Nord ¹⁴	16	12
Néerlandais/ Flamand	2	0

Tableau 9. Langues parlées par les parents des enfants

Pour avoir une vue complète du profil des parents, nous les avons différenciés par niveau scolaire en niveau 1, 2 et 3 (tableau 10). Au niveau 1 correspond un niveau de diplôme professionnel (CAP et BEP) ; au niveau 2 correspond un niveau de diplôme baccalauréat (BAC et BAC PRO) ; au niveau 3 un niveau post-bac avec ou sans certification (BTS, BAC+1, BAC+2, BAC+3, BAC+4, Licence, Maîtrise, DEUG, DEA). Comme nous pouvons le voir dans le tableau 10, la plupart des enfants ont des parents avec un niveau 3 d'étude selon notre classification, donc le plus haut.

	Nombre de parents	%
« Niveau 1 »	8	19,6
« Niveau 2 »	7	17
« Niveau 3 »	26	63,4

Tableau 10. Classification du niveau scolaire des parents

Nous avons comparé notre classement avec la nomenclature de professions et de catégories socio-professionnelles des emplois salariés d'entreprise INSEE (Desrosières & Thévenot, 1998). Cette classification distingue six groupes différents. En tenant compte de cela, nous distinguons les métiers des parents en quatre groupes :

14 Ce que on appelle ici « français du Nord » est identifié comme le sociolecte de la région dunkerquoise.

· **Groupe 1** - « parents cadres », le groupe 3 de la nomenclature INSEE appelé « cadres et professions intellectuelles supérieures » dont les professions considérées sont cadres de la fonction publique, professeurs et professions scientifiques, cadres administratifs et techniques, ingénieurs, professions des informations, des arts et du spectacle.

· **Groupe 2** - « parents intermédiaires », le groupe de la nomenclature INSEE nommé « professions intermédiaires » qui considère les professeurs des écoles, les professions intermédiaires et administratives et les techniciens.

· **Groupe 3** - « parents employés », le groupe 4 de la nomenclature INSEE qui compte les employés civils, administratifs et de commerce.

· **Groupe 4** - « parents ouvriers », le groupe 6 de la nomenclature INSEE où sont compris tous les types d'ouvriers et les chauffeurs.

Comme nous pouvons le voir dans le tableau 11, la plupart des enfants ont des parents employés. Pour toutes les catégories socio-professionnelles, la situation est assez équilibrée entre les enfants et dans les deux classes.

	Nombre d'enfants (CPA+CPB)	Nombre d'enfants	
		CPA	CP B
« parents cadres »	5	2	3
« parents intermédiaires »	9	4	5
« parents employés »	16	8	8
« parents ouvriers »	10	5	5

Tableau 11. Classification des catégories socio-professionnelles des parents des enfants au CP

Dans le questionnaire, nous avons aussi demandé aux parents s'ils pratiquent la lecture partagée avec leurs enfants et avec quelle fréquence. Les moments de lecture partagée, où les parents lisent un livre illustré avec leurs enfants, permettent aux enfants avant l'école primaire d'avoir un premier contact - visuel notamment - avec la langue écrite, d'avoir les moyens de former une conscience de l'espace entre les mots et d'établir les premières connexions graphème-phonème (voir chapitre 5). En effet, plusieurs études sur l'acquisition de l'écriture ont montré que ces pratiques aident les

literacy skills (voir chapitre 5). Cela veut dire qu'une exposition plus élevée à l'écrit, grâce à une fréquence élevée de moments de lecture partagée, peut avantager les enfants dans l'acquisition de l'écriture en leur montrant qu'il existe un codage spécifique de la langue à l'écrit. Nous avons établi une classification des enfants selon la fréquence de lectures partagées déclarée par les parents (Tableau 12).

Fréquence des activités de lecture partagée	Nombre de parents	% de parents
Une fois par jour	16	48,5
De 2 à 3 fois par semaine	9	27,3
1 fois par semaine	2	6,1
À la demande	2	6,1
Assez rarement	1	3
Aucune	3	9

Tableau 12. Fréquences des activités de lecture partagée déclarées par les parents

La plupart des parents pratiquent très fréquemment la lecture partagée avec leurs enfants (48,5 % des parents lisent une fois par jour avec leur enfant et 27,3 % lisent de 2 à 3 fois par semaine). Leurs enfants ont été exposés fréquemment à la forme visuelle des mots. Même s'ils ne les reconnaissent pas à ce stade, ils sauront que ce codage existe et qu'il correspond à une mise en forme de l'input oral. Nous devons considérer qu'il y a des parents qui ne lisent pas ou rarement avec leur enfant (6,1 % de parents qui lisent avec les enfants une fois par semaine ou seulement à la demande ; 9 % des parents ne font pas d'activités de lecture partagée avec leur enfant) ; pour ceux-là, le codage entre oral et écrit peut être de ce fait moins automatique. Nous proposons une comparaison des résultats sur les tâches d'écriture proposées par les maîtresses dans le livret de passation entre les enfants qui n'ont pas de moments de lecture partagée avec leurs parents (9 %) et les enfants qui partagent un moment de lecture tous les jours (48,5 %) (Tableau 13). Nous utilisons le codage à quatre couleurs déjà présenté : blanc (pas de réponse), rouge (difficulté), jaune (possibilité d'amélioration), vert (aucune difficulté). Nous présentons les pourcentages des résultats dans les tâches d'écriture du livret de passation selon les deux groupes d'enfants.

Fréquence lecture	Blanc	Rouge	Jaune	Vert
Une fois par jour	8 %	13 %	14,7 %	64 %
Aucune	20 %	15 %	25 %	40 %

Tableau 13. Résultats des tâches d'écriture du livret de passation selon la fréquence de la lecture partagée

Dans le tableau, les enfants qui lisent plus souvent avec leurs parents ont de meilleures performances (64 % de « Vert » vs. 40 %) pour les tâches liées à la langue écrite du livret de passation. En revanche, le pourcentage de difficulté (rouge) n'est pas particulièrement différent dans les deux groupes (13 % vs. 15 %).

Au moment de l'entrée à l'école, la fréquence de lecture partagée semble aider les performances dans les tâches d'écriture. Pour mieux comprendre le rôle de la lecture partagée dans le parcours scolaire, nous allons analyser dans les chapitres suivants cette variable.

7.2 La construction du matériel

Dans cette section, nous rendons compte de la manière dont nous avons construit le matériel expérimental proposé aux enfants pour tester nos questions de recherche.

7.2.1 La base de données pour les tâches orales : MANULEX

Le répertoire des mots que nous avons utilisé dans les tâches orales est construit à partir de la base de données MANULEX¹⁵ (Léte *et al.*, 2004) qui est établie à partir d'un corpus écrit de manuels scolaires. En l'absence d'une base de données orale pour les enfants français de 6 ans, le choix que nous avons fait part du principe que les mots écrits connus par les enfants à partir du CP sont plus ou moins les mots connus à l'oral.

MANULEX fournit les fréquences d'occurrences de 23900 lemmes et 48900 formes orthographiques. Le corpus utilisé comprend 54 manuels scolaires de lecture. Cette base de données rend aussi compte de l'exposition à l'écrit des enfants à trois

15 La base de données est téléchargeable depuis le site <http://www.manulex.org>

niveaux d'âge et de scolarisation : au CP (6 ans), au CE1 (7 ans) et au Cycle 3 (CE2-CM2, 8-11 ans) où se consolide et s'enrichit le stock lexical par exposition répétée à l'écrit. En effet, les indices de fréquence sont donnés selon les trois niveaux scolaires.

Dans MANULEX, différents types de valeurs sont associés à chaque mot :

- Fréquence cumulée (F) : la valeur indique la fréquence des items dans le corpus (les 54 manuels scolaires de lecture). Par exemple, un item qui a une fréquence cumulée égale à 1, n'est présent qu'une seule fois dans tout le corpus.
- Dispersion entre lectures (D) : cette valeur, comprise entre 0 et 1, montre la dispersion des occurrences des mots présents dans les différents manuels utilisés pour la construction du corpus. Si la valeur de D est 0,00, les occurrences d'un mot ont été repérées dans un seul manuel, quelle que soit la fréquence. Si la valeur est 1, les fréquences du mot sont réparties de façon égale sur tous les manuels.
- Estimation de la fréquence sur un million de mots (U) : cet indice nous permet de comprendre la fréquence relative du mot, calculée en rapport au total d'un million de mots. Il est dérivé de F au moyen d'un ajustement avec D. Quand la valeur de D est égale à 1, U est calculé comme la fréquence par million de mots. En revanche, si la valeur de D est inférieure à 1, U est ajusté à la baisse. Comme nous pouvons le lire sur le site de MANULEX (<http://www.manulex.org>) : « On considère que U est un meilleur estimateur de la « vraie » fréquence par million que l'on trouverait dans un corpus théorique de taille infinie ».
- Indice de fréquence standard (SFI) : il est calculé à partir de U, par transformation logarithmique au moyen de la formule en (24) :

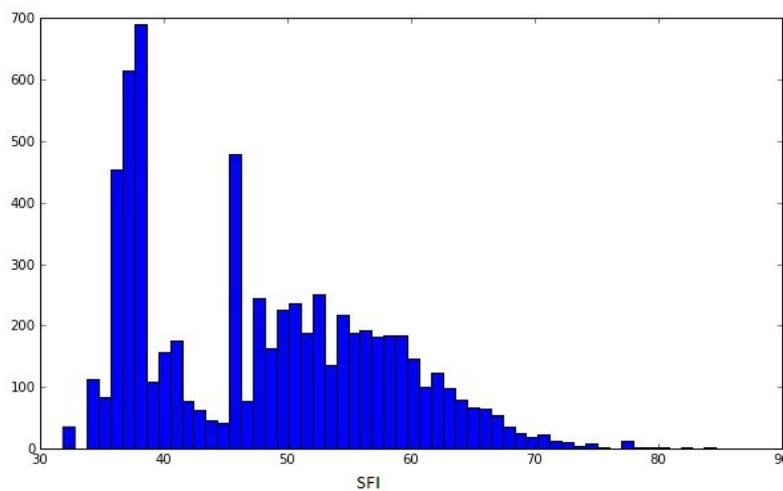
$$(24) \quad \text{SFI} = 10 * (\log_{10}(U) + 4)$$

SFI correspond à un indice standardisé. Pour mieux comprendre cet indice, nous donnons un exemple. Si un mot a un SFI de 90, nous le rencontrons une seule fois tous les 10 mots lus. Si le mot a un SFI égal à 80, il est présent une fois tous les 100 mots lus. Comme le rapport entre U et SFI est représenté par la formule en (24), nous devons considérer que si un mot a un SFI de 40, on le trouve une fois dans un million de mots lus.

Dans MANULEX, les lemmes et les formes orthographiques sont classifiés aussi selon leur nature syntaxique. Le corpus est organisé en 12 catégories grammaticales que nous présentons ici par ordre décroissant de pourcentage de présence : nom commun (NC), nom propre (NP), adjectif (ADJ), verbe (VER), adverbe (ADV), interjections (INT), pronom (PRO), préposition (PRE), abréviation (ABR), conjonction (CON), déterminant (DET), unités euphoniques (UEUPH).

7.2.1.1 La construction du répertoire de mots pour les tâches orales : le choix du Mot1 et du Mot2

Pour créer le répertoire, nous avons sélectionné dans la liste « lemmes » les noms communs commençant par une voyelle en fonction du niveau scolaire qui nous intéresse. Nous avons donc pris en considération les indices de fréquence uniquement au CP. Dans la liste de mots commençant par une voyelle, nous effaçons d'abord ceux qui ont D (au CP) égal à 0. De cette façon, nous sélectionnons seulement les mots apparus dans plusieurs manuels. En l'absence d'un modèle statistique précis et sans pouvoir tenir compte de médians ou d'écart-types qui ne sont pas présents dans la base de données pour les valeurs de fréquence, nous avons choisi de regarder la distribution de l'indice SFI. Cet indice donne le logarithme de la fréquence sur un million de mots, dans lequel on peut avoir foi parce qu'il est en rapport avec U (la fréquence estimée sur un million de mots) et parce que, par l'opération du logarithme, il a été statistiquement équilibré (Graphique 1). Comme nous pouvons le voir dans le graphique, l'intervalle de SFI entre 50 et 65 peut être choisi comme un intervalle correspondant à la fréquence moyenne. Les noms qui ont un SFI à l'intérieur de cet intervalle sont de fréquence moyenne. Les noms, repérés au-dessus de cet intervalle, avec un SFI plus élevé que 65, ont une haute fréquence. En revanche, les noms avec un SFI au-dessous de cet intervalle (inférieur à 50) sont des mots que nous considérons de basse fréquence.



Graphique 1. Distribution logarithmique des fréquences SFI dans MANULEX

Sur cette base, nous avons construit le répertoire de 30 Mots2 que nous présentons dans le tableau 14. Ici, nous donnons aussi une description phonologique, prosodique et fréquentielle des Mots1 pour les phénomènes consonantiques que nous avons testés dans les tâches orales. Si nous regardons la fréquence, la plupart des items ont une fréquence moyenne, sauf l'un d'eux qui a une fréquence haute (*ours*) et deux qui ont une fréquence basse (*artichaut* et *avance*). Nous étions obligés d'insérer dans notre répertoire des mots qui n'ont pas une fréquence moyenne à cause du choix des tâches. En effet, comme nous proposons une dénomination guidée au moyen d'images, il fallait choisir des mots pouvant être facilement représentés. Nous n'avions pas assez de mots de fréquence moyenne commençant par une voyelle et qui pouvaient être représentés par un dessin. Par conséquent, nous avons choisi un mot de fréquence plus élevée et pouvant être facilement dessiné. Pour l'autre tâche orale, de perception d'erreurs et de reformulation, nous avons essayé de garder la même voyelle initiale du Mot2 pour chaque phénomène ce qui nous a forcé à choisir deux mots ayant une fréquence plus basse.

Catégorie Grammaticale	Mot2	Forme Phonologique	Structure Syllabique	Phénomène consonantique	SFI
nom	agneaux	/a.ɲo/	V/CV	Liaison déterminant-nom	60.25
nom	anneaux	/a.no/	V/CV	Liaison déterminant-nom	58.74
nom	ananas	/a.na.na/	V/CV/CV	Liaison déterminant-nom	58.02
nom	effort	/e.fɔʁ/	V/CVC	Liaison déterminant-nom	52.6
nom	essai	/e.sɛ/	V/CV	Liaison déterminant-nom	53.21
nom	effet	/e.fɛ/	V/CV	Liaison déterminant-nom	55.3
adj	ennuyeux	/ã.nɥi.jø/	V/CG ¹⁶ V/CV	Liaison mot invariable-nom	60.25
nom	avance ¹⁷	/a.vãs/	V/CVC	Liaison mot invariable-nom	49,25
nom	ambulance	/ã.by.lãs/	V/CV/CVC	Liaison mot invariable-nom	58.02
nom	affaires	/a.fɛʁ/	V/CVC	Enchaînement déterminant-nom	56.77
nom	affiches	/a.fiʃ/	V/CVC	Enchaînement déterminant-nom	58.38
nom	agence	/a.ʒãs/	V/CVC	Enchaînement déterminant-nom	50.8
nom	armoire	/aʁ.mwaʁ/	VC/CGVC	Élision déterminant-nom	58,10
nom	habitudes	/a.bi.tyd/	V/CV/CVC	Liaison déterminant-nom	58.69

16 Nous indiquons les semi-voyelle avec en G pour « glide ».

17 Dans la locution adverbiale « être en avance ».

nom	habits	/a.bi/	V/CV	Liaison déterminant-nom	61.6
nom	hamburgers	^(h aspiré) /am.bœʁ.gœʁ/	VC/CVC/CVC	Liaison déterminant-nom	47,23
nom	hamsters	^(h aspiré) /am.stœʁ/	VC/CCVC	Liaison déterminant-nom	50.3
nom	haricots	^(h aspiré) /a.ʁi.ko/	V/CV/CV	Liaison déterminant-nom	53,46
nom	orange	/ɔ.ʁɑ̃ʒ/	V/CVC	Enchaînement adj-nom	58.64
nom	araignée	/a.ʁe.je/	V/CV/CV	Enchaînement adj-nom	57.88
nom	ogre	/ɔgʁ/	VCC	Enchaînement adj-nom	56.77
nom	ours	/uʁs/	VCC	Liaison déterminant-nom	68.08
nom	oie	/wa/	GV	Élision déterminant-nom	57.59
nom	oreiller	/o.ʁe.je/	V/CV/CV	Élision déterminant-nom	57.88
nom	arrosoir	/a.ʁɔ.zwaʁ/	V/CV/CGVC	Liaison déterminant-nom	56.59
nom	orage	/o.ʁɑ̃ʒ/	V/CVC	Élision déterminant-nom	58.28

Tableau 14. Description du répertoire utilisé dans les tâches orales

Les « adjectifs », utilisés comme Mot1 dans les cas d'enchaînement, n'ont pas pu être sélectionnés selon le critère de fréquence que nous avons utilisé pour la plupart des Mots2. Nous avons plutôt essayé de produire des séquences possibles et assez cohérentes pour les enfants comme « adjectif antéposé + nom » du type *la grosse araignée* ou *le minuscule ogre*. Ces séquences ont été validées par des locuteurs natifs et par les maîtresses. Dans le tableau 15, nous donnons une représentation des adjectifs utilisés comme Mot1 dans le cas d'enchaînement.

Catégorie grammaticale de Mot1	Mot 1	Forme Phonologique	Structure Syllabique
adjectif	petit	/pə.ti/	CV.CV
adjectif	grosse	/gʁos/	CCVC
adjectif	minuscule	/mi.nys.kyl/	CV.CVC.CVC

Tableau 15. Répertoire des adjectifs utilisés comme Mots1 dans les tâches orales

Utiliser les mêmes critères de choix pour les Mots1 dans la catégorie de MANULEX appelée « déterminants » a été impossible à cause du nombre réduit de déterminants (seulement 38 dans tout MANULEX). De plus, ils ne sont pas tous utilisables pour les trois phénomènes de resyllabation testés (19 déterminants dans toute la base). Enfin MANULEX donne seulement la fréquence d'un mot, pas de la séquence. Par exemple, dans le cas de la liaison, il nous fournit la fréquence du déterminant isolé, du nom isolé mais pas de la combinaison des deux. Pour remédier à ce manque, nous avons donc choisi, pour les mots à gauche « déterminant », « mot invariable » et « verbe », les mots les plus fréquents en réalisation de liaison dans chaque catégorie, en suivant les indications des données de la thèse de Mallet (2008, chapitre 3) sur le pourcentage de réalisations par catégorie grammaticale (Tableau 16).

Catégorie grammaticale	Mot 1	Forme phonologique	Structure syllabique	Contexte	% de réalisation (Mallet, 2008)
verbe	est	/ɛ/	V	verbe-nom	43 %
mot invariable	en	/ɑ̃/	V	mot invariable-nom	98 %
mot invariable	très	/tʁɛ/	CCV	mot invariable-nom	96,6 %
déterminant	les	/le/	CV	Déterminant-nom	100 %
déterminant	des	/de/	CV	Déterminant-nom	100 %
déterminant	un	/ɛ̃/	V	Déterminant-nom	100 %
verbe	peut	/pø/	CV	Verbe-nom	30,9 %

Tableau 16. Répertoire des Mots1 utilisés dans les tâches orales.

En ce qui concerne l'enchaînement déterminant-nom, nous avons utilisé comme déterminant l'article féminin *une*. Pour l'élision, nous avons utilisé l'article défini *l'*. Ces deux Mots1 sont très communs en français pour l'élision et l'enchaînement.

Description syllabique

Dans le tableau 14, nous avons 30 mots commençant par une voyelle : 29 noms et 1 adjectif. Si on regarde la structure syllabique, nous pouvons compter 3 mots monosyllabiques, 20 mots dissyllabiques et 7 mots trisyllabiques.

La resyllabation est une des caractéristiques principales des phénomènes consonantiques. Nous allons bien considérer la position prosodique entre Mot1 et Mot2 pour évaluer la dynamique d'acquisition.

Dans les items tests, nous avons deux types de syllabes initiales :

- Syllabe formée d'un seul noyau vocalique (V)
- Syllabe avec une coda consonantique simple (VC)

Pour les syllabes qui ne sont pas initiales, nous avons essayé de choisir des mots constitués, dans la plupart des cas, par des syllabes du type CV même si nous avons été obligé de choisir d'autres structures syllabiques internes étant donné la taille du lexique utilisable et contraint par la sémantique et la fréquence, surtout pour la tâche de jugement. En effet, pour faire une phrase compréhensible pour les élèves dans un contexte spécifique en faisant attention à la voyelle initiale impliquée dans la resyllabation, nous n'avons pas pris en compte l'homogénéité syllabique. De même, pour pouvoir satisfaire le critère de la fréquence, nous n'avons pas non plus toujours pu contrôler l'homogénéité syllabique du répertoire.

7.2.2 La base de données pour les tâches liées à la langue écrite

Pour les premières tâches, nous avons utilisé comme outil MANULEX. Cette base de données nous présente la fréquence des mots isolés et elle ne tient pas compte de la fréquence des séquences de mots. Pour les tâches d'écriture et de lecture, proposées à la fin de l'année scolaire, nous ne l'avons pas utilisée. Afin de tenir compte des vraies connaissances lexicales des enfants et pour avoir accès aux contextes

syntactiques dans lesquels les enfants ont appris les mots (savoir s'ils les ont rencontrés de façon isolée ou en contexte de liaison, d'élision ou d'enchaînement), nous avons créé notre propre base de données du lexique écrit des enfants du CP.

Les maîtresses nous ont fourni deux cahiers, utilisés depuis septembre, par les enfants dans les heures de langue française. Le premier cahier est utilisé par les enfants à l'école avec les maîtresses. Ce cahier contient des exercices qui aident à développer le rapport entre graphème et phonème (comme, par exemple, rechercher le même son dans différents mots écrits). L'autre cahier, utilisé à la maison avec les parents, rassemble des devoirs dont le but est la répétition de la leçon et le renforcement des connaissances apprises en classe. Dans ce cahier, nous trouvons des petites comptines qui poussent les élèves à travailler de manière autonome sur un son spécifique.

Pour créer notre base de données, nous avons sélectionné dans les deux cahiers seulement les noms communs. Nous les retrouvons sous 3 formes :

1. précédés par des déterminants (ex. *les amis* ou *une maison*),
2. précédés par des adjectifs (ex. *fausse statue*),
3. en isolation (ex. *igloo*).

Dans le premier exercice d'écriture, au début de l'année scolaire, nous retrouvons le plus souvent la dernière forme, qu'on a appelée « en isolation ». Petit à petit, les mots en isolation laissent la place aux noms accompagnés par le déterminant ou par l'adjectif, quelquefois associés à une image, quelquefois directement insérés dans un texte.

Pour construire une liste complète du lexique des enfants, nous avons traité les mots présentés aux enfants avec des déterminants différents comme des entrées différentes. Par exemple, nous avons considéré le même item *nez* comme deux formes différentes *le nez* et *un nez*, car elles sont précédées d'un déterminant différent. Ces différenciations sont très importantes pour les noms qui commencent par une voyelle car il est fondamental de voir si les mots ont été appris en contexte de liaison, d'élision ou d'enchaînement ou s'ils ont plutôt été stockés en isolation. Comme nous l'avons déjà dit, en présence de phénomènes de resyllabation, la consonne du Mot1 est resyllabifiée sur le Mot2. Les enfants qui ont vu le Mot2 seulement en forme isolée peuvent avoir des difficultés à l'enchaîner avec un déterminant spécifique en liaison, élision ou enchaînement. De plus, réussir à combiner avec un déterminant différent le nom appris avec un certain déterminant peut créer des problèmes. Par exemple, s'ils ont seulement vu l'orthographe de <l'éléphant> ils peuvent avoir des difficultés à écrire <un

éléphant>. Ils risquent plutôt de produire des formes comme <un l'éléphant>. Comprendre et classifier les contextes dans lesquels les enfants ont appris la forme écrite du Mot2 est fondamental pour notre recherche.

Sur la base de cette méthode, en considérant comme deux formes séparées deux mots utilisés avec des déterminants différents, notre base de données compte 855 formes. Les 855 formes ont été classées par ordre alphabétique (Figure 24). Dans la colonne adjacente, nous avons marqué la fréquence, basée sur le nombre d'occurrences des formes. La forme *madame* apparaît 10 fois dans les deux cahiers des enfants. En revanche, une séquence comme *une avenue* est retrouvée dans le lexique écrit seulement 2 fois. La fréquence moyenne des formes orthographiques dans notre base de données est de 1,3.

1	FORMES ORTHOGRAPHIQUES	Fréq.
2	aujourd'hui	1
3	une avenue	1
4	des ailes	3
5	ami	2
6	vrais amis	1
7	son ami	1
8	une amie	1
9	mon amie	1
10	un ami	1
11	d'avril	1
12	l'alphabet	1
13	un alphabet	1
14	les armoires	1
15	l'aiglon	1
16	l'accordéon	1
17	un arbre	1
18	une assiette	2
19	une addition	1
20	attention	1
21	un agneau	1
22	aspirateur	1
23	un animal	2
24	les animaux	1

Figure 24. Classification des mots écrits pour créer la base de données

Après avoir classifié les formes orthographiques d'un ou deux mots, nous avons regardé tous les noms commençant par des voyelles. Nous avons créé une liste d'items à voyelle initiale en associant une fréquence. La fréquence que nous avons indiquée est également la fréquence absolue, le nombre de fois où l'item a été repéré dans le cahier, sans tenir compte de la séquence Mot1 + Mot2 où il est contenu (voir figure 25). Par

exemple, l'item *ananas* a une fréquence de 2. Cela veut dire que nous pouvons trouver ce mot deux fois dans les deux cahiers des enfants.

ITEM	FREQ
ananas	2
oursons	3
ours	3
euros	1
essaim	2
épée	4
épi	1
yeux	3
uniforme	1
infirmier	1
enfants	1
éléphant	2
olives	1
iris	2
élèves	1
ailes	3
oiseau	3
oiseaux	2
escargot	2
œuf	1
amis	2

Figure 25: Liste de mots traités dans la base de données

La fréquence moyenne dans la liste des items est de 2. Pour choisir les mots à utiliser dans les tâches liées à la langue écrite, nous avons pris les items dont la fréquence moyenne est au-dessus de la moyenne (fréquence absolue supérieure à 2). Nous avons choisi de prendre des mots avec une fréquence supérieure à la moyenne pour rendre les tâches liées à la langue écrite plus accessibles au CP, car écrire et lire ces mots, qui n'ont pas une orthographe transparente, peut déjà être une source de problèmes. Proposer d'orthographier des mots de fréquence moyenne ou au-dessus de la moyenne peut donc combler le fait que nous n'avons pas pris en considération la transparence des séquences à écrire ou à lire.

Dans le tableau 17, nous récapitulons les mots que nous avons utilisés dans les tâches liées à la langue écrite. Nous avons également cherché à sélectionner les items qu'on peut facilement représenter par des dessins qui seront utilisés dans les tâches d'écriture.

Mot2 pour l'écrit	Fréquence
ours	3
épée	4
ailles	3
oiseau	3
âne	3
autruche	3
affiche	3
oie	7
ananas	2
éléphant	2
orage	2
igloo	2
orange	2
île	2
école	2

Tableau 17. Répertoire des Mots2 utilisés dans les tâches liées à la langue écrite et leurs fréquences

Nous avons décidé de choisir également des mots commençant par *h*, afin d'observer comment les enfants procèdent dans l'écriture. Pour les isoler, nous avons procédé de la même manière que pour les mots commençant par une voyelle. Nous avons aussi sélectionné les items les plus fréquents dans notre base des données (Tableau 18).

Mots à <i>h</i> initiale utilisés dans les tâches	Fréquence
hibou	3
hérisson	3
hippopotame	3
hirondelle	3

Tableau 18. Répertoire des Mots2 à initiale *h* utilisés dans les tâches liées à la langue écrite et leurs fréquences

Dans cette liste de mots, toutes les formes orthographiques commencent par le graphème *h*, mais ce *h* n'a pas la même valeur phonologique dans tous les cas. Dans les mots *hibou* et *hérisson*, le *h* est dit « *h* aspiré », ce qui interdit la liaison et l'élision. Dans *hippopotame* et *hirondelle*, le *h* est un « *h* muet » qui n'empêche ni l'élision, ni la liaison, ni l'enchaînement (voir chapitre 3).

Au total nous avons sélectionné 20 items contenus dans 20 séquences qu'on propose aux enfants, divisés en 5 catégories différentes :

- 4 séquences de liaison au singulier,
- 4 séquences de liaison au pluriel,
- 4 séquences d'élision,
- 4 séquences d'enchaînement,
- 4 séquences à *h* initial (2 séquences déterminant-nom avec *h* aspiré et 2 séquences avec *h* muet).

Dans les 4 séquences par phénomène, deux contextes sont connus par les enfants. Les deux autres contextes ne sont pas connus. La distinction des contextes « connus » et « inconnus » a été faite sur la base de la classification des formes orthographiques que nous avons présentée précédemment (voir Figure 24 et 25). Quand on parle de « contexte connu », on veut dire que, dans notre base de données (et donc dans les cahiers des enfants), les items à voyelle initiale apparaissent dans un contexte spécifique, avec un Mot1 spécifique qui les accompagnent. Par exemple, pour le Mot2 *ours*, nous proposons le contexte *les ours*, qui est un contexte connu par les enfants car nous retrouvons la suite Mot1 (*les*) + Mot2 (*ours*) dans notre base de données. En revanche, nous présentons l'item *igloo* dans le contexte *un igloo* qui est un contexte inconnu puisque dans la base de données, *igloo* apparaît toujours comme mot isolé. Les enfants ne l'ont donc jamais appris à l'école comme un Mot2 précédé par le Mot1 *un*. Dans le tableau 19, nous donnons les contextes, les niveaux de familiarité (connu vs. inconnu) et la séquence utilisée dans les tâches liées à la langue écrite. Le contexte syntaxique pour tous les phénomènes est le contexte déterminant + nom.

Phénomènes	Familiarité	Séquences tests
Liaison en [z]	Connu	Les ours [le.zuʁs]
		Des ailes [de.zɛl]
	Inconnu	Les oranges [le.zɔ.rɑ̃ʒ]
		Les affiches [le.za.fif]
Liaison en [n]	Connu	Un oiseau [ɛ̃.nwa.zo]
		Un orage [ɛ̃.nɔ.ʁaʒ]
	Inconnu	Un igloo [ɛ̃.ni.glu]
		Un ananas [ɛ̃.na.na.nas]
Élision	Connu	L'oise [lwa]
		L'école [le.kɔl]
	Inconnu	L'âne [lan]
		L'éléphant [le.le.fɑ̃]
Enchaînement	Connu	Une épée [y.ne.pe]
		Une autruche [y.not.ʁyʃ]
	Inconnu	Une ambulance [y.nɑ̃.by.lɑ̃s]
		Une île [y.nil]
<i>h</i> aspiré	Inconnu	Un hibou [ɛ̃.i.bo]
	Connu	Un hérisson [ɛ̃.e.ʁi.sɔ̃]
<i>h</i> muet	Connu	Un hippopotame [ɛ̃.ni.po.po.tam]
	Inconnu	Une hirondelle [y.ni.ʁɔ̃.dɛl]

Tableau 19. Séquences de Mots2 utilisées dans les tâches liées à la langue écrite et classification des contextes (connus vs. inconnus)

Pour la tâche d'amorçage, qui prévoit comme cible un nom écrit, nous utilisons la même base de données créée au moyen des cahiers. Nous choisissons les mots de fréquence absolue moyenne. De plus, nous vérifions que ces mots ont également une fréquence moyenne SFI dans MANULEX. Dans la tâche d'amorçage, nous proposons aussi une amorce acoustique. Cette amorce est une séquence déterminant-adjectif-nom. Le phénomène consonantique a lieu entre l'adjectif antéposé et le nom (ex. *un petit éléphant* [ɛ̃.pə.ti.te.le.fɑ̃]). Nous avons cherché à proposer des adjectifs présents soit dans la base de données basée sur les cahiers des enfants, soit dans MANULEX. Après avoir créé les suites adjectif-nom, nous avons demandé à des locuteurs natifs du français de les valider. Nous présentons en détail les répertoires dans le chapitre 10, après avoir introduit la tâche d'amorçage (§ 10.5).

7.2.3 Le traitement des données

Pour compiler et analyser nos résultats expérimentaux, nous avons créé un classeur *Excel*. La capture d'écran (Figure 26) explicite notre façon d'enregistrer les résultats. Pour chaque enfant est créée une feuille *Excel* contenant d'abord les données récoltées grâce aux questionnaires (les informations personnelles et familiales), les résultats des tâches expérimentales que nous avons fait passer dans notre étude et les évaluations réalisées par les maîtresses dans le cadre scolaire. Les enfants ont été anonymisés et chaque élève est identifié par une suite de lettres.

De plus, nous avons créés d'autres fichiers *Excel*, un pour chaque tâche, où nous avons noté les résultats finaux. Ces différents fichiers permettent de confronter les tâches entre elles.

Pour les tâches orales, les performances audio ont été enregistrées par un enregistreur au format *.wav* à une fréquence de 44000 Hz.

Pour découper les traces audio en morceaux, nous avons utilisé le logiciel *Audacity*. Le même logiciel a été utilisé pour préparer les séquences audio à écouter dans les tâches. Pour analyser les caractéristiques phonétiques des productions, nous avons utilisé le logiciel *PRAAT*. Pour présenter nos données et établir les relations entre variables, nous avons utilisé le logiciel *Excel* et le logiciel *STATISTICA*.

82	verb-nom		es(t)-on déjà arrivées?	TN	Y	
83	mot invariable		Il est parti à l'école en avance	TN	N - en avance	
84			Un(e) agence de voyage est un magasin où ils organisent des vacances	TN	Y	
85	Elision		La armoire de ma chambre est bien rangée	TN	N - la armoire vs. l'armoire	
86	H must (det-nom)		Le(s) habitudes de ce chien doivent changer, il n'est pas sage!	TN	Y	
87	H aspiré (det-nom)		Le(s) hamsters de mon cousin sont trop mignons	TY	Y	
88						
89	Liaison TOT		2			
90						
91	BONNE Consonne					
92	ERR		3			
93	BR		5			
94						
95	det-nom [s]		Thibeaut coupe des[]ananas en moceaux	TY	Y	
96	det-nom [n]		Ce médicament lui a fait un[n] effet très fort et il a guéri tout suite	TY	Y	
97	verb-nom		peut-on savoir tes secrets?	TY	N - "parce que elle a dit peut-on".	
98	mot invariable		Ce livre est très ennuyeux	TY	N - F:Ce livre B:Ce livre	
99	Enchainement		Maxime a vu une affiche du cirque	TY	Y	

Figure 26. Exemple d'une feuille *Excel* utilisée pour classer les données

Partie 3.
L'étude longitudinale

Chapitre 8

Première étape de l'étude longitudinale : le début de l'école primaire

8.1 La première étape de l'étude longitudinale

Comme nous l'avons annoncé, l'étude que nous avons conduite est une étude longitudinale. Pour étudier l'influence de l'apprentissage de l'écriture sur la connaissance et le traitement des phénomènes de resyllabation, nous testerons les enfants à trois moments :

- début du CP (octobre 2013),
- fin du CP (mai-juin 2014),
- milieu du CE1 (janvier 2015).

Dans ce chapitre, nous présenterons les résultats des premières tâches conduites en octobre 2013, au début du CP.

8.1.2 Les hypothèses pour le début du CP

Au début du CP, en France, la plupart des enfants ne sont pas encore capables de lire et d'écrire. Ainsi, toutes les compétences que nous avons décidé de tester sont uniquement des compétences orales.

Nous avons montré dans le chapitre 4 que les enfants après 6 ans, indépendamment de l'explication que nous voulons adopter, n'ont plus de problèmes évidents à établir et isoler les frontières des mots. Les données rapportées dans la

littérature existante montrent qu'ils semblent comprendre la différence entre une CL et une consonne lexicale (Dugua, 2006 ; Dugua & Chabanal, 2006 ; Basset 2000 ; Dugua & Baclesse, 2014). On s'attend donc à ce que, à ce stade de l'acquisition :

- ils aient une bonne maîtrise de la liaison, de l'élosion et de l'enchaînement,
- ils n'aient aucun problème dans la segmentation en mots, aussi bien en perception qu'en production.

Pour vérifier si les compétences de segmentation à l'oral sont acquises, nous proposons deux tâches :

- une tâche de perception d'erreurs apte à vérifier si les enfants ont des problèmes à identifier les erreurs de segmentation en cas de phénomènes de resyllabation (PERC^{CP1}).
- une tâche de production, pour évaluer la production de phénomènes de resyllabation à l'oral (DENa^{CP1}).

Dans la première question de recherche (Q1, chapitre 6) présentée dans le chapitre précédent, nous nous demandons quel modèle de l'acquisition de la liaison décrit le mieux le découpage en mots après 5 ans. Les deux modèles proposés pour l'acquisition de la liaison, que nous avons décrits en détail dans la section 4.2.3.1, sont différents surtout dans leur façon de traiter la question du statut phonologique et de la formalisation de la consonne qui se resyllabifie sur la frontière. Selon l'explication lexicale-constructionniste (Chevrot *et al.*, 2013), cette consonne n'est pas associée à une position prosodique et n'a aucune représentation sous-jacente. En revanche, pour l'explication morpho-phonologique (Wauquier & Braud, 2005 ; Wauquier, 2009), elle a une valeur phonologique et fait émerger une position prosodique sous-jacente associée.

Pour bien évaluer le traitement de cette consonne ainsi que l'existence d'une éventuelle position prosodique à l'attaque du Mot2, nous proposons une analyse fine des productions orales des enfants sur la frontière de mots au moyen du logiciel PRAAT. Par ailleurs, pour mieux comprendre le rôle du lexique et de la phonologie à ce stade de l'acquisition, nous évaluons le rapport entre la fréquence des items et les erreurs en production. Enfin, nous proposons une analyse des données orales qui tient compte de la variable sociolinguistique et de la fréquence des moments de lecture partagée entre les enfants et les parents (Lundberg, 2006).

8.1.3 La procédure des premières tâches : octobre 2013

Les deux classes du CP ont été testées pour la première fois en octobre 2013, au début de l'année scolaire. Les enfants ont été testés un par un dans une salle silencieuse pendant 40 minutes. L'ordre des tâches et de présentation du répertoire varient selon les enfants pour éviter des effets de facilitation ou de stratégies. Des pauses étaient prévues entre les différentes tâches et entre les blocs qui composent les tâches.

Nous avons testé les enfants sur deux tâches différentes : une tâche de perception d'erreurs et de reformulation (PERC^{CP1}) ainsi qu'une tâche de dénomination guidée (DENa^{CP1}). Des phases d'entraînement précédaient la passation proprement dite. Dans ces phases, l'expérimentateur expliquait à l'enfant la démarche et fournissait des exemples jusqu'à ce que la tâche soit comprise.

8.2 PERC^{CP1} : la perception d'erreurs et leur correction

8.2.1 Procédure de la tâche PERC^{CP1}

La tâche proposée mesure à la fois le jugement de perception d'une erreur « cible » insérée dans un énoncé à écouter et la correction que produit l'enfant sur sollicitation. D'abord, l'enfant, muni d'un casque audio, doit écouter une phrase prononcée par une locutrice de langue maternelle française. Après l'écoute de la phrase, l'enfant doit dire si la personne qui parle produit des fautes ou, au contraire, prononce correctement l'énoncé. La consigne est la suivante : « on va écouter une dame qui ne parle pas trop bien français. Quelquefois elle fait des fautes. Toi, tu dois me dire si la phrase est juste ». Les réponses possibles sont soit « oui », si la phrase écoutée est bien formée ; soit « non », si dans la phrase écoutée l'enfant repère une erreur. Dans ce cas, l'expérimentateur demande à l'enfant de corriger la faute repérée. La correction de la faute est très importante car la tâche que nous proposons est une tâche métalinguistique qui n'évalue pas seulement un type de compétence linguistique (ex. la seule compétence phonologique vs. d'autres compétences grammaticales). Cette phase de correction permet, en effet, d'évaluer quels critères ont conduit au jugement. Les erreurs-cibles que nous avons insérées dans les phrases sont des erreurs phonologiques concernant les phénomènes de resyllabation (ex. *le(s)[n]agneaux* ou *la armoire*).

Cependant, nous nous attendons aussi à ce que les enfants puissent considérer qu'une phrase est fautive pour des raisons sémantiques ou pragmatique plutôt que phonologiques. Par exemple, dans une phrase comme *le artichaut est un bon légume* (où l'erreur est dans la production de *le artichaut*) l'enfant peut dire que la phrase est fautive parce qu'il considère que les artichauts ne sont pas bons. La correction des fautes nous guide dans le traitement des données et permet de comprendre si l'erreur-cible du phénomène de resyllabation que nous avons explicitement insérée dans la phrase est bien celle qui a conduit à rejeter la phrase.

Les enfants ont écouté les phrases au moyen d'un casque audio *sennheiser hd 202 II* au niveau de volume 16 sur 100. Les réponses ont été enregistrées par l'enregistreur *Zoom H4* à une fréquence de 44000 Hz au format *.wav*. Les phrases ont été prononcées par une locutrice de langue maternelle française et enregistrées dans une chambre sourde à une fréquence de 44000 Hz.

8.2.2 Matériel de la tâche PERC^{CP1}

Nous avons proposé 48 phrases en français : 24 sont des phrases tests et 24 sont des *fillers* qui sont construits explicitement pour éviter que l'enfant comprenne les buts de notre recherche. Les phrases tests et les phrases *fillers* sont mélangées en différentes séquences de présentation et regroupées en 6 blocs de 8 phrases.

Dans le tableau 20, nous résumons la séquence de présentation des phrases dans les différents blocs. Comme nous pouvons le voir, nous avons varié l'alternance entre les phrases tests bien formées pour lesquelles nous attendons une réponse affirmative (signalée dans le tableau 20 comme « T, oui »), les phrases tests avec des erreurs pour lesquelles nous attendons une réponse négative (signalée comme « T, non »), les phrases *fillers* sans fautes pour lesquelles nous attendons une réponse affirmative (signalée comme « F, oui ») et les phrases *fillers* avec des fautes pour lesquelles nous attendons une réponse négative (signalée comme « T, non »). L'ordre de présentation des blocs de phrases a été varié pour chaque enfant.

Blocs	Phrases							
	1	2	3	4	5	6	7	8
1	T, non	T, non	F, oui	F, non	T, oui	F, non	F, non	T, non
2	T, non	F, oui	T, non	T, oui	F, oui	F, non	F, non	T, non
3	F, oui	F, non	T, oui	F, non	T, non	T, non	F, non	T, oui
4	F, oui	T, oui	T, non	F, oui	T, non	F, oui	T, non	F, non
5	F, oui	T, non	T, non	T, oui	F, oui	F, oui	F, oui	T, non
6	T, non	F, oui	T, oui	T, non	F, non	T, non	F, oui	F, oui

Tableau 20. Alternance de phrases tests et *fillers* avec les réponses attendues dans les 6 blocs de PERC

Les 24 phrases *fillers* (Annexe 4) sont des phrases grammaticales ou des phrases modifiées phonologiquement (ex. par les processus de métathèse, d'épenthèse ou de délétion de segments) ou syntaxiquement (ex. non respect de l'accord sujet-verbe).

Les phrases tests sont présentées dans 8 contextes phonologiques et syntaxiques. Les contextes syntaxiques où ont lieu les phénomènes de resyllabation sont au nombre de trois : déterminant-nom, verbe-nom et mot invariable-nom. L'élision et l'enchaînement sont toujours présentés dans le contexte syntaxique déterminant-nom. Les contextes phonologiques sont variés seulement pour la liaison. Pour cela nous avons créé quatre contextes phonologiques : liaisons en [z], en [n] et en [t] plus des liaisons avec nom qui commence par *h* muet ou *h* aspiré. L'élision est toujours présentée dans le contexte phonologique /l/ + V, alors que l'enchaînement se fait toujours dans un contexte en [n].

Les phénomènes de resyllabation dans les différents contextes phonologiques et syntaxiques sont présentés en trois conditions :

- Condition 1. Consonne erronée (CONDcons.err) : *Le(s)[n]agneaux [le.na.no] broutent l'herbe*

Entre les deux mots créant une resyllabation est insérée une consonne possible mais erronée (ex. [n] dans *le(s)[n]agneaux* [le.na.no] au lieu de [z] dans *les[z]agneaux* [le.za.no]). L'erreur-cible à identifier est une erreur d'insertion d'une consonne erronée.

Pour le cas de l'élision, nous n'avons pas pu faire cette insertion car le phénomène est d'une autre nature que la liaison et l'enchaînement. Pour cette raison, nous avons décidé de laisser dans la condition CONDcons.err un hiatus (ex. *le artichaut*).

- Condition 2. non-réalisation (CONDnon-réal) : *Le(s)[Ø]anneaux [leØa.no] de cette chaîne sont cassés*

Le phrase présente un contexte où le phénomène de resyllabation n'est pas réalisé (*Le(s)[Ø]anneaux [leØa.no]* au lieu de *les[z]anneaux [le.za.no]*). L'erreur à identifier est une erreur de non-réalisation du phénomène de resyllabation.

Pour l'élision, CONDnon-réal présente une phrase similaire à CONDcons.err, sauf que les voyelles en hiatus sont deux voyelles différentes.

- Condition 3. Réalisation (CONDréal) : *Thibaut coupe des[z]ananas [de.za.na.na] en morceaux*

Le phénomène de resyllabation est réalisé. Aucune erreur ne doit être détectée.

Dans le tableau 21, nous présentons le répertoire des phrases tests dans ces trois conditions. Pour presque tous les contextes, la phrase présentée dans la condition CONDréal est une phrase correcte tandis que dans les autres conditions elle présente des fautes. La seule exception est le contexte de liaison déterminant-nom avec *h* aspiré pour CONDnon-réal et CONDréal. Ici, dans la condition CONDnon-réal la phrase est correcte, car la liaison ne peut pas être réalisée (ex. pour *Le(s)[Ø]hamsters [leØam.stɛʁ] de mon cousin sont trop mignons*, les enfants ne doivent pas repérer de fautes). En revanche, dans la condition CONDréal, la phrase avec la réalisation de la liaison doit être considérée comme erronée car, en présence d'une liaison interdite (ex. dans la phrase *Paul et Marcel mangent des[z]hamburgers [de.zam.bœʁ.gœʁ]*, la séquence *des[z]hamburgers [de.zam.bœʁ.gœʁ]* doit être signalée comme erronée).

Cette tâche est précédée d'un entraînement constitué de 10 phrases qui sont présentées dans les annexes (Annexe 5).

Contextes	CONDcons.err	CONDnon-réal	CONDréal
Liaison en [z] déterminant-nom	Le(s)[n]agneaux [le.na.ɲo] broutent l'herbe.	Le(s)[Ø]anneaux [leØa.no] de cette chaîne sont cassés.	Thibaut coupe des[z]ananas [de.za.na.na] en morceaux.
Liaison en [n] déterminant-nom	Cet exercice est trop dur, encore u(n) [z]essai [ɛ.ze.sɛ] !	Je ferai u(n)[Ø]effort [ɛØe.fɔʁ] pour être là cet après-midi.	Ce médicament lui a fait un[n]effet [ɛ.ne.fɛ] très fort et il a guéri tout suite.
Liaison verbe- nom	Es(t)-[z]on [ɛ.zɔ̃] bientôt arrivés à la maison ?	Es(t)[Ø]-on [ɛØɔ̃] déjà arrivés ?	Peut[t]-on [pø.tɔ̃] savoir tes secrets ?
Liaison mot invariable -nom	Je suis rentré de l'hôpital en[z]ambulance [ɛ.zɑ̃.by.lɑ̃s].	Il est parti à l'école e(n)[Ø]avance [ɛØa.vɑ̃s].	Ce livre est très[z]ennuyeux [tʁɛ.zɑ̃.nɥi.jø].
Enchaînement dét-nom	Paul a oublié une[z]affaire [yn.za.fɛʁ] à l'école.	Un(e) agence [ynØa.ʒɑ̃s] de voyage est un magasin où ils organisent des vacances	Maxime a vu une affiche [y.na.fiʃ] du cirque.
Élision dét.-nom	Le artichaut [ləØɑʁ.ti.ʃo] est un bon légume.	La armoire [laØɑʁ.mwaʁ] de ma chambre est bien rangée.	J'ai découpé l'article [laʁ.tikl] sur les lions dans mon journal.
Liaison dét-nom avec <i>h</i> muet	Ma sœur doit s'acheter de(s) [n]habits [de.na.bi].	Le(s)[Ø]habitudes [leØa.bi.tyd] de ce chien doivent changer, il n'est pas sage !	Les[z]habitants [le.za.bi.tɑ̃] de Paris s'appellent les Parisiens.
Liaison dét-nom avec <i>h</i> aspiré	On a trouvé de(s) [n]haricots [de.na.ʁi.ko] dans le jardin.	Le(s)[Ø]hamsters [leØam.stɛʁ] de mon cousin sont trop mignons.	Paul et Marcel mangent des[z]hamburgers [de.zam.bœʁ.gœʁ].

Tableau 21. Répertoire de phrases tests dans les 3 conditions pour les contextes dans PERC

8.2.3 Le traitement de données pour la tâche PERC^{CPI}

Pour la tâche de perception d'erreurs, si nous présentons une phrase juste ou une phrase fautive, nous avons deux réponses possibles qui donnent lieu à deux évaluations possibles (correcte ou erronée), comme le montre le tableau 22 :

Phrase	Réponse enfant	Évaluation donnée
Juste	oui	Correcte
	non	Erronée
Fausse	non	Correcte
	oui	Erronée

Tableau 22. Réponses possibles et évaluations données dans PERC

Dans une première analyse « brute » de nos résultats, les réponses « non » à des phrases fausses et « oui » à des phrases justes doivent être évaluées comme des réponses correctes. En revanche, les réponses « oui » à des phrases fausses et « non » à des phrases justes doivent être comptées comme des erreurs. Ce sont les résultats de ce type d'analyse « brute » que nous présentons d'abord.

La tâche PERC^{CPI} ne se limite pas seulement à l'enregistrement des réponses « oui » ou « non ». Elle se compose aussi d'une partie de reformulation qui permet de rendre compte des critères utilisés par l'enfant dans le jugement et de vérifier l'identification de l'erreur-cible. Nous pouvons distinguer des réponses basées sur des critères plutôt sémantiques (où les enfants retrouvent des erreurs sémantiques dans la phrase) ou lexicaux (où les enfants considèrent comme fausse une phrase parce qu'elle contient un mot qu'ils ne connaissent pas). Les élèves peuvent aussi ne pas réussir à motiver leur jugement. La correction d'erreurs permet d'éliminer, dans la somme du total brut, les fautes qui ne nous intéressent pas, comme les fautes sémantiques-lexicales ou les fautes immotivées, et d'affiner l'analyse brute.

8.2.4 Résultats de la tâche PERC^{CPI}

Comme nous l'avons dit dans le paragraphe précédent, les premiers résultats que nous présentons sont des résultats bruts qui ne tiennent pas compte des reformulations produites par les enfants. Nous avons calculé les scores moyens de réponses correctes pour chaque condition (Tableau 23). Le score va de 0 (0 réponses correctes) à 8 (toutes les réponses sont correctes). Nous avons écarté des résultats les données de deux enfants (4,6 %) qui ne comprenaient pas ce qu'il fallait faire dans la tâche.

Conditions	Scores moyens de réponses correctes
CONDcons.err (<i>le(s)[n]agneaux</i>)	3,7
CONDnon-réal (<i>le(s)[Ø]agneaux</i>)	4,07
CONDréal (<i>les[z]ananas</i>)	6,50

Tableau 23. Résultats de PERC^{CP1} pour les trois conditions

L'analyse de la variance (ANOVA) des scores de réponses correctes dans trois modalités (types de conditions) utilisant les sujets et les items comme variables aléatoires montre un effet du facteur « condition » ($F_1(2,80) = 21,08, p < .05$; $F_2(2,14) = 7,490, p < .001$). L'analyse des contrastes montre que la CONDréal implique plus de réponses correctes que les deux autres conditions (CONDcons.err vs. CONDnon-réal $F_1(1,40) = .32, p = .57$; $F_2(1,7) = .017, p = .96$; CONDcons.err et CONDnon-réal vs. CONDréal $F_1(1,40) = 52,67 p < .05$; $F_2(1,7) = 14,402, p < .001$).

Pour mieux comprendre ce qui a guidé les enfants à faire leur choix, nous avons décidé d'analyser les reformulations en séparant les trois conditions dans les trois sections suivantes.

8.2.4.1 Résultats pour CONDcons.err (*le(s)[n]agneaux*)

Pour affiner le traitement des données pour CONDcons.err, nous avons analysé les reformulations et les avons classifiées en différents groupes :

- **Groupe « sémantique »** : dans ce groupe, nous considérons les enfants qui trouvent des erreurs sémantiques ou lexicales sur la phrase test.
- **Groupe « sans motivation »** : dans ce groupe, nous insérons les enfants qui ne motivent pas la réponse. Dans ce cas, les enfants trouvent que la phrase n'est pas bonne mais ils ne peuvent pas dire pourquoi.
- **Groupe « cibles »** : dans ce groupe, nous mettons les enfants qui trouvent l'erreur « cible » explicitement insérée.

Dans le tableau 24 nous présentons les résultats et la classification présentant le nombre d'enfants pour chaque groupe.

	Nombre d'enfants qui identifient une erreur du type :		
	« sémantique »	« sans motivation »	« cibles »
Liaison en [z]	5	2	18
Liaison en [n]	0	4	22
Liaison verbe-nom	7	2	9
Liaison mot invariable-nom	3	0	26
Enchaînement	0	2	32
Élision	3	2	12
Liaison avec <i>h</i> muet	3	2	24
Liaison avec <i>h</i> aspiré	3	2	28

Tableau 24. Classification des erreurs des enfants après l'analyse des corrections pour CONDcons.err dans PERC^{CP1}

La plupart des enfants arrivent à détecter l'erreur-cible. Cependant, nous remarquons des enfants qui signalent des erreurs sémantiques. Trois enfants jugent la première phrase « Le(s)[n]agneaux [le.na.ɲo] broutent l'herbe » comme fautive parce que ce ne sont pas les agneaux qui broutent l'herbe mais plutôt les moutons. L'enfant NB dit que les agneaux ne broutent pas l'herbe mais plutôt ils sentent l'herbe. Dans la phrase « Ma sœur doit s'acheter de(s)[n]habits [de.na.bi] », pour l'enfant EF, la faute n'est pas dans la CL erronée mais dans le fait qu'on n'achète pas des habits mais des vêtements. Enfin, pour donner un dernier exemple, l'enfant CDL nous dit qu'on ne trouve pas des haricots dans le jardin mais des abricots et que c'est pour ça que, selon lui, la phrase « On a trouvés *de(s)[n]haricots* [de.na.ʁi.ko] dans le jardin » est erronée.

Comme notre recherche considère les phénomènes de resyllabation, nous voulons voir si dans cette tâche au CP, les enfants sont capables de reconnaître les erreurs d'insertion d'une consonne erronée, insérées explicitement dans les différents contextes. Pour cette raison, nous proposons le pourcentage d'« identification » de l'erreur-cible (Tableau 25).

CONDcons.err		% d'identification de l'erreur-cible
Liaison	Déterminant-nom (en [z])	41,8
	Déterminant-nom (en [n])	46,5
	Verbe-nom	51,2
	Mot invariable-nom	55,8
Enchaînement		39,5
Élisions		53,4
Liaison	<i>h</i> muet	39,5
	<i>h</i> aspiré	13,5

Tableau 25. Identifications de l'erreur-cible pour CONDcons.err en PERC^{CP1}

Les contextes qui créent le plus de problèmes (moins de 50 % d'identifications) correspondent aux phrases où on présente une erreur de liaison entre déterminant et nom (en [z] 41,8 %, en [n] 46,5 %, *h* muet 39,5 %) et une erreur d'enchaînement (39,5 %). Pour les autres contextes, les enfants réussissent plus facilement à retrouver l'erreur-cible. Nous remarquons que, dans le dernier contexte « liaison *h* aspiré », 13,5 % des enfants reconnaissent la faute de CL erronée mais à la demande de reformulation, 60,7 % produisent une liaison avec *h* aspiré du type *le[z]haricots*. La reformulation de la phrase dans ce cas-là correspond à une production erronée car la liaison avec *h* aspiré est une liaison interdite.

8.2.4.2 Résultats pour CONDnon-réal (*le(s)[Ø]anneaux*)

La deuxième condition que nous avons traitée est CONDnon-réal (*le(s)[Ø]anneaux*), la condition de non-réalisation d'un phénomène de resyllabation. Dans cette section, nous proposons les mêmes analyses que celles proposées en 8.2.4. Cependant, il faut d'abord faire une distinction entre le contexte « liaison entre déterminant-nom avec *h* aspiré » et les autres contextes. Dans CONDnon-réal, la phrase avec Mot2 commençant par *h* aspiré est correcte. Les enfants qui n'identifient pas de fautes dans ce contexte (ex. dans la phrase « Le(s)[Ø]hamsters [leØam.stɛʁ] de mon cousin sont trop mignons ») répondent correctement. Au contraire, les enfants qui trouvent une faute dans ce contexte ne répondent pas correctement car la non-réalisation de liaison avec *h* aspiré est une production correcte et la phrase doit être jugée comme juste.

Nous allons voir plus en détail les différents types d'erreurs détectées en analysant les reformulations et les corrections que les enfants réalisent après l'écoute de la phrase (Tableau 26).

	Nombre d'enfants qui identifient une erreur du type :		
	« sémantique »	« sans motivation »	« cible »
Liaison en [z]	1	2	5
Liaison en [n]	2	1	17
Liaison verbe-nom	11	2	9
Liaison mot invariable-nom	2	3	19
Enchaînement	6	6	6
Élision	6	4	13
Liaison avec <i>h</i> muet	3	2	12
Liaison avec <i>h</i> aspiré	0	0	6

Tableau 26. Classification des erreurs après l'analyse des reformulations pour CONDnon-réal dans PERC^{CP1}

Dans CONDnon-réal, les erreurs sémantiques et lexicales sont plus fréquentes que dans CONDcons.err. La plupart des erreurs sémantiques et lexicales sont enregistrées dans le contexte « liaison verbe-nom » où beaucoup d'enfant changent la phrase « Es(t)[Ø]-on [εØɔ̃] déjà arrivés? » en proposant des réinterprétations comme « des gens est déjà arrivé ». Dans le contexte d'enchaînement déterminant-nom, l'agence de voyage devient pour un enfant « une agence du cinéma » ou pour un autre « une agence de voyage pour les vacances ». Dans la non-réalisation d'élision, *la armoire* bien rangée devient « pas rangée » ou « pas mangée ». En contexte de liaison déterminant-nom en *h* aspiré, les enfants SV et LT, nous disent que la phrase est fautive car les habitudes d'un chien ne peuvent pas changer et l'enfant IV nous répond « il y a une faute parce que si le chien n'est pas sage, on donne des coups de martinet ».

Nous présentons les identifications de l'erreur-cible de non-réalisation dans le tableau 27. Encore une fois, il faut faire attention au contexte avec *h* aspiré car la phrase en CONDnon-réal ne présente pas une erreur phonologique. Ici, les 6 enfants qui ont identifié une faute (soit 13,9 %) considèrent le contexte avec *h* aspiré comme le contexte avec *h* muet, en réalisant la liaison en [z] dans la phase de correction.

CONDnon-réal		% Identification de l'erreur-cible
Liaison	déterminant-nom (en [z])	34,9
	déterminant-nom (en [n])	49,5
	Verbe-nom	20,9
	Mot invariable-nom	41
Enchaînement		13,9
Élision		30,2
Liaison	<i>h</i> muet	27,9
	<i>h</i> aspiré	13,9

Tableau 27. Identifications de l'erreur-cible pour CONDnon.réal dans PERC^{CP1}

Pour cette condition, dans tous les contextes les enfants identifient très difficilement l'erreur-cible (moins de 50 % d'identifications). Les enfants acceptent plus facilement la non-réalisation dans les contextes : liaison verbe-nom (20,9 % d'identifications), enchaînement déterminant-nom (13,9 %), élision déterminant-nom (30,2 %) et liaison déterminant-nom avec *h* muet (27,9 %).

8.2.4.3 Résultats pour CONDréal (*des[z]ananas*)

La dernière condition test est CONDréal (*des[z]ananas*), la réalisation du phénomène de resyllabation. Dans ce cas, les enfants doivent juger la phrase comme correcte. La seule exception est la condition « liaison déterminant-nom avec *h* aspiré », où la réalisation de la liaison doit être considérée comme une erreur et la phrase correspondante doit être traitée comme fausse.

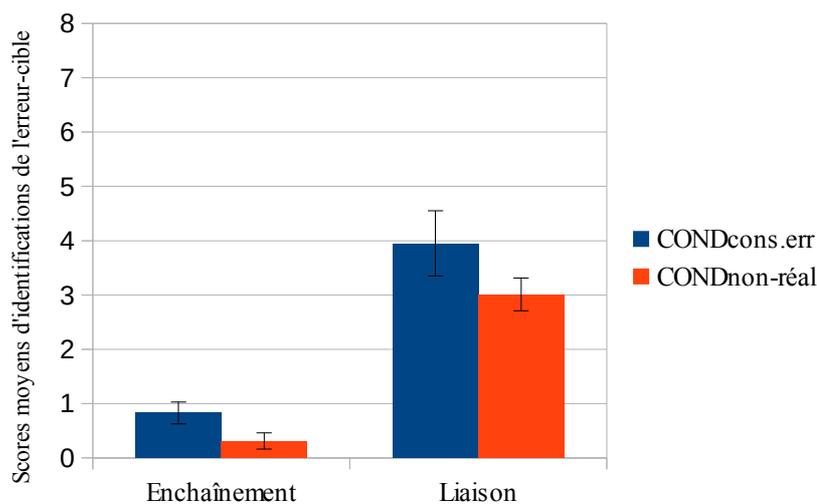
En dehors du contexte particulier avec *h* aspiré, la plupart des enfants qui trouvent une faute ne donnent aucune motivation ou trouvent des erreurs sémantiques ou lexicales. Par exemple, dans la phrase « Ce médicament lui a fait *un[n]effet* [ɛ̃.ne.fɛ] très fort et il a guéri tout suite », deux enfants nous disent que la phrase est erronée parce que les médicaments ne font pas un effet tout de suite mais qu'« il faut attendre ». Pour la phrase avec élision « J'ai découpé l'article [laʁ.tikl] sur les lions dans mon journal », les quatre enfants qui signalent des erreurs nous suggèrent qu'on ne découpe pas l'article mais les images. Dans la phrase « Ce livre est très[z]ennuyeux [tʁɛ.za ã.nɥi.jø] », quatre enfants ne considèrent pas la possibilité qu'un livre puisse être

ennuyeux. Puisque les enfants ne repèrent pas de problèmes phonologiques dans cette condition, nous ne proposons pas une analyse détaillée des identifications, contrairement à ce que nous avons fait pour CONDcons.err ou CONDnon-réal.

8.2.4.4 Comparaison entre les conditions : CONDnon-réal vs. CONDcons.err

Nous avons vu que les enfants ont des difficultés pour CONDnon-réal et CONDcons.err. Cela veut dire que, pour les enfants, certaines non-réalisations ou certaines utilisations d'une consonne erronée sont acceptables dans la langue orale. Nous allons comparer les cas où les enfants ont identifié d'erreurs phonologiques (erreurs-cibles) pour la CONDcons.err et pour la CONDnon-réal afin de comprendre comment ils jugent ces deux façons de traiter les frontières de mot en fonction du phénomène de resyllabation testé.

Nous comparons les scores d'identification de l'erreur-cible sur les deux CONDnon-réal et CONDcons.err (le score maximal est 8) et nous observons qu'il y a une différence statistiquement significative pour la liaison ($t_1(40) = 2,49$, $p < .05$, $t_2(5) = 0,30$, $p < .05$) et pour l'enchaînement ($t_1(40) = 5,92$, $p < .05$ ¹⁸) mais pas pour l'élision ($t_1(40) = 1,53$, $p = .13$). Le graphique 2 montre clairement que le score moyen d'identification de l'erreur-cible est plus élevé dans CONDcons.err pour ces deux phénomènes.



Graphique 2. Identification de l'erreur cible dans CONDcons.err et CONDnon-réal

18 Pour l'élision et l'enchaînement, nous avons seulement un item par condition. Nous ne pouvons pas calculer le t_2 .

Cela veut dire qu'à l'oral pour l'enchaînement et la liaison, les enfants au début du CP acceptent davantage une non-réalisation que l'ajout d'un segment erroné entre Mot1 et Mot2. Ils acceptent plus facilement une construction qui présente les mots dans leurs formes lexicales séparées et qui gardent leurs frontières intactes ([leØa.no]) qu'une construction qui altère ces formes lexicales.

Pour l'élosion, nous avons remarqué les mêmes difficultés pour les deux conditions. En effet, nous proposons dans les deux cas un jugement sur des séquences similaires, où nous ajoutons une pause entre déterminant et nom : soit *le artichaut* (CONDcons.err), soit *la armoire* (CONDnon-réal). Ces deux constructions sont perçues de la même façon et jugées avec la même difficulté, car il n'y a pas une véritable différence dans le traitement des frontières entre mots pour les deux. Ce résultat fait émerger une problématique liée à la construction de PERC^{CP1}. Nous allons décrire plus en détail les limites de cette tâche dans la section suivante.

8.2.5 Les limites de la tâche PERC^{CP1}

La tâche de perception d'erreurs et de reformulation est une tâche métalinguistique qui permet de tester les processus « de type C », les processus cognitifs où le participant doit être conscient pour traiter l'information (Jack & Shallice, 2001). Étant donné que le processus cognitif qui est testé dans la tâche PERC^{CP1} est un processus du type « conscient », cette tâche ne permet pas d'avoir des informations directes sur la forme sous-jacente. En revanche, elle permet de vérifier des compétences très larges associées à la compréhension et au jugement normatif qui à leur tour sont sujets à des variables individuelles.

Dans l'activité de compréhension, la mémoire à court terme et le niveau d'attention jouent un rôle. Le fait de ne pas avoir détecté une certaine erreur-cible peut être l'effet d'un niveau d'attention trop bas ou d'un problème de mémoire à court terme : la phrase que l'enfant écoute est déjà oubliée en phase de jugement. Par ailleurs, nous avons remarqué que certains enfants n'étaient pas capables de reformuler la phrase écoutée, soit parce qu'ils l'avaient oubliée soit parce qu'ils ne réussissent pas à se concentrer au moment de l'écoute.

De plus, puisque la tâche teste un processus cognitif conscient et qu'elle a pour but l'identification d'erreurs, les réponses qu'elle implique peuvent être le résultat d'un jugement normatif, sur la langue. Un enfant qui commence l'école primaire est déjà

familier avec le concept d'erreur, mais ce concept est très loin du concept linguistique de l'« agrammaticalité ». Il est plutôt proche de l'idée de quelque chose d'interdit, d'incorrect ou « qu'il ne faut pas faire ». Les résultats de cette tâche peuvent donc être influencés par le jugement normatif lié à une évaluation plutôt négative de l'erreur, propre aux enfants à cet âge. En plus, comme pour la compréhension, le jugement normatif est à son tour influencé par des variables individuelles comme le contexte social dans lequel l'enfant est inséré, ou l'éducation en famille et à l'école.

Enfin, PERC^{CP1} propose différents contextes et conditions qui ne sont pas bien équilibrés du point de vue expérimental. Par exemple, nous avons 4 contextes de liaison opposés à un seul contexte d'enchaînement et à un seul contexte d'élision. Dans l'élision, nous avons deux conditions équivalentes : pour la condition de non-réalisation nous proposons la faute *la armoire* et pour la condition de consonne erronée nous proposons une séquence similaire du type *le artichaut*.

Étant donné que pour vérifier nos hypothèses, nous voulons recueillir des informations sur la forme sous-jacente qui opère dans la segmentation, nous chercherons à créer une tâche qui permette d'évaluer le traitement non-conscient de l'information et qui soit aussi plus équilibrée du point de vue expérimental. La meilleure façon pour tester le processus non-conscient est une tâche psycholinguistique de *priming*, ou amorçage. Dans la littérature, les exemples de tâches de *priming* qui testent les phénomènes de resyllabation ont comme seuls participants des adultes. Aucun exemple de ce type d'expérience n'a été réalisé chez les enfants au début du CP. Pour cette raison, à ce point de la recherche, nous avons préféré faire référence à la littérature existante adaptée à l'âge de nos participants et nous avons décidé de construire une tâche qui teste la production. Nous avons repris une tâche de dénomination d'image utilisée pour tester la réalisation de la liaison (Braud & Wauquier, 2005 ; Dugua, 2006 ; Splendido 2014). Cela nous semble plus susceptible de mettre en évidence l'existence de représentations sous-jacentes des phénomènes de resyllabation mais aussi de vérifier si les problématiques de non-identification de certaines formes que nous avons trouvées dans la tâche de perception correspondent à des séquences réellement produites.

8.3 DENa^{CP1} : la tâche de dénomination guidée au début du CP

8.3.1 Procédure de la tâche DENa^{CP1}

Dix-huit couples de cartes avec des images colorées d'animaux et d'objets sont présentés un par un à l'enfant testé. La moitié représente les tests et l'autre moitié sont des *fillers*. Dans les 9 couples tests, 3 prévoient la production de liaisons, 3 la production d'élisions et 3 la production d'enchaînements. Dans chaque carte du couple, l'objet ou l'animal représenté est le même pour les deux cartes. La seule différence est qu'il y a un seul élément dans l'un (ex. un lapin) et que dans l'autre, il y en a plusieurs (ex. des lapins), de telle sorte que l'enfant doit, soit produire le singulier à partir du pluriel, soit produire le pluriel à partir du singulier.

L'expérimentateur montre la première carte du couple et donne comme input le nom précédé du déterminant, toujours l'article défini (ex. *l'ours*) et dans le cas de l'enchaînement, avec l'adjectif antéposé (ex. *la grosse araignée*) ; il présente ensuite la deuxième carte du couple et demande à l'enfant de dire ce qui est dessiné en tenant compte de ce qu'il avait prononcé pour la première carte. La cible est donc la dénomination de la représentation dans la deuxième carte précédée par l'article défini ou par l'adjectif associé (ex. *les ours* ou *les grosses araignées*).

Les enfants sont testés un par un dans une chambre silencieuse dans les bâtiments de l'école. Les productions orales sont enregistrées au moyen d'un enregistreur *Zoom H4* à la fréquence de 44000 Hz au format *.wav*.

8.3.2 Matériel de la tâche DENa^{CP1}

Dans les couples tests, les inputs sont des mots en contexte de liaisons, d'élisions et d'enchaînements. Les enfants doivent déconstruire les phénomènes consonantiques pour isoler les mots. Ensuite, pour produire les cibles, ils doivent ré-enchaîner ce qu'ils ont isolé. Les mots tests donnés par l'expérimentateur et les réponses attendues par les enfants sont présentés dans le tableau 28.

Input de l'expérimentateur : élision → Production enfant : liaison		
Contexte	Input	Réponse attendue
Déterminant-nom	l'ours [luʁs]	les ours [le.zuʁs]
Déterminant-nom	l'armoire [laʁ.mwaʁ]	les armoires [le.zaʁ.mwaʁ]
Déterminant-nom	l'arrosoir [la.ʁo.zwaʁ]	les arrosoirs [le.za.ʁo.zwaʁ]
Input de l'expérimentateur : liaison → Production enfant : élision		
Contexte	Input	Réponse attendue
Déterminant-nom	les oies [le.zwa]	l'oie [lwa]
Déterminant-nom	les orages [le.zɔ.ʁaʒ]	l'orage [lɔ.ʁaʒ]
Déterminant-nom	les oreillers [le.zɔ.ʁɛ.jɛ]	l'oreiller [lɔ.ʁɛ.jɛ]
Input de l'expérimentateur : liaison → Production enfant : enchaînement		
Contexte	Input : expérimentateur	Réponse attendue : enfant
Adjectif-nom	les minuscules ogres [le.mi.ny.sky.l.zɔgʁ]	le minuscule ogre [lə.mi.ny.sky.lɔgʁ]
Adjectif-nom	les petites oranges [le.pə.tit.zo.ʁɑ̃ʒ]	la petite orange [la.pə.ti.to.ʁɑ̃ʒ]
Adjectif-nom	les grosses araignées [le.gʁo.za.ʁɛ.jɛ]	la grosse araignée [la.gʁo.sa.ʁɛ.jɛ]

Tableau 28. Répertoire des séquences tests pour DENa

Cette tâche est précédée d'un long entraînement de 9 couples de cartes pour faire comprendre à l'enfant qu'il doit nommer la deuxième carte en tenant compte de ce que l'expérimentateur produit (Annexe 6). Les couples d'images tests proposés sont présentés en même temps que les 9 couples *fillers* (Annexe 7).

Nous avons créé 3 blocs de 6 productions avec 3 couples tests et 3 *fillers*. Les alternances entre les couples tests (T dans le tableau 29) et les couples *fillers* (F, dans le

tableau 29) sont variées dans les trois blocs comme il est montré dans le tableau 29. Dans chaque bloc apparaît systématiquement une production de liaison (dénommé « T, liaison »), une d'élision (dénommé « T, élision ») et une d'enchaînement (dénommé « T, enchaînement »). L'ordre de présentation des blocs varie selon les enfants.

BLOCS	Couple de cartes présenté					
	1	T, liaison	F	F	F	T, élision
2	T, liaison	T, élision	F	F	T, enchaînement	F
3	F	T, liaison	F	T, élision	F	T, enchaînement

Tableau 29. Ordre de présentation du répertoire pour DENa

8.3.3 Résultats de la tâche DENa^{CP1}

Nous avons écarté les productions d'un enfant (2 %) qui montrent qu'il n'a pas compris la tâche. Dans les résultats, nous avons des erreurs de tâches (soit la dénomination erronée de la carte, ex. *une abeille* pour *les[z]ours* ou l'absence de réponse) et des erreurs de segmentation qui montrent des difficultés à déterminer les frontières des mots à produire et à gérer les phénomènes de resyllabation.

Nous avons considéré les scores moyens d'erreurs en production (Tableau 30). Le score maximal total correspond à 18. Pour les items tests et pour les *fillers*, ce score peut arriver jusqu'à 9. Nous avons plus d'erreurs sur les productions d'items tests (4,3 vs. 1,4 sur les productions totales, $t_1(41) = 1,178$, $p < .001$, $t_2(8) = 4,811$, $p < .001$) et plus d'erreurs de segmentation que d'erreurs de la tâche (3,9 vs. 1,8 sur les erreurs totales $t_1(41) = 1,854$, $p < .001$; $t_2(17) = 3,13$, $p < .001$ et sur les item tests $t_1(41) = 2,546$, $p < .001$; $t_2(8) = 2,99$, $p < .05$).

	Scores moyens		
	erreurs de la tâche (ex. <i>une abeille</i> pour <i>les[z]ours</i> ou absence de réponse)	erreurs de segmentation (ex. <i>le(s)[n]ours</i> ou <i>les[Ø]ours</i> pour <i>le(s)[z]ours</i>)	TOTAL
Items <i>fillers</i>	0,7	0,7	1,4
Items tests	1,1	3,2	4,3
Tests + <i>fillers</i>	1,8	3,9	5,7

Tableau 30. Résultats généraux pour DENa^{CP1}

Nous analysons plus en détail les erreurs de segmentation (Tableau 31). Nous classifions les erreurs sous trois catégories :

- « non-réalisation » qui implique la non production du phénomène de resyllabation (ex *le(s)[Ø]ours* [leuʁs] pour *les[z]ours* [le.zuʁs] ; *la-armoire* [la.aʁ.mwaʁ] pour *l'armoire* [laʁ.mwaʁ] ; *la grosse-araignée* [la.gʁos.aʁɛne] pour [la.gʁo.sa.ʁɛ.ne]).
- « insertion d'une consonne erronée », où les enfants produisent une consonne différente de celle qui est attendue dans les contextes de liaison, d'élision et d'enchaînement (ex *le(s)[n]ours* [le.nuʁs] pour *les[z]ours* [le.zuʁs] ; *le [z]armoire* [lə.zaʁ.mwaʁ] pour *l'armoire* [laʁ.mwaʁ] ; *la grosse[z]araignée* [la.gʁo.za.ʁɛ.ne] pour *la grosse[s]araignée* [la.gʁo.sa.ʁɛ.ne]).
- « autres erreurs » qui sont plutôt des répétitions de l'input donné ou de mauvaises utilisations de l'article (ex. *le* pour *la* dans *le grosse araignée* [lə.gʁo.sa.ʁɛ.ne] pour [la.gʁo.sa.ʁɛ.ne]).

Nous présentons aussi les scores globaux de l'ensemble des différents types d'erreurs pour les trois phénomènes dans la deuxième colonne du tableau 31. Pour chaque phénomène de resyllabation, le score maximal est 3.

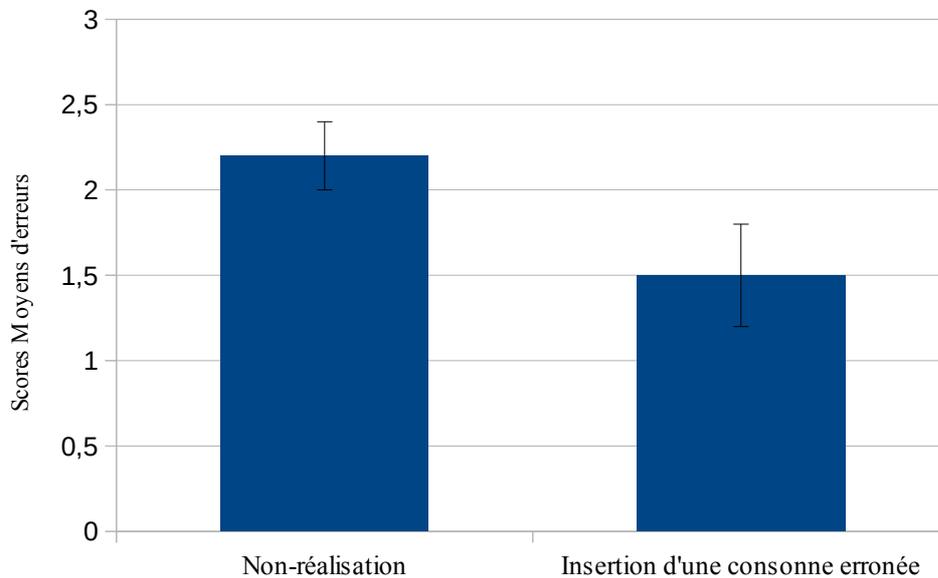
Phénomène	Scores moyens d'erreurs TOTAL	Types d'erreurs	Scores moyens par type d'erreur
Liaison	0,8	Non-réalisation (<i>le(s)[Ø]ours</i> [le.Øuɔʁs] pour [le.zuɔʁs])	0,5
		Insertion d'une consonne erronée (<i>les[n]ours</i> [le.nuɔʁs] pour [le.zuɔʁs])	0,3
Élision	1,8	Non-réalisation (<i>la-armoire</i> [la.aʁ.mwaʁ] pour [laʁ.mwaʁ])	1
		Insertion d'une consonne erronée (<i>la[n]armoire</i> [la.naʁ.mwaʁ] pour [laʁ.mwaʁ])	0,4
		Autres erreurs	0,4
Enchaînement	1,1	Non-réalisation (<i>la grosse-araignée</i> [la.gʁos.aʁɛ.ɲe] pour [la.gʁo.sa.ʁɛ.ɲe])	0,4
		Insertion d'une consonne erronée (<i>la grosse[z]araignée</i> [la.gʁo.za.ʁɛ.ɲe] pour [la.gʁo.sa.ʁɛ.ɲe])	0,3
		Autres erreurs	0,4

Tableau 31. Résultats pour chaque phénomène de resyllabation dans DENa^{CPI}

L'analyse de la variance (ANOVA) sur les scores d'erreurs pour trois modalités (trois phénomènes) utilisant les sujets et les items comme variables aléatoires montre un effet du facteur « phénomène » ($F_1(2, 82) = 15,335, p < .001$; $F_2(2, 4) = 24,167, p < .001$). L'analyse des contrastes montre que pour les erreurs, la différence est statistiquement significative pour les trois phénomènes (liaison vs élision $F_1(1,41) = 31,093, p < .001, F_2(1,2) = 52,56, p < .05$; liaison vs enchaînement $F_1(1,41) = 8,956, p < .001, F_2(1, 2) = 10,77, p < .05$; élision vs enchaînement $F_1(1,41) = 6,62, p < .05, F_2(1,2) = 8,89, p = .09$)¹⁹. Les erreurs sont plutôt concentrées sur les phénomènes d'élision et d'enchaînement. En analysant les scores d'erreurs par type (dernière colonne du tableau) nous observons une différence statistiquement significative entre les non-réalisations et les insertions de consonne erronée dans les

19 Dans ces résultats, F est calculé sur le nombre d'items testés par phénomène. Nous avons seulement 3 items par phénomènes, pour cette raison les résultats pour F_2 ne sont pas significatifs. Pour F_1 , qui est calculé sur le sujet, nous avons 42 scores. Il faut donc considérer les valeurs de F_1 qui ne sont pas influencées par le nombre très réduit des données.

trois phénomènes mélangés ($t_1(41) = 3,528$, $p < .001$, $t_2(8) = 2,34$, $p < .05$) et que la plupart des erreurs sont des non-réalisations du phénomène de resyllabation (Graphique 3).



Graphique 3. Non-réalisations vs. Insertions d'une consonne erronée dans DENa^{CPI}

L'analyse de la variance (ANOVA) pour les scores de non-réalisations pour trois modalités (trois phénomènes) utilisant les sujets et les items comme variables aléatoires montre un effet du facteur « phénomène » ($F_1(2,82) = 11,597$, $p < .001$; $F_2(2,4) = 19,137$, $p < .001$). L'analyse des contrastes montre qu'il existe une différence statistiquement significative entre l'élision et les autres phénomènes (liaison et enchaînement vs. élision $F_1(1,42) = 15,870$; $p < .001$; $F_2(1, 2) = 5,47$, $p < .05$) qui n'est pas observé pour la liaison et l'enchaînement (liaison vs. enchaînement $F_1(1,42) = 0,85$, $p = .35$; $F_2(1,2) = 0,37$, $p = .25$). Le phénomène d'élision implique plus d'erreurs de non-réalisation que les deux autres phénomènes.

L'analyse de la variance (ANOVA) pour les scores d'insertion d'une consonne erronée pour trois modalités (trois phénomènes) utilisant les sujets et les items comme variables aléatoires montre un effet du facteur « phénomène » ($F_1(2,82) = 13,444$; $p < .001$; $F_2(2,4) = 11,87$, $p < .05$). L'analyse des contrastes montre qu'il existe une différence statistiquement significative entre les trois phénomènes (liaison vs. élision $F_1(1,42) = 26,503$, $p = .35$; $F_2(1,2) = 3,57$, $p < .05$; liaison vs. enchaînement

$F_1(1,42) = 8,365$, $p < .001$; $F_2(1,2) = 6,78$, $p < .05$; élision vs. enchaînement $F_1(1,42) = 6,55$, $p < .05$; $F_2(1,2) = 5,53$, $p < .05$). La plupart des erreurs d'insertion d'une consonne erronée sont enregistrées dans ce qui aurait dû être les productions d'élision (Tableau 31). Ces erreurs produites dans le cas d'enchaînement sont moins nombreuses et elles sont encore plus rares dans la production de liaison.

En résumé, la plupart des erreurs de production sont enregistrées dans les contextes d'élision. Dans ces productions, les enfants ne réalisent pas le phénomène ou insèrent des consonnes erronées au début du Mot2. Nous analyserons ensuite ces séquences pour chercher à trouver une explication à ce comportement.

8.3.3.1 L'analyse phonétique des non-réalisations au début du CP

Comme nous l'avons souligné dans l'introduction de ce chapitre, nous voulons évaluer le traitement de la consonne qui se resyllabifie sur le Mot2 et l'éventuelle existence d'une position prosodique associée à sa réalisation. Cette évaluation vise à mieux décrire les représentations des phénomènes de resyllabation chez les enfants à ce stade de l'acquisition et à évaluer la pertinence des deux explications (lexicale-contructioniste et morpho-phonologique). Comme nous l'avons montré dans le chapitre 4, la grosse différence entre les deux explications sur l'acquisition de la liaison est l'association de cette position prosodique à une forme sous-jacente ou non. Pour évaluer l'existence de cette position sous-jacente, nous proposons une analyse phonétique de non-réalisations au moyen de PRAAT. Cette analyse fine nous permet d'abord d'évaluer phonétiquement l'espace sonore entre Mot1 et Mot2 et ensuite de valider ou non l'existence d'une possible position phonologique.

Nous avons décidé d'analyser phonétiquement les non-réalisations du type *le(s) [Ø]ours* [leØuʁs], *le-orage* [lə.o.ʁaʒ] ou *la petite-orange* [la.pə.tit.o.ʁãʒ] qui présentent la non-réalisation du phénomène consonantique entre Mot1 et Mot2. En analysant les spectrogrammes, nous pouvons regrouper les 63 non-réalisations enregistrées en quatre types différents :

- Continuum : où les enfants produisent un continuum entre la voyelle de l'article et celle du nom (comme dans la figure 27, pour le spectrogramme de [lə.o.ʁaʒ]).

- Coup de glotte : les enfants insèrent une occlusive glottale entre déterminant et nom, représentée par une barre d'explosion (Figure 28, spectrogramme de [lə.ʔə.ʁaʒ]).
- Ambiguë : les cas où l'enregistrement ne permet pas de bien évaluer la production (Figure 29).
- Pause : où il y a une pause phonétiquement entre Mot1 et Mot2, visible dans le spectrogramme comme « ligne de base », produite par l'appareil en l'absence de signal (Figure 30, spectrogramme de [le/ʁks]).

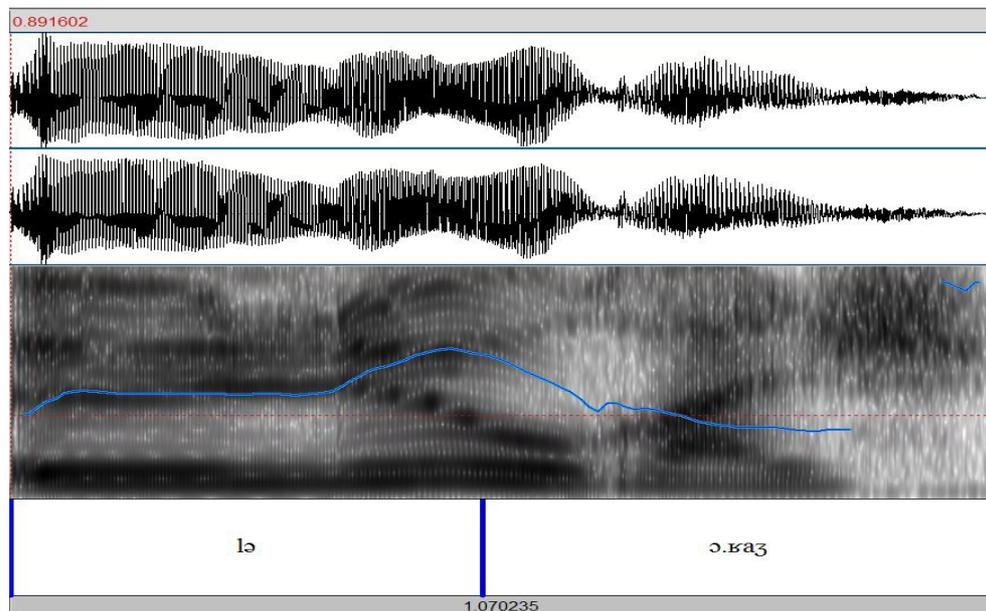


Figure 27. Spectrogramme pour la production de continuum

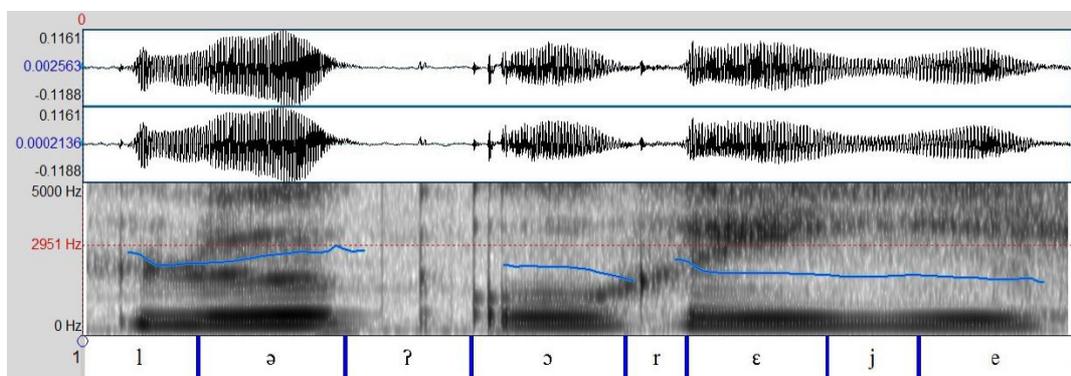


Figure 28. Spectrogramme pour l'insertion d'un coup de glotte

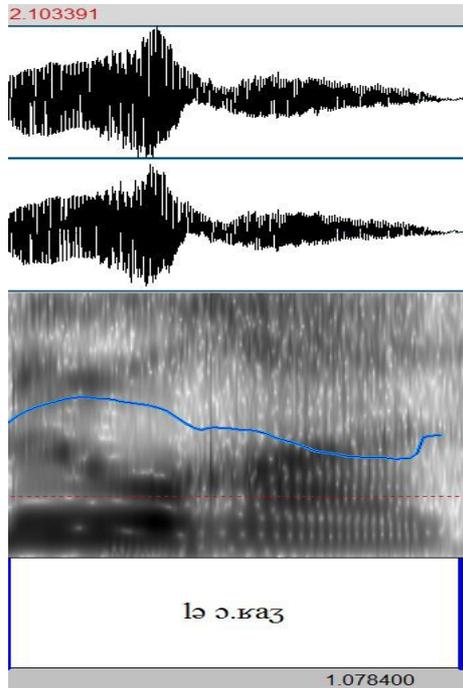


Figure 29. Spectrogramme pour une production ambiguë

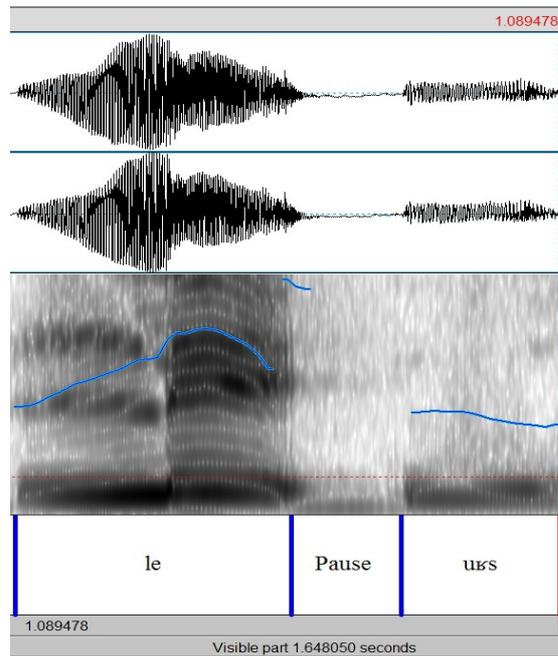


Figure 30. Spectrogramme pour la production d'une pause

Après avoir écarté les productions ambiguës (20,6 % de non-réalisations), nous considérons d'abord les productions de pauses, de continuums et de coups de glotte pour les trois phénomènes de resyllabation étudiés. Ensuite, nous concentrons notre analyse sur l'insertion des coups de glotte entre Mot1 et Mot2 pour mieux comprendre le traitement de l'espace sonore entre les deux mots et évaluer l'existence d'une position phonologique sous-jacente.

Pour la production de liaison, d'élision et d'enchaînement, les résultats de l'analyse phonétique sont présentés dans le tableau 32. Les pourcentages sont calculés sur le nombre de non-réalisations pour chaque phénomène dans la tâche DENa^{CP1}.

	%		
	liaison	élision	enchaînement
Coup de glotte ([le.ʔuʁs],[lə.ʔɔ.ʁε.je], [la.pə.tit.ʔo.ʁãʒ])	22,2	72,4	71,4
Pause ([le/uʁs], [lə/ɔ.ʁε.je], [la.pə.tit/o.ʁãʒ])	11,10	24,1	18,6
Continuum ([le.uʁs], [lə.ɔ.ʁε.je])	66,7	3,5	//

Tableau 32. Productions phonétiques dans les non-réalisations dans DENa^{CP1}

Nous pouvons voir que dans la plupart des non-réalisations en liaison, les enfants produisent un continuum (66,7 %) sans réaliser la CL qui doit normalement être insérée entre déterminant et nom.

Pour les contextes d'élision, dans la plupart des cas, les enfants qui ne produisent pas l'élision entre Mot1 et Mot2 introduisent un coup de glotte entre les deux mots (72,4 %).

En ce qui concerne l'enchaînement, nous avons seulement deux types de productions possibles²⁰ : insertion d'un coup de glotte ([la.pə.tit.ʔo.ʁãʒ]) ou pause ([la.pə.tit/o.ʁãʒ]). Dans les deux cas, la consonne d'enchaînement est bien réalisée mais il n'y a pas la resyllabation prévue par ce phénomène. Comme on peut le voir dans le tableau, dans 71,40 % des cas, les enfants insèrent un coup de glotte entre Mot1 et Mot2.

Nous avons décidé d'analyser les productions classifiées comme « coups de glotte », car elles sont des insertions d'une consonne entre Mot1 et Mot2 qui peuvent nous donner des informations sur l'existence d'une position sous-jacente, préalablement associée à la CL. Ces productions sont les seules que nous pouvons analyser en détail, car dans les productions de continuums et de pauses il n'y a pas de véritables réalisations sonores à analyser entre Mot1 et Mot2. Nous considérons les différents items à produire en relation avec le nombre d'insertions de coups de glotte (Tableau 33). Nous remarquons que la plupart des coups de glotte sont produits quand l'expérimentateur donne un input en liaison et demande à l'enfant de produire une

20 Pour l'enchaînement la production de « continuum » est une bonne production d'enchaînement, donc elle n'est pas présente dans la classification des « non-réalisations » de ce phénomène.

élision. La production d'élision était aussi la plus problématique dans la tâche de dénomination, comme nous l'avons vu dans la section 8.3.3. Dans les prochains paragraphes, nous proposons une analyse plus détaillée des non-réalisations d'élision.

Mot2	Phénomène à produire	Nombre de coups de glotte
ours	liaison	1
arrosoir	liaison	0
armoie	liaison	1
orage	élision	8
oreiller	élision	9
oie	élision	4
ogre	enchaînement	5

Tableau 33. Relation entre le phénomène à produire et la réalisation de coups de glotte dans DENa^{CP1}

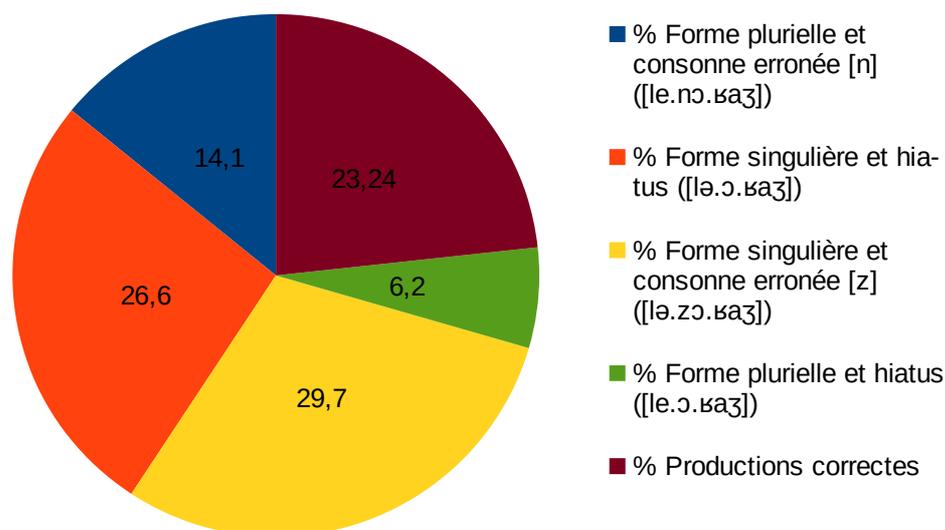
8.3.3.2 La production problématique d'élision

Nous avons vu que le phénomène qui crée le plus de problèmes aux enfants est l'élision (input expérimentateur : *les[z]orages* [le.zo.ʁaʒ] → production enfant : *l'orage* [lɔ.ʁaʒ]). Nous avons vu, dans la section 8.3.3, que le pourcentage de productions erronées est très élevé pour ce phénomène.

Les productions erronées d'élision sont réparties dans quatre catégories différentes :

1. forme plurielle et hiatus ([le.ɔ.ʁaʒ]),
2. forme plurielle et consonne erronée [n] ([le.nɔ.ʁaʒ]),
3. forme singulière et hiatus ([lə.ɔ.ʁaʒ]),
4. forme singulière et consonne erronée [z] ([lə.zɔ.ʁaʒ]).

Deux types de productions gardent l'article fléchi au pluriel comme il est donné dans l'input en liaison ([le.ɔ.ʁaʒ], [le.nɔ.ʁaʒ]). Les deux autres utilisent la forme de l'article au singulier mais montrent des erreurs sur les frontières : les enfants produisent un hiatus ([lə.ɔ.ʁaʒ]) ou insèrent une consonne erronée [z] ([lə.zɔ.ʁaʒ]) entre Mot1 et Mot2. Dans le graphique 4, nous présentons en détail les pourcentages pour ces quatre catégories de production.

Graphique 4. Pourcentages de formes produites pour l'élision dans DENa^{CPI}

En regardant le graphique, nous pouvons remarquer que dans les cas de deux types de productions ([lə.ɔ.ʁaʒ] et [lə.zɔ.ʁaʒ]), les pourcentages sont plus élevés que ceux des bonnes réponses (29,7 % et 26,6 % vs. 23,24 %). Étant donné les pourcentages très élevés, nous proposons une analyse phonétique de ces deux productions pour mieux comprendre leur nature. Nous proposons une comparaison entre ces deux types de séquence et :

1. les bonnes productions de liaison (ex. *les ours* [le.zuʁs]).

Nous allons comparer : la production de [z] dans *les ours* [le.zuʁs] à la production de [z] en [lə.zɔ.ʁaʒ] et à l'espace vide [Ø] entre Mot1 et Mot2 sans réalisation d'une CL dans [lə.ɔ.ʁaʒ] que nous pouvons transcrire comme [ləØɔ.ʁaʒ].

2. la production de *filler* (ex. *le zébre* [lə.zɛbʁ]).

Nous allons comparer la production de [ə] dans [lə.zɔ.ʁaʒ], [ləØɔ.ʁaʒ] et [lə.zɛbʁ].

8.3.3.2.1 L'analyse phonétique : [z]

Comme nous l'avons vu dans le chapitre 4, les analyses phonétiques montrent que les CL sont plus courtes que les consonnes lexicales (Spinelli *et al.*, 2003 ; Wauquier-Gravelines, 1996 ; Gaskell *et al.*, 2002). Au moyen du logiciel PRAAT, nous

avons donc décidé de comparer la durée de la consonne [z] dans deux types de productions enregistrées dans la tâche DENa^{CP1} : le [z] dans la production de liaison correcte ([le.zuʁs]) et le [z] dans la production d'élision erronée [lə.zɔ.ʁaʒ]. L'analyse de la longueur de [z] nous permet d'évaluer le statut de cette consonne dans [lə.zɔ.ʁaʒ]. Si elle a la même durée que le [z] dans [le.zuʁs], elle peut être traitée comme une CL ; si elle est plus courte, nous pouvons la traiter comme une consonne lexicale. Dans le tableau 34, nous présentons les valeurs de longueur moyenne (en millisecondes) que nous avons observées dans les deux productions.

	[z] en [lə.zɔ.ʁaʒ]	[z] CL dans [le.zuʁs]
Longueur moyenne (ms)	129	127
écart-type	24	40

Tableau 34. Longueur de [z] dans les deux productions (DENa^{CP1})

Nous pouvons dire qu'il n'y a pas de différence entre le [z] de [lə.zɔ.ʁaʒ] et le [z] de [le.zuʁs]. Dans [lə.zɔ.ʁaʒ], le [z] doit être considéré comme une CL plutôt qu'une consonne lexicale. Par conséquent, la mauvaise réalisation de l'élision dans [lə.zɔ.ʁaʒ] à partir de l'input [le.zuʁs] doit être traité comme un reflet de l'input en liaison donné par l'expérimentateur et des caractéristiques linguistiques de ce phénomène. Les enfants doivent effacer la CL pour passer de la séquence en liaison à la séquence en élision. Ils ne sont pas capables de le faire et la CL est gardée entre Mot1 et Mot2.

8.3.3.2.2 L'analyse phonétique : [Ø]

Nous allons analyser l'espace vide [Ø] dans la séquence [ləØɔ.ʁaʒ]. Cet espace est-il vraiment vide ou est-il rempli par un coup de glotte ?

L'analyse phonétique montre que cet espace n'est pas vraiment vide mais qu'il est rempli par une consonne occlusive glottale, marquée par la barre d'explosion entre Mot1 et Mot2 (Figure 31). Donc les séquences du type [ləØɔ.ʁaʒ] doivent plutôt être décrites comme [lə.ʔɔ.ʁaʒ]. L'occlusive glottale réalisée entre Mot1 et Mot2 empêche la rencontre des deux voyelles et la chute du noyau vocalique, qui donne lieu au processus d'élision.

Le [z] et le [ʔ] entre Mot1 et Mot2 peuvent être la réalisation phonétique d'une position phonologique qui ne peut pas être effacée dans l'input en liaison donné par l'expérimentateur. Cependant, vu l'article au singulier au début de ces deux formes, ils peuvent être aussi deux attaques du Mot2 ([ʔɔ.ʁaʒ], [zɔ.ʁaʒ]). Pour mieux comprendre si [z] et [ʔ] ont un statut similaire à celui de la CL, nous avons décidé d'analyser la voyelle [ə] qui précède ces deux segments.

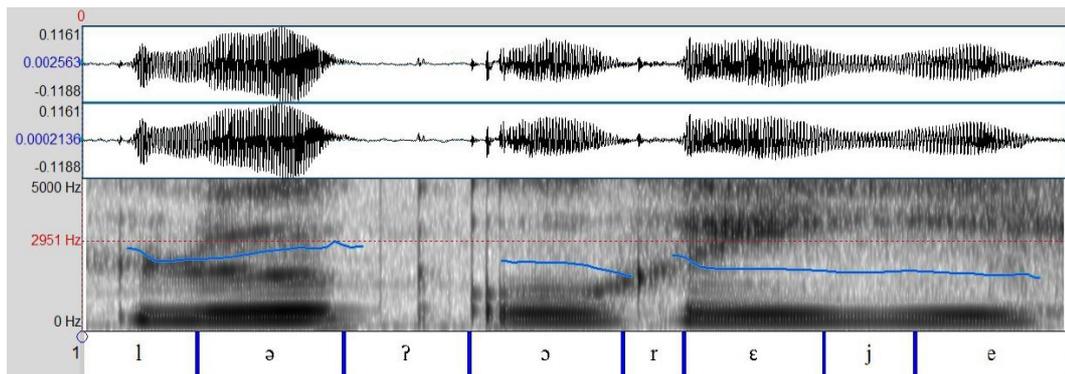


Figure 31. Spectrogramme pour l'insertion d'un coup de glotte en élision

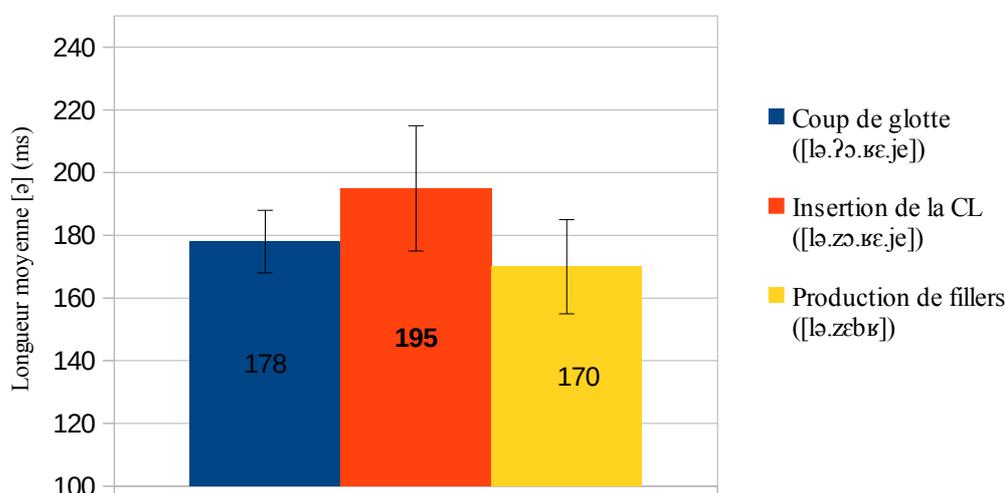
8.3.3.2.3 L'analyse phonétique de [ə]

Nous avons déjà vu que le [z] de [lə.zɔ.ʁaʒ] doit plutôt être traité comme une CL et qu'un coup de glotte est inséré entre Mot1 et Mot2 dans le cas de [lə.ʔɔ.ʁaʒ]. Pour confirmer le statut phonologique de [z] et mieux comprendre celui de [ʔ], nous proposons l'analyse phonétique de la voyelle [ə].

L'étude de Nguyen *et al.* (2007) montre que, dans le cas de liaison en [z], la voyelle qui précède une CL à la même longueur qu'une voyelle qui précède une autre consonne en position d'attaque mais est plus courte qu'une voyelle en syllabe fermée (voir chapitre 4). La voyelle peut donc fournir des indices sur le statut de la consonne qui suit.

Nous cherchons à voir si les voyelles [ə], dans les deux types de productions erronées d'élision ([lə.zɔ.ʁaʒ] et [lə.ʔɔ.ʁaʒ]), ont des caractéristiques phonétiques de durée similaires à celles d'une voyelle qui précède une CL. Pour cette raison, nous comparons la longueur de ces voyelles à la longueur de la même voyelle dans la production de *fillers* [lə.zɛbʁ], où [ə] est une voyelle qui précède une consonne [z] qui est attaque du Mot2.

Dans nos résultats (Graphique 5), le [ə] a la même longueur dans les productions [lə.ʔɔ.ʁɑʒ] et [lə.zɔ.ʁɑʒ] que dans la production [lə.zɛbɛ]. L'analyse de la voyelle [ə] nous permet de penser que [z] et [ʔ] peuvent avoir le statut d'une CL.



Graphique 5. Longueur de [ə] dans les trois productions analysées

8.3.3.2.4 La production d'élision comme une conséquence de l'analyse de la liaison

Quand nous demandons aux enfants de produire l'élision (ex. *l'orage* [lə.ʁɑʒ]), nous donnons toujours un input en liaison (ex. *les[z]orages* [lə.zɔ.ʁɑʒ]). La CL doit donc être effacée et la rencontre des deux voyelles doit causer la chute du noyau vocalique qui permet la production correcte de l'élision. Dans les mauvaises productions d'élision [lə.ʔɔ.ʁɑʒ] ou [lə.zɔ.ʁɑʒ], les enfants réussissent à produire la forme singulière dans l'article mais pas à effacer le contenu segmental entre Mot1 et Mot2, réalisé par [z] ou [ʔ]. De plus, l'analyse phonétique de la voyelle [ə] qui précède nous montre que ces consonnes peuvent avoir le statut d'une CL. Nous pouvons penser que [z] et [ʔ] sont les réalisations phonétiques d'une position phonologique présente dans les séquences input en liaison. Par conséquent, nous pouvons remarquer l'existence d'une position phonologique entre Mot1 et Mot2. Cette position phonologique, que les enfants ne réussissent pas à effacer, est remplie par le contenu segmental d'une CL, comme on peut le voir dans la production de [lə.zɔ.ʁɑʒ], où le [z] a le même statut d'une CL. La même position peut aussi être remplie par une consonne glottale dans le cas de productions comme [lə.ʔɔ.ʁɑʒ].

Comme nous l'avons déjà dit au début de ce chapitre, l'existence d'une position phonologique nous guide vers la validation de l'explication morpho-phonologique de l'acquisition de la liaison. Pour approfondir la validité éventuelle de cette affirmation, nous avons décidé d'analyser aussi le rôle de la fréquence et son influence sur le processus d'acquisition des frontières des mots.

8.3.3.3 Le rôle de la fréquence

Pour l'explication morpho-phonologique, les enfants réussissent à bien maîtriser la liaison car ils apprennent une forme sous-jacente et la représentation du double flottement. Pour l'explication lexicale-constructionniste, l'acquisition des frontières des mots a lieu grâce à une organisation de formes de surface basée sur la fréquence.

Nous pouvons évaluer le rôle de la fréquence grâce à la mise en relation des fréquences SFI des items tests dans MANULEX²¹ et le nombre d'erreurs dans la tâche de dénomination guidée DENa^{CP1} (Tableau 35).

Nom au singulier	SFI singulier	Nom au pluriel	SFI pluriel	Nombre d'erreurs de non-réalisations	Nombre d'erreurs de consonne initiale erronée		
					[z]	[n] ou [l]	Autres consonnes
orange	58,64	oranges	53,40	8	6	0	0
araignée	57,88	araignées	47,20	8	3	0	0
oie	57,59	oies	55,20	7	19	2	1
armoire	58,10	armoires	45,10	2	0	0	0
ogre	56,77	ogres	0	10	15	2	0
oreiller	57,88	oreillers	38,75	18	1	1	0
arrosoir	56,59	arrosoirs	37,95	7	0	0	0
orage	58,28	orages	38,75	15	0	1	0

Tableau 35. Indice SFI et nombre d'erreurs dans DENa^{CP1}

Le tableau montre la fréquence des items tests, au singulier et au pluriel. Dans les items tests nous ne considérons pas le mot *ours* car, dans la base de données utilisée, sa fréquence n'est pas donnée pour les deux formes que nous analysons²².

21 Pour une explication plus claire de la valeur de SFI dans la base de données MANULEX, nous invitons à la lecture du chapitre 6.

22 MANULEX nous donne la fréquence du lemme *ours* mais nous ne pouvons pas savoir si cette

Nous rappelons que les mots avec un SFI entre 50 et 65 sont traités comme des mots avec une fréquence moyenne. Les mots avec un SFI inférieure à 50 ont une fréquence peu élevée et ceux avec une valeur supérieure à 65 sont des mots à fréquence élevée. Seulement deux mots (*orange* et *oie*) ont des formes singulières et des formes plurielles avec des fréquences moyennes. En revanche, dans les autres couples, les items ont une fréquence moyenne au singulier et une fréquence basse au pluriel. C'est seulement la fréquence au pluriel qui change (de basse à moyenne).

Nous avons ensuite classifié les erreurs selon qu'elles étaient « non-réalisation » ou « consonne erronée ». Dans cette dernière catégorie, nous avons identifié trois sous-groupes, selon le type de consonne erronée utilisée à la frontière des mots : la consonne erronée [z], plutôt associée au pluriel dans la resyllabation par liaison ou enchaînement ; [n] et [l] qui apparaissent en général avec la forme singulière du nom en cas de liaison, d'élision et d'enchaînement ; les autres consonnes, présentes dans nos données mais qui ne peuvent pas être impliquées dans les phénomènes consonantiques testés (ex. [d]).

Si on considère l'explication lexicale décrite dans l'introduction, la fréquence de surface (donc la fréquence des deux formes séparément) influence les productions et la segmentation de l'enfant en cas de présence de phénomènes consonantiques. On s'attendrait donc à ce que l'enfant produise davantage de [z] au début du nom, lorsque le mot est plus fréquent au pluriel qu'au singulier et des non-réalisations ou de mauvaises consonnes de liaison (comme [n] et [l]) si le mot est plus fréquent au singulier. Si nous analysons nos données, il n'y a pas une corrélation entre les fréquences des formes au pluriel et le nombre de productions avec la consonne erronée [z] ($r(8) = .21$, $p = .60$).

Dans les couples où la fréquence au pluriel est basse et la fréquence au singulier moyenne, les erreurs de non-réalisations et de consonne erronée [n] ou [l] sont supérieures aux erreurs de consonne erronée en [z]. Cependant, nous trouvons une exception avec le mot *ogre*. MANULEX n'indique aucune occurrence pour sa forme plurielle au cours du CP, même si, dans sa forme singulière, il est présent en fréquence moyenne. Comme nous pouvons le voir dans le tableau, pour cet item, le nombre

fréquence fait référence à la forme plurielle ou singulière, car les orthographes pour les deux formes sont équivalentes et, en plus, la base de données ne donne pas d'informations sur les cooccurrences qui nous permettraient de comprendre le nombre.

d'erreurs de non-réalisation et de consonne erronée en [n] ou [l] est presque le même que le nombre d'erreurs de consonne erronée en [z]. Le cas du mot *ogre* montre donc une tendance différente des trois autres cas considérés.

Le rapport entre fréquence et nombre d'erreurs pour la tâche de dénomination DENa^{CP1} ne permet pas de considérer l'explication lexicale, basée sur le modèle exemplariste, comme la seule explication possible pour la détection des frontières des mots. Cependant, d'autres tâches où la fréquence des mots utilisés est plus contrôlée et où le nombre d'items tests est plus élevé doivent être conduites pour mieux évaluer le rôle de la fréquence dans les productions orales au CP.

8.4 Discussion des données de production et de perception au début du CP

Les études de Dugua (2006) et Dugua et Chabanal (2006) montrent que les enfants français entre 5 et 6 ans traitent correctement la CL et réussissent à la distinguer d'une consonne lexicale (§ 4.2.3). Le traitement correct de la CL implique la résolution des erreurs d'insertion d'une consonne erronée entre Mot1 et Mot2. De plus, les recherches de Basset (2000) et de Dugua et Baclesse (2014) nous suggèrent qu'après 7 ans, les enfants réalisent correctement plus de 97 % des liaisons obligatoires-catégoriques (§ 4.2.3).

Nous avons testé des enfants au début du CP (6 ans d'âge). Nous avons, conformément à la littérature, fait l'hypothèse qu'aucun problème de segmentation ne devait plus exister ni en perception ni en production sur les frontières resyllabées par des liaisons, élisions ou enchaînements (§ 8.1.2). Pourtant, contrairement aux hypothèses, au début du CP, nous enregistrons des difficultés à identifier les erreurs de non-réalisations dans les différents contextes obligatoire-catégoriques (PERC^{CP1}, § 8.2.4). Nous remarquons également un nombre important de non-réalisations dans les productions de liaison, d'élision et d'enchaînement (DENa^{CP1}, § 8.3.3). Néanmoins les séquences de Mot1 et Mot2 caractérisées par l'insertion d'une consonne erronée sont plus fréquemment repérées comme fautives et peu produites à l'oral, ce qui laisse à penser que les enfants savent qu'elles ne sont pas acceptables.

Nos résultats falsifient nos hypothèses basées sur la littérature de l'acquisition de la liaison. Les enfants à 6 ans ont encore des difficultés à détecter et à produire correctement les frontières des mots en présence des phénomènes de resyllabation.

La production non négligeable de non-réalisations de liaison, d'élision et d'enchaînement, en revanche, confirme les données de Chevrot *et al.* (2013) qui proposent une stabilité de ces erreurs après 5 ans (§ 4.2.3). Nos résultats sont aussi en accord avec Nardy et Dugua (2011) qui repèrent un nombre de non-réalisations très élevé pour la liaison en contexte adjectif-nom entre 5 et 6 ans (§ 4.2.3).

Le phénomène qui provoque le plus de problèmes en production (DENa^{CP1}) est l'élision. Dans notre procédure expérimentale, l'expérimentateur donnait un input en liaison (*les[z]orage* [le.zɔ.ʁaʒ]). L'enfant devait effacer la CL et réaliser une élision pour nommer correctement la carte (*l'orage* [lɔ.ʁaʒ]). Nous avons enregistré beaucoup de productions du type [ləθɔ.ʁaʒ] ou [lə.zɔ.ʁaʒ], où l'article est correctement produit à la forme singulière mais l'élision n'est pas réalisée. Ces productions incorrectes d'élision peuvent être une conséquence de l'input en liaison présenté par l'expérimentateur et peuvent nous donner des informations en ce qui concerne ce phénomène phonologique²³. Nous avons donc analysé phonétiquement ces productions. Nous découvrons que le [z] dans [lə.zɔ.ʁaʒ] a la même longueur qu'une CL et que les enfants insèrent un coup de glotte entre Mot1 et Mot2 dans [ləθɔ.ʁaʒ]. Nous avons donc décidé de transcrire les productions [ləθɔ.ʁaʒ] comme [lə.ʔɔ.ʁaʒ]. Ceci nous invite à interpréter le [z] et le coup de glotte comme des CL puisque l'un et l'autre ne causent pas l'allongement de la voyelle précédente. Nous proposons donc de les représenter comme un segment flottant avec remplissage segmental comme proposé dans le modèle autosegmental (Figure 32, dans Encrevé, 1988). Ces segments dans le palier syllabique, remplissent la position vide à l'attaque du Mot2, empêchant l'élision (et les résultats sont des séquences comme [lə.ʔɔ.ʁaʒ] ou [lə.zɔ.ʁaʒ]).

23 Dans le chapitre 9, nous proposons une deuxième tâche de dénomination (DENb^{CP2}) qui vérifie si ces productions d'élision sont une conséquence de l'input donné par l'expérimentateur. Les résultats montrent que, pour l'élision et l'enchaînement, la séquence-input a un effet d'amorçage qui implique la production de séquences incorrectes (§ 9.4).

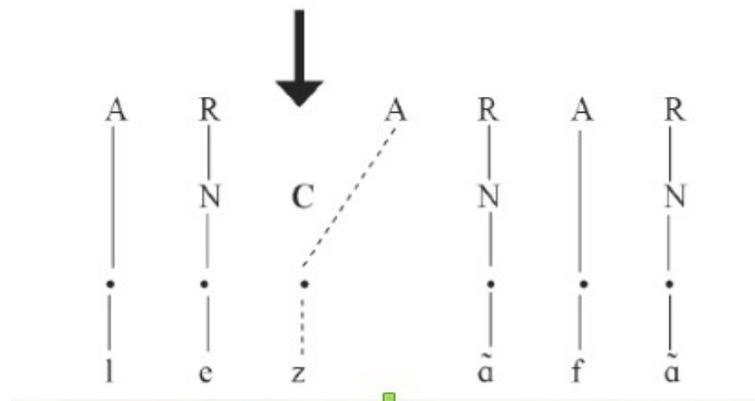


Figure 32. Représentation de la liaison par Encrevé (1988)

Nous pouvons dire que les deux productions [lə.ʔə.ʁaʒ] et [lə.zə.ʁaʒ] montrent une position prosodique appartenant à une forme sous-jacente entre Mot1 et Mot2. À partir de ces données, nous pouvons aussi défendre l'idée que l'acquisition de la liaison n'est pas seulement une acquisition de formes de surfaces stockées dans la mémoire, comme prévu par le modèle lexical-constructionniste, mais plutôt l'acquisition d'une forme sous-jacente et d'une représentation phonologique plus profonde.

Le modèle lexical-constructionniste explique les erreurs de non-réalisations de liaison comme la conséquence du choix de la variable à voyelle initiale du Mot2 /uʁs/. Par conséquent, nous nous attendons à enregistrer des productions du type /leuʁs/, où aucun segment n'apparaît entre Mot1 et Mot2 et la séquence est traitée comme un *chunk*. Notre analyse phonétique de non-réalisations de liaison nous montre, au contraire, qu'à l'oral, les enfants peuvent parfois insérer un coup de glotte entre Mot1 et Mot2 ([le.ʔuʁs]). La production de l'occlusive glottale, occupant la place de la CL, confirme l'existence d'une position prosodique sous-jacente entre les mots en liaison. De plus, dans les non-réalisations que nous avons enregistrées, les enfants produisent parfois une pause entre Mot1 et Mot2 où aucun segment n'est réalisé. La production de cette pause conduit à penser que les enfants traitent les mots en liaison comme des séquences [le] + [uʁs] et pas comme une construction du type /leuʁs/, proposé par l'explication lexicale-constructionniste.

Quel que soit le modèle de l'acquisition de la liaison considéré, les non-réalisations ne sont pas bien modélisées. Nous pouvons penser que ce manque de modélisation est dû au fait que les données dans la littérature pour ce type d'erreur ne sont pas claires et univoques. Comme l'affirment Chevrot *et al.* (2013 : 10) en traitant

les données concernant les non-réalisations "decrease or stagnation between 2 and 6 years depend on the study and social environment (Chevrot *et al.*, 2007, 2011; Dugua, 2006)". Cependant, les premières données de notre étude longitudinale mettent en évidence la question des non-réalisations. Nous pouvons penser que les non-réalisations peuvent être vues comme une étape intermédiaire dans l'acquisition de la liaison. De plus, leur structure nous permet de proposer une connexion entre les productions orales et l'orthographe. La production des séquences Mot1 + Mot2 sans réalisation du phénomène de resyllabation peut nous faire penser que les enfants cherchent à reproduire la structure de cette suite à l'écrit, où un espace blanc ou une apostrophe séparent les items. Cette hypothèse sera vérifiée dans les prochains chapitres traitant les étapes successives de notre étude longitudinale.

D'autres informations nous incitent à défendre l'interprétation morpho-phonologique de nos données : la fréquence des Mots2 ne semble pas jouer un rôle (§ 8.3.3), contrairement à ce que le modèle lexical-constructionniste propose, même si ces données confirment les résultats de Dugua *et al.* (2009) qui montrent que la fréquence perd son influence dans la réalisation des liaisons après 6 ans (§ 4.2.3.1).

Il faut noter que nous avons utilisé MANULEX comme base de données pour vérifier la fréquence des items. Cette base de données nous offre des indices de fréquence des items trouvés dans de nombreux livres et manuels pour les enfants. L'usage de MANULEX a des avantages et des inconvénients. Un des avantages est la possibilité de vérifier la fréquence des mots connus par les enfants dans la tranche d'âge que nous avons testée. En revanche, cette fréquence fait référence au lexique écrit et aux items pris isolément, pas aux cooccurrences dont la fréquence (et non pas celle ajoutée de fréq. Mot1 + fréq. Mot2) influence, selon Fougeron *et al.* (2001), les taux des réalisations de ces phénomènes. Par exemple, dans la séquence *les[z]ours*, nous n'avons pas la fréquence de la séquence *les + ours* mais les fréquences des deux items séparés (*les* et *ours*), ce qui est un indicateur très relatif de la cooccurrence. Malgré ces inconvénients, nous avons décidé de continuer à utiliser MANULEX au lieu des autres bases de données comme par exemple LEXIQUE.org (New *et al.*, 2001). Cette base de données créée sur le lexique de livres, d'articles de journaux et sous-titres de films prend en considération un lexique trop « adulte » et elle n'est pas adaptée à notre recherche²⁴. Nous avons aussi essayé de vérifier les fréquences des items sur CHILDES (MacWhinney & Snow, 1984) mais le processus de recherche n'a

24 LEXIQUE.org ne présente pas le mot *ours*.

pas donné de résultats car ce corpus n'a pas été créé pour ce type de consultation. Ces réflexions nous amènent à souligner au passage les lacunes existant actuellement en termes de bases de données tenant compte du lexique oral des enfants et des fréquences de cooccurrences entre les unités. Alors que de tels outils seraient très utiles notamment pour mieux évaluer le lexique oral des enfants et la fréquence de séquences de deux mots.

Malgré ces limites, nous avons décidé de continuer à vérifier le rôle de la fréquence de Mot2 dans les productions des enfants dans les autres étapes de notre étude longitudinale.

8.5 La variante sociolinguistique

Dans cette section, nous analyserons le rôle de la dimension sociolinguistique dans les productions des enfants en mettant en rapport la production de la liaison lors de DENa^{CP1} avec l'analyse des milieux sociaux des parents et de leur niveaux d'instruction. Nous concentrons notre analyse sur la liaison, car elle est le seul phénomène qui semble être conditionné par cette variable (§ 4.2.3). Les analyses que nous présentons dans cette section permettent exclusivement de décrire les données et elles n'ont pas pour but de valider des hypothèses sociolinguistiques en soi. L'expérience n'a pas été construite dans l'intérêt de recueillir des données homogènes et utilisables dans une analyse statistique fine qui permettrait de valider nos affirmations. Les données et leurs analyses doivent donc être lues comme une description qualitative de cette variante.

Comme nous l'avons précisé dans la section qui décrit les participants, les parents ont été classifiés en trois niveaux, selon les diplômes obtenus (§ 7.1.2). Les parents de Niveau 1 ont arrêté leurs études avant le baccalauréat (BAC) ; les parents de Niveau 2 ont un diplôme de baccalauréat et les parents du Niveau 3 ont continué leurs études après le baccalauréat (BAC +). Nous allons vérifier si les différents niveaux d'instruction des parents (Niveau 1, 2 ou 3) sont en rapport avec les erreurs de liaison obligatoire-catégorique dans la tâche DENa^{CP1}. Dans le tableau 36, nous présentons les scores d'erreurs des enfants dans la tâche DENa^{CP1} en rapport avec les trois niveaux scolaires des parents. Le score maximal est 3. Le nombre d'erreurs ne change pas selon

le niveau scolaire des parents entre le Niveau 2 et le Niveau 3 (0,6 vs. 0,4). En revanche, parmi les enfants dont les parents ont arrêté les études avant le baccalauréat, le score d'erreurs en production (1,2).

		Scores moyens d'erreurs
Enfants avec des parents de	« Niveau 1 »	1,2
	« Niveau 2 »	0,4
	« Niveau 3 »	0,6

Tableau 36. Scores moyens d'erreurs dans DENa^{CP1} selon le niveau scolaire des parents

D'après l'analyse, les enfants qui font le plus d'erreurs ont les parents ayant le niveau scolaire le plus bas. L'input qu'ils reçoivent est-il donc différent des autres enfants ? Mallet (2008), après l'analyse de son échantillon PFC, n'a trouvé aucune corrélation entre la production de liaison obligatoire ou facultative et les niveaux scolaires. Pour mieux comprendre l'influence de l'input sur les productions des enfants, il faudrait donc avoir plus de données sur la langue parlée par les parents. Même si ce type d'étude est très intéressant, cela n'est pas le but de la recherche menée dans cette thèse. Par ailleurs, les données présentées doivent être considérées plutôt comme une façon de décrire les données et non comme une forme d'explication finale des résultats²⁵.

Nous avons également comparé les erreurs et les types d'erreurs de notre échantillon en fonction de la profession des parents. Nous avons considéré l'étude longitudinale de Nardy et Dugua (2011). Celles-ci étudient l'acquisition de la liaison variable adjectif-nom²⁶ entre 2 et 6 ans en comparant les différents patterns développementaux des enfants selon la profession des parents. Comme on peut le voir dans les graphiques (Figure 33), si au début de l'acquisition de la liaison (2-3 ans) les enfants de cadres et les enfants d'ouvriers présentent un nombre plus élevé d'erreurs par remplacement (plus de 40 %), un peu moins d'erreurs de non-réalisations (entre

25 En effet, les trois groupes ne sont pas homogènes car ils contiennent un nombre différent de sujets. Pour cette raison, les données ne permettent pas une analyse statistique qui explique quantitativement ces résultats.

26 Nardy et Dugua traitent la liaison adjectif-nom (ex. *gros+nom*, *petit+nom*) comme une liaison variable. Dans le corpus PFC, ce type de liaison est considéré comme une liaison variable, même si dans ce contexte l'adjectif *petit* provoque invariablement une liaison.

30 % et 40 %), et un nombre moins élevé de réalisations justes (aux alentours de 20 %), la situation change entre 3 et 4 ans. Pour les enfants de cadres, les taux de non-réalisations et de réalisations justes augmentent et ils atteignent le même niveau à l'âge de 6 ans. Pour les enfants d'ouvriers, le nombre de non-réalisations augmente et le nombre d'erreurs de consonne erronée décroît plus tard, entre 4 et 5 ans. De plus, le nombre de réalisations justes ne subit pas de variations notables. L'explication donnée pour ces deux comportements est que la réalisation des enfants s'approche de la réalisation des liaisons facultatives de leur milieu (Chevrot *et al.*, 2011).

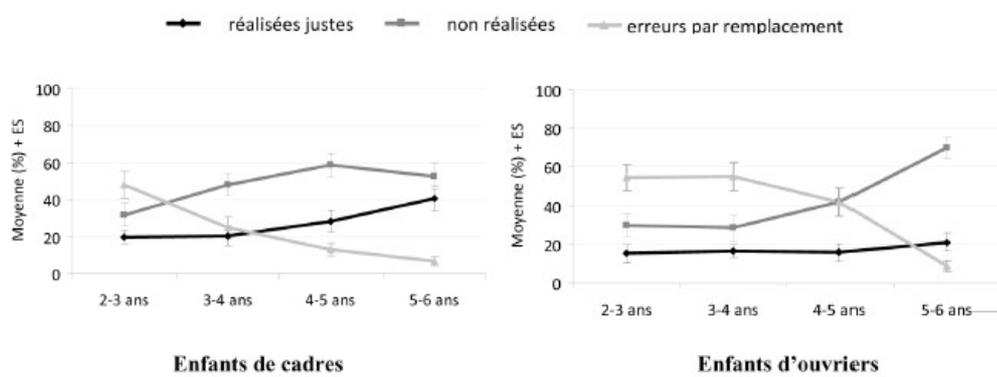
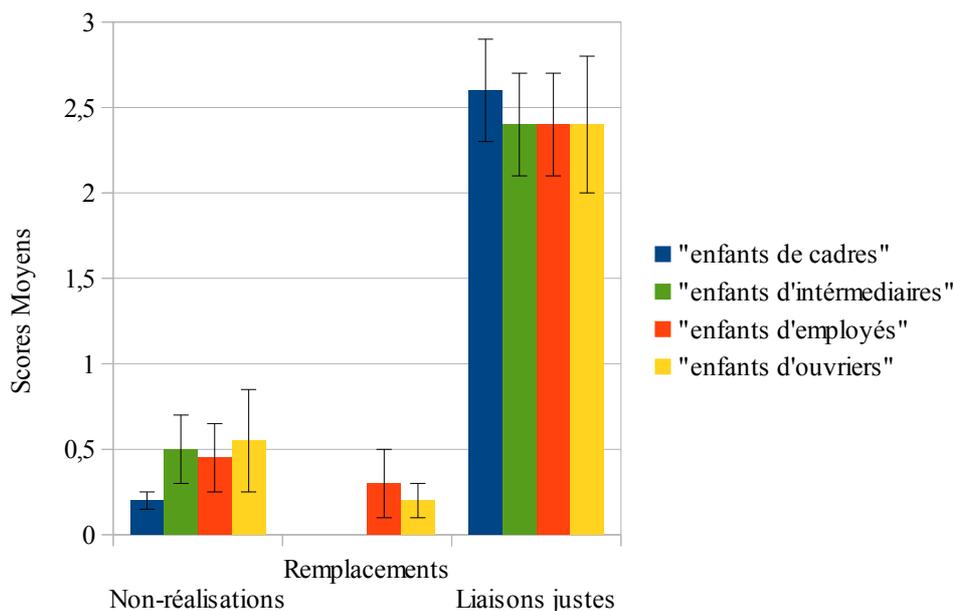


Figure 33. Graphique des résultats chez les enfants de cadres et d'ouvriers (Nardy & Dugua, 2011 : 114)

Nous avons donc décidé de regarder nos données en dénomination pour examiner la présence éventuelle d'une relation entre les erreurs de liaison et la profession des parents. Comme on l'a vu dans le chapitre 7, on a déjà divisé les enfants en quatre groupes selon les milieux sociaux des parents. Le graphique 6 nous montre les scores moyens pour les deux types d'erreurs (non-réalisations de liaison et remplacements) et pour les réalisations justes dans la tâche de dénomination guidée DENa^{CP1}. Le score maximal est 3.



Graphique 6. Scores des enfants dans DENa^{CP1} selon la profession des parents

Il n'y a de différence pour aucun des quatre groupes sur le nombre de liaisons justes. En ce qui concerne les erreurs, les « enfants de cadres » font moins d'erreurs de non-réalisations que les enfants des trois autres groupes qui ont des résultats homogènes pour ce type de faute. De plus, les « enfants de cadres » comme les « enfants d'intermédiaires » ne produisent pas de remplacements.

Ces données confirment les données sociolinguistiques de Nardy et Dugua. Les chercheuses enregistrent chez les « enfants de cadres » la tendance entre 5 et 6 ans à une diminution des erreurs de non-réalisation et de remplacement. Nos données après 6 ans confirment cette tendance : au début du CP, chez ce groupe d'« enfants de cadres », le nombre d'erreurs de non-réalisation est plus bas que celui qui a été enregistré dans les autres groupes. De même, les erreurs de remplacement disparaissent. Pour les « enfants d'ouvriers », les chercheuses enregistrent un nombre élevé de non-réalisation entre 5 et 6 ans. À partir de 6 ans, nos données montrent également que les erreurs de non-réalisation sont encore présentes mais qu'elles ne se différencient pas des erreurs des groupes « enfants intermédiaires » et « enfants d'employés ». Nos données montrent aussi que les erreurs de remplacement caractérisent les productions des enfants qui ont des parents avec des métiers moins qualifiés tels qu'employé ou ouvrier.

À la lumière de nos résultats, nous pouvons établir une certaine relation entre la production d'erreurs en liaison obligatoire-catégorique et le milieu social : les enfants

de cadres produisent moins d'erreurs (non-réalisation ou remplacement) que les enfants d'intermédiaires, d'ouvriers ou d'employés. La mise en évidence du milieu social sur la production de liaison dans un contexte catégorique (déterminant-nom) est toute nouvelle, car jusqu'à maintenant l'influence du milieu social sur la production de liaison était montré uniquement en contexte facultatif. Cependant, comme nous l'avons dit plus haut, le but de cette thèse n'est pas l'étude de la variante sociolinguistique. Les données que nous avons présentées ici ne permettent donc pas la mise en place d'un véritable traitement statistique. Pour cette raison, nous sommes obligés de considérer les analyses présentées comme une description qualitative des données.

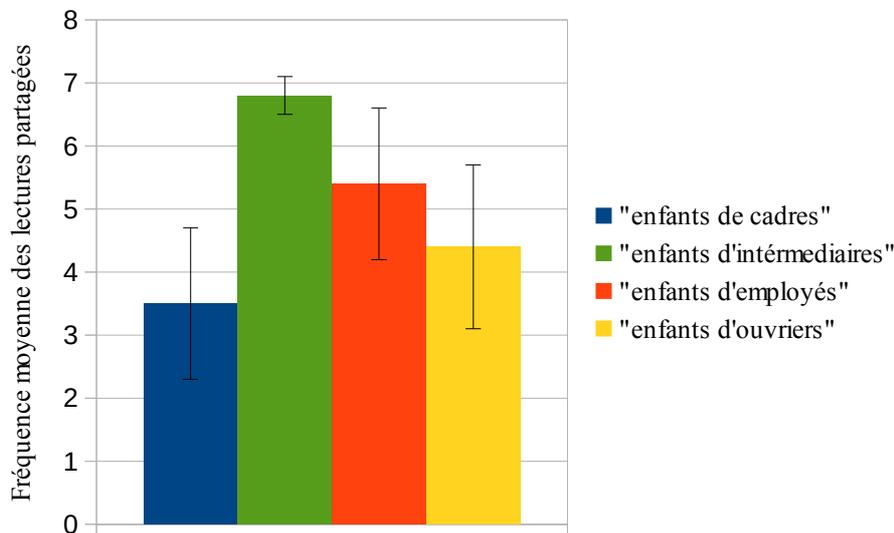
8.6 Perception d'erreurs, dénomination guidée et lecture partagée : y a-t-il une corrélation ?

Comme nous l'avons présenté dans le chapitre 5, plusieurs études ont montré une relation entre les compétences de lecture et d'écriture que les enfants développent à l'école et les activités de lecture partagée que ces enfants pratiquent chez eux avec leurs parents (§ 5.2.2). À partir de là, nous pouvons penser que si les activités de *literacy* influencent les productions orales dans le cas des phénomènes de resyllabation, les enfants qui pratiquent le plus la lecture avec les parents hors du contexte scolaire ont une plus grande capacité à segmenter les mots car ils ont été plus exposés à l'input écrit et à ses caractéristiques. Par conséquent, nous pouvons penser que ces enfants font moins d'erreurs en compréhension et production quand nous leur demandons d'identifier et de produire les frontières de mots dans les cas de liaison, d'élision et d'enchaînement.

Grâce aux questionnaires parentaux, nous avons des données sur la fréquence des lectures partagées. Quand nous demandons aux parents de nous dire avec quelle fréquence ils pratiquent la lecture partagée avec leurs enfants, tous les parents proposent spontanément un rythme hebdomadaire. Ainsi, pour classifier les fréquences de lecture partagée, nous avons construit une échelle à 7 degrés correspondant aux 7 jours de la semaine. Nous avons associé le degré 1 aux parents qui lisent une fois par semaine avec leurs enfants ; 2 aux parents qui lisent deux fois par semaine avec leurs enfants et ainsi de suite.

Nous avons d'abord cherché à vérifier si la fréquence de lecture partagée était en rapport avec le milieu social ou le niveau scolaire des parents.

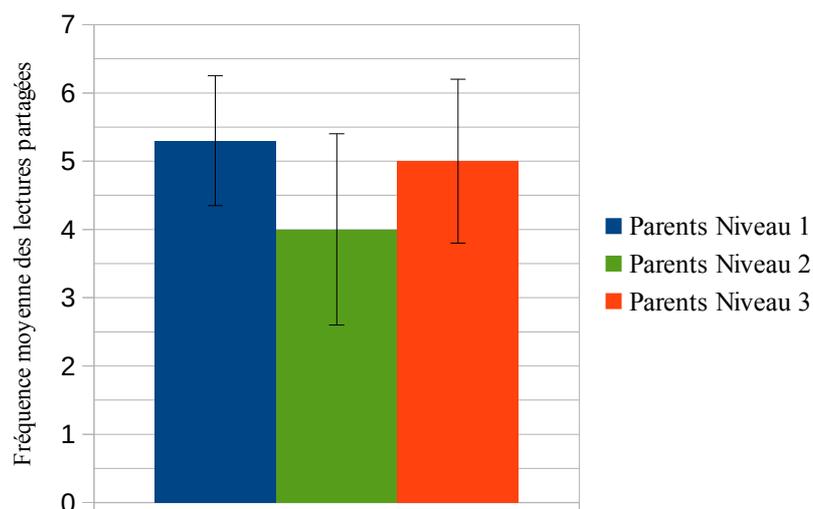
Pour évaluer l'influence du milieu social sur la pratique de la lecture partagée, nous considérons la fréquence moyenne de lecture partagée et les 4 groupes décrits dans le chapitre 7. On s'attend à ce que les enfants de cadres lisent plus souvent avec leurs parents et que les enfants d'ouvriers lisent moins fréquemment avec leurs parents. Les résultats présentés dans le graphique 7 montrent que les enfants de cadres lisent moins fréquemment avec leurs parents que les enfants d'ouvriers. Le milieu social ne semble pas être lié à la fréquence de lecture partagée²⁷.



Graphique 7. Fréquence des lectures partagées selon la profession des parents

Pour établir le rôle du niveau scolaire des parents sur la fréquence de lecture partagée, nous avons calculé la fréquence moyenne de lecture partagée pour les trois groupes de différents niveaux scolaires présentés dans la section 7.1.2. Il s'agit de vérifier que les parents avec un niveau scolaire plus élevé partagent plus fréquemment des moments de lecture avec leurs enfants. Les résultats dans le graphique 8 nous montrent qu'il n'y a pas de différence dans la fréquence de lecture partagée selon la classification du niveau scolaire des parents.

²⁷ Nous avons seulement analysé les réponses des parents aux questionnaires. Nous pouvons penser que les parents cadres se servent de nounous. Nous ne pouvons pas tenir compte des données sur la fréquence de lectures partagées avec du personnel employé pour la garde des enfants.



Graphique 8. Fréquence des lectures partagées selon le niveau scolaire des parents

En conclusion, ni le milieu social ni le niveau scolaire ne sont en rapport avec la fréquence de lecture partagée.

Nous considérons maintenant les résultats des deux tâches (PERC^{CP1} et DENa^{CP1}) en rapport à la fréquence des lectures partagées. Nous proposons les analyses mettant en rapport le nombre d'erreurs dans les tâches PERC^{CP1} ou DENa^{CP1} et la fréquence des activités de lecture partagée, comme nous l'avons vu au début de cette section. Nos données ne présentent pas une corrélation statistiquement significative entre la fréquence de la lecture partagée et les scores de réponses correctes dans la tâche PERC^{CP1} ($r(31) = .45, p = .12$) et dans la tâche DENa^{CP1} ($r(31) = .48, p = .23$).

Au vu des résultats, si l'acquisition des frontières de mots à l'oral est liée aux compétences orthographiques, elle ne peut pas être seulement en rapport avec la sensibilité à l'orthographe ou avec les compétences d'*early literacy* développées au moyen des moments de lecture partagée au niveau pré-scolaire. Il est plus plausible que les enfants aient besoin d'un enseignement spécifique du code écrit et que, grâce à cela, ils puissent bien gérer les frontières de mots à l'oral. Nous allons vérifier cette hypothèse dans les prochaines étapes de notre étude longitudinale.

Chapitre 9

Deuxième étape de l'étude longitudinale : la fin du CP

9.1 La deuxième étape de l'étude

Pour vérifier si les compétences orales des enfants sur les phénomènes de resyllabation évoluent pendant la première année de l'école primaire, nous avons testé à nouveau les 43 élèves du CP A et du CP B à la fin de l'année scolaire (mai 2014) au moyen de trois tâches orales :

- la tâche de perception d'erreurs et reformulation : PERC^{CP2} (mai 2014).
Nous avons proposé la même tâche qu'au début de l'année. Nous avons seulement testé un échantillon d'enfants considéré comme représentatif de tous les élèves ;
- la tâche de dénomination guidée : DENa^{CP2} (mai 2014).
Nous avons utilisé la même tâche qu'en octobre pour évaluer les compétences orales de tous les enfants parmi les deux classes ;
- une nouvelle tâche de dénomination DENb^{CP2} (mai 2014).
Nous avons décidé de créer une nouvelle tâche de dénomination pour vérifier si les résultats que nous avons enregistrés pour la première tâche de dénomination DENa^{CP1} découlent de la tâche en elle-même ou s'ils sont vraiment représentatifs des compétences orales des enfants.

Comme à la fin du CP, les enfants sont capables de lire et d'écrire, nous avons décidé de tester aussi les compétences écrites (juin 2014) au moyen de deux tâches liées à la langue écrite :

- une tâche d'écriture : PROD.ECR^{CP2} (juin 2014).

Nous évaluons les compétences à l'écrit, surtout en ce qui concerne la segmentation et la production écrite de la chaîne de mots liés à l'oral par les phénomènes de resyllabation ;

- une tâche de lecture : LECT^{CP2} (juin 2014).

Nous évaluons les compétences en lecture, en regardant surtout la lecture de séquences caractérisées par les phénomènes de resyllabation.

9.1.1 Les hypothèses pour la fin du CP

Comme nous l'avons vu dans le chapitre 5, les enfants réussissent à mieux produire les séquences caractérisées par les phénomènes de resyllabation après 6 ans (Dugua, 2006 ; Dugua & Chabanal, 2006 ; Basset, 2000 ; Dugua & Baclesse, 2014). De plus, lors de la première année de l'école primaire, les enfants ont acquis de nouvelles compétences qui leur permettent d'accéder aux formes écrites. Nous avons vu dans le chapitre 2 que la connaissance de la *literacy* influence les compétences phonologiques ainsi que celles de la langue orale. Pour ces raisons, nous pouvons penser qu'à la fin du CP, il devrait y avoir une amélioration générale dans les tâches orales (dénomination et perception).

Quand les enfants apprennent les formes écrites des mots, ils apprennent quelles sont leurs frontières lexicales, puisqu'elles sont marquées graphiquement par des espaces blancs. Les activités de *literacy* aident à comprendre que la séquence *les ours* est constituée de deux mots : *les* qui finit par <s> et *ours* qui commence par la voyelle <o>. Comme nous l'avons montré dans le chapitre 2, ces compétences peuvent influencer les compétences phonologiques et par conséquent les productions orales. Les données sur l'écriture et la lecture nous permettront d'établir un rapport entre les compétences orales et les compétences écrites des enfants. Nous pouvons penser qu'il y aura une influence de l'acquisition de la *literacy* sur la production orale des phénomènes de resyllabation. Plus précisément, on s'attend à ce que la connaissance de la *literacy* fasse diminuer les erreurs de production de consonne erronée à l'oral

(comme *les[n]ours*) mais ne modifie pas les non-réalisations (*le(s)[Ø]ours*). En effet, il pourrait exister un lien entre une position prosodique vide à l'oral et l'espace blanc à l'écrit. L'acquisition de la *literacy* permet de visualiser les séquences de mots comme des blocs séparés par des marques graphiques. Cette représentation à l'écrit peut conduire les enfants à associer un espace blanc à l'écrit à une position vide. On s'attend donc à trouver une stabilité du nombre de non-réalisations entre le début et la fin du CP.

Grâce aux tâches d'écriture et de lecture, nous pourrions également enrichir les données sur la langue écrite déjà présentes en littérature, surtout en ce qui concerne les phénomènes de resyllabation. Pour les tâches liées à la langue écrite, nous nous attendons à beaucoup d'erreurs dans la représentation des frontières des mots (ex. <les sours>, <un nange>) comme déjà documenté dans la littérature pour les premières années de l'école primaire (Soum, 1997 ; Soum-Favaro *et al.*, 2014 ; Sabio, 2000). Dans la lecture, nous nous attendons à un taux plus bas de réalisations des liaisons obligatoires-catégoriques que ce qui a été enregistré par Dugua & Baclesse (2014), car leurs données ont été récoltées au CE1.

Pour les productions orales, nous évaluerons le rôle de la fréquence et de la phonologie pour mieux comprendre quel modèle de l'acquisition des frontières des mots est le plus adapté à décrire ce stade (lexical-constructionniste, Chevrot *et al.*, 2013 ; morpho-phonologique, Wauquier & Braud, 2005 et Wauquier, 2009). Pour la tâche de dénomination guidée DENa^{CP2}, nous proposerons également l'analyse de la composante sociolinguistique, comme déjà fait au début du CP.

Pour les tâches orales comme écrites, nous cherchons à vérifier l'influence de la fréquence des moments de lecture partagée sur les compétences en langue orale et écrite. Nous chercherons à montrer que la fréquence de la lecture partagée aide le développement des connaissances orales et écrites chez les enfants, comme il a été évoqué dans la littérature à ce sujet (Crain-Thoreson & Dale, 1992 ; Purcell-Gares, 1996 ; Landry & Smith, 2006 ; Lundberg, 2006).

9.1.2 La procédure des tâches orales : mai 2014

Pour les tâches orales PERC^{CP2}, DENa^{CP2} et DENb^{CP2}, les deux classes ont été testées en mai 2014, vers la fin de l'année scolaire. Les enfants ont été testés un par un, dans une salle silencieuse, pendant 40 minutes par tâche. L'ordre des tâches et de la présentation du répertoire varie selon l'enfant pour éviter des effets de facilitation ou la mise en place de stratégies. Des pauses étaient prévues entre les différentes tâches et entre les blocs qui composent les tâches.

Des phases d'entraînement précédaient la mise en forme des tâches. Dans ces phases, l'expérimentateur expliquait la démarche à l'enfant et fournissait des exemples jusqu'à ce que la tâche soit comprise.

9.2. Tâche PERC^{CP2} : perception d'erreurs et reformulation

9.2.1 Procédure de la tâche PERC^{CP2}

La procédure pour la tâche PERC^{CP2} est la même procédure que nous avons utilisée pour la tâche PERC^{CP1}. Pour les détails, nous invitons à la lecture de la section 8.2.

Comme nous l'avons dit dans le § 8.2.4, la tâche de perception d'erreurs et de reformulations que nous avons proposée en octobre était considérée comme assez problématique. Il s'agit d'abord d'une tâche métalinguistique où plusieurs compétences sont mises en jeu. De plus, cette tâche propose différents contextes et conditions qui ne sont pas parfaitement équilibrés du point de vue expérimental. Par exemple, nous avons 4 contextes de liaison opposés à un seul contexte d'enchaînement et à un seul contexte d'élosion. Dans l'élosion, nous avons deux conditions équivalentes : pour la condition de non-réalisation nous proposons la faute *la armoire*, et pour la condition de consonne erronée nous proposons un cas de figure semblable au moyen d'une séquence similaire du type *le artichaut*. Cependant, pour évaluer l'amélioration des performances dans la perception d'erreurs entre octobre et mai, nous avons reproposé à certains enfants la même tâche.

Nous avons testé un échantillon de 12 enfants pour les deux classes (6 au CP A et 6 au CP B). Nous avons choisi l'échantillon d'enfants selon les résultats de la tâche en octobre qui ont permis la création de trois groupes de quatre enfants :

- Groupe 1 (AM, NB, CMM, CSW)²⁸ : 4 enfants qui ont eu des résultats au-dessus de la moyenne (2 dans le CP A et 2 dans le CP B) ;
- Groupe 2 (LT, ELF, CR, BW) : 4 enfants qui ont eu des résultats moyens (2 dans le CP A et 2 dans le CP B) ;
- Groupe 3 (EB, EM, COO, ERY) : 4 enfants qui ont eu des résultats au-dessous de la moyenne (2 dans le CP A et 2 dans le CP B) ;

Entre octobre et mai, nous avons respecté le même ordre de présentation du matériel, la procédure et la façon de traiter les données (voir chapitres 7 et 8).

9.2.2 Résultats généraux de la tâche PERC^{CP2}

Pour chacun des 12 enfants, nous avons une seule reformulation qui corresponde à la détection d'une erreur sémantique. Dans tous les autres cas, les enfants détectent l'erreur-cible. Pour évaluer les résultats de la tâche, nous tenons d'emblée compte de cette reformulation, sans proposer d'abord l'analyse brute des données comme nous l'avions fait pour PERC^{CP1} (§ 8.2.3 et § 8.2.4).

Nous avons calculé les scores moyens de réponses correctes pour chaque condition (Tableau 37). Le score va de 0 (0 réponse correcte) à 8 (toutes les réponses sont correctes). Le score de réponses correctes tient compte des identifications des erreurs-cibles dans la condition non-réalisation (CONDnon-réal, *le(s)[Ø]anneaux* [leØa.no]) ou dans la condition d'insertion d'une consonne erronée (CONDcons.err, *le(s)[n]agneaux* [le.na.no]) et des non-identifications d'une erreur dans la condition de réalisation (CONDréal) où il n'y en a pas (*les[z]ananas* [le.za.na.na]). Pour un rappel des conditions, nous invitons à la lecture du paragraphe 8.2.4.

Conditions	Scores moyens de réponses correctes
CONDcons.err (<i>le(s)[n]agneaux</i>)	7,25
CONDnon-réal (<i>le(s)[Ø]anneaux</i>)	5,25
CONDréal (<i>les[z]ananas</i>)	7,50

Tableau 37. Scores de réponses correctes pour les trois conditions dans PERC^{CP2}

28 Les sigles désignent les enfants de façon anonyme.

L'analyse de la variance (ANOVA) des scores de réponses correctes dans trois modalités (types de conditions) utilisant les sujets et les items comme variables aléatoires montre un effet du facteur « condition » ($F_1(2,22) = 6,82, p < .05$; $F_2(2,14) = 3,10, p < .05$). L'analyse des contrastes révèle que la condition de non-réalisation implique moins de réponses correctes comparée aux autres conditions (CONDcons.err vs. CONDréal $F_1(1,11) = .65, p = .43, F_2(1,7) = 1,80, p = .22$; CONDcons.err et CONDréal vs. CONDnon-réal $F_1(1,11) = 9,34, p < .05, F_2(1,7) = 6,30, p < .05$).

En général, les enfants à ce stade acceptent plus facilement la non-réalisation d'une consonne (comme la CL dans *le(s)[Ø]anneaux* [leØa.no]) que l'introduction d'une consonne erronée qui perturbe les frontières de mots (*le(s)[n]agneaux* [le.na.no]).

9.3 Tâche DENa^{CP2} : dénomination guidée

9.3.1 Procédure de la tâche DENa^{CP2}

Nous avons proposé la même tâche de dénomination guidée qu'en octobre : la procédure, le répertoire et le traitement de données ont été les mêmes (voir chapitre 7 et 8). Pour cette tâche, les 43 enfants ont été testés.

9.3.2 Résultats généraux de la tâche DENa^{CP2}

Comme nous l'avons déjà fait pour DENa^{CP1} (chapitre 7), nous montrons les résultats généraux de DENa^{CP2} dans le tableau 38²⁹. Nous observons qu'en mai nous avons plus d'erreurs sur les items tests que sur les items *fillers* (2,80 vs. 0,50 sur le score maximal de 18) et la plupart de ces erreurs sont des erreurs de segmentation (2,70 vs. 0,60 sur 9).

²⁹ Le score maximal total correspond à 18. Pour les items tests et pour les *fillers* ce score peut monter seulement jusqu'à 9.

	Scores moyens		
	erreurs de la tâche (ex. <i>une abeille</i> pour <i>les[z]ours</i> ou absence de réponse)	erreurs de segmentation (ex. <i>le(s)[n]ours</i> ou <i>les[Ø]ours</i> pour <i>le(s)[z]ours</i>)	TOTAL
Items fillers	0,40	0,10	0,50
Items tests	0,20	2,60	2,80
Tests + fillers	0,60	2,70	3,30

Tableau 38. Résultats généraux dans DENa^{CP2}

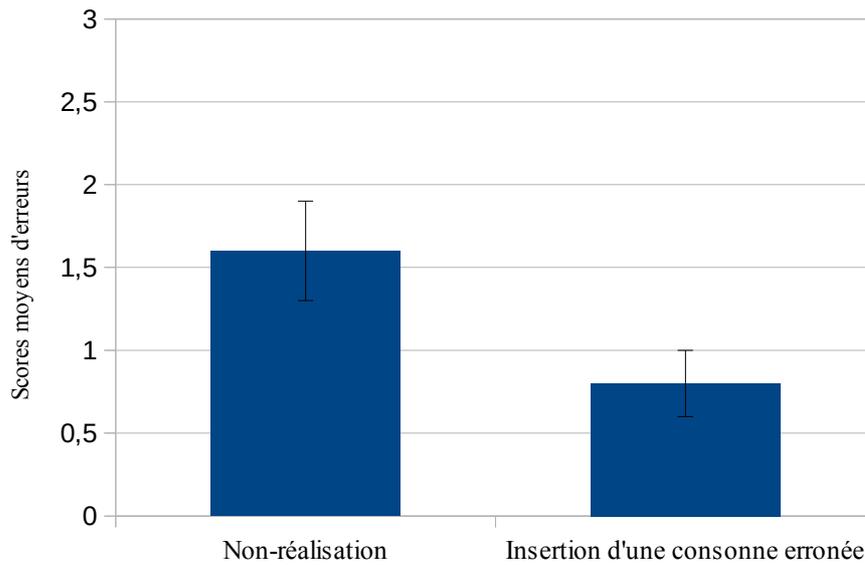
Nous nous concentrons maintenant sur les erreurs de segmentation pour les items tests. Les scores moyens d'erreurs sont décrits dans le tableau 39 où le score maximal est 3 pour chaque phénomène de resyllabation. Nous avons classifié les erreurs comme pour DENa^{CP1} (§ 8.3.3) : « non-réalisation » (*le(s)ours* [leØuʁs] pour [le.zuʁs]), « insertion d'une consonne erronée » (*les[n]ours* [le.nuʁs] pour [le.zuʁs]) et « autres erreurs » (*les grosse[s]araignées* [le.gʁo.sa.ʁe.ɲe] pour [la.gʁo.sa.ʁe.ɲe]). Le nombre d'erreurs de segmentation par phénomène (liaison, élision, enchaînement) est visible dans la deuxième colonne du même tableau.

Phénomène	Scores moyens d'erreurs TOTAL	Types d'erreurs	Scores moyens par type d'erreur
Liaison	0,7	Non-réalisation (<i>le(s)/Ø/ours</i> [leØuʁs] pour [le.zuʁs])	0,55
		Insertion d'une consonne erronée (<i>les[n]ours</i> [le.nuʁs] pour [le.zuʁs])	0,15
Élision	1,8	Non-réalisation (<i>la-armoire</i> [la.ɑʁ.mwaʁ] pour [laʁ.mwaʁ])	1,0
		Insertion d'une consonne erronée (<i>la[n]armoire</i> [la.nɑʁ.mwaʁ] pour [laʁ.mwaʁ])	0,8
Enchaînement	1,4	Autres erreurs	0,7
		Non-réalisation (<i>la grosse-araignée</i> [la.gʁos.ɑʁɛɲe] pour [la.gʁo.sa.ʁɛ.ɲe])	0,4
		Insertion d'une consonne erronée (<i>la grosse[z]araignée</i> [la.gʁo.za.ʁɛ.ɲe] pour [la.gʁo.sa.ʁɛ.ɲe])	0,3

Tableau 39. Scores par type d'erreur et pour chaque phénomène dans DENa^{CP2}

L'analyse de la variance (ANOVA) pour les scores d'erreurs pour trois modalités (trois phénomènes) utilisant les sujets et les items comme variables aléatoires montre un effet du facteur « phénomène » ($F_1(2,84) = 11,597$, $p < .001$; $F_2(2,4) = 10,02$, $p < .05$). L'analyse de contrastes révèle que la liaison est mieux maîtrisée que les deux autres phénomènes (élision vs. enchaînement $F_1(1,42) = 2,67$, $p = .10$, $F_2(1,2) = 1,17$, $p = .39$; élision et enchaînement vs. liaison $F_1(1,42) = 20,311$, $p < .001$, $F_2(1,2) = 405,582$, $p < .001$). Ces résultats sont très proches de ceux que nous avons vus dans le chapitre 8 : nous remarquons toujours un problème quand nous demandons aux enfants d'effacer la CL pour produire une élision ou un enchaînement.

La différence entre les non-réalisations et les insertions de consonnes erronées est statistiquement significative ($t_1(42) = 1,442$, $p < .001$; $t_2(8) = 2,14$, $p < .05$). Pour tous les phénomènes, la plupart des erreurs sont des non-réalisations (Graphique 9), comme nous l'avons déjà vu au début du CP.



Graphique 9. Erreurs d'insertion d'une consonne erronée ou de non-réalisation dans DENa^{CP2}

En conclusion, à la fin de l'année scolaire, les non-réalisations semblent être encore très présentes dans les productions des enfants à la fin du CP. De plus, si nous comparons le nombre de non-réalisations à la fin et au début du CP, il n'y a pas de différence significative ($t_1(41) = .84$, $p = .40$; $t_2(8) = .035$, $p = .21$)³⁰. Les non-réalisations restent donc stables au cours de l'année scolaire. Étant donné la persistance de ces non-réalisations, nous avons décidé de conduire une analyse phonétique de ce type de productions.

9.3.3 L'analyse phonétique des non-réalisations à la fin du CP

Au moyen du logiciel PRAAT, nous avons analysé phonétiquement les non-réalisations dans les cas de liaison, d'élision et d'enchaînement. Nous avons gardé la même classification des non-réalisations utilisée pour DENa^{CP1} (chapitre 8) : « continuum » (pour la liaison : [le.uks], sans réalisation de la CL ; pour l'élision : [lə.ɔ.ʁe.je]), « coup de glotte » (pour la liaison : [le.ʔuks], sans réalisation de la CL ; pour l'élision : [lə.ʔɔ.ʁe.je] et pour l'enchaînement : [la.pə.tit.ʔo.ʁãʒ]) et « pause » (pour la liaison : [le/uks], pour l'élision : [lə/ɔ.ʁe.je] et pour l'enchaînement : [la.pə.tit/o.ʁãʒ]). Pour l'enchaînement, nous ne pouvons pas avoir de productions du

³⁰ Pour une comparaison détaillée des résultats entre fin et début du CP, nous invitons à la lecture du chapitre 11.

type « continuum », car si les enfants produisent un continuum entre adjectif antéposé et nom, ils ne produisent pas une faute mais plutôt une bonne réalisation d'enchaînement ([la.pə.ti.to.ʁɑ̃ʒ]).

Nous présentons les résultats pour chaque phénomène consonantique dans le tableau 40. Les pourcentages sont calculés sur le nombre de non-réalisations pour chaque phénomène.

	%		
	liaison	élision	enchaînement
Coup de glotte ([le.ʁʊks],[lə.ʁʊ.ʁε.je], [la.pə.tit.ʁʊ.ʁɑ̃ʒ])	62	41,90	42,85
Pause ([le/ʊks], [lə/ʁ.ʁε.je], [la.pə.tit/o.ʁɑ̃ʒ])	8	35,5	57,15
Continuum ([le.ʊks], [lə.ʁ.ʁε.je])	30	22,5	//

Tableau 40. Analyse phonétique pour les non-réalisations dans DENa^{CP2}

Pour la liaison, les non-réalisations de la CL sont phonétiquement des productions où les enfants insèrent une occlusive glottale entre Mot1 et Mot2 (62 %). Les pourcentages de continuums et de véritables pauses sont moins élevés (8 % et 30 %).

Pour l'élision, il n'y a pas de grande différence entre les trois types de productions. Le pourcentage est légèrement supérieur pour l'insertion d'un coup de glotte entre Mot1 et Mot2 ([lə.ʁʊ.ʁε.je], 41,90 %).

En ce qui concerne l'enchaînement, il n'y a pas de grande différence entre l'insertion d'un coup de glotte ou l'insertion d'une pause (42,85 % vs. 57,15 %), même si les enfants semblent plutôt produire une pause.

Pour conclure, nous pouvons dire que, pour les trois phénomènes, il n'y a pas de différence importante entre les diverses stratégies que les enfants adoptent lors du traitement de l'espace entre les mots dans les cas de non-réalisations. L'espace sonore est parfois caractérisé par la production d'une occlusive glottale ou d'une pause. Parfois encore, les enfants enchaînent les deux mots l'un après l'autre sans prévoir de pause.

Étant donné la distribution très homogène de ces productions, nous avons décidé de ne pas nous arrêter à l'analyse d'une production particulière, contrairement à ce que nous avons fait pour DENa^{CP1}. En effet, cela ne nous apprendrait rien de plus. En revanche, nous avons préféré essayer d'établir une comparaison entre le traitement phonétique de l'espace sonore et l'acquisition des formes écrites pour évaluer les influences réciproques. Nous proposerons cette comparaison vers la fin de ce chapitre, après avoir présenté les données sur la langue écrite.

9.3.4 Le rôle de la fréquence

Comme nous l'avons fait dans la tâche DENa^{CP1} en octobre, nous avons décidé de mettre en rapport la fréquence SFI³¹ au pluriel avec le nombre d'erreurs d'insertion d'une consonne erronée [z] au début du Mot2 (chapitre 8). Si la fréquence des mots influence l'acquisition de leurs frontières, un Mot2 avec une haute fréquence au pluriel serait plus sujet à des erreurs d'insertion d'une consonne erronée [z]. Nous avons donc cherché à établir une corrélation statistique entre les erreurs d'insertion de [z] au début du Mot2 dans DENa^{CP2} et la fréquence SFI du Mot2. Le résultat du calcul de la corrélation montre qu'il n'y a pas de lien entre ces deux variables ($r(8) = -.21$, $p = .54$). À ce stade, la fréquence ne joue pas de rôle dans la dénomination guidée et donc dans les productions orales.

9.3.5 La variante sociolinguistique

Dans ce paragraphe, nous allons analyser l'influence de la composante sociolinguistique sur la tâche DENa^{CP2}. Nous nous concentrons ici sur les réalisations de la liaison obligatoire-catégorique, comme pour le début du CP (§ 8.5). Comme nous l'avons déjà dit dans le chapitre 8, les données récoltées permettent seulement une analyse qualitative de cette variable. Nous invitons donc à considérer cette section seulement comme une description de l'influence du milieu social et du niveau scolaire des parents sur les productions des enfants.

Nous évaluons le rôle du niveau d'instruction des parents qui permet de diviser

31 Nous invitons à la lecture du chapitre 7 pour une explication détaillée de la valeur de SFI.

les enfants en 3 groupes où le « niveau 1 » correspond au diplôme le moins élevé et le « niveau 3 » au diplôme le plus élevé. Nous invitons à la lecture de la section 8.5 et 7.1.2 pour une description plus détaillée de cette classification.

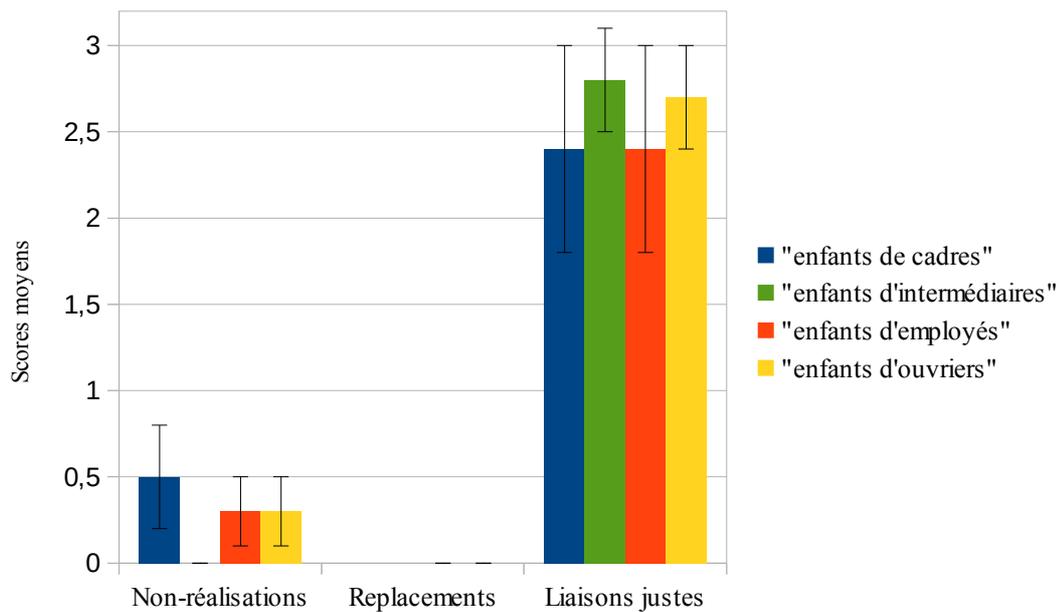
Dans le tableau 41, nous pouvons voir les scores moyens d'erreurs pour la tâche DENa^{CP2} dans les contextes de liaison. Le score maximal est 3.

		Scores moyens d'erreurs
Enfants avec des parents de	« Niveau 1 »	0,5
	« Niveau 2 »	0,71
	« Niveau 3 »	0,36

Tableau 41. Scores moyens d'erreurs des enfants dans DENa^{CP2} selon le niveau d'études des parents

À ce stade, nous ne remarquons pas de différence dans les réalisations de liaisons sur les trois groupes. Le niveau d'instruction des parents ne semble pas influencer les productions de liaisons obligatoires-catégoriques.

Nous évaluons également le milieu socioculturel des parents qui classifie les métiers parentaux en quatre groupes (pour plus de détails voir 7.1.3 et 8.5). Pour ces quatre groupes, nous présentons les scores moyens d'erreurs (remplacements et non-réalisations) et de productions correctes (liaisons justes) pour DENa^{CP2} dans le graphique 10. À ce moment de l'acquisition, les erreurs de remplacements disparaissent dans les trois groupes. De plus, les enfants qui ont les meilleurs résultats sont les « enfants d'intermédiaires », pas les « enfants de cadres », comme on l'avait vu au début du CP (§ 8.5). Cependant, que l'on analyse les erreurs ou que l'on tienne compte des liaisons justes, il n'y a pas de différence très nette selon le groupe d'appartenance.



Graphique 10. Les productions des enfants dans DENa^{CP2} selon le métier des parents

En conclusion, nous pouvons affirmer qu'à ce stade de l'acquisition de la langue, ni le milieu social ni le niveau d'instruction des parents n'influencent les productions de la liaison chez les enfants.

9.4 Tâche DENb^{CP2} : la tâche de dénomination guidée, une autre façon de produire

Dans la tâche de dénomination guidée DENa proposée en octobre et en mai, l'expérimentateur donne l'input en liaison, en élision et en enchaînement. L'enfant doit d'abord réussir à bien détecter les frontières de mots perturbées par les phénomènes consonantiques de l'input ; il doit ensuite enchaîner la forme lexicale à un Mot1 différent, pour nommer la carte. Par exemple, l'expérimentateur montre une carte où il y a deux oreillers et il produit *les[s]oreiller* [le.zɔ.ʁε.je]. Il montre ensuite une carte avec un seul oreiller dessiné et l'enfant doit produire *l'oreiller* [lɔ.ʁε.je]. Comme nous l'avons vu (§ 8.3.3), beaucoup d'enfants produisent la séquence *le oreiller* [lə.ɔ.ʁε.je] au lieu de [lɔ.ʁε.je]. Nous pouvons penser que ces productions sont causées par un effet

d'amorçage de l'input donné par l'expérimentateur et qu'elles sont donc créées par la tâche. Pour vérifier s'il y a ou non un effet de la tâche induisant un biais expérimental, nous avons construit une deuxième expérience de dénomination guidée DENb^{CP2}.

9.4.1 Procédure de la tâche DENb^{CP2}

Dans la tâche de dénomination guidée DENb^{CP2}, nous avons encore utilisé un couple de cartes mais ces cartes ne sont pas en rapport les unes avec les autres. Dans la première carte nous avons, par exemple, des verres et dans la deuxième carte nous avons des éléphants. L'expérimentateur montre la première carte et il la nomme en disant *les verres*. Ensuite, il montre la deuxième carte et l'enfant doit la nommer en produisant la liaison, par exemple *les[z]éléphants*. La procédure retenue pour cette tâche élimine l'effet d'amorçage de DENa^{CP2} car l'input donné par l'expérimentateur et la réponse attendue par l'enfant n'ont aucun lien phonologique.

Comme pour DENa^{CP2}, nous avons dans DENb^{CP2} 18 couples de cartes, 9 tests et 9 *fillers* (Annexe 8). Nous avons divisé les stimuli en 3 blocs de 6 couples de cartes (3 tests soit une production de liaison, une production d'élision et une production d'enchaînement mélangés à 3 *fillers*). Nous avons alterné la présentation de blocs comme dans DENa^{CP2} (voir chapitre 8). Le répertoire des items tests proposés est présenté dans le tableau 42. Tous les Mots2 choisis pour cette tâche ont une fréquence SFI moyenne dans le lexique des enfants MANULEX (voir chapitre 7). Nous avons utilisé les mêmes Mots2 que DENa^{CP2} pour construire les séquences.

Phénomène	Bloc	Expérimentateur : input	Enfants : cible
Liaison	1	Les lunettes	Les escargots [le.zɛs.kɑʁ.go]
	2	Les chaussures	Les oreilles [le.zɔ.ʁɛj]
	3	Les verres	Les éléphants [le.ze.le.fɑ̃]
Élision	1	Le soleil	L'abeille [la.bej]
	2	La voiture	L'avion [la.vjɔ̃]
	3	Le train	L'assiette [la.sjɛt]
Enchaînement	1	La grosse casserole	La grosse enveloppe [la.gʁo.sɑ̃v.lɑp]
	2	La minuscule coccinelle	La minuscule étoile [la.mi.nys.ky.le.twal]
	3	La petite fourmi	La petite aiguille [la.pə.ti.tɛ.gij]

Tableau 42. Répertoire pour DENb

9.4.2 Résultats de la tâche DENb^{CP2}

Nous présentons les erreurs de segmentation qui sont celles qui nous intéressent pour notre recherche. Nous avons écarté les « erreurs de la tâche » comme les non-réponses ou la dénomination incorrecte de la carte. Dans certains cas en effet (surtout avec la carte de l'abeille), beaucoup d'enfants n'ont pas été capables de nommer la carte car ils ne réussissaient pas à produire le nom de l'objet ou de l'animal dessinés. Nous avons aussi écarté, dans le compte des erreurs, les productions des enfants qui nomment la carte en utilisant un autre nom que la cible, comme par exemple *la guêpe* pour *l'abeille* ou *la petite épine* pour *la petite aiguille*.

Les scores moyens d'erreurs de segmentation sont présentés dans le tableau 43, où le score maximal est 9 pour les items test ou les *fillers* et 18 pour le total.

	Scores moyens d'erreurs de segmentation
Items Tests	1,60
Items <i>Fillers</i>	0,06
TOTAL	1,66

Tableau 43. Résultats généraux dans DENb^{CP2}

Dans ce type de dénomination guidée, les enfants font un nombre réduit de fautes sur les *fillers* (0,06). Cela veut dire qu'ils n'ont pas de problèmes à nommer les cartes là où il n'y a pas de phénomènes de resyllabation qui masquent les frontières de mots. C'est donc parmi les mots tests que nous avons enregistré la plupart des erreurs. Ces erreurs sont la conséquence d'une mauvaise segmentation. Dans le tableau 44, les scores sont divisés par phénomène consonantique (liaison, élision et enchaînement) et par type d'erreur enregistrée (nous retrouvons les trois types décrits pour DENa^{CP2} : non-réalisation, insertion d'une consonne erronée, autres erreurs). Le score maximal correspond à 3 pour chaque phénomène.

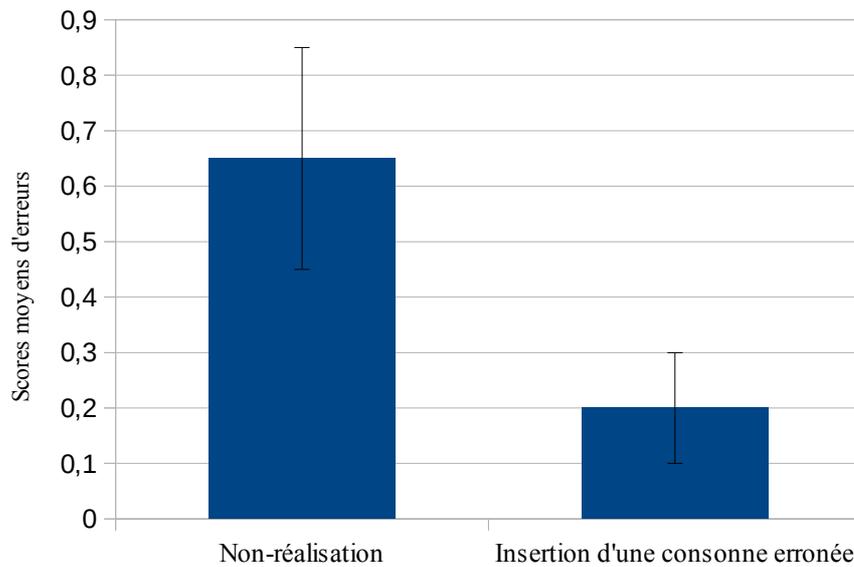
Phénomène	Scores moyens d'erreur TOTAL	Types d'erreurs	Scores moyens pour type d'erreurs
Liaison	0,3	Non-réalisation (ex. <i>le(s)[Ø]escargots</i> [le.Ø.ɛs.kɑ̃.go])	0,2
		Insertion d'une consonne erronée (ex. <i>le(s)</i> [n]escargots [le.nɛs.kɑ̃.go])	0,1
Élision	0,7	Non-réalisation (ex. <i>le-avion</i> [lə.a.vjɔ̃])	0,4
		Insertion d'une consonne erronée (ex. <i>le[n]avion</i> [lə.na.vjɔ̃])	0,3
Enchaînement	0,7	Non-réalisation (ex. <i>la grosse-enveloppe</i> [la.gʁos.ɑ̃v.lɔp])	0,4
		Insertion d'une consonne erronée (ex. <i>la grosse[z]enveloppe</i> [la.gʁo.zɑ̃v.lɔp])	0,3
		Autres erreurs (ex. <i>le grosse enveloppe pour la grosse enveloppe</i>)	0

Tableau 44. Classification des erreurs dans DENb^{CP2} selon le type d'erreur et le phénomène de resyllabation

L'analyse de la variance (ANOVA) pour les scores d'erreurs pour les trois modalités (trois phénomènes) utilisant les sujets et les items comme variables aléatoires ne montre pas d'effet du facteur « phénomène » ($F_1(2, 84) = 1,37, p = .87$; $F_2(2, 4) = 0,32, p = .15$). Produire une liaison, une élision ou un enchaînement implique la même problématique.

En ce qui concerne les types de fautes, il y a une différence statistiquement significative entre le nombre d'erreurs d'insertion d'une consonne erronée et le nombre d'erreurs par non-réalisation ($t_1(42) = 1,89, p < .05$; $t_2(8) = 2,25, p < .05$). De plus, pour tous les contextes, nous observons plus de non-réalisations du phénomène consonantique que d'insertion de consonne erronée (Graphique 11), même si dans les deux cas les scores restent bas³².

32 Le score maximal est 9.



Graphique 11. Erreurs d'insertion d'une consonne erronée et de non-réalisations dans DENb^{CP2}

Dans ce type de dénomination, pour l'enchaînement nous n'avons aucun autre type d'erreur, comme par exemple des erreurs sur le genre de l'article où, notamment, l'enfant produit *la* au lieu de *le*. Cela veut dire que cette erreur était liée à la présentation de l'input dans la tâche DENa^{CP2} et pouvait donc être attribuée à un effet d'amorçage. De plus, ces résultats envisagent un problème morphologique à ce stade de l'acquisition que nous allons mieux décrire dans la section suivante.

9.4.3 Une comparaison entre les deux tâches de dénomination en mai : DENa^{CP2} vs. DENb^{CP2}

Pour voir si les performances sont influencées par la procédure utilisée dans la tâche DENa^{CP2}, nous proposons une comparaison entre les deux tâches (DENa^{CP2} vs. DENb^{CP2}). Il faut se rappeler que, dans DENa^{CP2}, l'expérimentateur donne à l'enfant un input caractérisé par un phénomène consonantique et l'enfant doit produire un autre phénomène consonantique (par exemple passer d'une séquence en liaison *les[z]éléphants* [le.ze.le.fã] à une séquence en élision *l'éléphant* [le.le.fã]). Dans DENb^{CP2}, l'enfant doit simplement nommer la carte en produisant la séquence

caractérisée par le phénomène consonantique et l'expérimentateur n'influence pas sa réponse en donnant un input qui peut être interprété comme une amorce (par exemple l'expérimentateur dit *les verres* [le.vɛʁ] et l'enfant doit produire une séquence en liaison *les[z]éléphants* [le.ze.le.fã]).

Dans un premier temps nous pouvons confronter les résultats généraux pour les deux tâches de dénomination (Tableau 45).

	Scores moyens d'erreurs de segmentation	
	DENa ^{CP2}	DENb ^{CP2}
Test	2,60	1,60
Filler	0,10	0,06
TOTAL	2,70	1,66

Tableau 45. Comparaison des résultats généraux : DENa vs. DENb

Sur les erreurs de segmentation, il y a une différence statistiquement significative entre les deux tâches pour les items tests ($t_1(42) = 3,512$, $p < .001$; $t_2(8) = 3,472$, $p < .001$) et pour les *fillers* ($t_1(42) = 2,793$, $p < .001$; $t_2(8) = 3,944$, $p < .001$). Le premier type de dénomination guidée donne plus de difficultés aux enfants, pour les tests (2,60 vs. 1,60) comme pour les *fillers* (0,10 vs 0,06).

Nous présentons les résultats en détail pour les phénomènes consonantiques et les types d'erreurs sur les séquences tests dans le tableau 46, comme nous avons déjà illustré dans la section 9.3.2.

Phénomène	Scores moyens d'erreurs TOT		Type d'erreurs	Scores moyens d'erreurs	
	DENa ^{CP2}	DENb ^{CP2}		DENa ^{CP2}	DENb ^{CP2}
Liaison	0,7	0,3	Non-réalisation	0,55	0,2
			Insertion d'une consonne erronée	0,15	0,1
Élision	1,8	0,7	Non-réalisation	1,0	0,4
			Insertion d'une consonne erronée	0,8	0,3
Enchaînement	1,4	0,7	Non-réalisation	0,7	0,4
			Insertion d'une consonne erronée	0,4	0,3
			Autres erreurs	0,3	0

Tableau 46. Résultats par type d'erreur et par phénomène : DENa vs. DENb

Comme on peut le voir dans le tableau 46, pour tous les contextes et pour tous les types d'erreurs, nous avons des pourcentages plus élevés dans le premier type de dénomination guidée (DENa^{CP2}). De plus, dans DENa^{CP2}, nous remarquons des erreurs dans la catégorie « Autres erreurs » que nous n'enregistrons pas dans DENb^{CP2}.

En analysant les phénomènes un par un, on voit que pour l'élision ($t_1(42) = 1,214, p < .001$; $t_2(2) = 6,37, p < .05$) et l'enchaînement ($t_1(42) = 1,134, p < .001$; $t_2(2) = 5,40, p < .05$), il y a une différence statistiquement significative sur les scores d'erreurs pour les deux tâches qui n'est pas observée pour la liaison ($t_1(42) = 2,20, p = .12$; $t_2(2) = 1,88, p = .99$).

D'après cette analyse statistique, nous pouvons dire que la procédure de la tâche de dénomination guidée influence :

- la production des *fillers* ;
- la production des items tests en présence d'élision et d'enchaînement.

Cela veut dire que les résultats que nous avons pour la production d'élision et d'enchaînement en dénomination guidée sur la tâche DENa^{CP2} sont influencés par la procédure de la tâche elle-même. Si on a plus d'erreurs d'élision et d'enchaînement quand l'expérimentateur donne un input en liaison, c'est parce que cela implique une

compétence plus grande : réussir à segmenter la séquence de mots, faire abstraction de la CL et arriver à comprendre que le contexte vocalique prévoit la chute de la voyelle du déterminant ou qu'il faut utiliser une autre consonne pour enchaîner.

Le nombre d'erreurs plus élevé pour les *fillers* dans la tâche DENa^{CP2} (où on demande de passer du singulier au pluriel et vice versa) et la présence d'autres erreurs, dont l'utilisation du mauvais article pour l'enchaînement, sont aussi une conséquence de la tâche. La structure de la tâche fait émerger des difficultés qui sont d'origine morphologique, liées à l'acquisition du genre ou du nombre. Il est évident que si l'on donne un input comme *les maisons*, il est plus difficile pour les enfants de passer au singulier et de produire *la maison*, que de nommer directement la carte et produire la même séquence (*la maison*). La transformation, le calcul pour calculer le genre et le nombre mobilisent leurs connaissances morphologiques. On peut alors supposer que ceci rend la tâche plus complexe et entraîne des difficultés plus importantes. Les enfants doivent d'abord réactiver le trait de genre qui est neutralisé au pluriel. Ces résultats mettent en jeu la composante morphologique dans l'acquisition du français. Comme on l'a vu dans le chapitre 5 de cette thèse, plusieurs études montrent que la gestion complète de la composante morphologique du genre et du nombre est seulement effective après 6 ans et qu'elle est soutenue par l'apprentissage de l'écriture. Nous vérifierons le rapport entre la morphologie à l'écrit et les réalisations à l'oral dans la section 9.6.

Les performances des enfants en dénomination guidée ne sont pas influencées par la tâche dans le cas de la production de liaison. Nous avons vu (§ 9.3.2) que, dans le cas de la liaison obligatoire-catégorique, il y a plus de non-réalisations que d'insertions d'une consonne erronée. On peut donc considérer que les pourcentages plus élevés de productions de ces non-réalisations dans les cas de liaison doivent être le résultat du traitement des frontières de mots à ce stade de l'acquisition et non comme une conséquence de la procédure de la tâche DENa^{CP2}.

9.5 Les tâches liées à la langue écrite : juin 2014

Pour les tâches sur la langue écrite, les deux classes ont été testées en juin 2014. Nous avons proposé deux tâches : une tâche d'écriture (PROD.ECR^{CP2}) et une tâche de lecture (LECT^{CP2}). L'ordre des tâches et de présentation du répertoire varient par enfant pour éviter des effets de facilitation ou de stratégie. Des pauses sont prévues entre les différentes tâches et entre les blocs qui composent les tâches.

Pour la tâche d'écriture PROD.ECR^{CP2}, les enfants sont testés en groupe dans une classe vide du complexe scolaire. Pour la tâche de lecture LECT^{CP2}, les enfants sont testés un par un dans une salle silencieuse. Comme pour les productions orales, les lectures des enfants ont été enregistrées au format *.wav* avec une fréquence de 44000 Hz au moyen d'un enregistreur *Zoom H4*.

9.5.1 La tâche PROD.ECR^{CP2}

9.5.1.1 Procédure de la tâche PROD.ECR^{CP2}

Nous avons utilisé les 20 séquences tests déjà évoquées divisées en 3 blocs (Tableau 47). La division en blocs permet de mélanger les différents contextes à tester. Pour la construction du répertoire nous invitons à la lecture du chapitre 7. Nous rappelons que nous avons testé 5 phénomènes (liaison en [n], liaison en [z], élision, enchaînement, *h* initial) dans les contextes syntaxiques déterminant + nom. La moitié des contextes est connue de l'enfant ; l'autre moitié est classifiée comme inconnue (voir § 7.2.2 pour une classification).

Bloc	Mots tests à écrire
1	les ours, un oiseau, l'oie, une épée, un orage, un hippopotame, l'éléphant
2	un hérisson, les affiches, une ambulance, l'école, un ananas, une hirondelle
3	une autruche, l'âne, un hibou, les oranges, un igloo, une île, des ailes

Tableau 47. Bloc de séquences tests utilisées pour PROD.ECR^{CP2}

Nous avons divisé les enfants en 4 groupes : 2 groupes de 10, un groupe de 11 et un groupe de 12, en respectant la classe d'appartenance. Pour chaque groupe d'enfants nous avons changé l'ordre de présentation des blocs (Annexe 9).

Nous avons créé pour chaque élève un petit livret relié où il y a sur chaque page une image dans un carré. En dessous de l'image, il y a une ligne pointillée. C'est sur cette ligne pointillée que les enfants doivent écrire. Chaque bloc d'image est séparé par un papier coloré vide de façon à permettre aux enfants de faire une petite pause entre chaque bloc de mots à écrire. Un modèle de ce livret est visible en annexe (Annexe 10).

Les groupes d'enfants sont conduits un par un dans une salle de classe isolée. Sur chaque table il y a un crayon et une gomme pour écrire et effacer en cas de fautes. À chaque enfant est donné un livret. Sur la couverture, il y a une petite étiquette où ils doivent écrire leur prénom. Après que tous les enfants ont écrit leur prénom, ils sont invités, ensemble, à tourner la couverture pour découvrir la première image, par exemple « des ours ». L'expérimentateur nomme l'image en prononçant la séquence test, *les ours* [le.zuʁs]. Les enfants doivent donc écrire sur la ligne pointillée en dessous de l'image les deux mots *les ours*. Quand tout le monde a fini d'écrire la première forme orthographique, les enfants doivent tourner la page, regarder la deuxième image et écrire ce que l'expérimentateur dit, par exemple *un oiseau*. La même procédure est appliquée pour toutes les séquences de mots pour tous les blocs.

Les enfants sont invités à cacher avec leurs mains les formes orthographiques après les avoir écrites. D'abord nous pouvons ainsi comprendre quand l'enfant a fini ; ensuite, nous limitons les éventuels échanges entre eux. Pour éviter des facilitations, nous avons essayé de réduire la communication entre les enfants pendant la tâche et nous avons aussi évité de donner nous-mêmes de l'aide : quand les enfants nous disaient qu'ils n'étaient pas capables d'écrire certains mots, nous les avons poussés à essayer en leur expliquant qu'il ne s'agissait pas d'une évaluation. Cet encouragement a permis à tous les enfants d'écrire tous les mots dans leur livret.

La tâche PROD.ECR^{CP2} durait environ 45 minutes à l'issue de laquelle l'examineur ramassait les livrets.

9.5.1.2 Résultats de la tâche PROD.ECR^{CP2}

Tous les enfants écrivent les mots dictés en écriture liée et réalisent la tâche sans trop d'efforts. Nous avons considéré seulement 42 enfants dans les données car nous avons écarté les résultats d'un enfant (2,3 %) incapable de lire et d'écrire. Ses

productions écrites correspondent à des répétitions des mêmes lettres maîtrisées au niveau orthographique mais qui n'ont aucun rapport avec la séquence test et constituent des séquences impossibles dans la langue française (ex. <les moiam> pour *les affiches* ; <une menou> pour *une ambulance*, <le mou> pour *l'école* ou <une moue> pour *une hirondelle*).

9.5.1.2.1 La production écrite de noms

Les noms que nous avons choisis pour la tâche PROD.ECR^{CP2} sont des mots que les enfants ont déjà vus à l'école avec leurs maîtresses, mais ne sont pas des items transparents : les correspondances graphème-phonème ne sont pas un à un. Par exemple, dans la liste des mots tests, nous avons *autruche*, où le son [o] est orthographié avec le digramme <au> et le son [ʃ] avec le digramme <ch>. Cela fait de ce mot, un mot opaque. Le fait d'avoir introduit comme mots tests des mots non transparents implique la production d'erreurs, surtout en ce qui concerne l'orthographe des digrammes. Par exemple, les enfants substituent <o> avec <au> et vice-versa (<otruche> pour *autruche* ou <auranges> pour *oranges*). De plus, nous avons observé une confusion parmi les graphèmes <s> et <c> (<ambulanse> pour *ambulance* ou <érimon> pour *hérisson*), <j> et <g> (<oraj> pour *orange*) ou <s> et <z> (<oizo> pour *oiseau*) car, entre ces sons et les graphèmes correspondants, il n'y a pas de correspondance un à un.

Notre recherche est centrée sur les phénomènes consonantiques, sur l'espace entre Mot1 et Mot2 et sur le traitement des cas de non alignement entre frontières syllabiques et frontières lexicales. Par conséquent, nous avons décidé de ne pas analyser ce type d'erreurs orthographiques à l'intérieur des mots et nous nous sommes concentré sur l'espace entre frontières.

9.5.1.2.2 Liaison, élision, enchaînement dans PROD.ECR^{CP2} : frontières lexicales vs. frontières syllabiques

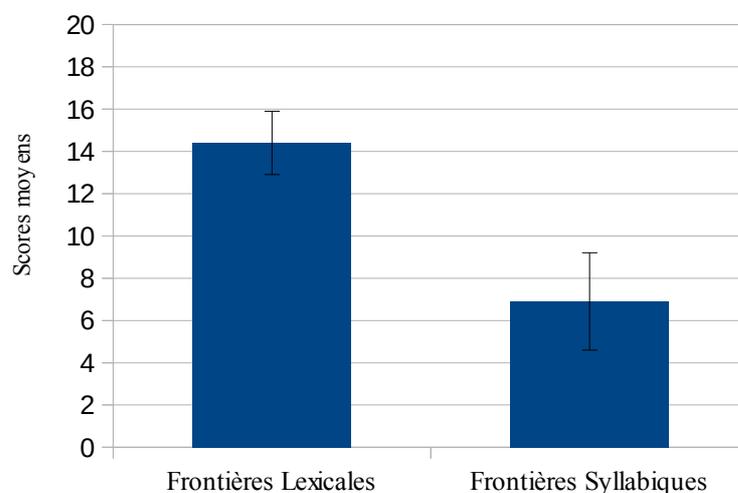
Nous avons remarqué plusieurs erreurs de segmentation où les enfants gardent la consonne de liaison, d'élision ou d'enchaînement attachée au nom. Nous avons ainsi des productions comme <les sailes> ou <des zailles> pour *les ailes*, qui montrant des

difficultés à orthographier le son [z], <une notruche> pour *une autruche* ou <lane> pour *l'âne*. Ces types de fautes sont le résultat d'un mauvais découpage et de l'utilisation de la syllabe comme unité pour segmenter et représenter les mots à l'écrit.

Nous tentons de classer les productions écrites des enfants selon deux modalités possibles :

1. Frontières lexicales : les séquences en liaison, élision et enchaînement sont écrites en respectant les frontières lexicales des mots (ex. pour la liaison <les ours>, pour l'élision <l'éléphant>, pour l'enchaînement <une autruche>).
2. Frontières syllabiques : la consonne qui resyllabifie sur le Mot2 est traitée comme la consonne initiale du Mot2, en respectant les frontières syllabiques (ex. pour la liaison <les sours>, pour l'élision <le léléphant>, pour l'enchaînement <une nautruche>).

Il y a une différence statistiquement significative entre le nombre des productions écrites à frontières lexicales et les productions écrites à frontières syllabiques ($t_1(37) = 6,501$, $p < .001$; $t_2(20) = 3,145$, $p < .001$). La plupart des enfants réussissent à écrire correctement les mots, en respectant les frontières lexicales perturbées par les phénomènes de resyllabation à l'oral (Graphique 12³³). Les productions écrites où les enfants favorisent les frontières syllabiques au lieu des frontières lexicales sont moindres. Nous pouvons donc conclure que les productions des enfants à la fin du CP respectent plutôt bien les frontières lexicales et que les séquences en liaison, élision et enchaînement sont écrites correctement, sans aucun problème de segmentation en modalité écrite.



Graphique 12. Scores moyens de productions écrites (frontières lexicales vs. frontières syllabiques) dans PROD.ECR^{CP2}

33 Le score maximal est 20.

Dans la prochaine section, nous analyserons ces productions écrites majoritaires respectant les frontières des mots en présence de phénomènes de resyllabation.

9.5.1.2.3 Les productions écrites qui respectent les frontières lexicales

Dans cette section, nous décrivons les productions écrites correctes à « frontières lexicales ». Dans le tableau 48, nous présentons les pourcentages de productions écrites correctes selon le phénomène de resyllabation testé (liaison au pluriel, liaison au singulier, élision et enchaînement). Les pourcentages sont calculés sur la base des productions totales de tous les enfants (soit 16 séquences fois 42 enfants). Il faut souligner que, dans cette première analyse, nous n'avons pas pris en considération les séquences où le nom s'écrit avec la lettre *h* en initiale. Ces productions seront traitées en détail dans la section 9.5.1.2.6.

Phénomènes	% des productions écrites qui respectent les frontières lexicales (<i>les ours, l'âne, une ambulance</i>)
Liaison au pluriel	86,90
Liaison au singulier	90,70
Élision	85,47
Enchaînement	96,80

Tableau 48. Pourcentage des productions écrites qui respectent les frontières lexicales dans PROD.ECR^{CP2}

En général, plus de 85 % des productions écrites ne montrent pas de problème de segmentation. Le phénomène qui a l'air d'être le mieux maîtrisé à l'écrit est l'enchaînement (avec 96,80 % de productions correctes) ; le plus complexe est l'élision (85,47 % de productions écrites correctes). À ce niveau scolaire, les élèves n'ont pas trop de problèmes à bien orthographier les noms à voyelle initiale, même si l'input donné à l'oral est caractérisé par une resyllabation.

Nous pouvons penser que la forme orthographique des noms tests est déjà présente dans le lexique orthographique des enfants. En présence d'une séquence comme *les ours*, ils orthographient correctement l'article et traitent le nom comme un mot qui commence par une voyelle. Le problème n'est pas de traiter le Mot2 comme un mot qui commence par une voyelle, mais plutôt de bien orthographier ce segment

vocalique et donc de maîtriser les correspondances graphème-phonème. Seulement 5 % des productions écrites ne présentent pas d'erreurs dans les correspondances graphème-phonème. Comme nous l'avons déjà dit, il est très facile pour les enfants de s'appuyer sur la phonologie pour orthographier le [o] au début du mot *autruche* et ainsi d'utiliser <o> au lieu du digramme <au>. Dans ce cas spécifique, 59,50 % des enfants orthographient <au> avec <o> contre 35,71 % des enfants qui orthographient correctement <au> à l'initiale (les 4,79 % restant cherchent à reproduire le digramme en produisant <une eautruche> ou <une oetruche>). Dans le cas du mot *ails*, 95,20 % des enfants orthographient le digramme <ai> avec un <e>, en appuyant leur écriture sur la représentation du son. Seulement 2,4 % des enfants réussissent à bien reproduire <ai> au début du mot. Ces résultats montrent qu'à ce stade de l'acquisition de l'orthographe, les règles de correspondance graphème-phonème guident l'accès à la forme lexicale du mot. Cependant, les enfants représentent correctement Mot1 et Mot2 en présence de phénomènes de resyllabation à l'oral.

9.5.1.2.4 Des questions morphologiques

Dans le chapitre 5, nous avons traité la question de l'acquisition de l'écriture et la difficulté que les enfants francophones ont à acquérir les marques du nombre et du genre à l'écrit. Au début de l'acquisition de la *literacy*, les enfants utilisent les correspondances graphème-phonème pour écrire. En français, dans la plupart de cas, les <s> finaux qui marquent le pluriel à l'écrit ne sont pas prononcés à l'oral. Sauf quelques cas (ex. *les ours*), la correspondance graphème-phonème n'a pas lieu pour la marque du pluriel. Cette absence cause beaucoup de problème dans les premières productions écrites des enfants qui, dans la plupart des cas, n'écrivent tout simplement pas le <s> à la fin du mot au pluriel (§ 5.6.2.1).

Nous avons décidé d'analyser les productions écrites des séquences en liaison au pluriel pour voir comment les enfants traitent le morphème du pluriel à l'écrit. Dans le tableau 49, nous présentons les pourcentages de <s> qui apparaissent à la fin du Mot2 dans ces formes écrites³⁴. Les pourcentages sont calculés sur la totalité de séquences en liaison au pluriel (soit 42 fois 4).

34 Les <s> à la fin du Mot1 *les* est orthographié dans 100 % des cas.

Séquences	% de marquages du pluriel
Les oranges	9,52
Des ailes	19,04
Les ours	61,9
Les affiches	7,19
TOTAL	24,40

Tableau 49. Pourcentages du marquage du pluriel à l'écrit dans les séquences en liaison en [z] dans PROD.ECR^{CP2}

Nous pouvons voir que dans 24,40 % des cas, les enfants arrivent à bien marquer les séquences au pluriel, en ajoutant à l'écrit un <s> final au Mot2. Si nous analysons en détail les résultats pour chaque séquence, nous pouvons voir que les enfants ont moins de problèmes à marquer le pluriel pour l'item *ours*. L'explication est très simple : à l'oral, ils entendent un son [s] à la fin du mot, car, dans la forme singulière comme dans la forme plurielle, il se termine par un [s] réalisé ([uks]). Dans cette séquence, une correspondance graphème-phonème a lieu et les enfants se servent d'elle pour bien orthographier le Mot2. Pour les autres séquences, où cette correspondance n'a pas lieu, la mise en œuvre d'un processus de marquage est plus difficile. Les résultats nous montrent que les enfants ajoutent le <s> du pluriel à la fin du Mot2 dans moins de 20 % des cas. Cette tendance montre qu'à ce stade de l'acquisition de l'écriture, les enfants n'ont pas encore bien compris que le <s> à la fin du mot correspond à la marque du pluriel. L'écriture est encore une écriture phonologique, où ce sont plutôt les sons produits qui sont bien orthographiés.

9.5.1.2.5 Les productions écrites qui respectent les frontières syllabiques

Nous avons analysé les productions écrites qui ne respectent pas les frontières lexicales mais les frontières syllabiques où les enfants représentent à l'initial du Mot2 la consonne qui resyllabifie sur ce mot à l'oral (ex. <les sours> pour *les ours*).

Pour cette première analyse, nous avons seulement pris en considération les séquences caractérisées par les phénomènes consonantiques de liaison, élision et enchaînement sans tenir compte des contextes déterminant + nom à *h* initial. Dans le tableau 50, nous résumons les pourcentages des productions écrites qui respectent les

frontières syllabiques pour chaque phénomène de resyllabation. Les pourcentages sont calculés sur le nombre de productions écrites totales multiplié par le nombre total d'enfants (16 séquences fois 42 enfants).

Phénomène	% de productions écrites à « frontières syllabiques »
Liaison au pluriel	12,79
Liaison au singulier	9,50
Élision	14,53
Enchaînement	3,20

Tableau 50. Pourcentages de productions écrites à "frontières lexicales" dans PROD.ECR^{CP2}

Les résultats montrent que les enfants réussissent à mieux écrire les séquences en enchaînement sans être perturbés par la resyllabification présente à l'oral (3,20 % des productions écrites à « frontières syllabiques »).

Nous avons décidé de traiter en détail chaque phénomène de resyllabation et d'examiner la manière dont il a été produit à l'écrit, pour mieux comprendre les différences ou les affinités entre eux. Les pourcentages présentés ne font pas référence à la totalité des productions écrites chez les enfants. Il y a plutôt un calcul lié aux 4 séquences tests proposées pour chaque phénomène dans la tâche.

La liaison en [z] au pluriel et la liaison en [n] au singulier à l'écrit

Pour les séquences déterminant-nom avec liaison en [z], nous avons classifié les mauvaises segmentations en trois catégories :

1. <s> au début du nom (ex. <les sours> pour *les ours*),
2. <z> au début du nom (ex. <les zours> pour *les ours*),
3. hyposegmentation (ex. <lesaffiches> pour *les affiches*).

Dans les catégories 1 et 2, la CL est traitée comme une consonne lexicale qui devient initiale du Mot2. Dans le troisième type de production, que nous avons appelée « hyposegmentation », déterminant et nom sont considérés comme une seule unité et non comme deux mots séparés. Il n'y a pas de véritable segmentation en cours.

	% d'erreurs sur les productions totales pour la liaison en [z]
<s> au début du nom	5,95
<z> au début du nom	5,95
hyposegmentation	1,19

Tableau 51. Segmentations incorrectes à l'écrit pour la liaison en [z] dans PROD.ECR^{CP2}

En ce qui concerne les 3 types d'erreurs (Tableau 51), il y a un équilibre entre les productions écrites avec un <s> ou un <z> au début du nom (5,95 %). Le fait que ce soit un [z] ou un [s] ne change pas grand chose (même pourcentage dans le tableau 51), c'est seulement la démonstration d'une difficulté à bien gérer les correspondances entre les graphèmes et les phonèmes. Le pourcentage de production, quant à lui, d'hyposegmentations est très faible (seulement 1,19 %).

Pour la liaison en [n], au singulier, nous remarquons deux types de productions écrites qui laissent apparaître un problème dans le traitement des frontières des mots :

1. <n> au début du nom (ex. <un noiseau> pour *un oiseau*),
2. <n'> + nom (ex. <un n'oiseau> pour *un oiseau*).

Dans le premier cas, la CL est la consonne initiale du Mot2 ; dans le deuxième cas, la consonne est insérée entre Mot1 et Mot2 mais séparée du Mot2 par une apostrophe. Les enfants qui écrivent <un n'oiseau> pour *un oiseau* comprennent que [n] n'est pas une consonne lexicale initiale du Mot2 mais ne sont pas capables de la représenter comme seule consonne finale du Mot1. Ils insèrent une apostrophe pour la séparer du nom et, peut être, rendre compte de son statut, différent d'une consonne lexicale normale.

Dans le tableau 52, nous présentons les pourcentages de ces deux types de productions écrites. Cette fois-ci, le pourcentage de productions où la CL est traitée comme initiale du Mot2 est très élevé (26,79 %) mais il y a un faible pourcentage de productions écrites où la CL est traitée comme une consonne qui reste au milieu, avant une apostrophe (2,38 %).

	% d'erreurs sur les productions totales pour la liaison en [n]
<n> au début du nom	26,79
<n'> + nom	2,38

Tableau 52. Segmentations incorrectes à l'écrit pour la liaison en [n] dans PROD.ECR^{CP2}

Le tableau nous montre qu'il y a beaucoup de productions écrites où la consonne de liaison [n] est utilisée comme initiale du mot. Le pourcentage est très élevé comparé aux productions en [z] ou [s] initiaux dans les cas de liaison au pluriel (Tableau 51). Ces résultats différents sont peut-être dus à une moindre quantité de mots commençant par le son [z] que de mots commençant par le son [n]. Si l'on regarde le lexique orthographique que nous avons créé en utilisant les cahiers des enfants, nous pouvons voir qu'il y a seulement 7 items commençant par le son [z] contre 18 items qui commencent par le son [n]. Si nous élargissons cette comparaison à la base de données MANULEX, nous trouvons, pour le CP, 759 noms communs qui commencent par le son [n] contre seulement 11 noms qui commencent par [z]. Un son [n] au début du nom est donc plus probable qu'un son [z]. En suivant ce raisonnement, nous pouvons dire que, si les enfants ne sont pas encore sûrs de la nature du segment initial du Mot2, choisir le bon serait plus facile dans le cas où la CL est [z]. Dans ce cas, les enfants peuvent plus facilement orthographier le Mot2 comme un mot qui commence par une voyelle et qui est en contexte de liaison car ce dernier cas de figure est plus probable qu'un mot commençant par le son [z]. Pour les séquences au singulier, la tâche serait plus difficile car il y a plus de mots dans le lexique orthographique qui commencent par la consonne <n> et le son [n]. Il y a donc une probabilité plus élevée d'orthographier Mot2 avec le graphème <n> à l' initiale.

L'élision à l'écrit

En analysant les productions écrites pour l'élision, nous pouvons créer quatre catégories :

1. hyposegmentation (ex. <loie> pour *l'oie*),
2. déterminant + <l> comme consonne lexicale du nom (ex. <une/un/le/la loie> pour *l'oie*³⁵),

35 Dans PROD.ECR, l'enfant doit produire à l'écrit la séquence prononcée par expérimentateur, donc dans ce cas, *l'oie*.

3. <n> comme consonne lexicale du nom (ex. <noie> pour *l'oie*),
4. hyper-segmentation (ex. <la ne> pour *l'âne*).

Pour l'élision, nous enregistrons deux phénomènes : l'hyposegmentation, où déterminant et nom sont traités comme une seule unité et la consonne <l> devient l'initiale du Mot2 ; l'hyper-segmentation, où les enfants segmentent la séquence en deux morceaux qui correspondent aux unités syllabiques. L'hyper-segmentation peut avoir deux explications différentes. Pour ce qui est de la première explication, nous pouvons penser que les enfants cherchent à reproduire la structure syllabique de la séquence. Pour cela, ils séparent les deux syllabes grâce à l'insertion d'un espace blanc. Une deuxième explication peut être liée à une compétence plutôt lexicale. Les enfants comprennent que la séquence en élision est en réalité composée de deux mots séparés (Mot1 et Mot2). Pour rendre compte de ces deux mots (deux unités lexicales), ils insèrent un espace blanc ; le problème est que cet espace blanc n'est pas inséré au bon endroit : la segmentation finale n'est pas la bonne.

En ce qui concerne l'erreur en 2, du type <une/un/le/la loie> pour *l'oie*, l'article est intégré dans la forme lexicale du nom et la consonne de l'article est perçue comme une consonne initiale du Mot2. Enfin, dans 3 les enfants gardent la CL [n] dans la séquence orale *un[n]oie* comme initiale du nom, en la traitant comme une consonne lexicale.

	% d'erreurs sur les productions totales pour l'élision
hyposegmentation	39,27
déterminant + <l> comme consonne lexicale	5,38
<n> comme consonne lexicale	9,52
hyper-segmentation	4,17

Tableau 53. Segmentations incorrectes à l'écrit pour l'élision dans PROD.ECR^{CP2}

Dans le tableau 53, nous enregistrons les différents types de productions écrites. Nous pouvons noter que le pourcentage d'hyposegmentations est très élevé (39,29 %). Les enfants ont tendance à traiter la séquence en élision comme une seule séquence où le déterminant est collé au nom. Les hyper-segmentations sont moins fréquentes (4,17 %).

En ce qui concerne l'utilisation d'une mauvaise consonne comme consonne initiale, la CL [n] est maintenue dans 9,52 % des cas comme consonne lexicale et initiale du Mot2. Dans seulement 5,38 % des cas, la consonne [l] du Mot1 est gardée comme consonne initiale du Mot2 et un autre déterminant est antéposé.

L'enchaînement à l'écrit

Pour les productions écrites de l'enchaînement, nous pouvons établir deux catégories :

1. <n> comme consonne lexicale du nom (ex. <une nautruche> pour *une autruche*),
2. <l> comme consonne lexicale du nom (ex. <lautruche/une lautruche> pour *une autruche*³⁶).

Dans le premier cas, la consonne d'enchaînement qui est resyllabifiée dans Mot2 est traitée comme une consonne lexicale initiale. Dans le deuxième cas, l'attaque vide du Mot2 est remplie avec la consonne trouvée dans les séquences en élision (*l'autruche*) et elle est traitée comme consonne lexicale et initiale du nom.

Le tableau 54 montre les pourcentages de ces deux types de productions écrites par rapport à toutes les productions des enfants pour les quatre séquences en enchaînement (soit 42 fois 4). La plupart des erreurs sont du type 1, où la consonne d'enchaînement est gardée comme consonne initiale du Mot2 (8,72 %).

	% d'erreurs sur les productions totales pour l'enchaînement
<n> comme consonne lexicale du nom	8,72
<l> comme consonne lexicale du nom	4,07

Tableau 54. Segmentations incorrectes à l'écrit pour l'enchaînement dans PROD.ECR^{CP2}.

9.5.1.2.6 H comme initiale à l'écrit

Pour analyser les séquences écrites où Mot2 commence par le graphème *h*, il faut d'abord vérifier si, dans les noms dictés, *h* muet ou *h* aspiré sont orthographiés à l'initiale. Il faut se rappeler que dans le cas du *h* muet, la consonne *h* est seulement

³⁶ Pour l'enchaînement, l'enfant doit écrire la séquence dictée par l'expérimentateur, donc *une autruche*.

graphique et elle n'a donc pas de réalisation au niveau phonétique. En revanche, nous n'avons pas non plus de correspondance graphème-phonème dans les mots qui commencent par *h* aspiré, même si le *h* aspiré montre une position phonologique qui empêche les phénomènes consonantiques d'avoir lieu (voir chapitre 3). Nous présentons d'abord les pourcentages d'écriture de *h* muet ou aspiré comme initiale du nom dans le tableau 55, sur la totalité des productions des enfants (42 fois 4).

	% de <h> au début du nom
Contexte <i>h</i> muet (<i>hippopotame</i> [i.pɔ.pɔ.tam])	5,95
Contexte <i>h</i> aspiré (<i>hibou</i> [ibu])	3,57

Tableau 55. Pourcentage de <h> orthographié au début du Mot2 dans PROD.ECR^{CP2}

Que l'on considère *h* muet ou *h* aspiré, les pourcentages de réalisations écrites d'un *h* graphique en initial sont très faibles (3,57 % et 5,95 %). Ce faible pourcentage nous amène à penser que les enfants n'ont pas encore totalement acquis le lexique orthographique et qu'à ce stade, ils préfèrent plutôt utiliser la phonétique et la phonologie pour écrire. Le graphème <h> n'a pas de production sonore correspondante, il est donc plus difficile à produire à l'écrit car il n'est présent que dans la représentation visuelle du mot. Contrairement au cas des mots commençant par [z] (§ 9.5.1.2.5), il y a de nombreux mots à <h> initial dans le lexique orthographique dans notre base de données (14 mots) et dans MANULEX (891 mots). La non production à l'écrit de *h* au début du nom n'est donc pas seulement une question de fréquence de ce graphème au début du mot dans le lexique, mais aussi d'une ambiguïté de la réalisation sonore.

Pour les séquences à « frontières syllabiques » en présence d'un Mot2 à *h* initial, nous avons trouvé 3 types de productions écrites :

1. hyposegmentation (ex. <unippopotame> pour *un hippopotame*),
2. <n> comme consonne lexicale (ex. <un nibou> pour *un hibou* ou <un nippopotame> pour *un hippopotame*),
3. <l> comme consonne lexicale (ex. <un lippopotame> pour *un hippopotame*).

Les types 1 et 3 sont présents seulement dans le cas d'un Mot2 à *h* muet initial. Dans 1, les enfants ne segmentent pas les deux mots et ils traitent Mot1 et Mot2 comme une seule unité. Dans 3, au lieu du *h* au début du Mot2, nous trouvons le <l> qui peut être l'effet d'un contexte d'éliision (<l'hippopotame> → <l'ippopotame> → <lippopotame>). Le type 2 a lieu soit avec *h* muet soit avec *h* aspiré. Les enfants produisent comme consonne initiale la CL ou la consonne d'enchaînement qu'ils trouvent dans ces deux phénomènes, même si la liaison et l'enchaînement n'ont pas lieu dans le cas du *h* aspiré. Dans le tableau 56, nous présentons les pourcentages de ces 3 catégories de productions écrites.

	% des productions écrites	
	<i>h</i> muet	<i>h</i> aspiré
Hyposegmentation	2,38	0
<n> comme consonne lexicale	14,29	7,14
<l> comme consonne lexicale	1,19	0

Tableau 56. Segmentations incorrectes à l'écrit en présence d'un Mot2 à *h* initial dans PROD.ECR^{CP2}

Les résultats ne sont pas étonnants. Nous trouvons plus de problèmes de segmentation à l'écrit pour les séquences avec des Mots2 à *h* muet initial. Dans le cas du *h* muet, les phénomènes consonantiques sont toujours réalisés, les erreurs d'hyposegmentation ou de <n> et <l> comme consonnes lexicales sont donc plus prévisibles. Dans le contexte du *h* aspiré initial, nous avons moins d'erreurs (7,14 %) car les Mots2 sont toujours réalisés sans resyllabation.

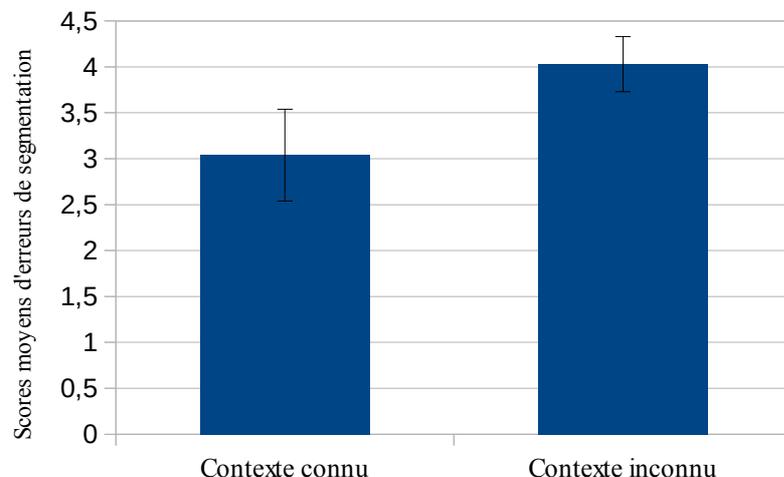
Il faut souligner que nous avons dicté les séquences. Nous avons donc fourni un input où les phénomènes consonantiques ne sont pas réalisés dans le cas de séquences à *h* aspiré initial. L'input donné dans la tâche peut avoir influencé ces résultats. Pour voir si les enfants connaissent vraiment la différence entre liaison ou enchaînement dans ces deux contextes, nous devons analyser les données de lecture traitées dans la section 9.5.2.2.2.

9.5.1.2.7 Le rôle de la fréquence et de la familiarité dans PROD.ECR^{CP2}

Familiarité. Contextes inconnus et contextes connus : y a-t-il une différence dans les productions écrites ?

Comme nous l'avons expliqué dans la présentation du répertoire pour la tâche PROD.ECR^{CP2} (§ 7.2.2), la moitié des contextes est connue par les élèves et l'autre moitié n'est pas connue par les élèves. Les contextes connus sont traités en classe par les maîtresses qui proposent des travaux de lecture et d'écriture sur ces séquences. Les contextes inconnus ne sont pas sujets à un apprentissage explicite à l'école. Le choix de ces deux contextes fait émerger la variable « familiarité » : dans les contextes connus, le niveau de familiarité des enfants avec les séquences est très élevé car les enfants ont déjà vu la suite Mot1 + Mot2 à l'écrit. Cela n'est pas le cas pour les contextes inconnus parce que les élèves n'ont jamais vu dans le contexte scolaire ces deux mots l'un après l'autre.

En évaluant les productions écrites sur ces deux contextes, nous pouvons mieux comprendre le rôle de ce que nous avons appelé « familiarité du contexte » dans les productions écrites. Nous avons vérifié si la différence d'erreurs dans les productions écrites (erreurs de segmentation à l'écrit) dans les deux types de contextes (connus vs. inconnus) est significative. Les résultats montrent que cette différence est significative ($t_1(41) = 2,139$, $p < .001$; $t_2(9) = 1,55$, $p < .05$) : nous observons plus d'erreurs de segmentation à l'écrit pour les contextes inconnus (Graphique 13³⁷).



Graphique 13. Scores d'erreurs à l'écrit : contexte connu vs. contexte inconnu

37 Le score maximal est 18.

Nous pouvons dire que si les enfants ont déjà vu une certaine séquence déterminant-nom, caractérisée par un phénomène de resyllabation à l'écrit et qu'ils ont, par conséquent, une représentation disponible de cette séquence, ils ont moins de problèmes à déterminer sur papier les frontières des deux mots qui la composent. La familiarité avec les séquences d'items est donc déterminante pour la segmentation à l'écrit à ce stade.

Il faut remarquer que nous avons classifié les contextes selon la base de données de l'écrit créée au moyen des items utilisés dans les activités d'écriture et de lecture réglées par les maîtresses. Nous ne pouvons pas tenir compte des activités externes à la classe où les enfants peuvent accéder aux formes écrites (ex. lecture partagée avec leurs parents).

La fréquence des mots écrits

Quand nous avons présenté le répertoire, nous avons décidé de choisir les mots les plus fréquents dans le lexique des enfants après avoir créé notre base de données sur leurs cahiers (§ 7.2.2). Même si nous avons choisi les noms les plus fréquents, il faut noter que nous avons des items qui ont plus d'occurrences (comme *oie*, avec 7 occurrences) et des mots qui ont moins d'occurrences (comme *île*, avec 2 occurrences).

Dans un premier temps, nous évaluons le rapport entre la valeur des occurrences des Mots2 dans notre base de données (§ 7.2.2) et les scores de productions écrites totalement correctes (où le Mot2 est correctement orthographié et les frontières lexicales sont bien identifiées). Le calcul de la corrélation entre ces deux variables montre qu'il n'y a pas de relation entre eux ($r(19) = -.016$, $p = .384$).

Dans un deuxième temps, nous avons traité les erreurs de segmentation à l'écrit et les occurrences des items dans notre base de données. Au moyen de cette analyse, nous voulons voir si la capacité à connaître l'initiale du Mot2 à l'écrit et par conséquent à traiter les consonnes de liaison, d'élision et d'enchaînement telles qu'elles sont, est liée à la quantité d'inputs écrits reçus. Un Mot2 plus lu ou écrit à l'école peut être mieux identifié et mieux représenté à l'oral dans une séquence caractérisée par des phénomènes de resyllabation. Un mot moins écrit peut devenir difficile à orthographier et à séparer correctement du déterminant dans des séquences où il devrait y avoir un phénomène de resyllabation à l'oral. Au moyen du calcul du r de Pearson, nous avons évalué si le nombre d'erreurs de segmentation à l'écrit (productions écrites à

« frontières syllabiques » du type <les sours>) est corrélé avec la fréquence des mots dans notre base de données. Les résultats nous montrent qu'il n'y a pas de corrélation entre le nombre d'occurrences des Mots2 et les erreurs de segmentation à l'écrit ($r(19) = .53$, $p = .06$). La fréquence des Mots2 à l'écrit ne résout pas les problèmes d'identification des frontières de mots à l'écrit.

Pour vérifier ces corrélations, nous allons proposer les mêmes corrélations en considérant une autre valeur de la fréquence : l'indice SFI des Mots2 au CP dans MANULEX (voir chapitre 7). La fréquence dans MANULEX tient compte de plus de variables que la nôtre et elle a aussi été normalisée par les créateurs de la base de données. Nous avons calculé la valeur de r en mettant en rapport : les productions écrites correctes et le SFI d'une part et les erreurs de segmentation en écriture (« frontières syllabiques ») et le SFI d'autre part. Les résultats montrent que dans aucun des deux cas, il n'y a de corrélation entre ces variables ($r_1(19) = .30$, $p = .19$; $r_2(19) = .29$, $p = .20$).

En conclusion nous pouvons dire que la fréquence intrinsèque des mots (dans notre base de données et dans MANULEX) n'a pas de rapport avec les productions écrites correctes des séquences des mots. De plus, la fréquence n'aide pas à résoudre les problèmes de segmentation dans le cas des phénomènes de resyllabation à l'oral. Nous pouvons conclure que l'écriture à ce stade est encore plutôt guidée par le décodage graphème-phonème et qu'elle n'est pas influencée par la fréquence des items dans le lexique orthographique des enfants.

9.5.2 La tâche LECT^{CP2} de lecture à voix haute

9.5.2.1 Procédure et matériel de la tâche LECT^{CP2}

Nous avons créé 20 cartes (5cm x 10cm) où nous avons imprimé 20 séquences de mots à lire, écrites en minuscules, police *Times New Roman*, taille 25. Les séquences écrites sont les mêmes que celles que nous avons proposées pour la tâche PROD.ECR^{CP2} d'écriture (voir § 9.5.1.1).

La tâche de lecture a eu lieu une semaine après la tâche d'écriture PROD.ECR^{CP2}. Il faut se souvenir que, dans la tâche PROD.ECR^{CP2}, nous demandions aux enfants d'écrire les mots prononcés par l'expérimentateur mais sans toutefois

donner de retours sur la forme écrite correcte qu'ils doivent orthographier. Par conséquent, nous n'avons pas mis à disposition des enfants des informations sur les productions écrites correctes qui peuvent influencer la tâche de lecture et leur laisser des traces en mémoire. De plus, le délai d'une semaine entre les deux tâches nous garantit d'annuler les éventuels effets d'amorçage (Ferrand, 2009).

Dans cette tâche, l'expérimentateur montre une carte à l'enfant et l'invite à lire la séquence écrite, à voix haute. Dès que l'enfant a lu la séquence, il passe à la deuxième carte jusqu'à ce que toutes les cartes soient lues. Nous avons varié l'ordre de présentation des cartes pour éviter l'effet de facilitation.

Les enfants sont testés un par un dans une salle silencieuse et les productions orales sont enregistrées au format *.wav* à la fréquence de 44000 Hz au moyen d'un enregistreur *Zoom H4*.

9.5.2.2 Résultats généraux de la tâche LECT^{CP2}

Dans cette première classification des performances des enfants, nous n'avons pas pris en compte les problèmes de segmentation en lecture sur les phénomènes de resyllabation (comme par exemple *les[Ø]affiches* pour *les[z]affiches*). Cette analyse permet une classification des lectures des participants sous trois catégories :

1. Séquences lues correctement.

La lecture des items est bien réalisée. Les enfants lisent correctement les articles et les noms proposés dans la tâche.

2. Lecture syllabique.

Les enfants nous proposent une lecture comme [le/le/fã] pour la séquence *l'éléphant*, ou [le/za/fi] pour *les affiches*.

3. Problèmes de lecture (lecture erronée/inventée).

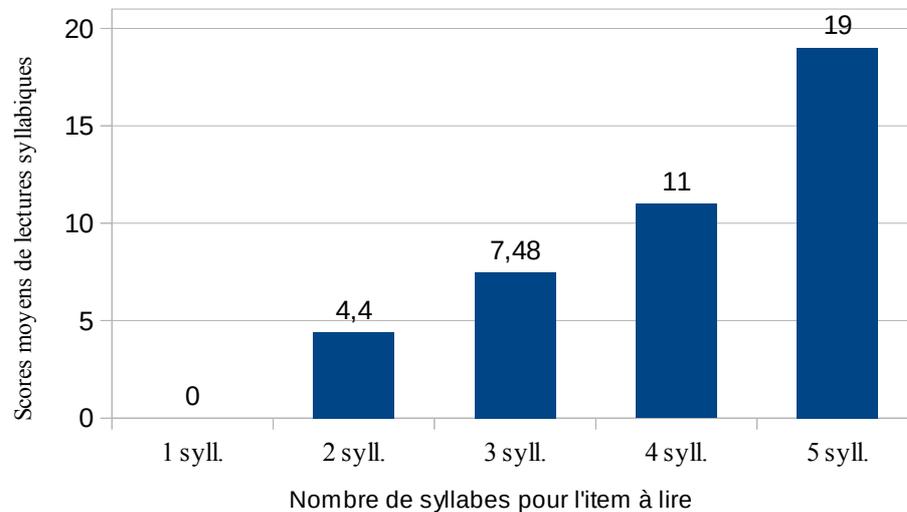
Les enfants ne réussissent pas à lire les mots ou se trompent dans la lecture. Par exemple, ils lisent *une abeille* pour *une ambulance*.

Nous présentons en pourcentages les résultats des 42 enfants dans le tableau 57. Nous avons écarté 2 enfants (4,6 %) qui n'ont pas été capables de lire les mots.

	%
Séquences lues correctement	81,3
Lecture syllabique	14,1
Problèmes de lecture (lecture erronée/inventée)	4,6

Tableau 57. Pourcentages des lectures selon les trois catégories dans LECT^{CP2}

La plupart des séquences sont lues correctement (81,3 %). Il y a néanmoins des pourcentages de lectures syllabiques et qui montrent encore des problèmes de lecture à la fin du CP (14,1 %). Nous pouvons penser que la lecture syllabique est liée à la longueur du nom que les enfants doivent lire et, en effet, nos données confirment cette intuition. Comme nous le montrons dans le graphique 14, le score de lectures syllabiques est plus élevé lorsque les noms à lire ont un nombre plus grand de syllabes.

Graphique 14. Scores moyens de lectures syllabiques dans LECT^{CP2} selon la longueur de Mot2

Dans les données présentées jusqu'à maintenant, nous n'avons pas tenu compte du traitement des frontières de mots en lecture. Nous analyserons cela dans la section suivante.

9.5.2.2.1 Les productions des phénomènes de resyllabation dans LECT^{CP2}

Dans cette section, nous proposons l'analyse des séquences lues correctement. Nous avons écarté les lectures syllabiques (*les-a-fi-ches*) et les lectures inventées/erronées (*une abeille* pour *une ambulance*).

Nous analysons les réalisations correctes et les réalisations incorrectes des phénomènes de resyllabation en proposant deux catégories :

- REAL.LECT^{CP2} : les phénomènes de resyllabation sont bien réalisés (la lecture correspond à une production orale adulte sans faute) ;
- ERR.LECT^{CP2} : les phénomènes de resyllabation ne sont pas réalisés en lecture (ex. *les[Ø]affiches* [leØa.fiʃ] pour *les[z]affiches* [le.za.fiʃ], *éléphant* [Øe.le.fã] pour *l'éléphant* [le.le.fã], *une-île* [yn.il] pour *une[n]île* [y.nil]) ou les enfants produisent une pause après Mot1 et réalisent ensuite un Mot2 à attaque remplie (ex. *les-pause-[z]affiches* [le/za.fiʃ] pour *les[z]affiches* [le.za.fiʃ], *le-pause-léléphant* [lə/le.le.fã] pour *l'éléphant* [le.le.fã]).

Nous remarquons plus de REAL.LECT^{CP2} (61,5 %) que de ERR.LECT^{CP2} (38,5 %). Pour vérifier quel phénomène consonantique cause le plus de difficultés, nous proposons une comparaison entre les pourcentages de ERR.LECT^{CP2} pour chaque phénomène de resyllabation testé (Tableau 58). Pour calculer les pourcentages, nous avons considéré les résultats pour les phénomènes de resyllabation (donc sur 4 séquences fois 41 enfants). Nous ne considérons pas dans ce calcul le contexte à *h* initial, qui sera traité plus tard.

	% ERR.LECT ^{CP2}
Liaison en [z]	36,61
Liaison en [n]	35,77
Enchaînement	23,12
Élision	2,37

Tableau 58. ERR.LECT^{CP2} pour les phénomènes de resyllabation en lecture

Nous pouvons voir que c'est la lecture en élision qui cause le moins de problème. À l'inverse, les séquences en liaison et enchaînement donnent plus de difficultés aux enfants : il y a un segment (latent, dans la liaison) visiblement attaché au Mot1 au niveau orthographique qui doit devenir à l'oral l'attaque du Mot2.

Nous présentons l'analyse de ERR.LECT^{CP2} pour chaque phénomène de resyllabation dans les paragraphes suivants.

La liaison

Nous classifions les erreurs de lecture (ERR.LECT^{CP2}) selon deux catégories :

- NON-REAL.LECT^{CP2} : erreur de non-réalisation (ex. *les[Ø]affiches* [leØa.fiʃ] pour *les[z]affiche* [le.za.fiʃ]) ;
- PRA.LECT^{CP2} : erreur de Pause + Remplissage de l'Attaque, où les enfants, après avoir lu l'article, proposent une pause et ensuite un Mot2 qui commence par la CL (ex. *les-pause-[z]affiches* [le/za.fiʃ] pour *les[z]affiche* [le.za.fiʃ]).

Dans le tableau 59, nous pouvons voir qu'il n'y a pas de différence pour les deux contextes de liaison sur les deux types d'erreurs (liaison en [z] 2,9 % et 33,36 % vs. liaison en [n] 2,38 % et 33,36 %).

	%	
	NON-REAL.LECT ^{CP2}	PRA.LECT ^{CP2}
Liaison en [z]	33,36	2,9
Liaison en [n]	33,9	2,38

Tableau 59. Pourcentages de NON-REAL et PRA pour la liaison dans LECT^{CP2}

La plupart des erreurs sont des NON-REAL.LECT^{CP2} (plus ou moins 33 %) où la CL n'est pas réalisée. Ceci n'est pas étonnant. Les mots, même s'ils sont liés par la liaison à l'oral, sont bien séparés par un espace blanc à l'écrit. Les enfants traitent les deux mots comme deux blocs séparés, deux unités qu'il faut lire en produisant une pause entre eux. Nous allons donner une description plus détaillée des non-réalisations en lecture au moyen d'une analyse phonétique présentée dans la section 9.5.2.2.3.

En ce qui concerne les erreurs PRA.LECT^{CP2}, les pourcentages sont plus bas que ceux des non-réalisations. Dans ce type d'erreur, l'attaque vide est remplie par la CL traitée comme consonne initiale du Mot2 et une pause est insérée entre Mot1 et Mot2. Il faut dire que visuellement, le graphème associé à la CL est attaché au Mot1, mais dans ces productions, il est lu comme s'il était au début du Mot2. Pour quelle raison peut-on trouver ce type d'erreurs ? Une réponse peut être donnée par la théorie que nous avons présentée sur la syllabe en lecture. Dans le chapitre 5, nous avons expliqué que la syllabe est utilisée comme unité de base par les enfants français qui apprennent à lire. Les lectures PRA.LECT, où la CL prend la place d'une attaque vide, seraient donc une conséquence des compétences des enfants à ce stade : ils chercheraient à reproduire la structure syllabique CV au début du Mot2. En plus, ce type de lecture montre que la phonologie et sa représentation influencent encore les activités de *literacy*. Mais cette consonne qui est placée au début de Mot2 est-elle une CL ou plutôt une consonne lexicale ? Les résultats des analyses phonétiques de la lecture nous donneront une réponse (§ 9.5.2.2.3.3).

L'enchaînement

Pour l'enchaînement, les seules erreurs enregistrées sont les erreurs NON-REAL.LECT^{CP2} (23,12 %) où la consonne d'enchaînement n'est pas resyllabifiée sur le Mot2 (*une-île* [yn.il] pour *une[n]île* [y.nil]).

Même si, à l'oral, il y a un phénomène de resyllabation qui permet le lien, à l'écrit les deux mots sont plutôt représentés comme séparés. Dans la lecture, où la forme écrite est oralisée, comme nous pouvons nous y attendre, les enfants lisent les deux mots séparément en respectant l'espace blanc entre les deux mots. En conséquence, les enfants en lecture produisent une non-réalisation de l'enchaînement. Cela veut dire que, lorsque les enfants doivent lire la séquence *une[n]île* [y.nil], ils produisent *une-île* [yn.il]. La consonne d'enchaînement est produite dans la lecture de l'article ([yn]) et le nom est prononcé sans altération de ses frontières.

L'élision

Nous avons essayé la même classification pour l'élision dans le peu d'erreurs réalisées par les enfants dans ces séquences (environ 2 %). Nous avons : des NON-REAL.LECT^{CP2} (*éléphant* [Øe.le.fã] pour *l'éléphant* [le.le.fã]) et des productions PRA.LECT^{CP2} (*le léléphant* [lə/le.le.fã] pour *l'éléphant* [le.le.fã]). Les résultats sont résumés dans le tableau 60.

	%	
	NON-REAL.LECT ^{CP2}	PRA.LECT ^{CP2}
Élision	0,59	1,78

Tableau 60. Pourcentages de NON-REAL et PRA pour l'élision dans LECT^{CP2}

En cas d'élision à l'oral, un Mot2 à voyelle initiale devient un mot où l'attaque est remplie par la consonne de l'article qui resyllabifie après la chute du [ə]. Ces erreurs en lecture peuvent donc être le résultat d'un problème avec les frontières des mots : les enfants n'ont pas encore un lexique orthographique stable et peuvent penser que le mot [e.le.fã] est plutôt le mot [le.le.fã] et que pour lire correctement la séquence, il faut ajouter l'article [lə]. Comme pour la liaison, on peut dire qu'à ce stade, la représentation phonologique influence la *literacy*.

9.5.2.2.2 Les séquences à *h* muet et *h* aspiré au début du nom

Dans cette analyse, nous écartons les lectures « syllabiques » ([ɛ̃/ni/bu] ou [ɛ̃ ɛ̃/i/bu]) ou non-correctes (*une abeille* pour *un hibou*).

Dans les quatre Mots2 tests à *h* initial, deux commencent par *h* aspiré (*hibou* et *hérisson*) et deux par *h* muet (*hirondelle* et *hippopotame*). Dans les séquences commençant par *h* aspiré, pour lire correctement la séquence, les enfants doivent ne pas réaliser la liaison. Nous utilisons la classification utilisée pour les autres phénomènes et nous proposons trois catégories :

- PRA.LECT^{CP2} : les enfants lisent l'article et, après une pause, le Mot2 avec une réalisation de la CL comme consonne initiale du Mot2 (*un-pause-[n]hibou* [ẽ/ni.bu]). La séquence n'est pas lue correctement.
- REAL.LECT^{CP2} : les enfants lisent la séquence en produisant la liaison (*un[n]hibou* [ẽ.ni.bu]). La séquence n'est pas lue correctement, car la liaison est réalisée.
- NON-REAL.LECT^{CP2} : les enfants lisent les deux mots comme deux formes séparées. La liaison n'est pas réalisée (*un[Ø]hibou* [ẽØi.bu]) et la séquence est lue correctement.

Nous présentons les pourcentages de ces trois types de productions pour chaque séquence testée (Tableau 61) et dans la deuxième colonne le total. Nous présentons les pourcentages de ces types de lecture sur la totalité des lectures de séquences avec *h* aspiré, soit 41 fois 2.

	% TOTAL		%
PRA.LECT ^{CP2}	9,8	un-pause-[n]hérisson	7,4
		un-pause-[n]hibou	12,2
REAL.LECT ^{CP2}	12,2	un[n]hérisson	14,6
		un[n]hibou	9,8
NON-REAL.LECT ^{CP2}	62,2	un[Ø]hérisson	65,8
		un[Ø]hibou	58,6

Tableau 61. Pourcentages de PRA, REAL et NON-REAL dans les contextes où Mot2 commence par *h* aspiré dans LECT^{CP2}

La plupart des enfants lisent les séquences correctement, en évitant de réaliser la liaison mais en introduisant une pause entre Mot1 et Mot2 (62,20 % de NON-REAL.LECT^{CP2} *un[Ø]hérisson* ou *un[Ø]hibou*). Nous ne savons pas si ce pourcentage de lecture est lié au fait que les enfants sont conscients que les noms dans ces séquences commencent par des *h* aspirés (et que la liaison ne doit pas avoir lieu) ou s'ils sont plutôt liés au fait que les mots à l'écrit sont séparés par un espace blanc (et

que l'orthographe influence la lecture et la non-réalisation de la liaison à l'oral). Nous allons mieux comprendre ces lectures au moyen de l'analyse phonétique des productions en lecture dans une autre section (§ 9.5.2.2.3).

En ce qui concerne les séquences déterminant + nom qui commencent par *h* muet, les résultats de lecture sont présentés dans le tableau 62. Nous avons utilisé la même classification que pour le *h* aspiré : PRA.LECT^{CP2} (*un-pause-[n]hippopotame*, *une-pause-[n]hirondelle*), REAL.LECT^{CP2} (*un[n]hippopotame*, *une[n]hirondelle*) et NON-REAL.LECT^{CP2} (*un[Ø]hippopotame*, *une[Ø]hirondelle*). Il faut toutefois préciser que dans le cas du *h* muet, la lecture de ces séquences sans réalisation de la liaison ou de l'enchaînement (NON-REAL.LECT^{CP2} *un[Ø]hippopotame*, *une[Ø]hirondelle*) doit être traitée comme une erreur. En revanche, la réalisation de la liaison ou de l'enchaînement (REAL.LECT^{CP2} *un[n]hippopotame*, *une[n]hirondelle*) est considérée comme licite.

	% TOTAL		%
PRA.LECT ^{CP2}	7,34	une-pause-[n]hirondelle	7,34
		un-pause-[n]hippopotame	7,34
REAL.LECT ^{CP2}	18,2	une[n]hirondelle	14,6
		un[n]hippopotame	21,8
NON-REAL.LECT ^{CP2}	36,5	une[Ø]hirondelle	29,2
		un[Ø]hippopotame	33,8

Tableau 62. Pourcentages de PRA, REAL et NON-REAL dans les contextes où Mot2 commence par *h* muet dans LECT^{CP2}

Nous pouvons remarquer qu'il n'y a pas de différence entre les lectures selon le phénomène consonantique qui s'installe entre Mot1 et Mot2. Dans l'enchaînement et dans la liaison, les enfants ont tendance à ne pas réaliser les phénomènes consonantiques (pourcentage total de non-réalisations 36,50 %) : les deux mots sont traités comme deux blocs séparés. Dans ce contexte, le graphème *h* peut empêcher la

réalisation des phénomènes de resyllabation. Nous analysons ci-dessous les non-réalisations de la liaison et de l'enchaînement dans le cas du *h* muet afin d'observer à quoi correspond phonétiquement ce <h> graphique.

9.5.2.2.3 Analyses phonétiques de la tâche LECT^{CP2}

9.5.2.2.3.1 Analyse phonétique des non-réalisations (NON-REAL.LECT^{CP2})

Dans la section 9.3.3, nous avons analysé phonétiquement les non-réalisations en DENa^{CP2}. Nous avons remarqué qu'entre Mot1 et Mot2, les enfants ajoutent parfois une occlusive glottale. Nous analysons maintenant les non-réalisations dans la tâche de lecture LECT^{CP2} à l'aide du logiciel PRAAT pour pouvoir évaluer ce qui se passe dans l'espace sonore entre déterminant et nom quand le phénomène de resyllabation n'est pas réalisé.

Après nos analyses, nous avons trouvé trois types de traitement différents de l'espace sonore entre Mot1 et Mot2 en lecture :

1. Les enfants proposent une véritable pause entre Mot1 et Mot2, où rien n'est produit : pour la liaison *le(s)-pause-affiches* [le/a.fiʃ] ou *un-pause-oiseau* [ɛ̃/wa.zo] pour *un[n]oiseau* [ɛ.nwa.zo] ; pour l'enchaînement *une-pause-île* [yn/il]). Un exemple est donné dans la figure 34.
2. Les enfants produisent au début du Mot2 un coup de glotte. Donc les productions en lecture doivent être traitées comme *le(s)[ʔ]affiches* [le.ʔa.fiʃ], *un[ʔ]oiseau* [ɛ̃.ʔwa.zo ; pour l'enchaînement *une[ʔ]île* [yn.ʔil] (Figure 35).
3. Entre déterminant et nom, ni pause, ni coup de glotte ne sont produits. Les deux mots sont, au contraire, prononcés comme un « continuum » du type [ɛ̃.wa.zo] mais la CL n'est pas réalisée. Une production [ɛ̃.wa.zo] est exemplifiée dans la spectrogramme en figure 36³⁸.

38 Comme nous l'avons déjà dit, la production d'un continuum dans le cas d'enchaînement n'est pas un erreur mais une bonne réalisation.

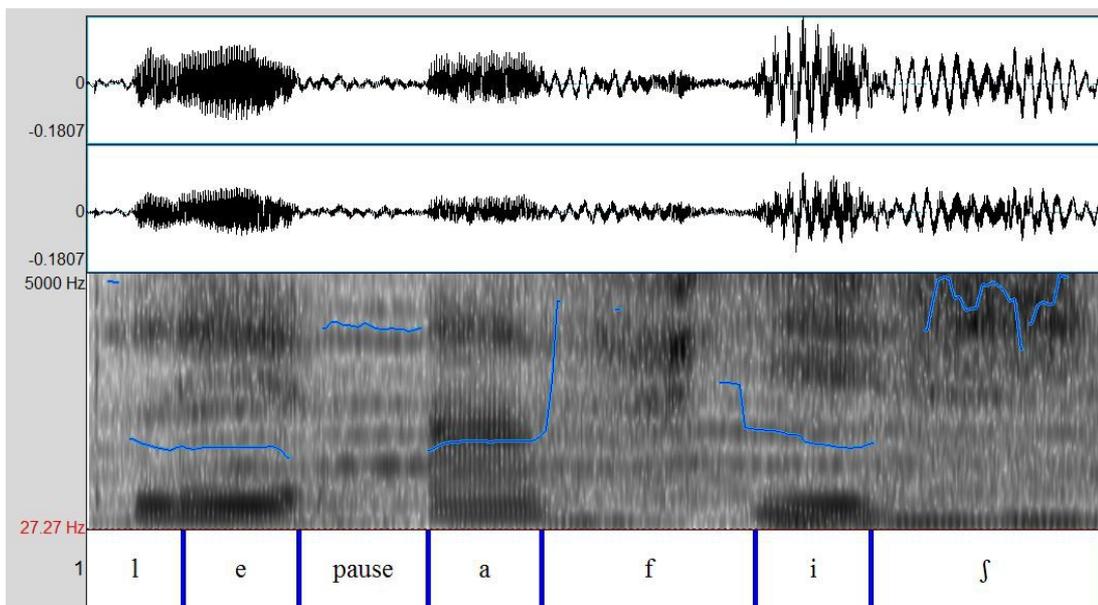


Figure 34. Spectrogramme pour la production en lecture de [le/a.fi]

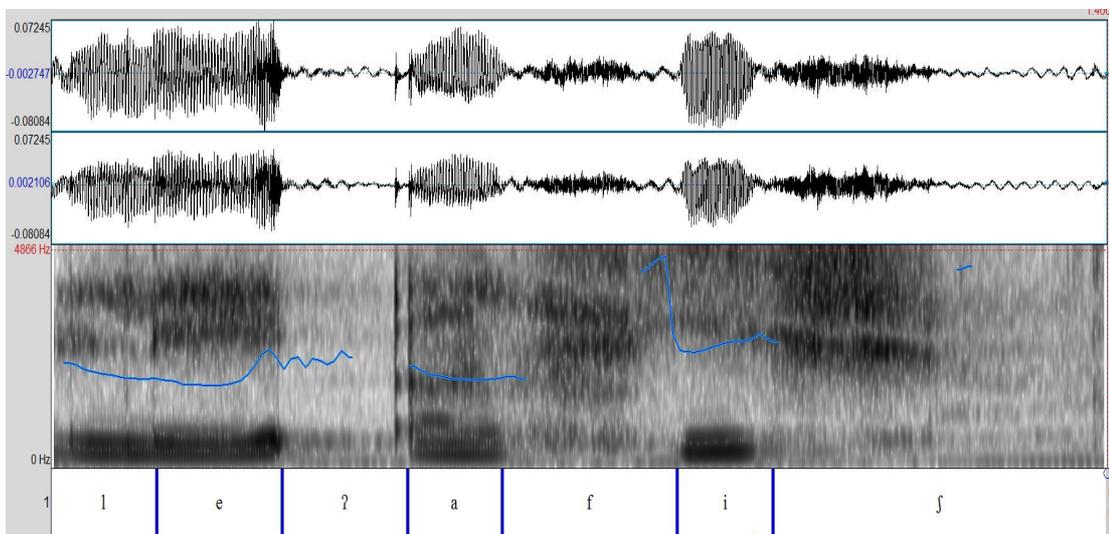


Figure 35. Spectrogramme pour la production en lecture [le.ʔa.fi]

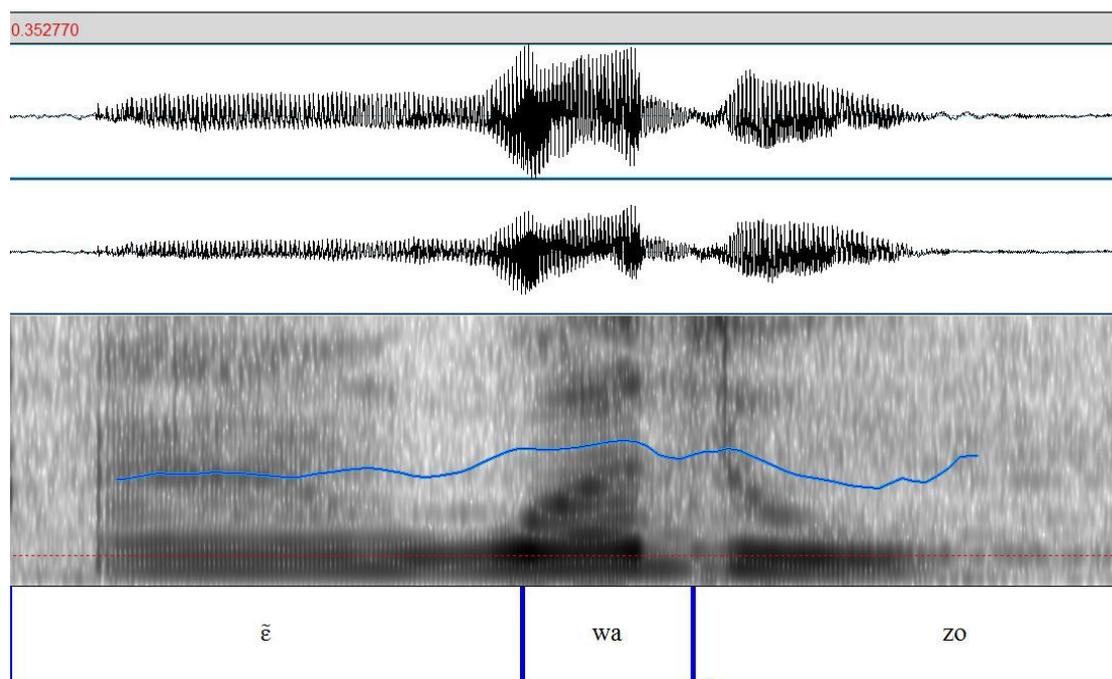


Figure 36. Spectrogramme pour la production en lecture [ɛ̃.wa.zo]

Nous avons calculé les pourcentages de ces deux productions (pause, coup de glotte, continuum) dans la totalité de non-réalisations pour chacune des 4 séquences tests proposé pour chaque phénomène. Nous avons écarté de l'analyse 12 productions (7,5 %) où la qualité de l'enregistrement ne permettait pas une analyse phonétique fine. Les résultats sont résumés dans le tableau 63.

	% NON-REAL. LECT ^{CP2}		
	liaison en [z]	liaison en [n]	enchaînement
pause (ex. [le/a.fiʃ], [ɛ̃/wa.zo], [yn/il])	47	35,4	52,94
coup de glotte (ex. [le.ʔa.fiʃ], [ɛ̃.ʔwa.zo], [yn.ʔil])	53	58,8	47,06
continuum (ex. [ɛ̃.wa.zo])	//	5,8	//

Tableau 63. Analyse des non-réalisations pour LECT^{CP2}

Pour la liaison en [z] et en [n], dans le total des non-réalisations, il n'y a pas de grande différence entre les deux types de productions. Nous pouvons donc penser que lorsque les enfants lisent les mots sans réaliser la liaison, soit ils produisent une pause, en respectant l'espace blanc entre Mot1 et Mot2, soit ils remplissent l'attaque vide au début du Mot2 avec une occlusive glottale. En produisant une occlusive glottale au début du Mot2, ils produisent une suite CV qui représente aussi la structure syllabique typique de la langue orale dans les cas de réalisation de la liaison. Le remplissage de l'attaque vide du Mot2 peut être expliqué par le fait que les enfants qui apprennent à lire s'appuient sur l'unité syllabique de la langue orale pour lire les mots, ce qui souligne l'utilisation de la phonologie dans les processus de *literacy*. Nous allons approfondir le rôle de la syllabe par des analyses phonétiques sur les autres contextes.

Pour l'enchaînement, d'un point de vue général, les enfants produisent au total plus de coups de glotte (52,94 %) et moins de pauses (47,06 %). Il faut dire que la différence de pourcentages entre les productions de coups de glotte et de pauses n'est pas assez significative. Nous pouvons conclure que, dans la production d'enchaînement, les enfants choisissent une fois sur deux la production d'une pause et le reste du temps d'une occlusive glottale au début du Mot2.

Puisque nous testons les enfants au début de l'acquisition de la lecture, l'insertion du coup de glotte entre Mot1 et Mot2 peut marquer une hésitation et avoir seulement une valeur phonologique. Contrairement à l'insertion de l'occlusive glottale au CP, les données ne permettent pas d'exclure le statut phonétique de cette consonne.

9.5.2.2.3.2 Les non-réalisations en présence de *h* aspiré et de *h* muet

Nous avons déjà dit qu'en présence de *h* aspiré, les phénomènes de resyllabation n'ont pas lieu (chapitre 3). Les productions du type NON-REAL.LECT^{CP2} représentent donc des lectures correctes. Pour bien comprendre le traitement de l'espace sonore entre Mot1 et Mot2, nous proposons donc une analyse phonétique des séquences déterminant + nom à *h* aspiré initial pour voir si dans les cas de non-réalisation de la CL, les enfants produisent vraiment une pause entre les mots. Nous écartons 7 NON-REAL.LECT^{CP2} (3,5 %) où l'enregistrement ne permet pas une analyse phonétique.

	% non-réalisations avec Mot2 à <i>h</i> aspiré initial
coups de glotte (ex. [ɛ̃.ʔi.bu])	42,87
pauses (ex. [ɛ̃/i.bu])	53,06
continuum (ex. [ɛ̃.i.bu])	4,07

Tableau 64. Analyse des non-réalisations pour le Mot2 à *h* aspiré initial en lecture dans LECT^{CP2}

Le tableau 64 montre qu'il n'y a pas de grande différence entre les productions de coups de glotte et de pauses (53,06 % vs 42,87 %). En revanche, le pourcentage de traitement des séquences comme un continuum est très bas (environ 4 %).

Nous avons mené la même analyse pour les NON-REAL.LECT^{CP2} dans les séquences déterminant + mot à *h* muet initial du type (*unØhippopotame* [ɛ̃Øi.po.po.tam]), où la non-réalisation du phénomène consonantique doit être considérée comme erronée (Tableau 65). Dans le calcul, nous avons écarté 4 enregistrements (6 %), car l'analyse phonétique n'était pas possible.

	% non-réalisations avec Mot2 à <i>h</i> muet initial
coups de glotte (ex. [ɛ̃.ʔi.po.po.tam])	43,24
pauses (ex. [ɛ̃/i.po.po.tam])	56,76

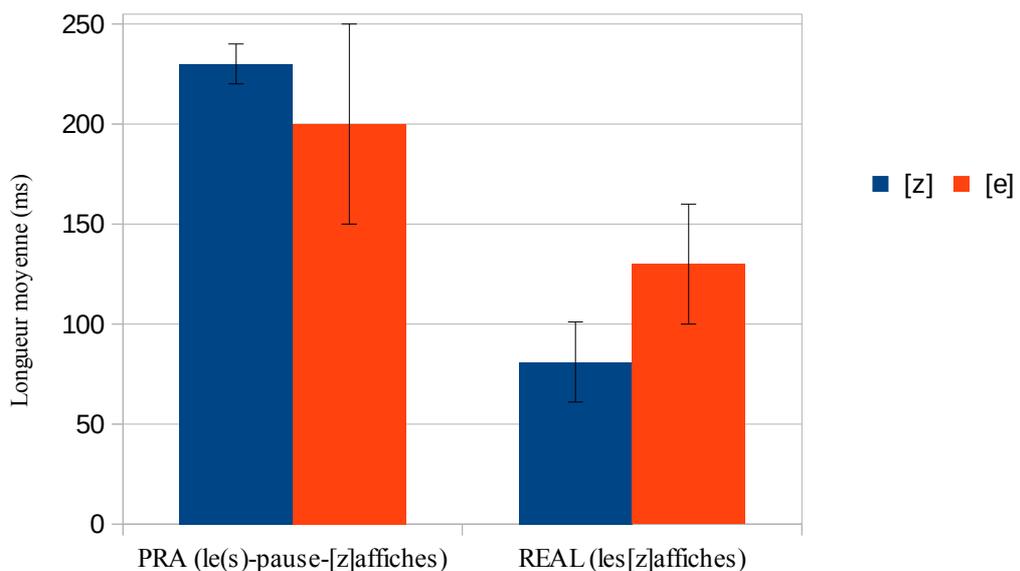
Tableau 65. Analyse de non-réalisations pour le Mot2 à *h* muet initial dans LECT^{CP2}

Comme pour le contexte à *h* aspiré initial, il n'y a pas de grande différence entre les pourcentages de productions de coups de glotte ou de pauses (43,24 % vs. 56,76 %). Mais cette présence d'un coup de glotte est-elle liée à la présence d'un *h* graphique ? Dans les paragraphes qui mettent en rapport écriture et lecture, à la fin de ce chapitre, nous tenterons de donner une explication à la présence d'un coup de glotte en lecture en rapport avec l'existence d'un *h* graphique à l'écrit.

9.5.2.2.3.3 Analyse phonétique de PRA.LECT^{CP2}

Comme nous l'avons montré dans la section 9.5.2.1, les enfants produisent en lecture des séquences comme *les-pause-[z]affiches* [le/za.fiʃ], où nous pouvons trouver une pause entre déterminant et nom et l'attaque du nom (Mot2) remplie par une consonne qui peut apparaître comme une reproduction de la CL. Nous avons appelé ce type de lectures PRA.LECT^{CP2}. Mais cette consonne produite à l'attaque du Mot2, après la pause, a-t-elle le même statut que la CL ?

Pour répondre à cette question, nous avons mis en œuvre des analyses phonétiques. Pour la liaison en [z], nous avons calculé la longueur de la consonne [z] et de la voyelle [e] qui la précède dans les deux productions PRA.LECT^{CP2} ([le/za.fiʃ]) et REAL.LECT^{CP2} ([le.za.fiʃ]) chez les mêmes enfants. Les résultats sont présentés dans le graphique 15.



Graphique 15. Longueur de [z] et [e] dans les deux productions dans LECT^{CP2}

Pour la voyelle [e], il n'y a pas de différences dans les deux types de productions en lecture. Cela implique que dans les deux cas la voyelle a le même statut qu'une voyelle en syllabe ouverte. En revanche, il y a une différence pour la consonne [z]. Dans le cas de la lecture [le/za.fiʃ], la consonne est en moyenne plus longue que dans le cas de la liaison [le.za.fiʃ]. Comme on l'a dit dans la section 4.3.2.1, les analyses phonétiques montrent que les CL sont plus courtes que les consonnes

lexicales (Spinelli *et al.*, 2003 ; Wauquier-Gravelines, 1996 ; Gaskell *et al.*, 2002). Nos données permettent de dire que la consonne [z] en lecture [le/za.fif] est plutôt traitée comme une consonne lexicale à l'attaque du Mot2.

Même si la consonne n'est pas présentée à l'écrit, les enfants la produisent et la traitent phonologiquement comme une consonne lexicale à l'attaque du Mot2. Pourquoi ce type de traitement ? Cela peut être expliqué par le fait que les enfants qui commencent à lire doivent s'appuyer, dans le processus, sur l'unité syllabique perçue à l'oral, et ils remplissent alors l'attaque du nom pour reproduire la suite non-marquée CV.

9.5.2.2.4 Le rôle de la fréquence et de la familiarité dans la tâche LECT^{CP2}

Familiarité : contexte connu vs contexte inconnu ?

Nous avons vu que pour la tâche d'écriture PROD.ECR^{CP2}, il y a une différence statistiquement significative entre les productions écrites des contextes connus et les productions écrites des contextes inconnus. En effet, quand nous demandons aux enfants d'écrire la séquence déterminant + nom dans un contexte déjà appris à l'école, ils l'écrivent mieux qu'une séquence qu'ils n'avaient jamais vue en classe³⁹. Demander aux enfants de lire des séquences déterminant + nom vues pendant l'année scolaire aide-t-il la lecture et la réalisation des phénomènes de resyllabation ? Les contextes connus sont-ils mieux lus et les mots mieux enchaînés que dans les contextes inconnus ?

39 Nous avons dit, que la tâche de lecture a lieu une semaine après la tâche d'écriture, c'est-à-dire quand les contextes que nous avons appelés « inconnus » étaient déjà connus par les enfants dans la tâche d'écriture PROD.ECR. Cependant, comme nous l'avons déjà expliqué dans la procédure, nous n'avons jamais fourni de *feed-back* aux enfants pour leur montrer comment bien orthographier les séquences dictées. Nous pouvons dire qu'ils n'ont jamais vu ces séquences à l'écrit et n'ont jamais eu l'occasion de les lire. Au moment de la tâche de lecture, les contextes sont inconnus car c'est la première fois que les enfants voient à l'écrit ces séquences correctement orthographiées et c'est aussi la première fois qu'il se retrouvent à les lire.

Le résultat du t-test montre que pour la tâche LECT^{CP2}, il n'y a pas de différence entre les ERR.LECT^{CP2} pour les contextes connus et pour les contextes inconnus ($t_1(36) = 0,90$, $p = .92$; $t_2(9) = 0,46$, $p = .66$). La familiarité à l'écrit avec un certain type de contexte n'intervient pas dans le processus de la lecture des séquences en liaison, en élision ou en enchaînement.

L'effet de la fréquence en lecture

Nous avons déjà remarqué que les mots de notre répertoire ont un nombre d'occurrences (fréquences absolues) différent dans notre base de données. De plus, nous avons remarqué (§ 9.5.1.2.7) que la fréquence du Mot2 n'influence pas l'identification des frontières des mots à l'écrit. Cependant, la fréquence des Mots2 influence-t-elle la lecture ?

Nous avons d'abord vérifié s'il existe ou non une corrélation entre les fréquences des Mots2 à lire et les lectures correctes, où les phénomènes de resyllabation sont bien réalisés et les séquences sont parfaitement lues. Pour la fréquence, nous avons considéré deux indices : la fréquence absolue dans notre base de données (le nombre d'occurrences) et la valeur SFI pour le CP dans MANULEX⁴⁰ qui donne des informations plus précises sur la fréquence des mots dans le lexique des enfants (§ 7.2.2).

Le calcul du r de Pearson permet d'établir une corrélation positive entre les scores de lectures correctes et les deux valeurs de la fréquence (lectures correctes - SFI : $r_1(19) = .60$, $p < .05$; lectures correctes - Fréq. : $r_2(19) = .45$, $p < .05$). Dans les deux cas, à une augmentation de la fréquence correspond une augmentation des scores de lectures correctes. Cela veut dire qu'un Mot2 plus fréquent est mieux lu qu'un mot moins fréquent et qu'il est aussi plus facile de produire les phénomènes de resyllabation quand ce mot est fréquent.

Pour mieux analyser la lecture des phénomènes de resyllabation en rapport avec la fréquence, nous avons également calculé la corrélation entre le nombre d'erreurs dans la lecture des phénomènes consonantiques ($ERR.LECT^{CP2} = PRA.LECT^{CP2} +$

40 Pour une description détaillée de la base de données MANULEX, nous invitons à la lecture du chapitre 7.

NON-REAL.LECT^{CP2}) et les deux indices de fréquence des Mots2, afin de vérifier si les Mots2 les plus fréquents sont mieux enchaînés aux Mots1 en présence de resyllabation. Les ERR.LECT^{CP2} ne corrèlent pas avec la fréquence des items dans notre base de données ($r(19) = -.37$, $p = .09$) mais corrèlent négativement avec l'indice SFI en MANULEX ($r(19) = -.53$, $p < .05$). Nous observons une diminution d'erreurs en rapport avec l'augmentation de la fréquence des items à lire.

Pour résumer nos analyses, nous pouvons dire que la fréquence de Mot2 est corrélée à la quantité de lectures correctes. Elle est également en rapport avec le traitement des frontières de mots en présence des phénomènes de resyllabation à lire : les Mots2 les plus fréquents sont plus facilement lus.

9.5.3 Le rapport entre les productions écrites et les productions en lecture à la fin du CP

L'écriture et la lecture sont deux activités très différentes qui impliquent diverses capacités cognitives et motrices (voir chapitre 5). Cependant, ces deux activités ont un point commun : la forme écrite. Par la forme écrite, lecture et écriture entrent en relation. Ce-ci nous suggère d'analyser le lien entre les lectures et les productions en écriture pour vérifier s'il existe une relation entre elles, liée au traitement des formes orthographiées.

- **Les non-réalisations en lecture (NON-REAL.LECT^{CP2} *le(s)[Ø]ours* [*leØuɔʁs*]) et les productions écrites correctes où les frontières des Mot1 et Mot2 sont bien identifiées (productions écrites à « frontières lexicales » du type <les ours>).**

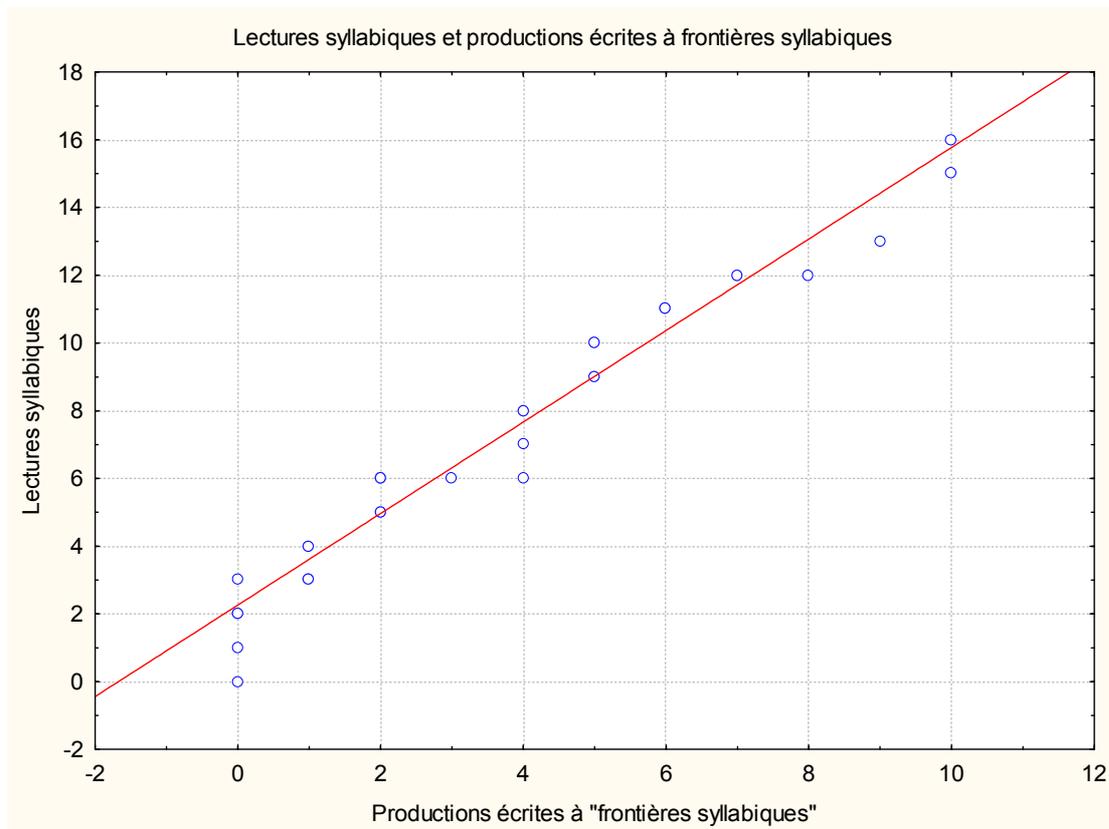
D'abord, nous allons vérifier s'il existe une corrélation entre les scores de non-réalisations en lecture (NON-REAL.LECT^{CP2} *le(s)[Ø]ours* [*leØuɔʁs*]) et les scores de productions écrites correctes (« frontières lexicales » <les ours>), où les enfants écrivent les mots en respectant leurs frontières lexicales. Nous cherchons à voir s'il y a un rapport entre le traitement à l'écrit de Mot1 et Mot2 comme deux formes visuellement séparées et la lecture de ces deux formes comme s'il n'y avait pas de phénomène de resyllabation entre eux. Le calcul du

r de Pearson montre que nous ne pouvons pas établir une corrélation statistiquement significative entre les productions écrites correctes (où les mots sont bien séparés par un espace blanc) et NON-REAL.LECT ($r(40) = -.21$, $p = .17$).

- **Les productions écrites à « frontières syllabiques » (<les zaffiches>) et les lectures syllabiques ([le/za/fiff]).**

Dans les lectures syllabiques, les enfants ne réussissent pas à lire la séquence de deux mots comme un continuum mais lisent les séquences syllabe par syllabe. Comme nous l'avons vu dans les premiers chapitres, les enfants français qui apprennent à lire et à écrire utilisent la syllabe comme unité fondamentale dans le processus d'apprentissage. C'est pourquoi les lectures syllabiques sont tout-à-fait attendues en tant qu'étapes intermédiaires dans l'acquisition de la *literacy*. Pour cette raison, nous pouvons aussi penser que les productions écrites à « frontières syllabiques » (ex. <les zaffiches> pour <les affiches>) soient l'effet d'une *literacy* encore en construction qui utilise la syllabe comme unité de base pour repérer les frontières lexicales masquées par la resyllabation à l'oral. Nous allons comparer ces productions écrites proposant une segmentation orthographique basée sur la syllabe (syllabe CV au début du Mot2, comme à l'oral) et les lectures syllabiques qui montrent l'usage de cette unité pour passer des formes écrites aux formes acoustiques. Si le résultat donne une corrélation statistiquement remarquable, nous pourrions dire que le remplissage de l'attaque en écriture (<les zaffiches>) est liée à des compétences de la *literacy* encore en développement et que ces compétences utilisent comme unité fondamentale la syllabe pour segmenter et représenter les mots à l'écrit.

Dans le calcul du r de Pearson, nous remarquons une corrélation positive entre les scores de lectures syllabiques et les scores de productions écrites à « frontières syllabiques » ($r(40) = .96$, $p < .05$). Les enfants qui utilisent le plus la syllabe pour lire produisent à l'écrit le plus de séquences à « frontières syllabiques » (Graphique 16). Pour cette raison, à ce stade de l'acquisition, il faut rendre compte de l'importance de la syllabe dans la construction du lexique orthographique.



Graphique 16. Relation entre les scores de productions écrites à "frontières syllabiques" et les scores de lectures syllabiques

- **Les productions écrites correctes où les frontières sont bien identifiées (« frontières lexicales » ex. <les ours>) et les lectures correctes où les phénomènes de resyllabation sont bien réalisés (REAL.LECT^{CP2} les[z]ours [le.zuʁs]).**

Existe-t-il à la fin du CP une relation entre le bon traitement des frontières de mots à l'écrit et la lecture correcte de séquences de mots avec la réalisation des phénomènes de resyllabation ? Nous proposons une comparaison entre les scores de productions écrites correctes à « frontières lexicales », où les deux mots sont bien séparés à l'écrit (<les ours>) et les scores de lectures où les phénomènes consonantiques sont bien réalisés REAL.LECT^{CP2} ([le.zuʁs]). Le résultat du calcul de r de Pearson montre qu'il y a une corrélation positive entre ces deux variables ($r(40) = .39, p < .05$) : à l'augmentation des scores de productions écrites correctes correspond une augmentation des lectures

correctes REAL.LECT^{CP2} et vice-versa. Ce qui signifie que la maîtrise écrite des séquences consonantiques resyllabifiées implique une meilleure capacité à les lire correctement et vice-versa.

Cela peut vouloir dire que les compétences de base dans les deux activités font référence aux mêmes composantes linguistiques. Cependant, le rapport entre, d'une part, les productions écrites complètement correctes (sans fautes d'orthographe à l'intérieur des mots) et, d'autre part, les lectures correctes (où Mot1 et Mot2 sont bien lus et les phénomènes de resyllabation bien réalisés) ne peut pas s'établir car, à la fin du CP, les mots écrits sont encore caractérisés par de nombreuses fautes d'orthographe même s'ils sont correctement lus. Apprendre à écrire en français est, en effet, plus difficile qu'apprendre à lire et le travail cognitif et linguistique de ces deux activités n'a pas un développement complètement symétrique (§ 5.6).

9.5.3.1 Le rapport entre productions écrites et lecture pour *h* muet et *h* aspiré

Nous avons montré que, dans la lecture de séquences à *h* aspiré ou à *h* muet initiaux, la plupart des enfants ne réalisent pas les phénomènes de resyllabation (§ 9.5.2.2.2). De plus, les analyses phonétiques montrent que la moitié des enfants qui produisent des non-réalisations, en lecture (NON-REAL.LECT^{CP2}) ajoutent un coup de glotte entre Mot1 et Mot2 (§ 9.5.2.2.3.1). Nous allons mieux analyser les rapports entre l'insertion de ces coups de glotte en lecture et la production écrite de la lettre *h* au début du nom : y a-t-il une corrélation entre les productions de coups de glotte entre Mot1 et Mot2 et les représentations écrites de ces noms ? Plus précisément, le coup de glotte en lecture correspond-il à la réalisation d'une position orthographiée au moyen du graphème *h* ?

Nous allons vérifier si les enfants qui produisent en lecture un coup de glotte entre déterminant et nom avec *h* aspiré ou *h* muet initiaux sont aussi les enfants qui écrivent les mêmes mots avec un *h* graphique à l'écrit. Au moyen du calcul du coefficient de corrélation *r* de Pearson, nous voyons qu'il n'y a pas de corrélation entre les insertions de coups de glotte dans la lecture et les productions écrites des mots avec

un *h* graphique initial ($r(26) = .11$, $p = .61$). À partir de nos données, nous ne pouvons pas dire que la connaissance d'une position graphique au début du nom empêchant la liaison est liée à la production phonétique d'un coup de glotte en lecture.

9.6 Le rapport entre la *literacy* et les productions orales à la fin du CP

Comme nous l'avons dit dans les premiers chapitres, la phonologie et l'écriture ont un lien très étroit, surtout au premier stade de l'apprentissage, lorsque la *literacy* fait appel aux compétences phonologiques uniquement. De plus, comme nous l'avons déjà remarqué dans le chapitre 2, il y a toujours une influence réciproque entre les compétences phonologiques et les compétences d'écriture et de lecture. Les problèmes de segmentation à l'écrit pourraient-ils donc être liés aux problèmes que les enfants montrent à l'oral ? Nous avons essayé d'établir différentes corrélations pour réussir à comprendre le plus précisément possible les rapports entre écrit et oral dans les questions de segmentation. Pour commencer, nous cherchons à établir des corrélations entre les tâches liées à la langue orale (DENa^{CP2}, DENb^{CP2} et PERC^{CP2}) et les tâches liées à la langue écrite (PROD.ECR^{CP2} et LECT^{CP2}).

- **Les productions écrites correctes (« frontières lexicales » ex. <les ours>) dans PROD.ECR^{CP2} et les non-réalisations à l'oral (*le(s)[Ø]ours* [leØuʁs]) dans DENa^{CP2} et DENb^{CP2}**

Quand les enfants produisent des mots à l'oral avec des non-réalisations, ils produisent les deux mots séparés, sans ajouter de consonne à l'attaque du Mot2 (*le(s)[Ø]ours* [leØuʁs]). Ce type de productions peut être la conséquence de l'acquisition de l'orthographe. Quand les enfants apprennent à écrire, des informations graphiques (comme les espaces blancs, les tirets ou l'apostrophe) peuvent les aider à isoler les frontières des mots. Nous pouvons donc penser qu'il y a une corrélation entre les productions écrites correctes des mots (à « frontières lexicales » <les ours>) et la non-réalisation de la resyllabation à l'oral (*le(s)[Ø]ours* [leØuʁs]). À l'espace blanc de l'écrit vont correspondre la non-réalisation du phénomène consonantique à l'oral et un traitement des Mot1

et Mot2 comme une séquence composée de deux formes séparées et isolées. Au moyen d'une analyse statistique de corrélation (r de Pearson), nous comprenons qu'il n'y a pas de corrélation entre les scores de productions écrites correctes (« frontières lexicales ») et les scores de non-réalisations dans les tâches de dénomination guidée DENa^{CP2} ($r_1(41) = .16$, $p = .50$) et DENb^{CP2} ($r_2(41) = .17$, $p = .28$).

- **Les productions écrites correctes (« frontières lexicales » *les ours*) dans PROD.ECR^{CP2} et les non-identifications de non-réalisations à l'oral (*le(s) [Ø]ours [leØuʁs]*) dans PERC^{CP2}**

Dans la tâche de perception d'erreurs PERC^{CP2}, nous pouvons remarquer une persistance des non-identifications des non-réalisations des phénomènes de resyllabation à la fin du CP. Cette persistance peut-elle être corrélée à l'apprentissage de l'écriture et au fait que les enfants ont appris à séparer Mot1 et Mot2 par l'orthographe ?

En comparant les productions écrites des 12 enfants échantillonnés et les résultats de perception d'erreurs, nous pouvons voir qu'il n'y a pas de corrélation entre les scores de non-identifications de non-réalisations et les scores de productions écrites correctes à frontières lexicales ($r(11) = .15$, $p = .33$) : un score plus élevé de non-identifications de non-réalisations ne peut pas être associé à une meilleure segmentation à l'écrit.

- **Les productions écrites correctes (« frontières lexicales », ex. <les ours>) dans PROD.ECR^{CP2} et les productions orales correctes où les phénomènes de resyllabation sont réalisés (*le(s)[z]ours [le.zuʁs]*) dans DENa^{CP2}**

Aux premiers stades de l'acquisition de la lecture et de l'écriture, les enfants utilisent leurs compétences phonologiques pour construire leur lexique orthographique et segmenter en mots la chaîne orale. Nous pouvons donc essayer d'établir un rapport entre les productions correctes à l'oral dans DENa^{CP2}, qui impliquent une connaissance de la segmentation en mots à l'oral, et les productions correctes à l'écrit dans PROD.ECR^{CP2}, qui impliquent une

connaissance de la segmentation à l'écrit. Le résultat du calcul statistique montre qu'il y a une corrélation positive entre les scores de productions écrites correctes et de productions orales correctes ($r(41) = .440$, $p < .001$)⁴¹. Une meilleure production à l'écrit dans PROD.ECR^{CP2} est associée à une meilleure production à oral en DENa^{CP2}.

- **Liaison en [z] : corrélation entre marquage du pluriel <s> en PROD.ECR^{CP2} et réalisation correcte à l'oral dans DENa^{CP2}**

Après avoir établi le lien entre oral et écrit, nous cherchons à analyser le phénomène de liaison pour comprendre si le marquage du pluriel à l'écrit influence les réalisations de liaisons en [z] à l'oral. Nous proposons une comparaison des résultats de la tâche de DENa^{CP2} et de la tâche écrite PROD.ECR^{CP2} pour les séquences tests « liaison en [z] ». Nous calculons la valeur du r de Pearson entre les scores de productions correctes à l'oral (où la liaison en [z] est bien réalisée) et les scores du marquage à l'écrit (les enfants ajoutent des <s> graphiques à la fin de Mot1 et de Mot2). Le résultat montre qu'il n'y a pas de corrélation entre ces deux variables ($r(41) = .11$, $p = .47$).

- **Les non-réalisations à l'oral dans DENa^{CP2} (*le(s)/Ø*jours [*leØuɔ̃s*]) et les non-réalisations en LECT^{CP2} (NON-REAL.LECT *le(s)/Ø*jours [*leØuɔ̃s*])**

Nous avons décidé de comparer les non-réalisations dans LECT^{CP2} et dans DENa^{CP2} pour vérifier s'il existe un rapport entre les deux. Même si les résultats de la lecture et de la dénomination guidée sont en apparence les mêmes (des productions acoustiques), ces deux activités sont très différentes. L'activité de la lecture est liée à une oralisation de la forme écrite. Surtout au début de l'apprentissage, lire revient à associer un graphème à un phonème. En revanche, l'activité de dénomination implique la réalisation d'une représentation mentale éventuelle qui ne fait pas référence à une séquence de lettres. Cependant, nous avons décidé de vérifier s'il était possible d'établir une relation pour les non-

41 Pour calculer cette corrélation, nous avons réduit les contextes de PERC aux contextes : liaison (en [z]), élision et enchaînement qui sont les mêmes contextes que nous avons dans DENa. Nous écartons le contexte liaison en [n] et à *h* initial. Nous avons utilisé les mêmes contextes pour le calcul de la corrélation qui met en rapport les productions orales et la lecture LECT.

réalisations entre l'oralisation d'une représentation écrite et la production orale spontanée. Nous comparons les scores de non-réalisations dans DENa^{CP2} et LECT^{CP2}. La valeur du r de Pearson suggère qu'il n'y a pas de relation entre ces deux types de non-réalisations ($r(40) = .20, p = .20$).

- **Les réalisations des phénomènes de resyllabation en LECT^{CP2} (REAL.LECT^{CP2} *le(s)/z]ours* [le.zuʁs]) et les productions orales correctes dans DENa^{CP2} (*le(s)/z]ours* [le.zuʁs])**

Pour mieux établir le rapport entre les performances à l'écrit et à l'oral, nous calculons le rapport de corrélation entre deux variables : les réalisations des phénomènes de resyllabation dans LECT^{CP2} (REAL.LECT^{CP2} *le(s)/z]ours* [le.zuʁs]) et lors de DENa^{CP2} (*le(s)/z]ours* [le.zuʁs]). Comparer ces deux variables permet de comprendre le rapport entre compétences écrites et compétences orales en ce qui concerne le traitement des frontières de mot et de la réalisation des phénomènes de resyllabation.

Le calcul de l'indice de corrélation r de Pearson montre qu'il y a une corrélation positive entre les scores de réalisations des phénomènes consonantiques dans LECT^{CP2} (REAL.LECT^{CP2} *le(s)/z]ours* [le.zuʁs]) et les scores de productions orales correctes dans DENa^{CP2} (*le(s)/z]ours* [le.zuʁs]) ($r(40) = .47, p < .05$) : à l'augmentation de la moyenne de réalisations du phénomène consonantique en LECT^{CP2} (REAL.LECT^{CP2}) correspond une augmentation des scores de productions correctes en dénomination et vice-versa.

- **Les insertions d'une consonne erronée dans DENa^{CP2} (*le(s)/n]ours* [le.nuʁs]) et les productions écrites correctes (« frontière lexicale » <les ours>) dans PROD.ECR^{CP2}**

Nous cherchons à établir le même rapport pour les insertions des consonnes erronées. Le résultat du calcul de r de Pearson montre qu'il n'y a pas de corrélation entre les insertions d'une consonne erronée et les scores des productions écrites correctes ($r(41) = 0,40, p = 0,57$).

- **Les insertions d'une consonne erronée dans DENa^{CP2} (*le(s)/n/jours* [le.nuɪs]) et les lectures correctes dans LECT^{CP2} (REAL.LECT *les/z/jours* [le.zuɪs])**

Comme nous l'avons déjà fait pour l'écriture, nous cherchons à comprendre si les erreurs d'insertion de consonne erronée sont en rapport avec les lectures correctes (REAL.LECT). Au moyen du calcul du r de Pearson, nous pouvons observer qu'il n'existe pas de relation entre ces deux variables ($r(40) = 0,37$, $p = 0,64$).

Nous pouvons résumer ainsi les résultats obtenus :

- Il y a un rapport entre les productions écrites et les lectures d'une part et les productions orales d'autre part, dans le cas des séquences caractérisées par des phénomènes de resyllabation.
La phonologie et l'orthographe ont des interactions et des influences réciproques. Une relation existe entre « bien produire » à l'oral et « bien écrire-lire ». La *literacy* influence la production orale et permet de mieux comprendre les frontières des mots : bien identifier les frontières des mots à l'écrit et bien les enchaîner dans la lecture signifie une meilleure production orale. De même, les enfants qui montrent de meilleures productions orales sont aussi capables de mieux écrire et de mieux lire les séquences Mot1 + Mot2 en présence des phénomènes de resyllabation.
- Il n'y a pas de relation entre l'identification des frontières de mots à l'écrit et les non-réalisations à l'oral. La *literacy*, à ce stade, n'influence pas négativement le traitement des phénomènes de resyllabation en production.
- Nous n'observons pas de corrélations entre les erreurs d'insertion d'une consonne erronée en dénomination et les compétences d'écriture et de lecture. Les meilleures performances en lecture et en écriture n'impliquent pas une diminution de ces erreurs à l'oral.
- De plus, connaître la morphologie du pluriel à l'écrit n'aide pas à réaliser correctement les séquences en liaison en [z].

9.7 Le rapport entre la fréquence des activités de lecture partagée et les résultats des tâches à la fin du CP

Nous avons déjà largement expliqué dans le chapitre 5 que plus les parents partagent de moments de lectures avec leurs enfants et plus la *literacy* est influencée. Après avoir analysé toutes les données sur la lecture et sur l'écriture nous cherchons à établir une corrélation entre la fréquence de lecture partagée et les résultats des tâches DENa^{CP2}, LECT^{CP2} et PROD.ECR^{CP2}. Nous nous attendons à ce que plus les lectures et les dénominations sont correctes, plus la fréquence de lecture partagée est haute. De plus, plus les enfants réussissent à bien identifier les frontières des mots dans PROD.ECR^{CP2} plus ils ont sans doute eu des moments de lecture partagée. Nous considérons la classification de la fréquence présentée au chapitre 8 pour les 32 enfants⁴².

Le résultat montre qu'il n'y a pas de corrélation entre les scores de REAL.LECT^{CP2} et la fréquence de lecture partagée ($r(31) = .021$, $p = .083$). Par conséquent, nous ne pouvons pas établir un rapport entre la fréquence de lecture partagée et la bonne réalisation des phénomènes de resyllabation en lecture.

Nous avons calculé la corrélation aussi pour l'écriture, en regardant la fréquence des lectures partagées et les scores de productions écrites à « frontières lexicales » (<les ours>) dans PROD.ECR^{CP2}. Le calcul du coefficient r de Pearson ne montre aucune corrélation entre la fréquence de lecture partagée et les segmentations correctes à l'écrit ($r(31) = -.08$ $p = .87$).

Nous cherchons à voir si cette fréquence peut être mise en rapport avec les résultats de la tâche DENa^{CP2}, la tâche de dénomination guidée. Le calcul montre qu'il n'existe pas de rapport entre les productions orales correctes (où les phénomènes de resyllabation sont bien réalisés) et la fréquence des lectures partagées ($r(31) = .10$, $p = .65$).

42 Seulement 32 parents ont complété le questionnaire.

La fréquence des lectures partagées ne semble influencer ni la dénomination guidée, ni la lecture et l'écriture au CP en ce qui concerne les phénomènes de resyllabation. L'apprentissage explicite de la lecture et de l'écriture conditionne les capacités des enfants. Une simple exposition passive à de l'information visuelle n'est pas suffisante pour les aider à segmenter correctement.

9.8 Discussions des hypothèses à la fin du CP

9.8.1 Les productions orales des phénomènes de resyllabation à la fin du CP

D'après la littérature existante sur la liaison, les enfants français n'ont plus de difficultés en production après 6 ans (Basset, 2000 ; Dugua & Baclesse, 2014). Pourtant, à la fin comme au début du CP, nous avons constaté encore des difficultés à détecter les non-réalisations dans la tâche de perception d'erreurs et reformulation (PERC). En production, nous enregistrons également une stabilité des non-réalisations des phénomènes de resyllabation. Cette stabilité confirme les données récoltées par Chevrot *et al.* (2013) et Nardy et Dugua (2011).

Au moyen d'un deuxième type de tâche de dénomination guidée (DENb), nous avons vérifié les données en production pour évaluer la présence d'un effet d'amorçage lors de DENa. La comparaison entre les résultats de DENa et de DENb nous suggère qu'il y a un effet de la tâche pour la production d'élision et d'enchaînement. Cela veut dire que dans DENa, l'input en liaison donné par l'expérimentateur influence la production d'élision et d'enchaînement de l'enfant. Dans le chapitre 8, nous avons déjà donné une interprétation des non-réalisations d'élision comme la conséquence d'un problème de segmentation des séquences en liaison (§ 8.4). Cette explication nous a permis de démontrer, au moyen des analyses phonétiques, l'existence d'une position sous-jacente associée à la CL dans sa représentation. L'analyse des non-réalisations de l'enchaînement nous amène-t-elle à la même conclusion ? Considérons à nouveau l'analyse acoustique des non-réalisations d'enchaînement dans DENa^{CP1} (§ 8.3.3.1). Pour induire à produire *la petite orange* [la.pə.ti.tɔ.rɑ̃ʒ], l'expérimentateur donne un input en liaison, du type *les petites[z]oranges* [le.pə.tit.zɔ.rɑ̃ʒ]. Les productions des enfants sont caractérisées par

des non-réalisations de l'enchaînement du type *la petite-orange* [la.pə.tit.ɔ.rãʒ], où le [t] de *petit* n'est pas resyllabifié sur Mot2. Nous avons analysé l'espace sonore entre adjectif et nom. Les résultats montrent que la plupart des enfants insèrent une occlusive glottale entre Mot1 et Mot2 en produisant *la petite[ʔ]orange* [la.pə.tit.ʔɔ.rãʒ]. Ces productions sont aussi le reflet de l'input en liaison donné par l'expérimentateur. Les enfants qui entendent la liaison [le.pə.tit.zɔ.rãʒ] ne réalisent pas la CL [z] mais ils produisent à sa place un coup de glotte [ʔɔ.rãʒ]. La réalisation de cette consonne à l'attaque du Mot2 confirme l'existence d'une position prosodique associée à la CL qui est toujours présente dans la non-réalisation d'enchaînement. Cette position prosodique entre Mot1 et Mot2 est une sorte de position « fantôme » et se trouve remplie par une occlusive glottale qui, en occupant l'attaque du Mot2, empêche la resyllabation de la consonne d'enchaînement [t] sur le mot suivant (Figure 37).

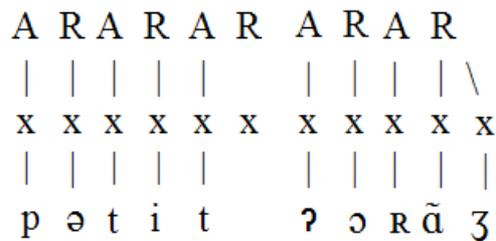


Figure 37. Représentation des résultats pour la production d'enchaînement

Contrairement à l'élision où nous disposons de plusieurs indices phonétiques, pour l'enchaînement nous pouvons analyser seulement l'espace sonore entre Mot1 et Mot2. De plus, dans la dénomination en enchaînement nous proposons une séquence adjectif + nom qui est syntaxiquement différente des autres séquences tests déterminant + nom. Pourtant ces raisons, nous ne pouvons pas exclure une explication pragmatique à l'apparition d'un coup de glotte dans la séquence : l'occlusive glottale peut être insérée entre Mot1 et Mot2 pour marquer une séparation entre l'adjectif et le nom.

La comparaison entre les deux tâches de dénomination guidée a fait émerger également que les productions de liaison dans DENa^{CP2} ne sont pas influencées par la procédure. L'input en élision donné par l'expérimentateur n'influence pas la production de la liaison chez l'enfant. Pour ce phénomène, nous remarquons de nombreuses non-

réalisations (§ 9.3.2). Nous avons proposé une analyse acoustique de ces non-réalisations qui montre que, dans la plupart des cas, les enfants insèrent un coup de glotte entre Mot1 et Mot2. Dans les séquences en liaison, si la CL n'est pas réalisée, sa position prosodique est occupée par une occlusive glottale (Figure 38). Ces données confirment l'existence de cette position prosodique et la représentation phonologique de Mot1 + Mot2 pour les séquences en liaison.

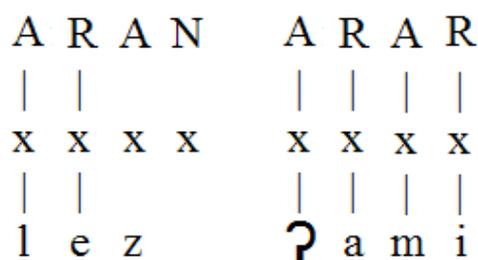


Figure 38. Représentation de l'insertion d'un coup de glotte en liaison

Les analyses phonétiques de non-réalisations en élision et en enchaînement dans DENa^{CP1} et en liaison dans DENa^{CP2} permettent de vérifier l'existence d'une position vide sous-jacente entre Mot1 et Mot2 dans les séquences en liaison et donc de considérer comme meilleure explication possible l'explication donnée par le modèle morpho-phonologique (Wauquier, 2009). Ces données sont confirmées par l'analyse de la fréquence de mots qui à la fin du CP ne joue pas de rôle dans les productions en contexte de liaison des enfants⁴³, confirmant les données de Dugua *et al.* (2009) qui montrent qu'après 6 ans la fréquence n'influence pas les productions des séquences des Mot1 + Mot2.

La comparaison entre les deux tâches de dénomination fait aussi émerger un problème morphologique en production. Dans la tâche DENa^{CP2}, les enfants entendent un input au pluriel et doivent produire une séquence au singulier et vice-versa. Dans le deuxième type de tâche, DENb^{CP2}, cette transformation n'est pas demandée. Nous remarquons des erreurs de mauvais article (*le petit orange* pour *la petite orange*, *le noix* pour *la noix*) seulement dans le premier type de tâche. Les données montrent des difficultés morphologiques à repérer le genre à partir du nom, car dans la très grande majorité des cas en français c'est l'article qui porte la marque de genre, et à accorder le genre entre constituants (accorder le genre de l'adjectif au nom). Ils ont également des

43 Pour une analyse détaillée de la question de la fréquence nous invitons à la lecture du § 9.3.4.

difficultés à réaliser l'accord entre le nom et l'adjectif. Parisse et Maillart (2004 : 28) nous expliquent que ces erreurs de genre, « qui sont en général produites chez l'enfant à développement normal à partir de l'âge de 4 ans, reflètent un mécanisme d'appropriation des règles grammaticales de la langue maternelle ».

9.8.2 Les productions écrites des phénomènes de resyllabation à la fin du CP

Les résultats de recherches précédentes remarquent que les enfants dans les premières années de l'école primaire ont des difficultés à orthographier des mots en séquence de liaison et d'enchaînement (§ 5.6.2.2). Les études de Soum Favaro (1997), de Sabio (2000) et de Soum-Favaro *et al.* (2014) montrent que la plupart des enfants écrivent la consonne qui est prononcée resyllabifiée sur Mot2, comme si elle était l'initiale de ce mot (ex. <un nange> pour *un ange*). À l'inverse, notre tâche d'écriture PROD.ECR^{CP2} montre que déjà au CP les enfants traitent plus ou moins correctement à l'écrit les frontières de mots perturbées par les phénomènes de resyllabation à l'oral. Nous enregistrons davantage de productions écrites correctes qui respectent les frontières lexicales (ex. <les affiches>, <l'éléphant>, <une ambulance>) que de productions écrites incorrectes qui respectent les frontières syllabiques (ex. <les saffiches>, <le lélephant>, <une nambulance>).

L'analyse des productions écrites des enfants montre une utilisation encore importante de la stratégie phonologique dans les activités d'écriture : à la fin du CP les enfants utilisent encore les correspondances graphème-phonème pour écrire. Pour cette raison, les digrammes et les mots plus opaques sont orthographiés plus difficilement (§ 5.6.2). De plus, les graphèmes qui ne sont pas associés à des phonèmes sont rarement orthographiés, comme par exemple le <h> graphique au début du nom à *h* muet qui n'a pas de véritable réalisation à l'oral (§ 9.5.1.2.6). En revanche, les graphèmes qui ont des réalisations dans la langue orale sont fréquemment écrits, comme par exemple la marque du pluriel du mot *ours* (§ 9.5.1.2.4). Nous observons aussi que les items les plus fréquents ne sont pas mieux orthographiés que les items les moins fréquents (§ 9.5.1.2.7). La fréquence des items à orthographier n'influence pas l'exactitude des productions écrites. Nos résultats sont en accord avec les données de Sprenger-Charolles & Serniclaes (2003) qui montrent que les enfants au début de l'apprentissage de l'écrit s'appuient sur la langue orale pour lire et écrire. De plus, ils confirment le

problème de la consistance pour la langue française dans les processus de *literacy*, évoqué par Ziegler et Goswami (2005).

Chez les enfants que nous avons testés, la stratégie phonologique reste d'importance primaire dans l'écriture mais également dans la lecture. Les données sur la lecture nous montrent un taux très haut de non-réalisations des phénomènes de resyllabation (ex. *le(s)[Ø]affiches* [le.Øa.fiʃ] pour *les[z]affiches* [le.za.fiʃ]). Nous avons analysé phonétiquement ces non-réalisations et nous avons remarqué que les enfants insèrent un coup de glotte au début du Mot2 ([le.ʔa.fiʃ]). Cette insertion nous suggère que les enfants cherchent à reproduire la structure des Mots2 qu'ils entendent à l'oral, où un segment est resyllabifié à son début. Ils ne réalisent pas la CL mais ils ajoutent une occlusive glottale à sa place, qui permet de conserver la structure CV au début du nom de l'oral. Le coup de glotte sert aussi à marquer la frontière, donc à bien séparer les mots. La phonologie et les fortes tendances syllabiques du français les aident à bien séparer les mots : ils les prononcent bien séparément comme le leur indique le blanc entre Mot1 et Mot2 mais en respectant les frontières syllabiques

Dans la tâche de lecture, nous avons enregistré des productions du type *les-pause-[z]affiches* [le/za.fiʃ]. Dans ces productions, Mot1 est bien lu et séparé par une pause de Mot2 où l'attaque syllabique est remplie. Au moyen des analyses phonétiques, nous avons vu que le [z] produit au début du Mot2 est une CL. De la même façon que les non-réalisations, les enfants ajoutent un coup de glotte au début du Mot2, ils ajoutent une CL dans ce type de lectures. Pour les mêmes raisons que nous avons données plus haut, les enfants cherchent à reproduire en lecture la structure syllabique du Mot2 qui caractérise la langue orale (CV). Simultanément, la contrainte orthographique de l'espace blanc suggère de produire une pause qu'ils ajoutent entre [le] et [za.fiʃ].

Ces données montrent l'utilisation des unités phonologiques de la langue orale dans le processus de lecture : les enfants qui ne réussissent pas à enchaîner correctement les mots proposent pour Mot2 la structure phonologique CV caractéristique des productions orales (ex. [za.fiʃ] ou [ʔa.fiʃ]). Les frontières syllabiques de la langue orale sont respectées en lecture. De plus, au moment de nos analyses, nous enregistrons des lectures du type syllabique (ex. [le/za/fiʃ]). Plus le Mot2 est long, plus les enfants utilisent l'unité syllabique en l'épelant (§ 9.5.2.2). L'ensemble des résultats en lecture montre que, à la fin du CP, l'unité syllabique a

encore une importance centrale car elle est la première unité utilisée pour accéder à l'écrit, comme souligné dans la littérature sur l'acquisition de la *literacy* en français (Ziegler & Goswami, 2005 ; Goswami, 1993, 1999 ; Goswami & Bryant, 1990 ; Goswami & East, 2000). De plus, ces nouvelles données confirment l'importance de la syllabe en lecture de mots isolés bien décrite par la littérature (Colé *et al.*, 1999 ; Doignon & Zagar, 2006 ; Chetail & Mathey, 2012). Pour mieux comprendre l'importance de la structure syllabique à l'oral et ses relations avec l'écrit, nous proposerons dans le chapitre 10, deux tâches d'amorçage.

Dans les processus d'écriture et de lecture à la fin du CP, la stratégie phonologique est utilisée en parallèle avec une stratégie lexicale comme le soulignent Sprenger-Charolles *et al.*, (1997). Nous remarquons une stratégie lexicale pour les productions écrites de PROD.ECR^{CP2}. Nous comptons moins de productions écrites où la CL [z] est insérée au début du Mot2 (ex. <les jours>) que des productions écrites où la CL [n] apparaît au début du Mot2 (ex. <un norage>). Nous pouvons penser que ces résultats sont influencés par la fréquence de Mots2 dans le lexique écrit des enfants. Dans notre base de données et dans MANULEX, il y a plus de Mots2 qui commencent par le son [n] que des Mot2 commençant par le son [z]. Puisque le son [n] est plus probable au début d'un mot, dans les cas de liaison, les enfants orthographient plus fréquemment le Mot2 à voyelle initiale comme un mot à <n> initial. Ces résultats peuvent confirmer également une sensibilité aux régularités orthographiques de la langue, comme montré par Pacton *et al.* (1996 ; 1998). Les enfants comprennent qu'en français il y a moins de mots qui commencent avec le son [z] orthographié par <z> et ils produisent moins d'erreurs du type <les jours> pour les séquences en liaison.

Nous avons aussi remarqué que la fréquence des Mots2 influence le taux de lectures correctes des séquences resyllabifiées en LECT^{CP2}. Ces résultats sont en accord avec les résultats de Sprenger-Charolles *et al.* (1997) confirmant, à nouveau, la mise en place d'une procédure lexicale. Cependant, dans la tâche d'écriture PROD.ECR^{CP2}, nous ne trouvons pas les mêmes résultats. La fréquence des Mots2 à écrire n'est pas en rapport avec le taux de productions écrites correctes. Cette différence entre la lecture et l'écriture n'est pas la seule. Dans la tâche d'écriture PROD.ECR^{CP2}, les contextes les mieux écrits sont les contextes connus, où la frontière vocalique gauche du Mot2 est identifiée plus facilement. Dans la tâche de lecture LECT^{CP2}, nous n'enregistrons pas de différences dans le traitement de ces deux contextes. Ces résultats confirment l'idée que

lecture et écriture sont deux activités différentes qui font référence à des stratégies différentes (Cossu & Marshall, 1985). Par conséquent, l'apprentissage doit être structuré selon l'activité que l'enseignant veut renforcer. Pour bien apprendre à orthographier les frontières gauches des Mots2 en présence des phénomènes de resyllabation, il est préférable d'utiliser une méthode d'apprentissage qui propose la visualisation et la production écrite des suites Mot1 + Mot2 plutôt qu'une liste de mots en isolation. Pour bien lire les phénomènes de resyllabation, il faut plutôt construire du matériel didactique qui alterne la présence des Mots2 isolés aux suites Mot1 + Mot2.

Dans notre tâche d'écriture nous remarquons également que les enfants n'arrivent pas encore à ajouter le <s> à la fin du Mot2 pour marquer le pluriel. Hormis le cas spécifique de *ours*, que nous avons déjà présenté dans les paragraphes précédents, nous enregistrons moins de 25 % de <s> orthographiés à la fin des Mots2. Ces résultats ne sont pas en accord avec les données de Keeney & Xolf (1972) et Nicolaci-da Costa & Harris (1983) qui disent que la morphologie du nombre est une des premières contraintes orthographiques à être maîtrisée à l'écrit par l'enfant en français. Nos données ne permettent pas non plus de confirmer les données de Thévenin *et al.* (1999) car les enfants au CP que nous avons testés n'ont pas encore appris à l'écrit le <s> final nominal.

Même si certains résultats montrent que la lecture et l'écriture doivent être traités comme deux activités différentes, nous pouvons remarquer certains points en commun. Quand les enfants identifient correctement les frontières des mots à l'écrit, ils réussissent également à bien produire les phénomènes de resyllabation en lecture. Une meilleure connaissance de l'orthographe semble induire des lectures plus justes. De plus, les enfants qui réalisent des lectures syllabiques ([le/za/fiʃ]) sont aussi les enfants qui écrivent les Mots2 en remplissant l'attaque vide avec la consonne qui resyllabifie sur ce mot (<les saffiches>). Les données montrent que les enfants qui utilisent la syllabe pour lire, l'utilisent aussi pour segmenter les mots à l'oral et pour orthographier les mots. Ces derniers résultats portent à nouveau l'attention sur le constituant syllabique.

9.8.3 Les relations entre productions écrites et productions orales

Dans le chapitre 6, nous avons fait l'hypothèse (H2) que les non-réalisations restent stables à cause de l'apprentissage des formes orthographiques des mots. Selon notre hypothèse, les enfants produisent, à l'oral, une pause entre Mot1 et Mot2 pour reproduire les deux formes orthographiques séparées dans la langue écrite. Nous avons vu qu'il n'y a aucun rapport entre les non-réalisations à l'oral et l'identification correcte des frontières des mots à l'écrit. En revanche, identifier correctement les frontières des mots à l'écrit implique des productions orales correctes (où les phénomènes de resyllabation sont bien réalisés). Pour la lecture, nous remarquons la même tendance. Nous ne pouvons pas établir un rapport entre les non-réalisations en lecture et en production orale mais les enfants qui lisent correctement les séquences Mot1 + Mot2 en lecture, sont aussi les enfants qui réussissent à les produire correctement en dénomination guidée.

Ces résultats montrent des interactions entre les compétences orales et les compétences de la langue écrite. Nous ne pouvons pas vraiment dire si ce sont les compétences orales qui influencent les compétences écrites ou vice-versa. Apprendre la langue écrite aide à bien identifier les frontières des mots qui sont séparés par des marques graphiques et à comprendre également, qu'en présence des phénomènes de resyllabation, un Mot1 qui termine par une certaine consonne à l'écrit (ex. le <s> dans *les*) s'enchaîne à un Mot2 à voyelle initiale. De plus, grâce à l'enseignement explicite, les enfants fixent la distinction entre les lettres qui représentent des consonnes et les lettres qui représentent des voyelles. Cette distinction permet de comprendre les régularités contextuelles des phénomènes de resyllabation à l'oral : quand le Mot2, souvent un nom, commence par une voyelle et le Mot1 (souvent un article) se termine par une certaine consonne, il faut produire une resyllabation. L'activité de lecture aide à son tour à améliorer la production orale car elle permet de s'entraîner dans le processus de resyllabation et d'enchaîner deux formes écrites (Mot1 à consonne finale à l'écrit et Mot2 à voyelle initiale) séparés sur papier par une marque graphique. Pour toutes ces raisons, nous pouvons penser que la connaissance de la *literacy* influence positivement les compétences et les productions orales, confirmant les idées de Ehri et Wilce (1982-1985), Chevrot (1998), Morgan (1998-1999) qui montrent, dans leurs travaux, que la connaissance de l'orthographe influence les réalisations orales.

Une autre interprétation de ces relations entre oral et écrit permet de dire que la langue orale influence la langue écrite. Comme nous l'avons vu dans le chapitre 5, plusieurs études ont montré que des compétences phonologiques influencent l'apprentissage de la langue écrite (Ecalte & Magnan, 2008 ; Goswami, 2001 ; Bertelson et al, 1985 ; Casalis & Louis-Alexandre, 2000 ; Ecalte, 2000). Par conséquent, nous pouvons penser que les enfants qui produisent mieux à l'oral ont aussi des productions écrites et des lectures plus justes : des compétences orales (et phonologiques) plus développées aident à mieux écrire et lire en présence de phénomènes de resyllabation, confirmant donc ce qui a déjà été vu dans la littérature. Nous pouvons également penser que, comme plusieurs études l'ont montré (Calfée *et al.*, 1972 ; Lewkowitz, 1980 ; Liberman *et al.*, 1974) au cours de l'apprentissage de la *literacy*, les enfants qui lisent et écrivent mieux accèdent aux unités plus petites de la langue orale (les phonèmes). Cet accès permet d'améliorer les productions orales et explique les meilleurs résultats dans la tâche de dénomination guidée. La conscience des phonèmes, unités de base du codage de notre système alphabétique, est initialement plus difficile à acquérir que la conscience des syllabes et se développe plus tard, au cours de l'apprentissage de la lecture.

Enfin, nous ne remarquons pas de relation entre la connaissance du marquage du pluriel à l'écrit et la réalisation des liaisons en [z]. Ces données sont en accord avec le résultats de Masquida *et al.* (2010), qui montrent que les <s> à la fin des mots pour indiquer le pluriel ne sont pas mieux orthographiés dans le contexte de liaison.

Chapitre 10

Troisième étape de l'étude longitudinale : l'année au CE1

10.1 Troisième étape de l'étude longitudinale : janvier 2015

Comme l'explique le ministère de l'éducation sur son site Internet, le CP et le CE1 représentent le cycle des apprentissages fondamentaux. Dans ces deux classes, l'un des buts définis par le ministère est l'apprentissage de la lecture et de l'écriture de la langue française.

Cette idée est bien résumée de la façon suivante⁴⁴ :

« Au cours préparatoire, l'apprentissage de la lecture passe par le décodage et l'identification des mots, par l'acquisition progressive des connaissances et compétences nécessaires à la compréhension des textes. Les apprentissages de la lecture et de l'écriture, qu'il s'agisse des mots, des phrases, des textes, menés de pair, se renforcent mutuellement tout au long du cycle. Ces apprentissages s'appuient sur la pratique orale du langage et sur l'acquisition du vocabulaire. Ils s'accompagnent d'une première initiation à la grammaire et à l'orthographe. »

Nous avons donc prolongé notre étude longitudinale pendant le cours élémentaire première année, CE1.

Au moyen de la tâche orale de dénomination guidée (DENa^{CE1}), déjà présentée en CP (voir tâche DENa^{CP1}), nous avons évalué les compétences de segmentation dans les deux classes de CE1. À cette tâche, nous avons ajouté les deux tâches qui

44 http://www.education.gouv.fr/cid38/presentation-des-programmes-et-des-horaires-a-l-ecole-elementaire.html#Programmes_du%20cycle%20des%20apprentissages%20fondamentaux, consulté le 4 Février 2015, 15h25

concernent la langue écrite, déjà proposées à la fin du CP et que nous appelons PROD.ECR^{CE1} et LECT^{CE1}. Les enfants ont été testés pour les trois tâches (DENa^{CE1}, PROD.ECR^{CE1} et LECT^{CE1}) en janvier 2015, vers la moitié du CE1.

Les résultats de la tâche orale de dénomination guidée (DENa^{CE1}) et des tâches d'écriture (PROD.ECR^{CE1}) et de lecture (LECT^{CE1}), peuvent nous fournir un aperçu indirect sur le lien entre les compétences orales et les compétences écrites. Pour chercher à établir un rapport plus direct entre eux, nous proposons deux tâches d'amorçage multimodal (*cross-modal priming*) en partant de l'expérience conduite chez les adultes par Gaskell *et al.* (2002) et que nous avons résumée dans la section 4.3.2. Ces tâches permettent non seulement de mieux comprendre les rapports entre écrit et oral en présence de phénomènes de resyllabation mais également de pouvoir évaluer le traitement cognitif de ces séquences à l'oral et à l'écrit. Pour ces deux tâches (AMOa et AMOb), nous avons testé un groupe contrôle d'adultes francophones.

Nous avons décidé de ne pas répéter la tâche de jugement de grammaticalité (PERC^{CP2}), car comme nous l'avons déjà dit, celle-ci était la cible de plusieurs critiques dues à un problème de contexte et de planification de la tâche en elle-même.

10.1.1 Les hypothèses au CE1

Nous pouvons penser que pendant le CP, les compétences des enfants ont été renforcées par le travail en classe. Nous nous attendons donc à une évolution des résultats au CE1, en ce qui concerne le niveau de segmentation en mots dans la tâche orale (DENa^{CE1}) et la connaissance de la langue écrite dans les tâches de langue écrite (PROD.ECR^{CE1} et LECT^{CE1}) proposées. Pour la langue orale, nous pouvons penser à une résolution des non-réalisations en accord avec les résultats de Dugua et Baclesse (2014) et Basset (2000). Pour la langue écrite, nous nous attendons à de meilleures résultats en lecture et en écriture grâce à l'intégration d'une stratégie lexicale (Sprenger-Charolles *et al.*, 2003) et d'une connaissance croissante du marquage du pluriel à l'écrit (Thevenin *et al.*, 1999). Cependant, comme montré par Mesquida *et al.* (2010), nous supposons qu'orthographier le graphème <s> du pluriel n'implique pas une meilleure réalisation de la liaison en [z] au pluriel.

Dans le chapitre précédent, nous avons vu que les compétences de la langue orale étaient liées aux compétences de la langue écrite. Plus précisément, nous avons vu que des meilleures performances en lecture et en écriture sont associées à des productions orales plus justes (§ 9.8.3). Nous nous attendons à trouver la même corrélation au CE1 et, en plus, que la résolution de non-réalisations soit en accord avec les meilleures compétences en *literacy*.

En ce qui concerne les tâches d'amorçage (AMOA et AMOB), nous attendons chez les enfants des résultats différents de ceux qui étaient enregistrés chez les adultes par Gaskell *et al.* (2002) : cela s'explique par une différence de compétences de la langue écrite. En revanche chez le groupe d'adultes, nous nous attendons à trouver les mêmes résultats que Gaskell *et al.* (2002) et Wauquier-Gravelinnes (1996).

Au CE1 nous continuons à évaluer le rôle de la fréquence et de la phonologie (au moyen de l'analyse phonétique) pour la tâche DENa^{CE1}. Enfin nous analysons la variable sociolinguistique pour les productions orales de liaison et la fréquence des lectures partagées.

10.1.2 Participants

Nous avons isolé dans les deux sections de CE1 (CE1 A et CE1 B) les mêmes enfants que nous avons déjà testés au CP (voir chapitre 7). Il ne nous reste au CE1 que 37 enfants (21 filles ; 16 garçons ; âge moyen = 7;5) parmi les 43 de l'année précédente à cause des redoublements et de certains changements d'école. De plus, les enfants que nous avons testés au CP ont été mélangés dans les deux sections. Non seulement la structure des deux classes a été modifiée, mais les maîtresses en charge des cours ont aussi changé.

10.2 DENa^{CE1} : la tâche de dénomination guidée au CE1

10.2.1 Procédure de la tâche DENa^{CE1}

Pour la tâche de dénomination guidée DENa^{CE1}, nous avons adopté la même procédure que celle que nous avons utilisée pour les tâches DENa^{CP1} et DENa^{CP2}. Pour une explication détaillée de la procédure et du répertoire, nous invitons à la lecture des sections 8.3 et 9.3.

10.2.2 Résultats de la tâche DENa^{CE1}

Nous présentons ici les résultats de la tâche DENa^{CE1}. Le tableau 66 nous résume les résultats généraux. Comme nous l'avons déjà dit dans le chapitre précédent, nous classifions comme « erreurs de la tâche » la dénomination non-correcte de la carte (p.ex. *une abeille* pour *les[z]ours* ou l'absence de réponse). Les erreurs de segmentation concernent le traitement non-correct des frontières des mots (ex. *l'oie* pour *la noix* ou *le(s)[l]ours* pour *de les[z]ours*). Le score maximal pour les mots tests ou pour les items *fillers* est 9. Le score maximal de tests + *fillers* est 18.

Dans le tableau 66, nous observons un score d'erreurs plus élevé pour les erreurs de segmentation (2,05 vs. 0,58 $t_1(36) = 28,431$, $p < .001$; $t_2(8) = 2,30$, $p < .05$). Pour les items tests, la plupart des erreurs sont des erreurs de segmentation (2 vs. 0,13 $t_1(36) = 36,889$, $p < .001$; $t_2(8) = 3,03$, $p < .05$). Pour les *fillers*, il n'y a pas de différence entre le nombre d'erreurs de la tâche et le nombre d'erreurs de segmentation ($t_1(36) = 3,64$, $p = .06$; $t_2(8) = 4,56$, $p = .80$) mais le score moyen total reste très bas (0,50).

	Scores moyens		
	erreurs de la tâche (ex. <i>une abeille</i> pour <i>les[z]ours</i> ou absence de réponse)	erreurs de segmentation (ex. <i>le(s)[n]ours</i> ou <i>les[Ø]ours</i> pour <i>le(s)[z]ours</i>)	TOTAL
Items <i>fillers</i>	0,45	0,05	0,50
Items tests	0,13	2	2,13
Tests + <i>fillers</i>	0,58	2,05	2,63

Tableau 66. Résultats généraux de la tâche DENa^{CE1}

Dans le tableau 67, nous analysons plus en profondeur les erreurs de segmentation pour la dénomination d'items tests. Nous classifions les erreurs de segmentation par phénomène consonantique et par type d'erreur (non-réalisation et insertion d'une consonne erronée) en suivant la classification déjà donnée pour DENa^{CP1} et DENa^{CP2} (§ 8.3.3 et § 9.3.2). Le score maximal pour ces catégories est 3.

La catégorie « autres erreurs » présente dans les analyses des données pour DENa^{CP1} et DENa^{CP2} n'est pas présente dans ce tableau car au CE1 nous n'avons aucune erreur qui n'appartienne à aucune des deux autres catégories de classification.

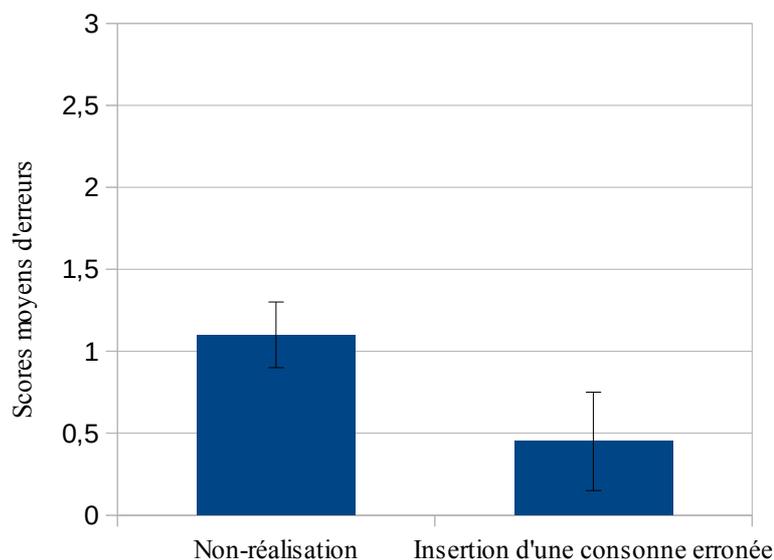
Phénomène	Scores moyens d'erreurs TOTAL	Types d'erreurs	Scores moyens d'erreur par type
Liaison	0,15	Non-réalisation (<i>le(s)[Ø]ours</i> [leØuʁs] pour [le.zuʁs])	0,13
		Insertion d'une consonne erronée (<i>les[n]ours</i> [le.nuʁs] pour [le.zuʁs])	0,02
Élision	0,53	Non-réalisation (<i>la-armoire</i> [la.aʁ.mwaʁ] pour [laʁ.mwaʁ])	0,40
		Insertion d'une consonne erronée (<i>la[n]armoire</i> [la.naʁ.mwaʁ] pour [laʁ.mwaʁ])	0,13
Enchaînement	0,99	Non-réalisation (<i>la grosse-araignée</i> [la.gʁos.aʁɛɲe] pour [la.gʁo.sa.ʁɛ.ɲe])	0,70
		Insertion d'une consonne erronée (<i>la grosse[z]araignée</i> [la.gʁo.za.ʁɛ.ɲe] pour [la.gʁo.sa.ʁɛ.ɲe])	0,29

Tableau 67. Scores par type d'erreur et pour chaque phénomène dans DENa^{CE1}

L'analyse de la variance (ANOVA) pour les scores d'erreurs pour trois modalités (trois phénomènes) utilisant les sujets et les items comme variables aléatoires montre un effet du facteur « phénomène » ($F_1(2,74) = 23,347$, $p < 0,01$;

$F_2(2,4) = 9,72$, $p < .05$). L'analyse de contrastes révèle que la différence n'est pas significative entre la liaison et l'élision ($F_1(2, 74) = 2,01$, $p = .16$, $F_2(1, 2) = 0,86$, $p = .30$) mais elle est statistiquement significative entre la liaison et l'élision d'une part et l'enchaînement d'autre part ($F_1(2, 74) = 5,763$, $p < .001$; $F_2(1, 2) = 39,00$, $p = .11$)⁴⁵. Liaison et élision sont les deux phénomènes qui causent le moins d'erreurs. Le phénomène consonantique qui se différencie des deux autres et qui donne le plus de problèmes est donc l'enchaînement.

En analysant les types d'erreurs, nous remarquons une différence statistiquement significative entre les erreurs d'insertion d'une consonne erronée et les erreurs de non-réalisation ($t_1(36) = 3,19$, $p < .05$, $t_2(8) = 3,10$, $p < .05$). Le type d'erreur le plus courant est la non-réalisation, comme illustré dans le graphique 17.



Graphique 17. Erreurs de non-réalisations et d'insertion d'une consonne erronée dans DENa^{CE1}

Nous observons qu'au CE1, les non-réalisations semblent être encore les plus présentes dans la tâche de dénomination guidée, surtout pour l'enchaînement. Cependant l'analyse des données entre CP et CE1 montre que le nombre de non-réalisations baisse au CE1⁴⁶. Nous avons décidé d'analyser phonétiquement l'espace

⁴⁵ Comme nous avons déjà dit, le F_2 est calculé sur les items et nous avons seulement 3 items par condition. Pour cette raison parfois la significativité de F_1 et celle de F_2 ne correspondent pas. Quand nous sommes en présence de cette situation, nous avons décidé de considérer seulement la valeur de F_2 et de sa significativité statistique.

⁴⁶ Pour une comparaison détaillée des résultats entre fin et début du CP nous invitons à la lecture du chapitre 11.

sonore entre Mot1 et Mot2 pour vérifier si cette diminution des non-réalisations est en rapport avec un traitement phonétique préférentiel.

10.2.3 Analyse phonétique des non-réalisations au CE1

Comme nous l'avons déjà fait pour le CP, nous avons décidé d'analyser au moyen du logiciel PRAAT l'espace sonore entre Mot1 et Mot2 dans le cas des non-réalisations pour les 3 phénomènes de resyllabation. Nous avons toujours fait référence à la classification sur trois types de productions : « continuum », « coup de glotte » et « pause ». Nous ne pouvons pas inclure l'enchaînement dans une production du type « continuum », car si les enfants produisent un continuum entre adjectif antéposé et nom, ils produisent une bonne réalisation d'enchaînement. Nous invitons à la lecture de la section 8.3.3.1 pour la procédure d'analyse et la description phonétique de ces trois productions. Dans le tableau 68, nous présentons les pourcentages calculés sur le nombre total de non-réalisations pour chaque phénomène de resyllabation.

	%		
	liaison	élision	enchaînement
Coup de glotte ([le.ʔuʁs],[lə.ʔo.ʁε.je], [la.pə.tit.ʔo.ʁɑ̃ʒ])	40	8,4	30
Pause ([le/uʁs], [lə/o.ʁε.je], [la.pə.tit/o.ʁɑ̃ʒ])	0	25	70
Continuum ([leuʁs], [lə.o.ʁε.je])	60	66,40	//

Tableau 68. Analyse phonétique pour les non-réalisations dans DENa^{CE1}

Pour la liaison, la plupart des enfants qui ne produisent pas la CL entre Mot1 et Mot2 réalisent des séquences qui ressemblent à des « continuums » entre les deux mots (60 %), sans antéposer une véritable pause (0 %). À ce stade, la production de coups de glotte est inférieur (40 %).

Pour l'élision, la plupart des enfants produisent un « continuum » (66,40 %). Ils n'insèrent donc ni pause ni coup de glotte entre Mot1 et Mot2 mais préfèrent produire

la séquence comme si elle était composée de deux mots collés, ce qui est en effet le cas dans une simple élision.

En ce qui concerne l'enchaînement, les enfants produisent une pause entre adjectif et nom (70 %) plutôt qu'un coup de glotte (30 %).

En résumé, nous pouvons dire que pour l'élision et la liaison, même si les enfants ne produisent pas le phénomène consonantique (la CL ou l'élision), ils essaient de rendre compte de la production orale et ils prononcent Mot1 et Mot2 comme une séquence unique, un continuum. Nous chercherons à mettre en rapport ces résultats avec l'acquisition de la lecture qui aide à créer un continuum sonore entre deux mots normalement séparés par une marque graphique (§ 10.4).

En ce qui concerne l'enchaînement, nous pouvons dire que les enfants préfèrent introduire une pause entre adjectif et nom. Il faut souligner que 50 % des pauses sont enregistrées dans la production de la séquence *un minuscule ogre*. L'insertion d'une pause entre *minuscule* et *ogre* peut être causée par la longueur de l'adjectif (4 syllabes) qui devient difficile à enchaîner dans le processus de dénomination guidée.

10.2.4 Le rôle de la fréquence

Nous avons cherché à établir un rapport entre la fréquence SFI des items tests (la fréquence dans la base de données MANULEX) et les erreurs d'ajout d'une mauvaise consonne [z] au début du Mot2 dans DENa^{CE1}, comme on l'avait déjà fait pour DENa^{CP1} et DENa^{CP2} (§ 8.3.3.3 et § 9.3.4). Le rapport entre ces deux variables permet d'évaluer le rôle de la fréquence dans l'acquisition des frontières des mots : si la fréquence joue un rôle dans cette acquisition, les valeurs de SFI au pluriel sont en rapport avec les erreurs d'insertion d'une consonne erronée [z] au début du Mot2. Pour les données au CE1, la valeur du r de Pearson nous montre qu'il n'y a pas de corrélation entre la fréquence SFI et la production des consonnes erronées [z] ($r(8) = .63$, $p = .06$). L'absence d'une corrélation permet d'exclure l'influence de la fréquence dans le processus de l'acquisition de la liaison à ce moment.

10.2.5 La variante sociolinguistique

Nous analysons l'influence de la variante sociolinguistique sur la production des liaisons au CE1, comme proposé pour le CP (§ 8.5 et § 9.3.5). Comme nous l'avons déjà expliqué dans les chapitres précédents, le traitement des données qui prennent en considération la variante sociolinguistique est purement qualitatif.

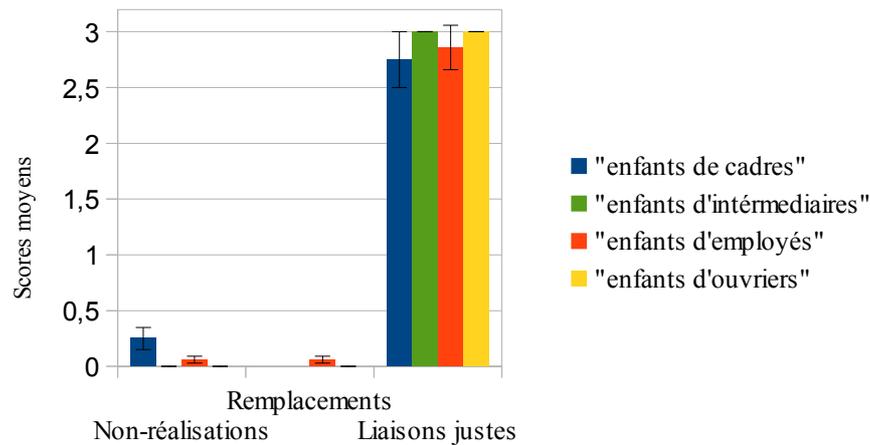
D'abord nous cherchons à évaluer l'importance du niveau d'instruction des parents sur les productions des enfants. Nous reproposez ici les 3 groupes définis au chapitre 6. Dans le tableau 69, nous classifions les scores moyens d'erreurs pour DENa^{CE1} dans les contextes de liaison, par rapport aux trois niveaux. Le score maximal est 3.

		Scores moyens d'erreurs
Enfants avec des parents de	« Niveau 1 »	0
	« Niveau 2 »	0,60
	« Niveau 3 »	0,08

Tableau 69. Scores moyens des erreurs des enfants dans DENa^{CE1} selon le niveau d'études des parents

Les résultats nous indiquent que ce ne sont pas les enfants avec des parents qui ont un niveau d'instruction plus élevé qui produisent le moins d'erreurs de liaison. En revanche ce sont les enfants qui ont des parents avec un niveau scolaire moins poussé (niveau 1) qui ont des meilleurs résultats.

Nous évaluons également le milieu socioculturel des parents qui considère la classification des métiers parentaux (classification en 4 groupes, § 8.5). Pour les 4 groupes de cette classification, nous analysons les scores moyens d'erreurs (d'insertion d'une consonne erronée et non-réalisations) et de productions correctes dans la tâche DENa^{CE1} pour la liaison. Le score maximal est 3.



Graphique 18. Scores des enfants dans DENa^{CE1} selon la profession des parents

Dans le graphique 18, nous pouvons voir qu'à ce stade de l'acquisition les « enfants de cadres » et les « enfants d'employés » ont les moins bons résultats, car ils montrent encore des erreurs dans la production de liaison. Dans les autres groupes, les erreurs sont résolues. Ces résultats permettent de voir que, contrairement aux attentes, à un niveau socioculturel meilleur ne correspondent pas des meilleures productions de liaisons et vice-versa.

10.3 Les tâches liées à la langue écrite : janvier 2015

Pour les tâches sur la langue écrite, les deux classes de CE1 ont été testées en janvier 2015. Nous avons proposé deux tâches : une tâche d'écriture PROD.ECR^{CE1} et une tâche de lecture LECT^{CE1}.

Pour la tâche d'écriture PROD.ECR^{CE1}, les enfants sont testés en groupes (deux groupes de 13, un groupe de 11) dans une classe vide du complexe scolaire. Pour la tâche de lecture LECT^{CE1}, les enfants sont testés un par un dans une salle silencieuse. L'ordre de présentation des blocs varie selon les enfants pour éviter des effets de facilitation ou des stratégies. Comme pour les productions orales, les lectures des enfants ont été enregistrées au format *.wav* avec une fréquence de 44000 Hz au moyen d'un enregistreur *Zoom H4*.

10.3.1 La tâche d'écriture PROD.ECR^{CE1}

10.3.1.1 Procédure de la tâche PROD.ECR^{CE1}

Nous reprenons les 20 séquences tests et la division en 3 blocs que nous avons utilisée pour la tâche PROD.ECR^{CP2} (§ 9.5.1.1). Nous y avons ajouté 9 séquences *fillers* où le Mot2 commence par une consonne qui peut aussi avoir le statut de la CL (<z> ou <n>) ou être impliquée dans l'élision (<l>). Dans le tableau 70, nous pouvons voir les séquences *fillers* que nous avons ajoutées à la version précédente de la tâche et qui ont été réparties dans ces 3 blocs de façon aléatoire (Tableau 71).

Séquences tests	Types de séquences
Les zéros	À <z> initiale
Les zèbres	
Les zoos	
Un nez	À <n> initiale
Un nid	
Un navire	
Le lit	À <l> initiale
Le lapin	
Le livre	

Tableau 70. *Fillers* pour PROD.ECR^{CE1}

Bloc	Séquences tests
1	les ours, un nid, un oiseau, l'oie, une épée, les zoos, un orage, un hippopotame, l'éléphant, le lit
2	le lapin, un hérisson, les affiches, une ambulance, les zèbres, un navire, l'école, un ananas, une hirondelle
3	une autruche, les zéros, l'âne, un nez, un hibou, les oranges, un igloo, le livre, une île, des ailes

Tableau 71. Insertion des *fillers* dans le bloc de séquences tests dans PROD.ECR^{CE1}

Nous avons divisé les enfants en 3 groupes : deux groupes de 13 et un groupe de 11. Pour chaque groupe d'enfants, nous avons changé l'ordre de présentation des blocs comme montré dans l'annexe 11.

Avec ce matériel, nous avons créé un livret, comme nous l'avions déjà fait en juin pour PROD.ECR^{CP2}. Nous avons aussi appliqué la même procédure expérimentale (§ 9.5.1). La tâche d'écriture au CE1 s'est déroulée plus rapidement qu'au CP, malgré le fait qu'il y ait une augmentation de la quantité d'items. Sa durée moyenne est d'environ 30 minutes.

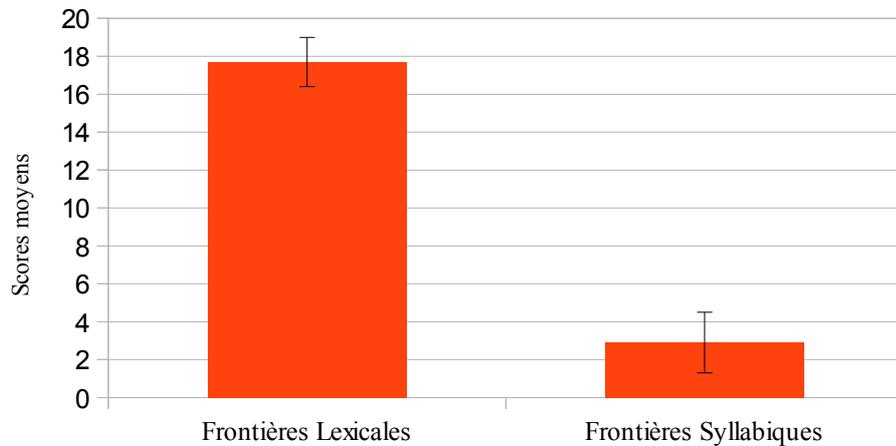
10.3.1.2 Résultats de la tâche PROD.ECR^{CE1}

Nous avons déjà analysé les productions écrites des enfants en proposant une classification en deux catégories (§ 9.5.1.2) :

- des productions écrites que nous avons appelées à « frontières lexicales » où les séquences en liaison, en élision et en enchaînement sont écrites en respectant les frontières lexicales des mots (ex. <les ours>, <l'éléphant>, <une autruche>);
- des productions écrites que nous avons appelées à « frontières syllabiques » où la consonne qui resyllabifie sur le Mot2 est traitée comme consonne initiale du Mot2, en respectant les frontières syllabiques perturbées à l'oral par la resyllabation (ex. pour la liaison <les sours>, pour l'élision <le léléphant>, pour l'enchaînement <une naustruche>).

Nous avons vu que les enfants au CP préfèrent plutôt écrire les mots en respectant les frontières lexicales. Est-ce encore le cas au CE1 ? Nous observons une différence statistiquement significative entre les productions écrites à « frontières syllabiques » et celles à « frontières lexicales » ($t_1(36) = 4,709$, $p < .05$; $t_2(19) = 2,79$, $p < .05$). Le graphique 19 montre qu'au CE1, les enfants préfèrent aussi écrire les séquences en respectant les frontières lexicales⁴⁷.

47 Le score maximal est 20.



Graphique 19. Scores de productions écrites au CE1 : « frontières syllabiques » vs. « frontières lexicales »

10.3.1.2.1 Les productions écrites qui respectent les frontières lexicales

D'abord nous traitons les productions écrites où il y a un respect des frontières lexicales.

Comme nous l'avons déjà dit dans le chapitre 9, nous présentons les données pour les séquences caractérisées par les phénomènes consonantiques de liaison, d'élision et d'enchaînement sans tenir compte des contextes déterminant + nom à *h* initial ou des *fillers* (du type *les zéros*). Dans le tableau 72, nous présentons les pourcentages d'erreurs pour chaque phénomène consonantique. Les pourcentages sont calculés sur le nombre de productions écrites totales pour tous les enfants (16 séquences fois 37 enfants).

Phénomènes	% des productions écrites qui respectent les frontières lexicales (<i>les ours, l'âne, une ambulance</i>)
Liaison au pluriel	95,27
Liaison au singulier	95,78
Élision	96,62
Enchaînement	98,99

Tableau 72. Pourcentages de productions écrites à « frontières lexicales » dans PROD.ECR^{CE1}

Pour tous les phénomènes, les enfants arrivent à bien segmenter les séquences à l'écrit : le pourcentage de segmentations écrites correctes est supérieur à 95 %.

Ces bons résultats en segmentation sont-ils aussi associés à des productions écrites globalement correctes ? Pour répondre à cette question, nous analysons les productions écrites globalement correctes, où les enfants réussissent à bien orthographier tous les mots de la séquence sans fautes ou incertitudes. Le tableau 73 montre les résultats sur la totalité des productions écrites par phénomène (37 fois 16). Nous pouvons voir que plus de 80 % des productions écrites sont bien orthographiées.

Phénomènes	% de productions écrites globalement correctes
Liaison au pluriel	86,3
Liaison au singulier	80,3
Élision	86,1
Enchaînement	86,3

Tableau 73. Pourcentages des productions écrites globalement correctes dans PROD.ECR^{CE1}

En comparant les résultats des deux tableaux (Tableau 72 et tableau 73), nous pouvons dire que les enfants qui réussissent à bien identifier les frontières des mots à l'écrit en présence des phénomènes de resyllabation à l'oral réussissent aussi à bien orthographier globalement le Mot1 et le Mot2 contenus dans la séquence. Maîtriser l'orthographe du français en dictée de mots est une compétence qui est acquise au CE1.

10.3.1.2.2 Des questions morphologiques

Comme nous l'avons déjà fait au CP (chapitre 9), nous analysons le marquage du pluriel du Mot2 à l'écrit pour la liaison en [z]. Dans le tableau 74, nous présentons les pourcentages de <s> orthographiés à la fin du Mot2. Les résultats sont calculés sur la totalité de séquences en liaison au pluriel (soit 37 fois 4).

Séquences	% de marquage du pluriel à l'écrit
Les oranges	35,13
Des ailes	40,50
Les ours	67,56
Les affiches	29,72
TOTAL	43,24

Tableau 74. Pourcentages du marquage du pluriel à l'écrit dans la liaison en [z] dans PROD.ECR^{CE1}

Nous pouvons voir que presque la moitié des séquences sont caractérisées par un marquage du pluriel du Mot2 (43,24 %). On constate de nouveau que le mot le plus facile à écrire avec un <s> final est *ours* (67,56 % de marquage du pluriel à l'écrit). Comme nous l'avons déjà expliqué pour les données au CP, dans la production de ce Mot2, une correspondance graphème-phonème a lieu pour le <s> final. Cette correspondance aide les enfants à bien marquer le <s> du pluriel. Pour les autres séquences nous ne remarquons pas de grande différence de pourcentages. Nous allons comparer les résultats au CP et au CE1 dans le chapitre 12, pour voir s'il y a une amélioration dans le marquage du pluriel entre ces deux moments.

10.3.1.2.3 Les productions écrites qui respectent les frontières syllabiques

Dans cette section, nous allons analyser les productions écrites « syllabiques » du type <les jours>, où la consonne de liaison, d'élision ou d'enchaînement est écrite comme si elle était la consonne initiale du Mot2, en respectant la structure syllabique de l'oral.

Nous considérons les trois différents phénomènes consonantiques. Nous ne traitons pas le contexte à *h* initial qui sera développé ensuite. Les pourcentages des productions écrites qui respectent les frontières syllabiques sont présentés dans le tableau 75.

Contextes	% productions écrites « à frontières syllabiques »
Liaison au pluriel	4,73
Liaison au singulier	4,20
Élision	3,38
Enchaînement	1,01

Tableau 75. Pourcentages de productions écrites à « frontières lexicales » dans PROD.ECR^{CE1}

Les résultats montrent que les productions écrites qui respectent les frontières syllabiques ont presque disparu dans tous les contextes testés (ce pourcentage est inférieur à 5 % et dans le cas de l'enchaînement nous avons un pourcentage proche de 1 %). Cela signifie que pour les séquences tests très peu d'enfants ont encore des doutes sur la place à l'écrit de la consonne qui intervient dans les phénomènes consonantiques. Ils réussissent à orthographier correctement le Mot2 comme un mot à voyelle initiale.

Puisque ce type d'erreurs reste marginal à cet âge, nous n'allons pas analyser en détail tous les phénomènes, comme nous l'avons fait pour le CP (chapitre 9).

10.3.1.2.4 Les productions écrites des *fillers*

Pour cette tâche d'écriture, nous avons décidé d'insérer des séquences *fillers* où le Mot2 commence par une consonne qui peut être aussi une consonne impliquée dans les phénomènes de resyllabation (ex. *les zéros*). De cette façon, nous pouvons évaluer si les enfants réussissent à écrire correctement les mots en traitant la consonne lexicale comme initiale du Mot2, sans la confondre avec une possible consonne de liaison, d'enchaînement ou d'élision. Dans le tableau 76, nous proposons une analyse des pourcentages des mauvaises segmentations à l'écrit des *fillers* en relation avec les trois types de consonne initiale du Mot2 (<l>, <n>, <z>).

Les enfants montrent n'avoir aucun problème avec le traitement de la consonne lexicale à l'initiale (Tableau 76). La consonne lexicale est confondue avec la CL dans moins de 2 % des productions écrites (ex. <un avir> pour <un navire> ou <les eaus> pour <les zoos>).

Contextes	% d'erreurs de segmentation à l'écrit
Consonne lexicale <l>	0
Consonne lexicale <n>	1,50
Consonne lexicale <z>	1,80

Tableau 76. Erreurs de segmentation à l'écrit pour les *fillers* au CE1 dans PROD.ECR^{CE1}

Pour la séquence du Mot2 à <l> initial, il n'y a aucun problème à identifier ce <l> comme une consonne lexicale. Les problèmes sont plutôt pour les séquences où le Mot2 présente un <n> ou un <z> initiaux, c'est-à-dire une consonne de liaison ou d'enchaînement. L'absence des problèmes dans les mots qui commencent par la consonne lexicale <l> est probablement due au fait que, dans les séquences article-nom, elle n'est jamais présente en finale de mot car elle est toujours en position d'attaque. C'est le cas de l'élision où la consonne <l> n'est pas visuellement séparée du Mot2 par un espace blanc et est phonologiquement attaque de ce même mot.

10.3.1.2.5 Orthographier les séquences où Mot2 commence par *h*

Nous concentrons à présent notre attention sur les séquences tests où le Mot2 commence par un <h> graphique, comme nous l'avons déjà fait au CP (§ 9.5.1.2.6). Le tableau 77 montre que plus ou moins la moitié des séquences sont orthographiées avec un <h> à l'initial des Mots2. Nous remarquons une grosse amélioration par rapport au CP⁴⁸. Cependant nous n'observons pas de différence entre les contextes à *h* muet et à *h* aspiré initiaux.

		% de <h> au début de Mot2
Contextes	<i>h</i> muet (<i>hippopotame</i> [i.pɔ.pɔ.tam])	49
	<i>h</i> aspiré (<i>hibou</i> [ibu])	50

Tableau 77. Pourcentage de <h> orthographié au début du Mot2 au CE1 dans PROD.ECR^{CE1}

Nous considérons la classification des productions écrites proposée au CP, qui tient compte des 3 catégories en tableau 78 (hyposegmentation, ex. <unippopotame> pour *un hippopotame* ; <n> comme consonne lexicale, ex. <un nibou> pour *un hibou* ou <un nippopotame> pour *un hippopotame* ; <l> comme consonne lexicale, ex. <un lippopotame> pour *un hippopotame*).

48 Nous invitons à la lecture du chapitre 11 pour une comparaison plus détaillée.

	% des productions écrites	
	<i>h</i> muet	<i>h</i> aspiré
Hyposegmentation	0	0
<n> comme consonne lexicale	4,6	4,6
<l> comme consonne lexicale	0	0

Tableau 78. Mauvaises segmentations en contexte *h* aspiré ou *h* muet dans PROD.ECR^{CE1}

La seule erreur qui reste est le traitement de <n> comme consonne lexicale. À nouveau, nous observons que les enfants ont les mêmes difficultés avec les séquences à *h* muet initial qu'avec les suites à *h* aspiré initial.

10.3.1.2.6 Le rôle de la fréquence et de la familiarité dans la PROD.ECR^{CE1}

Familiarité : contextes inconnus et contextes connus : y a-t-il une différence dans leurs productions écrites ?

Nous avons déjà vu au CP que les frontières des mots écrits sont plus facilement identifiées sur les contextes connus que sur les contextes inconnus (voir § 9.5.1.2.7).

Nous avons comparé les scores d'erreurs de segmentation à l'écrit (ex. <les sours> pour *les ours*) dans les deux conditions (contextes connus et inconnus) au CE1. Les résultats du *t*-test montrent qu'il n'y a pas de différence statistiquement significative pour les erreurs de segmentation à l'écrit entre les productions écrites des séquences dans des contextes connus et les séquences dans des contextes inconnus ($t_1(36) = 1,05, p = .30$; $t_2(9) = 1,04, p = .05$).

Contrairement aux résultats du CP, au CE1, la familiarité avec les contextes n'implique pas l'augmentation de séquences correctement orthographiées. Ces résultats peuvent être expliqués par notre classification de contextes connus et inconnus. Cette classification a été établie au début du CP. Nous n'avons pas de données consécutives à cette période qui pourraient nous permettre de contrôler exactement les items connus et inconnus pour les enfants. On peut supposer que certains contextes qui au CP leur étaient inconnus, sont familiers en CE1. À ce stade, la classification peut avoir subi un changement qui n'a pas été contrôlé. De plus, nous avons accès seulement aux données

fournies par l'école, et par conséquent nous écartons les possibles acquisitions à l'écrit qui sont faites en dehors du contexte scolaire et qui peuvent influencer la classification des séquences dans ces deux contextes. Pour ces raisons, il vaut mieux considérer l'analyse de la « familiarité » comme une analyse plus pertinente au CP qu'au CE1.

La fréquence des mots écrits

Nous avons déjà vu dans le chapitre 7 que dans la base de données du CP que nous avons créée, les mots que nous avons utilisés pour la tâche écrite ont des occurrences différentes (fréquences absolues différentes). Nous nous attendions normalement à observer un effet de la fréquence : un mot plus fréquent doit être mieux écrit par les enfants, car la forme visuelle du mot est plus stable et disponible. Pour vérifier l'effet de la fréquence, nous avons décidé d'étudier le lien entre les scores moyens de productions écrites correctes et la fréquence dans la base de données que nous avons créée. Nous avons traité comme « productions écrites correctes » les séquences bien écrites, sans erreurs de segmentation (ex. <les sours>) et sans erreurs orthographiques (ex. <une otruche> pour *une autruche* ou <un net> pour *un nez*). Le calcul de l'indice de corrélation r de Pearson entre ces deux variables montre qu'il n'y a pas de corrélation entre les productions écrites correctes et la fréquence des items dans notre base de données de l'écrit ($r(19) = -.40, p = .09$).

La base de données que nous avons créée prend en considération les cahiers des enfants au CP, vers la fin de l'année scolaire. Entre le moment de la création de notre base de données et celui de la passation de cette tâche, 8 mois ont passé, dont 4 mois d'école. Il est donc sûr que le lexique orthographique des enfants a évolué et la fréquence des items changé. Pour mieux évaluer le rapport entre les productions écrites d'items et la fréquence orthographique des mots, il vaut mieux prendre en considération la base de données MANULEX, car elle présente non seulement la fréquence des mots écrits au CP, mais aussi au CE1 (voir chapitre 7, pour les détails). Dans MANULEX nous avons sélectionné pour chaque item (Mot2 dans les séquences) la fréquence SFI au CE1. Ensuite, nous avons cherché à calculer le rapport de corrélation entre la fréquence d'item SFI en MANULEX et les scores de productions écrites correctes. Le résultat de ce calcul montre qu'il n'y a pas de corrélation entre ces deux variables ($r(19) = .37, p = .09$).

Ensuite nous avons mis en rapport ces deux fréquences (fréquence absolue dans notre base de données et SFI au CE1) et les erreurs de segmentation à l'écrit, soit les productions écrites à « frontières syllabiques » où les enfants n'identifient pas correctement les frontières des mots à l'écrit (ex. <les saffiches>). Les résultats du calcul de r de Pearson montre qu'il n'y a pas de corrélation entre les erreurs de segmentation à l'écrit et la fréquence dans notre base de données ($r(19) = 0,29$, $p = .60$) et également entre ces erreurs et l'indice SFI au CE1 ($r(19) = 0,27$, $p = .35$). Cela veut dire que la fréquence des Mots2 à écrire n'influence la représentation des frontières des mots à l'écrit en présence des phénomènes de resyllabation à l'oral.

10.3.2 La tâche LECT^{CE1} de lecture à voix haute

10.3.2.1 Procédure et matériel pour la tâche LECT^{CE1}

Pour la procédure de la tâche de lecture à voix haute LECT^{CE1} nous invitons à la lecture de la procédure pour LECT^{CP2} (§ 9.5.2.1). Le répertoire utilisé est le même que celui que nous avons utilisé pour la tâche d'écriture PROD.ECR^{CE1}, y compris pour l'introduction des *fillers*.

10.3.2.2 Résultats généraux de la tâche LECT^{CE1}

Comme nous l'avons déjà montré dans LECT^{CP2}, nous avons classifié les lectures des enfants selon trois types : séquences lues correctement, lectures syllabiques, problèmes de lecture (lecture inventée/erronée). Dans cette classification, lire correctement la séquence signifie bien lire Mot1 et Mot2, même si le phénomène de resyllabation n'est pas réalisé (lecture, *les[0]affiches* pour *les[z]affiches*). Les données présentées dans le tableau 79 ne prennent pas en considération les problèmes causés par les phénomènes de resyllabation en lecture. Cela sera le cœur de l'analyse des prochains paragraphes.

	%
Séquences bien lues	94,50
Lectures syllabiques	2,75
Problèmes de lecture (lecture erronée/inventée)	2,75

Tableau 79. Pourcentages pour les trois types de lecture chez les enfants dans LECT^{CE1}

La plupart des séquences sont bien lues (94,50 %). Les pourcentages de lectures syllabiques ou de problèmes sérieux en lecture sont très bas (2,75 % et 2,75 %).

10.3.2.2.1 Les productions de phénomènes de resyllabation en LECT^{CE1}

Nous proposons l'analyse plus détaillée des séquences lues correctement. Nous avons écarté les lectures syllabiques (*les-a-fi-ches*) et les lectures inventées/erronées (*une abeille* pour *une ambulance*). De plus nous n'analysons pas les séquences à *h* initial qui seront décrites en détail dans une section dédiée.

Comme pour le CP, nous avons établi une comparaison entre les pourcentages de lectures où le phénomène de resyllabation est bien réalisé (REAL.LECT^{CE1}) et les pourcentages de lectures qui font émerger un problème dans la réalisation du phénomène de resyllabation (ERR.LECT^{CE1}). La plupart des phénomènes consonantiques sont bien réalisés en lecture (60,06 % de REAL.LECT^{CE1}). Il reste encore un pourcentage assez élevé (34,94 %) des ERR.LECT^{CE1}.

Pour vérifier quel phénomène consonantique cause le plus de problèmes, nous proposons les pourcentages de ERR.LECT^{CE1} pour chaque phénomène consonantique testé (Tableau 80). Les pourcentages sont calculés pour chaque phénomène consonantique (donc sur 4 séquences fois 37 enfants).

	% ERR.LECT ^{CE1}
Liaison en [z]	19,82
Liaison en [n]	15,38
Enchaînement	19,67
Élision	0

Tableau 80. ERR.LECT^{CE1} pour les 4 phénomènes

Nous pouvons voir que la lecture de séquences en élision ne cause aucun problème. Les enfants lisent sans problème le [l] du déterminant comme attaque du Mot2 car, pour ces séquences, la forme orthographique ne complique pas la tâche : le [l] est positionné plutôt à l'initial du Mot2 séparé seulement par une apostrophe.

Pour les autres phénomènes consonantiques, c'est-à-dire la liaison en [n], en [z] ou l'enchaînement, la situation est presque la même : nous constatons entre 15 % et 20 % d'erreurs en lecture. Nous pouvons dire qu'entre ces trois phénomènes, il n'y a pas un qui soit plus difficile à lire : tous les trois déclenchent la même quantité de fautes. Dans les prochains paragraphes, nous analysons en détail chaque phénomène pour rendre compte des problématiques spécifiques en lecture. Comme l'élision ne cause aucun problème en lecture, elle ne sera pas traitée.

La liaison

Nous utilisons une classification qui tient compte des trois catégories pour analyser la lecture des liaisons en [z] et en [n]. Deux de ces trois catégories étaient déjà utilisées pour décrire la lecture au CP :

1. NON-REAL.LECT^{CE1}. La CL n'est pas réalisée entre Mot1 et Mot2 qui sont lus comme deux items séparés (ex. *les[Ø]affiches* [leØa.fiʃ] pour *les[z]affiches* [le.za.fiʃ]) ;
2. PRA.LECT^{CE1}. La CL va remplir l'attaque vide au début du Mot2. De plus, après la lecture du déterminant il y a toujours une pause et le nom à voyelle initiale est lu comme un mot à consonne initiale (ex. *les-[z]affiches* [le/za.fiʃ] pour *les[z]affiches* [le.za.fiʃ]) ;
3. NONEnch.LECT^{CE1}. La CL est réalisée à la fin du Mot1 et, après une pause, les enfants lisent le Mot2 qui est traité comme un mot à voyelle initiale (ex. *les[z]affiches* [lez/a.fiʃ] pour *les[z]affiches* [le.za.fiʃ]).

Nous présentons les pourcentages pour ces trois catégories dans le tableau 81. Les pourcentages font référence aux 4 séquences tests de chaque phénomène multiplié par les 37 enfants.

	%		
	NON-REAL.LECT ^{CE1}	PRA.LECT ^{CE1}	NONEnch.LECT ^{CE1}
Liaison en [z]	34,45	6,75	0,67
Liaison en [n]	29,05	3,35	2,02

Tableau 81. Pourcentages de NON-REAL^{CE1}, PRA^{CE1} et NONEnch^{CE1}

La situation est assez équilibrée pour la liaison en [z] et la liaison en [n] : dans les deux contextes, la plupart des erreurs sont des non-réalisations de la CL (% NON-REAL.LECT^{CE1} 34,45 % et 29,05 %). Nous pouvons aussi voir dans le tableau que les pourcentages de PRA.LECT^{CE1} sont très bas (6,75 % et 3,35 %), tout comme les productions NONEnch.LECT^{CE1} (0,67 % et 20,2 %). Les résultats nous suggèrent que les enfants ont plus de difficultés à réaliser la CL en lecture plutôt que de décider si elle est coda du Mot1 ou attaque du Mot2. Ces résultats ne sont pas surprenants car, à l'écrit, deux mots liés par une liaison sont séparés par un espace blanc. Nous proposons dans la section 10.3.2.2.2 une analyse phonétique fine de ces non-réalisations pour voir si les enfants produisent une véritable pause entre Mot1 et Mot2.

L'enchaînement

Le seul problème qui émerge en lecture est la non-réalisation de ce phénomène consonantique (NON-REAL.LECT^{CE1}, 19,67 %). Dans ce type de lecture, la consonne d'enchaînement est produite mais elle n'est pas resyllabifiée sur Mot2 (*une-île* [yn.il] pour *une[n]île* [y.nil]). Nous allons mieux analyser l'espace sonore entre Mot1 et Mot2 pour voir quelles sont les caractéristiques de ces lectures.

Les séquences à h muet et h aspiré au début du nom

Nous analysons d'abord les lectures de séquences où Mot2 commence par un *h* aspiré. Nous écartons dans l'analyse les lectures erronées/inventées (*une abeille* pour *un hibou*) et les lectures du type syllabique (*un-hi-bou* [ẽ/i/bu] ou [ẽ/ni/bu]). Pour bien

lire la séquence, les enfants ne doivent pas réaliser la liaison. Nous classifions les lectures enregistrées au CE 1 en 2 types⁴⁹ :

- REAL.LECT^{CE1} : les enfants lisent la séquence en produisant la liaison (*un[n]hibou* [ɛ.ni.bu]). La séquence n'est pas lue correctement.
- NON-REAL.LECT^{CE1} : les enfants lisent les deux mots comme deux formes séparées. La liaison n'est pas réalisée (*un[Ø]hibou* [ɛ̃Øi.bu]) et la séquence est lue correctement.

Nous présentons les pourcentages pour ces deux types de lectures pour chaque séquence testée (Tableau 82). Les pourcentages sont calculés sur la totalité des lectures, soit 37 fois 2.

	% TOTAL		%
REAL.LECT ^{CE1}	54	un[n]hérisson	53
		un[n]hibou	55
NON-REAL.LECT ^{CE1}	46	un[Ø]hérisson	47
		un[Ø]hibou	46

Tableau 82. Réalisation et non-réalisation des phénomènes de resyllabation dans LECT^{CE1} pour les séquences à *h* aspiré initial

La situation présentée dans le tableau est assez équilibrée (% TOTAL : 54 % vs. 46 %). Nous pouvons voir que le phénomène de resyllabation est réalisé dans la moitié des lectures.

Pour les séquences déterminant + nom à *h* muet initial, nous avons proposé la même classification. Nous présentons, dans le tableau 83, deux types de lectures possibles : des lectures REAL.LECT^{CE1} où liaison et enchaînement sont bien réalisés (*un[n]hippopotame, une[n]hirondelle*) et des lectures NON-REAL.LECT^{CE1} qui

⁴⁹ Contrairement à ce qu'on a vu dans le chapitre 9 pour la lecture de ces séquences au CP, nous n'avons que deux catégories. La catégorie PRA.LECT n'est pas utilisée dans cette analyse car aucune lecture ne présente les caractéristiques propres de cette catégorie (Mot1 + pause + Mot2 à attaque remplie).

proposent la non-réalisation de ces deux phénomènes (*un[Ø]hippopotame*, *une[Ø]hirondelle*). Dans le cas du *h* muet, la lecture des séquences sans réalisation de liaison ou d'enchaînement est une erreur.

	% TOTAL		%
REAL.LECT ^{CE1}	51,50	une[n]hirondelle	53
		un[n]hippopotame	50
NON-REAL.LECT ^{CE1}	48,50	une[Ø]hirondelle	47
		un[Ø]hippopotame	50

Tableau 83. Réalisation et non-réalisation des phénomènes de resyllabation dans LECT^{CE1} pour les séquences à *h* muet initial

Le tableau 83 est très similaire au tableau 82. Les enfants produisent la moitié des phénomènes de resyllabation (51,50 %). Les pourcentages ne changent pas selon la séquence que les enfants ont lue ou le phénomène consonantique impliqué (autour de 50 %).

En ce qui concerne la non-réalisation du phénomène, les enfants savent-ils qu'il y a un *h* graphique et qu'il ne faut pas réaliser les phénomènes ? Réussissent-ils à comprendre la différence entre *h* muet et *h* aspiré ? Les résultats des analyses phonétiques des non-réalisations nous aideront à répondre à ces questions.

10.3.2.2.2 Analyse phonétique des non-réalisations (NON-REAL.LECT^{CE1})

Nous avons proposé la même analyse de non-réalisations que celle de la section 9.5.2.2.3 (NON-REAL.LECT^{CP1}) au moyen du logiciel PRAAT. Nous avons écarté 11 productions (11,9 %), car la qualité de l'enregistrement ne permettait pas une analyse phonétique fine. D'après nos analyses, il y a trois modalités du traitement de l'espace entre Mot1 et Mot2 en lecture : la production d'une pause (ex. [le/a.fi]), [ɛ/wa.zo], [yn/il]), la production d'un coup de glotte (ex. [le.ʔa.fi]), [ɛ.ʔwa.zo], [yn.ʔil]) et la production d'un « continuum » ([ɛ.wa.zo], [le.a.fi]).

Nous avons calculé les pourcentages de ces trois types de productions pour chaque phénomène (Tableau 84).

	% NON-REAL. LECT ^{CE1}		
	liaison en [z]	liaison en [n]	enchaînement
pause (ex. [le/a.fiʃ], [ɛ̃/wa.zo], [yn/il])	62,3	60	81,5
coup de glotte (ex. [le.ʔa.fiʃ], [ɛ̃.ʔwa.zo], [yn.ʔil])	35,5	30	18,5
continuum (ex. [ɛ̃.wa.zo], [le.a.fiʃ])	2,2	20	//

Tableau 84. Analyse phonétique des non-réalisations dans LECT^{CE1}

Pour la liaison en [z] et en [n], la plupart des enfants produisent de véritables pauses entre Mot1 et Mot2, où l'espace sonore reste vide. Les productions de coups de glotte sont moins fréquentes et encore moins fréquentes sont les productions du type continuum.

Pour la lecture des séquences en enchaînement, les résultats montrent également que les enfants produisent toujours une véritable pause entre Mot1 et Mot2. Le pourcentage de coups de glotte pour ce phénomène est très bas et n'est pas spécialement notable (18,5 %).

10.3.2.2.3 Les non-réalisations en présence de *h* aspiré et de *h* muet

Nous proposons une analyse phonétique des non-réalisations de liaison dans la lecture de séquences déterminant + nom à *h* aspiré initial (*un[Ø]hibou* [ɛ̃Øi.bu]). Dans ces séquences, la non-réalisation du phénomène de resyllabation doit être traitée comme le résultat d'une lecture correcte. Cependant, nous voulons vérifier la nature de l'espace sonore entre Mot1 et Mot2 pour voir s'il y a une trace phonétique du *h* graphique au début du Mot2. Nous avons écarté de l'analyse quatre NON-REAL.LECT^{CE1} (10 %) où la qualité audio était trop mauvaise pour une analyse

phonétique fine. Nous avons classifié les NON-REAL.LECT^{CE1} selon 3 catégories comme nous l'avons déjà vu au CP, présentés dans le tableau 85.

	% non-réalisations avec Mot2 à <i>h</i> aspiré initial
pause (ex. [ẽ/i.bu])	54,5
coup de glotte (ex. [ẽ.ʔi.bu])	30,4
continuum (ex. [ẽ.i.bu])	15,20

Tableau 85. Analyse phonétique de non-réalisations dans LECT^{CE1} pour le contexte à *h* aspiré initial

Nous observons qu'en général, la plupart des productions de non-réalisations dans les séquences déterminant + nom à *h* aspiré initial sont des productions de pauses (54,4 %). Le pourcentage de productions de coups de glotte et de continuums sont plus bas (30,4 % et 15,20 %).

Nous comparons ces données avec l'analyse NON-REAL.LECT^{CE1} pour les séquences où le nom a un *h* muet à l'initiale (*un*[Ø]*hippopotame* [ẽØi.po.po.tam]). Dans ce cas, la non-réalisation du phénomène de resyllabation en lecture doit être considérée comme une erreur. Nous avons écarté 4 enregistrements (12 %), car l'analyse phonétique n'était pas possible. Au moyen de la même classification utilisée pour les autres analyses⁵⁰, nous présentons les pourcentages dans le tableau 86.

	% non-réalisations avec Mot2 à <i>h</i> muet initial
pause (ex. [ẽ/i.po.po.tam])	83,9
coup de glotte (ex. [ẽ.ʔi.po.po.tam])	12,9
continuum (ex. [ẽ.i.po.po.tam])	3,2

Tableau 86. Analyse phonétique de non-réalisations dans LECT^{CE1} pour le contexte à *h* muet initial

50 La production de « continuums » est enregistrée seulement dans la séquence *un hippopotame*. Dans l'autre séquence test *une hirondelle*, prononcer un continuum signifie bien enchaîner Mot1 et Mot2 (et donc nous n'avons pas d'erreur de non-réalisation).

Nous pouvons voir que la plupart des productions de non-réalisations dans le cas du *h* muet initial sont des productions de pauses (83,9 %). Les productions de coups de glotte et les productions d'un continuum sont en pourcentage inférieur (12,9 % et 3,2 %).

Nous pouvons conclure qu'au CE1, pour tous les phénomènes testés, quand les enfants ne réalisent pas le phénomène en lecture, ils ne remplissent pas l'espace sonore entre Mot1 et Mot2 qui reste vide. Peut-on supposer que cet espace resté vide à l'oral reflète l'espace blanc entre les mots ? Nous vérifierons ce lien dans des paragraphes dédiés à la fin du chapitre (§ 10.4).

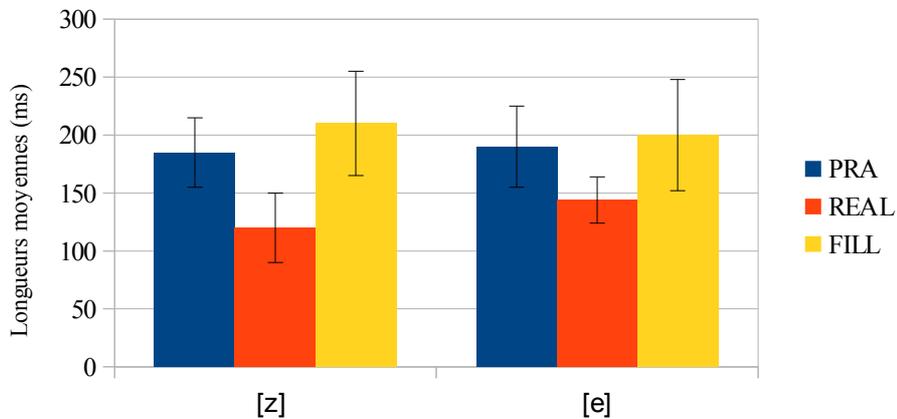
10.3.2.2.4 L'analyse phonétique de PRA.LECT^{CE1}

Comme nous l'avons déjà fait pour les données au CP, nous proposons une analyse des lectures que nous avons appelées PRA.LECT^{CE1} où les enfants lisent le Mot1, produisent une pause et lisent ensuite le Mot2 en remplissant l'attaque avec la consonne qui normalement resyllabifie sur Mot2 ([le/za.fiʃ]). Au moyen du logiciel PRAAT, nous proposons une analyse de ces lectures pour vérifier, dans les séquences en liaison en [z], si la consonne produite au début du nom est une CL ou plutôt une consonne lexicale.

Nous comparons la réalisation de [z] et de [e] (la voyelle qui précède [z]) dans PRA.LECT^{CE1} et dans deux autres types de lecture :

1. REAL.LECT^{CE1} : les lectures correctes de liaison [le.za.fiʃ], où la CL est bien réalisée et resyllabifiée ;
2. FILL.LECT^{CE1} : les lectures des *fillers* où le [z] est la consonne lexicale du Mot2 (ex. *les zoos* [le.zo]).

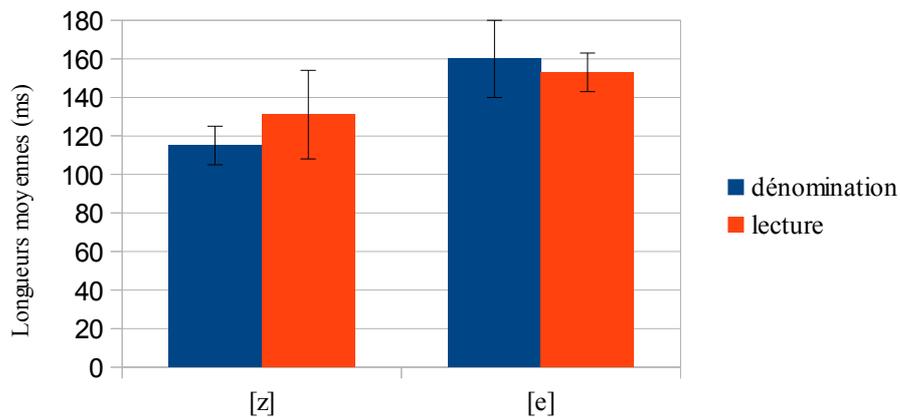
Le graphique 20 montre la représentation des longueurs moyennes (en millisecondes) du [z] et du [e] dans ces trois catégories. Nous observons qu'il n'y a pas de différence entre les longueurs de [z] et [e] dans les cas PRA.LECT^{CE1} et FILL.LECT^{CE1}. Cela implique une réalisation plus courte de [z], qui peut dès lors être considéré comme une CL.



Graphique 20. Longueur de [z] et de [e] dans les trois types de lecture au CE1

Au vu de ces résultats phonétiques, nous pouvons dire que la consonne [z] dans les productions PRA.LECT^{CE1} ([le/za.fiʃ]) est traitée plutôt comme une consonne lexicale que comme une CL.

Cependant les résultats de longueur que nous avons obtenus peuvent être influencés par le fait que ces productions sont véhiculées par l'écrit. Il s'agit d'une tâche « artificielle » et elle est donc moins naturelle qu'une production orale spontanée. Les allongements d'une voyelle ou d'une consonne peuvent être la conséquence d'une hésitation. Nous avons décidé d'analyser phonétiquement les productions orales en dénomination guidée (DENa) des mêmes enfants pour la même séquence en liaison (*les[z]ours*) pour vérifier si la production de liaison à l'oral (sans input écrit) peut être comparée à la production de liaison en lecture (REAL.LECT^{CE1}). Les résultats sont présentés dans le graphique 21.



Graphique 21. Longueur de [z] et de [e] en lecture et en dénomination guidée

Nous n'observons pas de différences en longueur pour [z] et [e] dans les deux modalités. Même si nous avons dans la lecture un support écrit et la représentation des deux mots séparés par un espace, ces deux facteurs ne causent pas de différence acoustique dans les productions au CE1.

Par conséquent, ces dernières données nous permettent d'affirmer que dans la lecture à attaque remplie PRA.LECT^{CE1} [le/za.fif], la consonne [z] au début du Mot2 est plutôt la réalisation d'une consonne lexicale initiale et que les enfants qui lisent ces séquences traitent le Mot2 comme un nom à consonne initiale.

10.3.2.2.5 Le rôle de la fréquence et de la familiarité de la tâche LECT^{CE1}

Familiarité : contexte connu vs. contexte inconnu

Comme nous l'avons fait pour l'écriture dans ce chapitre, nous proposons une comparaison entre contextes connus et contextes inconnus. Nous vérifions si les contextes connus (avec lesquels les enfants ont plus de familiarité à l'écrit) donnent moins d'erreurs en lecture que les contextes inconnus (jamais vus par les enfants en classe). La classification en contextes connus et inconnus est celle présentée au chapitre 7.

Nous comparons le nombre d'ERR.LECT^{CE1} sur les phénomènes de resyllabation et nous observons qu'il n'existe pas de différence statistiquement significative sur les deux contextes ($t_1(36) = 2,00, p=.09$; $t_2(9) = 0,44, p =.32$). Comme au CP, il n'y a pas au CE1 plus d'erreurs en lecture sur les contextes inconnus que sur les contextes connus.

L'effet de la fréquence sur la lecture

Comme pour les productions écrites, nous avons établi un rapport entre les scores de lectures correctes (où Mot1 et Mot2 sont bien lus et où les phénomènes de resyllabation entre eux sont bien réalisés) et la fréquence des mots dans notre base de données (fréquence absolue calculée sur le nombre d'occurrences, voir § 7.2.2). Les résultats du calcul de r de Pearson montrent une corrélation positive entre la fréquence des items dans notre base de données et les scores de lectures correctes ($r(19) = .45, p<.05$) : les scores de lectures correctes augmentent en accord avec l'augmentation de la fréquence des Mots2 à lire.

Nous avons décidé de vérifier ce rapport en utilisant la fréquence d'items SFI au CE1 dans MANULEX⁵¹. Le calcul de la valeur de l'indice de corrélation r de Pearson montre une corrélation positive entre les scores de lectures correctes et la fréquence SFI des Mots2 dans MANULEX ($r(19) = .45, p<.05$) : la fréquence SFI est plus élevée lorsque le score de lectures correctes est plus élevé et vice-versa.

Nous avons ensuite établi des rapports entre les indices de fréquence des items et les scores d'ERR.LECT^{CE1} (= PRA.LECT^{CE1} + NON-REAL.LECT^{CE1} + NONEnch.LECT^{CE1}), donc les erreurs en lecture sur les frontières des mots en présence de phénomènes de resyllabation. Nous nous attendions à ce qu'un item vu à l'écrit plus souvent (donc avec une fréquence plus élevée) implique moins d'erreurs dans le traitement de ses frontières en lecture. Nous remarquons une corrélation seulement quand nous prenons en considération SFI ($r(19) = -.51, p<.05$) et pas quand nous analysons notre variable « fréquence absolue » ($r(19) = -.60, p = .80$). La corrélation avec SFI est négative : la diminution des ERR.LECT^{CE1} est en accord avec

51 Nous invitons à la lecture du chapitre 7 pour plus de détails.

l'augmentation de la fréquence SFI. Les items plus fréquents à l'écrit sont mieux lus dans les séquences en resyllabation et le traitement de leurs frontières est mieux maîtrisé en lecture.

10.3.3 Le rapport entre la lecture et l'écriture au CE1

Nous cherchons à établir plusieurs relations entre la lecture et l'écriture, comme nous avons déjà fait au CP (§ 9.5.3).

- **Les non-réalisations en lecture (NON-REAL.LECT^{CE1} *le(s)/[Ø]ours* [*leØours*]) et les productions écrites correctes où les frontières des Mot1 et Mot2 sont bien identifiées (productions écrites à « frontières lexicales » du type <les ours>).**

D'abord nous mettons en relation les scores de productions écrites correctes « à frontières lexicales », où les enfants identifient correctement les frontières des mots à l'écrit et les scores de NON-REAL.LECT^{CE1} où les enfants traitent Mot1 et Mot2 comme deux items séparés. Le calcul de la corrélation montre que nous ne pouvons pas établir de rapport entre ces deux variables ($r(36) = -.21$, $p = .22$).

- **Les productions écrites à « frontières syllabiques » (<les zaffiches>) et les lectures syllabiques ([*le-za-fi*]).**

Dans les productions écrites à « frontières syllabiques », les enfants orthographient la consonne qui resyllabifie sur Mot2 comme l'initiale de ce mot, en gardant la structure syllabique que nous avons dans les séquences orales. Ils utilisent donc l'unité syllabique pour représenter les frontières des mots à l'écrit. Ces productions guidées par la syllabe sont-elles en relation avec les lectures syllabiques ? Au CE1, nous pouvons répondre « non » à cette question, car il n'y a pas de corrélation entre ces deux variables ($r(36) = .07$, $p = .68$).

- **Les productions écrites correctes où les frontières sont bien identifiées (« frontières lexicales », ex. <les ours>) et les lectures correctes où les phénomènes de resyllabation sont bien réalisés (REAL.LECT^{CE1} *les/z/jours* [le.zuʁs]).**

Nous établissons un rapport entre le score d'identification des frontières de mots à l'écrit et la réalisation correcte des phénomènes de resyllabation en lecture. Le résultat montre qu'il n'y a aucun rapport entre ces deux variables ($r(36) = .03$, $p = .86$). Orthographier correctement les séquences Mot1 + Mot2 n'implique pas de mieux lire ces séquences en enchaînant les deux mots.

10.4 Le rapport entre la *literacy* et les productions orales au CE1

Comme nous l'avons déjà fait à la fin du CP (§ 9.6), nous analysons les rapports entre les productions orales et les productions écrites en établissant plusieurs corrélations.

- **Les productions écrites correctes (« frontières lexicales », ex. <les ours>) dans PROD.ECR^{CE1} et les non-réalisations à l'oral (*le(s)/Ø/jours* [leØuʁs]) dans DENa^{CE1}**

Nous pouvons penser qu'apprendre à écrire aide à séparer les mots et que cette séparation peut impliquer une augmentation du nombre de non-réalisations à l'oral. Y a-t-il une relation entre les productions écrites correctes (« frontières lexicales », *les ours*) où les frontières des mots sont bien orthographiées et les non-réalisations à l'oral ? Nous calculons l'indice de corrélation r de Pearson entre le nombre de non-réalisations à l'oral et les scores de productions écrites correctes pour les 37 enfants. Il y a une corrélation négative entre le traitement correct des frontières des mots à l'écrit et les non-réalisations à l'oral ($r(36) = -.546$, $p < .001$)⁵² : mieux les frontières sont traitées à l'écrit, plus le

52 Pour calculer ce indice de corrélation, nous avons considéré les productions écrites des contextes présentés dans les deux tâches : liaison (en [z]), élision et enchaînement, pour avoir le même type de score que nous pouvons avoir sur la tâche DENa. Nous avons écarté de l'analyse les séquences en

nombre de non-réalisations est bas. Contrairement aux attentes, apprendre à écrire aide à bien produire à l'oral les phénomènes de resyllabation.

- **Les productions écrites correctes (« frontières lexicales », ex. <les ours>) dans PROD.ECR^{CE1} et les productions correctes à l'oral avec la réalisation du phénomène de resyllabation (*le(s)/z]ours* [le.zuʁs]) dans DENa^{CE1}**

Étant donné le lien entre *literacy* et phonologie, nous pouvons penser qu'il y a un lien entre les productions écrites correctes (« frontières lexicales », ex. <les ours>) dans PROD.ECR^{CE1}, où les enfants arrivent à bien délimiter les frontières de Mot1 et Mot2 à l'écrit, et les productions correctes à l'oral où les phénomènes de resyllabation sont réalisés correctement (*le(s)/z]ours* [le.zuʁs]) dans DENa^{CP2}. Le calcul de l'indice de corrélation montre que nous pouvons établir une corrélation positive ($r(36) = .43, p < .05$) entre ces deux variables : à des productions orales correctes correspondent des productions écrites correctes où la frontière gauche de Mot2 est bien déterminée et vice-versa.

- **Liaison en [z] : corrélation entre marquage du pluriel en PROD.ECR^{CE1} et réalisation correcte à l'oral dans DENa^{CE1}**

Nous prenons en considération les contextes de liaison en [z]. Nous cherchons à comprendre si le marquage du pluriel à l'écrit (c'est-à-dire orthographier correctement le <s> à fin de Mot1 et de Mot2) corrèle avec la réalisation correcte de la liaison en [z] à l'oral. Le résultat du calcul de l'indice de corrélation entre ces deux variables montre qu'il n'y a pas de relation entre les deux ($r(36) = .16, p = .34$).

- **Les non-réalisations à l'oral dans DENa^{CE1} (*le(s)/Ø]ours* [leØuʁs]) et les non-réalisations en LECT^{CE1} (NON-REAL.LECT^{CE1} *le(s)/Ø]ours* [leØuʁs])**

Nous cherchons à voir si les enfants qui ne réalisent pas les phénomènes de resyllabation en lecture sont aussi les enfants qui continuent à ne pas réaliser ces phénomènes à l'oral, dans la tâche de dénomination guidée. Le résultat de le

liaison en [n] ou avec *h* initiale. La même procédure a été appliquée pour la corrélation entre les résultats de la tâche de lecture et ceux de la tâche de dénomination guidée.

calcul de l'indice r révèle qu'il n'y a pas de corrélation entre ces deux variables ($r(36) = .02, p = .98$).

- **Les réalisations des phénomènes de resyllabation en LECT^{CE1} (REAL.LECT^{CE1} *le(s)/z/jours* [le.zuʁs]) et les productions orales correctes dans DENa^{CE1} (*le(s)/z/jours* [le.zuʁs])**

Lire correctement les séquences en réalisant les phénomènes de resyllabation est-il en rapport avec les productions orales correctes dans DENa^{CE1} (où Mot1 et Mot2 sont bien enchaînés) ? Le résultat du calcul de r de Pearson montre qu'il y a une corrélation positive entre ces deux variables ($r(36) = .39, p < .05$) : à un nombre plus élevé de lectures correctes correspond le plus de réalisations orales correctes en dénomination et vice-versa.

- **Les insertions d'une consonne erronée dans DENa^{CE1} (*le(s)/n/jours* [le.nuʁs]) et les productions écrites correctes (« frontières lexicales », ex. <les ours>) dans PROD.ECR^{CE1}**

Nous avons vu que le score de productions écrites correctes corrèle avec le nombre de non-réalisations : plus les enfants identifient correctement à l'écrit les frontières des mots, moins ils font d'erreurs de non-réalisation. Pouvons-nous établir le même rapport pour les insertions des consonnes erronées ? Le résultat du calcul de r de Pearson montre que nous ne pouvons pas établir de corrélation entre les deux ($r(36) = 0,21, p = 0,47$).

- **Les insertions d'une consonne erronée dans DENa^{CE1} (*le(s)/n/jours* [le.nuʁs]) et les lectures correctes dans LECT^{CE1} (REAL.LECT *les/z/jours* [le.zuʁs])**

Nous avons vu que le nombre de lectures correctes, où les phénomènes de resyllabation sont bien réalisés, corrèle avec le nombre de non-réalisations. Nous cherchons à établir le même rapport avec les erreurs d'insertion de consonne erronée. Nous n'observons pas de corrélation entre les insertions d'une consonne erronée et les REAL.LECT^{CE1} ($r(36) = 0,018, p = 0,954$).

10.5 Les tâches d'amorçage (AMOA et AMOB)

Les tâches d'amorçage (AMOA et AMOB) permettent d'établir un rapport encore plus direct entre l'écrit et l'oral en présence des phénomènes de resyllabation. Avant d'expliquer en détail AMOA et AMOB, nous avons décidé de clarifier la nature d'une tâche d'amorçage et d'en expliquer le fonctionnement cognitif. Dans son ouvrage *Psychologie cognitive de la lecture* (2009 : 38), Ludovic Ferrand nous donne une définition : « la technique d'amorçage [...] consiste à présenter un premier stimulus (appelé « amorce ») suivi par un second stimulus (appelé « cible ») que le sujet doit traiter. La tâche du sujet est de traiter le second mot seulement et il peut s'agir d'une tâche de décision lexicale, de production, d'identification perceptive ou de catégorisation sémantique ».

Pour étudier différents types d'interaction, la cible et l'amorce peuvent être liées visuellement (IFKO-JERQ), orthographiquement (CHAT-CHAR), phonologiquement (FOIS-FOIE), morphologiquement (MURAL-MURET), sémantiquement (PONEY-CHEVAL) ou syntaxiquement (UNE-TABLE). De plus, l'amorce et la cible peuvent être présentées dans la même modalité (ex. visuelle-visuelle) ou dans différentes modalités (ex. acoustique-visuelle). Dans ce dernier cas nous sommes en présence d'une tâche d'amorçage multimodal.

Dans sa structure la tâche d'amorçage peut être associée ou non à un « masquage ». Dans la tâche d'amorçage avec masquage, l'expérimentateur introduit entre amorce et cible une image ou une suite de caractères (comme #####). Le masquage permet d'effacer les effets résiduels de l'amorce qui peuvent avoir des influences sur le traitement de la cible. Pour éviter des effets résiduels, dans les tâches sans masquage, l'amorce est présentée pour une longue durée (supérieure à 150 millisecondes).

Au moyen de ce type de tâche nous pouvons tester l'influence de l'amorce sur la cible. L'amorce peut faciliter ou inhiber le traitement de la cible. Au moyen de ce type de tâche nous pouvons vérifier, par exemple, si la présentation d'une amorce du type TABLE peut faciliter la lecture de la cible CHAISE plus qu'une autre amorce qui n'a rien à voir au niveau sémantique (l'amorce contrôle) comme par exemple SPORT. Si le temps pour lire le mot CHAISE après la présentation de l'amorce TABLE est plus court qu'avec l'amorce SPORT, il y a un effet.

Nous avons proposé deux expériences d'amorçage multimodal sans masquage⁵³, inspirées par Gaskell *et al.* (2002). Pour construire les tâches, nous avons utilisé le logiciel *OpenSesame* (Mathôt *et al.*, 2012). Nous avons testé les 37 enfants du CE1 un par un dans une pièce calme, pendant environ 25 minutes pour chaque tâche. Les tâches sont précédées d'une période d'entraînement qui continue jusqu'à ce que l'enfant comprenne la procédure. L'ordre des deux tâches (AMOA et AMOb) est varié pour éviter les effets de facilitation et de stratégies.

Pour ces tâches, nous avons testé aussi un groupe d'adultes pour avoir un groupe « contrôle ». Les adultes ont été testés un par un entre octobre et novembre 2015. Chez les adultes, l'ordre des deux tâches (AMOA et AMOb) est également varié pour éviter des effets de facilitation ou de stratégies. La durée de chaque tâche est d'environ 15 minutes par sujet. Nous présentons dans la prochaine section le groupe d'adultes testés.

10.5.1 Le groupe d'adultes

Nous avons testé 37 adultes (12 hommes et 25 femmes) entre 18 et 35 ans d'âge ($M = 25;7$, $SD = 3,6$). Pour recueillir des données sociolinguistiques, nous avons proposé aux participants de compléter un questionnaire (Annexe 12). À partir de l'analyse de ces questionnaires, nous pouvons décrire les niveaux scolaires des participants, comme les diplômes obtenus (Tableau 87). Comme on peut le voir, la plupart des adultes ont un niveau d'étude supérieur au BAC + 5 (59,5 % des étudiants).

	% d'adultes
BAC	5,4
Dès BAC + 1 à BAC + 4	35,1
BAC + 5 et plus	59,5

Tableau 87. Classification des diplômes du groupe d'adultes testés

Nous avons classifié les adultes selon les catégories socio-professionnelles que nous avons utilisées pour classier les parents des enfants (voir chapitre 7). Nous avons ajouté à cette classification la catégorie « étudiants » qui comprend les étudiants

⁵³ Nous n'avons pas prévu un masquage car la cible est présentée pour une longue durée sur l'écran (supérieur à 150ms)

en licence, en master et en doctorat. Dans le tableau 88, nous avons marqué le pourcentage pour chaque catégorie de participants. Comme on peut le voir, la plupart des adultes sont des étudiants (56,75 % de participants).

	% de participants
« cadres »	24,32
« intermédiaires »	5,42
« employés »	13,51
« ouvriers »	0
« étudiants »	56,75

Tableau 88. Classification des métiers des adultes testés

En ce qui concerne les données linguistiques, tous les adultes testés sont des francophones natifs scolarisés en France depuis l'école maternelle.

Dans le questionnaire, nous avons aussi posé une question sur la fréquence des activités de lecture et d'écriture pendant la journée pour évaluer leur rapport avec la *literacy*. Tous les participants lisent et écrivent tous les jours et plusieurs fois par jour. Nous observons donc une interaction régulière avec la langue écrite.

10.5.2 Procédure

La procédure des deux tâches d'amorçage (AMOA et AMOB) est équivalente chez les enfants et chez les adultes. AMOA et AMOB ont la même structure. Cette structure est bien exemplifiée dans la figure 39. Au début de la tâche d'amorçage, nous avons décidé d'afficher sur l'écran une image, où nous avons marqué le mot *GO !*. Pour commencer la tâche le sujet doit appuyer sur la touche *espace*. Dès qu'il a appuyé sur *espace*, le *trial* commence. Dans chaque tâche nous comptons 84 *trials* que nous avons divisés en 6 blocs de 14 *trials*. Après chaque bloc de 14 *trials* nous avons décidé de faire afficher sur l'écran une image inspirée par un dessin-animé. Cette image permet au participant (surtout aux enfants) de faire une pause avant de continuer la tâche. Le participant peut regarder l'image pour un temps illimité et quand il se sent prêt à recommencer avec les *trials* suivants, il doit appuyer sur la touche *espace*. Après 3 blocs de *trials* (donc à la moitié de la tâche), nous avons proposé un petit dessin-animé de 30 secondes. Après ce dessin-animé, nous retrouvons l'image du début où il était

marqué *GO !*. Le participant doit encore appuyer sur *espace* et un autre bloc de *trial* commence. À la fin de la tâche nous proposons une image où est marqué *BRAVO !* qui fait comprendre qu'ils ont terminé.

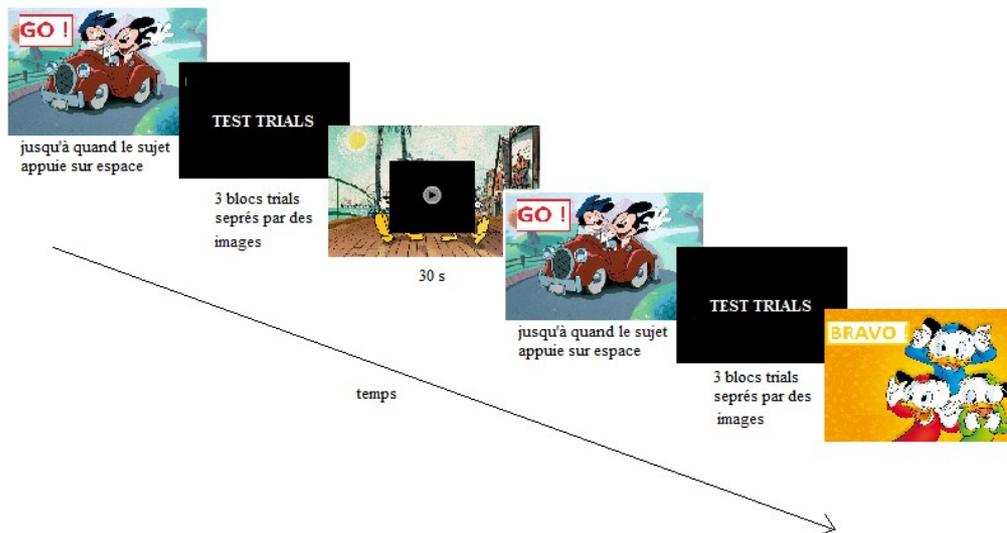


Figure 39. Structures de AMOa et AMOb

Pour AMOa et AMOb, le *trial* est composé de 4 parties (Figure 40) :

1. Fixation

Au centre de l'écran apparaît un point. Les participants sont invités à le regarder. Le point reste affiché 1000 millisecondes sur l'écran.

2. Amorçage acoustique

Un autre point s'affiche sur l'écran. Pendant que le point reste affiché les participants entendent la production d'une séquence sonore. La séquence sonore est l'amorce.

3. Cible visuelle

À la fin de l'input acoustique, le mot écrit cible apparaît sur l'écran. Les participants doivent lire ce mot pour accomplir la tâche. Le mot reste affiché sur l'écran jusqu'à ce que le sujet donne une réponse en appuyant sur un des deux boutons possibles sur le clavier.

4. Réponse

La tâche prévoit deux réponses possibles : « oui » ou « non ». Les participants doivent appuyer sur la touche *m* pour répondre « oui » et *z* pour répondre « non »⁵⁴. Des étiquettes « oui » ou « non » étaient collées aux touches.

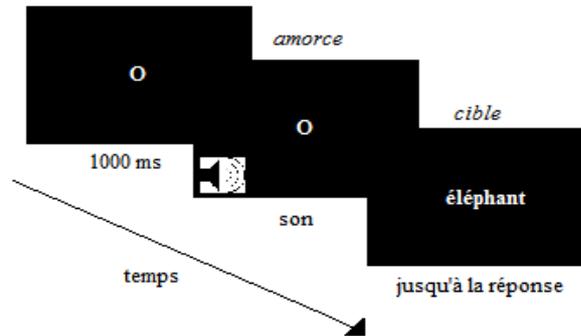


Figure 40. Structure d'un *trial* dans AMOa et AMOb

Le logiciel *OpenSesame* enregistre les temps de réponses (TR) et les réponses (oui ou non). Les TR sont calculés à partir de l'affichage de la cible visuelle sur l'écran jusqu'au moment de la réponse (et donc dès que les participants appuient sur une des deux touches).

Les séquences sonores amorces ont été enregistrées dans une chambre sourde par une locutrice de langue maternelle française avec une fréquence de 44000 Hz et au format *.wav*. Elles sont écoutées à la même fréquence au moyen d'un casque audio *sennheiser hd 202 II* au niveau de volume 16 sur 100.

Les mots écrits sont présentés au centre de l'écran en minuscules, taille 38, police « mono » d'*OpenSesame*. Nous utilisons pour toute la tâche un arrière plan noir et un premier plan blanc.

L'ordre de présentation des *trials* dans les blocs est randomisé par le logiciel pour éviter des processus de facilitation ou de stratégies.

La procédure que nous venons de décrire est commune à AMOa et AMOb. Cependant les deux tâches testent deux compétences différentes. AMOa est une tâche de détection des mots (*word detection task*). Les participants doivent donc entendre ce

⁵⁴ Le clavier utilisé était un clavier italien du type *qwerty*. Sur ce type de clavier, la touche *z* est la plus à droite et la touche *m* est la plus à gauche parmi les touches qui identifient les lettres.

qui était prononcé dans la phrase amorce et dire si le mot cible affiché sur l'écran était présent dans la phrase. S'il était présent, les participants doivent cliquer sur « oui ». S'il n'était pas présent, ils doivent cliquer sur « non ». En revanche, dans AMOb, nous sommes en présence d'une tâche de jugement orthographique où, après avoir écouté un amorçage acoustique, les participants doivent dire si le mot cible qui s'affiche sur l'écran est bien orthographié. Si le mot cible est bien orthographié, les participants doivent cliquer sur « oui ». Si le mot cible n'est pas bien orthographié, ils doivent cliquer sur « non ».

Nous allons d'abord décrire le répertoire et les résultats d'AMOA pour ensuite introduire la tâche AMOb.

10.5.3 La tâche d'amorçage AMOA

10.5.3.1 Répertoire pour AMOA

Comme nous l'avons déjà dit, AMOA est une tâche de détection des mots (*word detection task*). L'amorce de cette tâche est un syntagme nominal déterminant + adjectif + nom. La cible est un nom que nous proposons dans sa forme écrite canonique.

Nous avons choisi les noms-cibles dans la base de données de l'écrit des enfants que nous avons créée au moyen des cahiers (voir § 7.2.2). De plus, nous avons vérifié la fréquence en MANULEX et nous avons sélectionné seulement les noms ayant une fréquence SFI moyenne au CE1 (voir § 7.2.1). Nous avons créé ensuite les séquences d'amorces déterminant + adjectif + nom. La probabilité de ces séquences a été vérifiée par des locuteurs de langue maternelle française non impliqués dans l'étude.

La tâche prévoit 3 conditions d'amorçage :

- Condition 1 : liaison (amorce : *un petit éléphant* [ɛ̃.pə.ti.te.le.fã] - cible : ELEPHANT⁵⁵). L'adjectif est resyllabifié dans le nom. De cette façon, les frontières syllabiques ne correspondent pas aux frontières lexicales. La séquence d'amorçage est donc *un petit éléphant* [ɛ̃.pə.ti.te.le.fã] et la cible visuelle est ELEPHANT. La réponse attendue par les enfants est « oui » car la cible est contenue dans la séquence d'amorçage.

55 Nous avons utilisé l'annotation habituellement utilisée pour la cible dans les tâches de *priming*. Cependant elle est toujours orthographiée en minuscule.

- Condition 2 : alignée (amorce : *un gentil éléphant* [ẽ.ʒã.ti.e.le.fã] - cible : ELEPHANT). L'adjectif n'est pas resyllabifié dans le nom et les frontières du nom correspondent à ses frontières syllabiques. Un exemple est la séquence *un gentil éléphant* [ẽ.ʒã.ti.e.le.fã]. Les enfants doivent reconnaître la cible visuelle ELEPHANT. La réponse attendue par les enfants est « oui », car dans ce cas la cible est aussi contenue dans la séquence acoustique.
- Condition 3 : contrôle (amorce : *un grand arôme* [ẽ.grã.ta.ʁom] - cible : ELEPHANT). Dans cette condition l'amorce n'a aucun rapport avec la cible. Par exemple nous proposons une amorce comme *un grand arôme* et une cible comme ELEPHANT. Les enfants doivent répondre « non », car le mot-cible visuel n'est pas contenu dans la séquence acoustique d'amorçage.

Nous comptons 84 amorces totales, soit 42 séquences tests et 42 séquences *fillers*. Dans les séquences tests, nous avons 4 types de liaisons : en [z], en [t], en [n] et avec *h* muet à l'initiale. Ces liaisons caractérisent non seulement les séquences dans la condition en liaison mais aussi les amorces dans la condition de contrôle (ex. *un grand[t]arôme*). Pour les *fillers*, nous avons proposé des noms à consonne initiale ou à *h* aspiré initial, où la liaison n'est donc pas possible (Annexe 13). Nous présentons dans le tableau 89 les amorces et les cibles (tests).

CIBLE	Amorçage : conditions		
	1 : liaison	2 : alignée	3 : contrôle
ELEPHANT	un petit éléphant [ɛ̃.pə.ti.te.le.fã]	un gentil éléphant [ɛ̃.ʒã.ti.e.le.fã]	un grand arôme
OURSON	un petit ourson [ɛ̃.pə.ti.tuʁ.sõ]	un gentil ourson [ɛ̃.ʒã.ti.tuʁ.sõ]	un grand appartement
ESCARGOT	un petit escargot [ɛ̃.pə.ti.tes.kaʁ.go]	un gentil escargot [ɛ̃.ʒã.ti.es.kaʁ.go]	un excellent apéritif
ENFANT	un petit enfant [ɛ̃.pə.ti.tã.fã]	un gentil enfant [ɛ̃.ʒã.ti.ã.fã]	un excellent employé
AGNEAU	un gros agneau [ɛ̃.gʁo.za.põ]	un mini agneau [ɛ̃.mi.ni.a.põ]	un mauvais élève
ASPIRATEUR	un gros aspirateur [ɛ̃.gʁo.zas.pi.ka.tœʁ]	un mini aspirateur [ɛ̃.mi.ni.as.pi.ka.tœʁ]	un mauvais abri
OISEAU	un gros oiseau [ɛ̃.gʁo.zwa.zõ]	un mini oiseau [ɛ̃.mi.ni.wa.zõ]	un curieux animal
ANANAS	un gros ananas [ɛ̃.gʁo.za.na.na]	un mini ananas [ɛ̃.mi.ni.a.na.na]	un curieux appareil
ORAGE	aucun orage [o.kœ.nõ.ʁaʒ]	un vrai orage [ɛ̃.vʁe.õ.ʁaʒ]	un ancien ouvrier
ARBRE	aucun arbre [o.kɛ.naʁbʁ]	un vrai arbre [ɛ̃.vʁe.aʁbʁ]	un ancien avion
AMI	aucun ami [o.kɛ.na.mi]	un vrai ami [ɛ̃.vʁe.a.mi]	un commun accord
IGLOO	aucun igloo [o.kɛ.ni.glu]	un vrai igloo [ɛ̃.vʁe.i.glu]	un commun espace
HIPPOPOTAME	un gros hippopotame [ɛ̃.gʁo.zi.po.po.tam]	un mini hippopotame [ɛ̃.mi.ni.i.po.po.tam]	un affreux accident
HOMME	un grand homme [ɛ̃.gʁã.tõm]	un vrai homme [ɛ̃.vʁe.õm]	un petit abricot

Tableau 89. Répertoire pour AMOa

10.5.3.2 Les hypothèses pour AMOa

Chez les enfants

Pour les conditions « liaison » et « alignement », nous nous attendions à des TR différents. Vu qu'un input en liaison est plus difficile à segmenter à cause de la resyllabation entre Mot1 et Mot2, nous pensons que reconnaître la cible écrite après la séquence acoustique est plus difficile dans la condition de liaison. En revanche reconnaître la cible après une amorce alignée va être plus facile.

TR : *un petit éléphant* – ELEPHANT > *un gentil éléphant* – ELEPHANT

Ces résultats sont très différents de ceux trouvés par Wauquier-Gravelines (1996) avec une tâche de reconnaissance lexicale ou Gaskell *et al.* (2002) avec une tâche de *cross-modal priming* chez les adultes. Gaskell *et al.* et Wauquier-Gravelines montrent qu'il n'y a pas de TR différents entre les deux conditions. Nous pensons que les TR ne changent pas car les adultes ont déjà un lexique orthographique. Ce lexique orthographique est activé en présence d'un input acoustique, comme nous avons vu dans le § 2.3.2. Les formes orthographiques correctes de Mot1 et de Mot2 activées par la séquence orale aident à reconnaître tout de suite la cible. Les connaissances de *literacy* peuvent donc jouer un rôle sur les TR. Les enfants au CE1, qui ont moins de connaissances de l'orthographe que celles des adultes, peuvent avoir en présence d'un input non-aligné plus de difficultés à reconnaître la cible car ils utilisent encore les informations de la langue orale pour segmenter. Notamment au début de l'acquisition de l'écrit, la syllabe est l'unité sur laquelle les enfants s'appuient pour reconnaître les mots orthographiés. Pour cette raison, nous nous attendons à des TR plus courts après une amorce alignée, car dans cette condition le découpage syllabique à l'oral correspond au découpage syllabique à l'écrit. Étant donné que dans la condition d'amorçage en liaison le découpage lexical à l'oral ne correspond pas au découpage à l'écrit, nous nous attendons à des TR plus longs. Pour la même raison, nous nous attendons un taux de réponses correctes inférieur pour la condition « liaison ».

Pour la condition de contrôle, nous nous attendons à des TR pour la réponse « non » plus élevés que les TR pour la réponse « oui » des deux autres conditions. Normalement, répondre « non » dans une tâche d'amorçage implique des temps plus longs (Brouillet & Syssau, 2005). De plus, les TR plus élevés pour la condition contrôle peuvent être dus au fait que l'enfant pense d'abord devoir répondre « oui », vu que les amorces dans les deux autres conditions sont plus nombreuses. Le facteur d'entraînement causerait un ralentissement des réponses négatives.

Chez les adultes

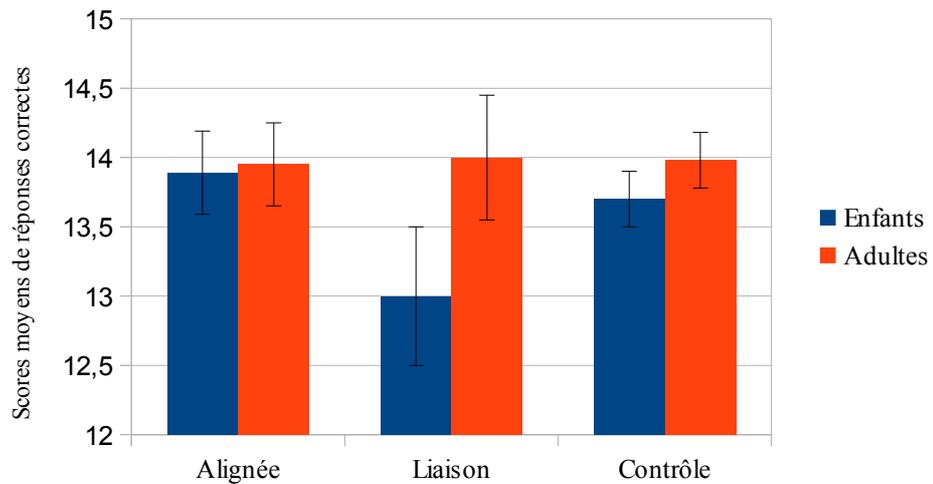
Dans AMOa chez les adultes, nous nous attendions à trouver les mêmes résultats que ceux observés par Wauquier-Gravelines (1996) et Gaskell *et al.* (2002) c'est-à-dire qu'ils répondent correctement dans les trois conditions et qu'ils ont des TR équivalents pour la condition « liaison » et pour la condition « alignée ».

On s'attend à ce que le TR ne change pas selon la condition d'amorçage car les adultes connaissent les formes orthographiques de Mot1 et Mot2. Ces formes sont activées par la séquence orale d'amorçage et empêchent que l'amorce en liaison ralentisse les TR. Pour la condition « contrôle », nous nous attendions à des TR plus longs car elle implique des réponses négatives qui, comme nous l'avons déjà dit dans le paragraphe précédent, suscitent des ralentissements à la réponse.

10.5.3.3 Les résultats de la tâche AMOa

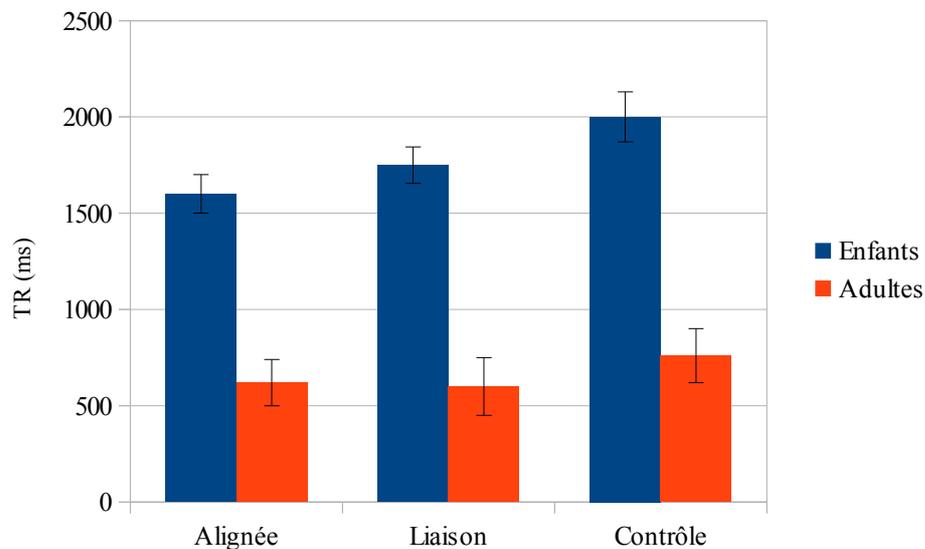
Nous avons calculé la moyenne totale des TR par condition. Nous avons ensuite écarté des résultats les performances de 7 enfants (18,2 %) et de 7 adultes (18,2 %) ayant des TR inférieurs ou supérieurs à 3 écarts-types à la moyenne. Les variables dépendantes traitées sont les TR moyens par condition et les scores des réponses correctes. Nous considérons la réponse comme correcte quand l'enfant reconnaît bien la cible après la présentation de l'amorce, donc quand il répond « non » dans la condition « contrôle » et « oui » dans les deux autres conditions. Les TR sont calculés pour toutes les conditions à partir de l'apparition de la cible écrite jusqu'à la réponse.

Nous avons réalisé une analyse ANOVA à mesures répétées utilisant les sujets et les items comme variables aléatoires. Nous avons testé l'interaction entre deux facteurs : l'âge et la condition. L'interaction « âge*condition » est statistiquement significative ($F_1(2,116) = 13,071$, $p < .001$, $F_2(2, 52) = 6,78$, $p < .05$). Il y a un effet principal du facteur « âge » ($F_1(1, 58) = 7,490$, $p < .001$, $F_2(1, 26) = 3,34$, $p < .05$) mais il n'y a pas d'effet principal pour le facteur « condition » ($F_1(2, 116) = 3,438$, $p = .035$, $F_2(2, 52) = .78$, $p = .55$). Comme nous pouvons le voir dans le graphique 22, les enfants ont des scores de réponses correctes plus bas dans la condition « liaison » comparée aux autres conditions et aux scores des adultes.



Graphique 22. Réponses correctes dans AMOa pour les trois conditions chez les enfants et les adultes

Nous avons analysé les TR. L'analyse nous montre qu'il y a un effet principal du facteur « âge » ($F_1(1, 58) = 229,341, p < .001, F_2(1, 26) = 504,765, p < .001$), pour le facteur « condition » ($F_1(2, 116) = 34,441, p < .001, F_2(2, 52) = 24,400, p < .001$) et un effet de l'interaction « âge*condition » ($F_1(2,116) = 13,071, p < .001, F_2(2, 52) = 9,314, p < .001$). Comme nous pouvons le voir dans le graphique 23, les enfants ont des TR plus longs que pour toutes les conditions que les adultes. Les enfants ont également des TR différents pour les 3 conditions.



Graphique 23. TR dans AMOa pour les trois conditions chez les enfants et les adultes

Chez les enfants, nous observons l'effet du facteur « condition » ($F_1(3,27) = 228,813$, $p < .001$, $F_2(3,10) = 356,419$, $p < .001$). Cependant, les 3 conditions n'impliquent pas le même type de réponses : dans le cas des conditions « alignée » et « liaison », les enfants doivent répondre « oui » tandis que dans la condition « contrôle » ils doivent répondre « non ». Comme nous l'avons déjà dit, répondre « non » implique des TR plus longs. Les deux types de réponses (« oui » et « non ») sont différents et par conséquent les TR ne sont pas comparables. Nous pouvons quand même comparer les TR dans les conditions « alignée » et « liaison » pour vérifier s'il y a une différence significative en cas de réponse affirmative. Pour évaluer cette différence, nous avons réalisé une analyse des contrastes. Les résultats montrent qu'il y a une différence statistiquement significative pour les TR dans la condition « liaison » et « alignée » par enfants ($F_1(1,29) = 5,64$, $p < .05$) mais pas par items ($F_2(1,13) = 1,29$, $p = .26$). Comme nous pouvons le voir dans le graphique 23, après la présentation d'un input acoustique en condition de liaison, reconnaître la cible visuelle qui est le Mot2 de cette séquence implique un TR plus long que de reconnaître une cible présentée dans une séquence acoustique alignée :

TR : *un petit éléphant* – ELEPHANT > *un gentil éléphant* – ELEPHANT

Chez les adultes, il n'y a pas de différence statistiquement significative entre les TR dans ces deux conditions ($F_1(1,31) = 3,173$, $p = 0,845$; $F_2(1,12) = 1,496$, $p = .244$). Nous pouvons donc conclure que, pour les adultes, détecter un mot-cible dans une séquence en liaison n'implique pas un coût cognitif plus élevé que dans la condition « alignée » :

TR : *un petit éléphant* – ELEPHANT = *un gentil éléphant* – ELEPHANT

10.5.3.4 Discussion sur les résultats de AMOa

Nous observons une différence significative pour le groupe d'enfants sur les TR. Chez les enfants, reconnaître le mot ELEPHANT après avoir écouté la séquence

un petit éléphant implique des TR plus longs que de le détecter dans la séquence *un gentil éléphant*. En revanche, chez les adultes, nous n'enregistrons aucune différence sur les TR pour les deux conditions.

Les résultats montrent que la perception d'un input en liaison et un input aligné change la vitesse du traitement d'une cible visuelle chez les enfants. Nous pouvons donc dire qu'à la différence des adultes, le processus cognitif derrière le découpage en mots en cas de liaison implique des coûts supplémentaires à ce stade de l'acquisition. La resyllabation de l'input acoustique cause des coûts de traitement dans le processus de reconnaissance des frontières de mots au CE1.

Les séquences d'amorçage activent les formes écrites de Mot1 et de Mot2. L'activation de ces formes permet aux adultes de reconnaître avec la même facilité le mot ELEPHANT dans les deux conditions. Les enfants ont encore besoin de passer par la séquence orale pour reconnaître les mots, car ils sont encore en train de construire leurs connaissances orthographiques et utilisent une stratégie phonologique pour découper au moyen de l'unité syllabique. Les enfants réussissent à reconnaître de façon plus rapide un input écrit qui a le même découpage syllabique à l'oral et où la resyllabation ne perturbe pas les frontières lexicales. En revanche, reconnaître un input écrit qui est présenté dans une séquence resyllabifiée est un processus plus difficile et implique un traitement cognitif plus long car il n'y a pas de correspondance entre les frontières syllabiques et les frontières lexicales. À ce stade, ce sont les frontières syllabiques qui sont favorisées au moment du découpage au niveau cognitif. Pour cette raison, une amorce en liaison cause un ralentissement dans la reconnaissance du mot-cible à l'écrit.

Pour confirmer le rôle de la syllabe et de l'importance d'un input aligné avec cette unité pour la reconnaissance des mots visuels aux premiers stades de l'acquisition de la *literacy*, nous avons décidé de proposer une deuxième tâche d'amorçage. Nous allons présenter cette tâche dans les prochains paragraphes.

10.5.4 La tâche d'AMOb

La tâche AMOb a la même procédure qu'AMOa (§ 10.5.2). Pour des questions de temps, nous n'avons testé dans cette tâche que 22 enfants sur 37. Pour le groupe d'adultes, nous avons testé les 37 participants.

10.5.4.1 Le répertoire pour la tâche AMOb

La tâche AMOb est une tâche de décision orthographique. Après l'écoute d'une phrase amorce, un mot apparaît au milieu de l'écran. Les participants doivent dire si ce mot est bien orthographié ou pas. Nous avons toujours trois conditions d'amorçage : « liaison », « alignée » et « contrôle ».

- Condition 1 : liaison (amorce : *un petit éléphant* [ɛ̃.pə.ti.te.le.fɑ̃] - cible : TELEPHANT). L'adjectif est resyllabifié dans le nom. La cible visuelle est un calque de l'amorce car nous présentons un mot qui commence par la CL [t] : TELEPHANT. La cible proposée est un mot qui n'est pas bien orthographié. La réponse que nous attendons des enfants est « non ».
- Condition 2 : alignée (amorce : *un gentil éléphant* [ɛ̃.ʒɑ̃.ti.e.le.fɑ̃] - cible : TELEPHANT). L'adjectif n'est pas resyllabifié dans le nom et les frontières du nom correspondent à ses frontières syllabiques. Nous proposons comme cible un mot écrit où l'attaque est remplie par une CL possible. La cible est mal orthographiée, la réponse attendue est donc « non ».
- Condition 3 : contrôle (amorce : *un grand arôme* [ɛ̃.gʁɑ̃.ta.ʁom] - cible : TELEPHANT). Dans cette condition l'amorce n'a aucun rapport avec la cible visuelle, qui de surcroît n'est pas orthographiée correctement. Les enfants doivent répondre « non ». Cette condition permet de vérifier l'effet d'amorçage des autres conditions de la tâche.

Nous avons toujours 84 *trials*, 42 *fillers* (Annexe 14) et 42 tests (Tableau 90). L'amorce est la même que celle que nous avons présentée pour la tâche AMOa. Les cibles sont différentes. Nous pouvons voir que, pour les cibles tests, la forme visuelle du mot n'est pas correctement orthographiée (TELEPHANT), car elle commence par une possible consonne de liaison tandis que dans le lexique orthographique elle est normalement à voyelle initiale (*éléphant*).

CIBLE	Amorçage : conditions		
	1 : liaison	2 : alignée	3 : contrôle
TELEPHANT	un petit éléphant [ɛ̃.pə.ti.te.le.fɑ̃]	un gentil éléphant [ɛ̃.ʒɑ̃.ti.e.le.fɑ̃]	un grand arôme
TOURSON	un petit ourson [ɛ̃.pə.ti.tuʁ.sɔ̃]	un gentil ourson [ɛ̃.ʒɑ̃.ti.tuʁ.sɔ̃]	un grand appartement
TESCARGOT	un petit escargot [ɛ̃.pə.ti.tɛs.kɑʁ.go]	un gentil escargot [ɛ̃.ʒɑ̃.ti.ɛs.kɑʁ.go]	un excellent apéritif
TENFANT	un petit enfant [ɛ̃.pə.ti.tɑ̃.fɑ̃]	un gentil enfant [ɛ̃.ʒɑ̃.ti.ɑ̃.fɑ̃]	un excellent employé
ZAGNEAU	un gros agneau [ɛ̃.gʁo.za.ɲo]	un mini agneau [ɛ̃.mi.ni.a.ɲo]	un mauvais élève
ZASPIRATEUR	un gros aspirateur [ɛ̃.gʁo.zas.pi.ʁa.tœʁ]	un mini aspirateur [ɛ̃.mi.ni.as.pi.ʁa.tœʁ]	un mauvais abri
ZOISEAU	un gros oiseau [ɛ̃.gʁo.zwa.zo]	un mini oiseau [ɛ̃.mi.ni.wa.zo]	un curieux animal
ZANANAS	un gros ananas [ɛ̃.gʁo.za.na.na]	un mini ananas [ɛ̃.mi.ni.a.na.na]	un curieux appareil
NORAGE	aucun orage [o.kɛ.nɔ.ʁaʒ]	un vrai orage [ɛ̃.vʁɛ.ɔ.ʁaʒ]	un ancien ouvrier
NARBRE	aucun arbre [o.kɛ.naʁbʁ]	un vrai arbre [ɛ̃.vʁɛ.aʁbʁ]	un ancien avion
NAMI	aucun ami [o.kɛ.na.mi]	un vrai ami [ɛ̃.vʁɛ.a.mi]	un commun accord
NIGLOO	aucun igloo [o.kɛ.ni.glu]	un vrai igloo [ɛ̃.vʁɛ.i.glu]	un commun espace
ZHIPPOPOTAME	un gros hippopotame [ɛ̃.gʁo.zi.po.po.tam]	un mini hippopotame [ɛ̃.mi.ni.i.po.po.tam]	un affreux accident
THOMME	un grand homme [ɛ̃.gʁɑ̃.tɔ̃m]	un vrai homme [ɛ̃.vʁɛ.ɔ̃m]	un petit abricot

Tableau 90. Répertoire pour AMOb

10.5.4.2 Les hypothèses pour AMOb

Chez les enfants

En ce qui concerne les réponses correctes, nous nous attendions à ce que les enfants n'aient pas de difficultés à répondre correctement, car ils savent déjà que les mots présentés sont des noms à voyelle initiale. Comme nous l'avons vu pour la tâche PROD.ECR^{CE1}, au CE1, les enfants segmentent correctement les séquences en liaison et identifient sans problèmes la frontière gauche du Mot2 à l'écrit.

Pour les TR, nous nous attendions pour la condition « contrôle » à des TR plus longs et à plus d'erreurs, car refuser la cible après une amorce qui n'a aucun rapport avec elle doit être plus facile.

Ensuite, nous pouvons faire des hypothèses sur les deux autres conditions. Nous avons observé dans la tâche AMOa que la congruence entre le découpage syllabique de l'amorce et de la cible influence le TR. Nous nous attendions à ce qu'une correspondance entre les frontières orales et écrites cause des difficultés dans le refus d'un mot mal orthographié comme TELEPHANT. Nous pensons trouver de TR plus longs dans la condition « liaison » que dans la condition « alignée ».

TR : *un petit éléphant* – TELEPHANT > *un grand arôme* – TELEPHANT

Chez les adultes

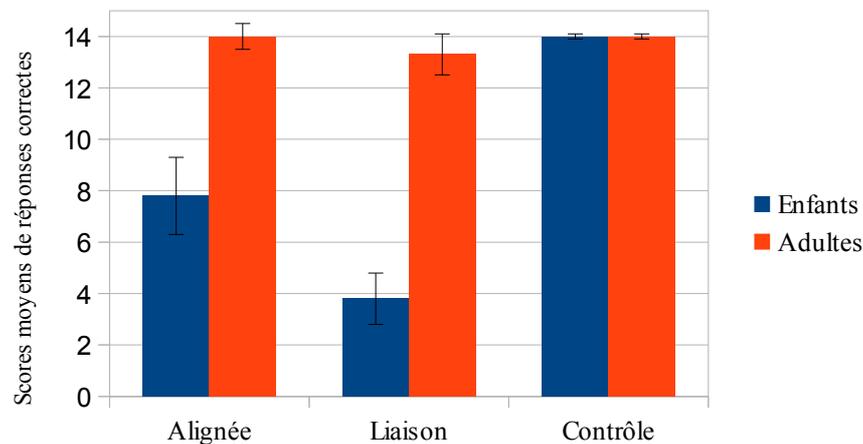
Pour le groupe d'adultes, nous nous attendons à ne pas enregistrer d'erreurs dans toutes les conditions car les adultes connaissent très bien la forme orthographique des mots. En ce qui concerne les TR, nous nous attendons à avoir les mêmes TR dans les trois conditions, sans enregistrer d'effet de *priming*. Les séquences orales d'amorçage activent la forme orthographique correcte du Mot2. Les adultes ne doivent pas découper la séquence orale au moyen d'une stratégie phonologique pour repérer les frontières des mots et juger la cible écrite. Pour ces raisons :

TR : *un gentil éléphant* – TELEPHANT = *un petit éléphant* – TELEPHANT = *un grand arôme* – TELEPHANT

10.5.4.3 Les résultats de la tâche AMOb

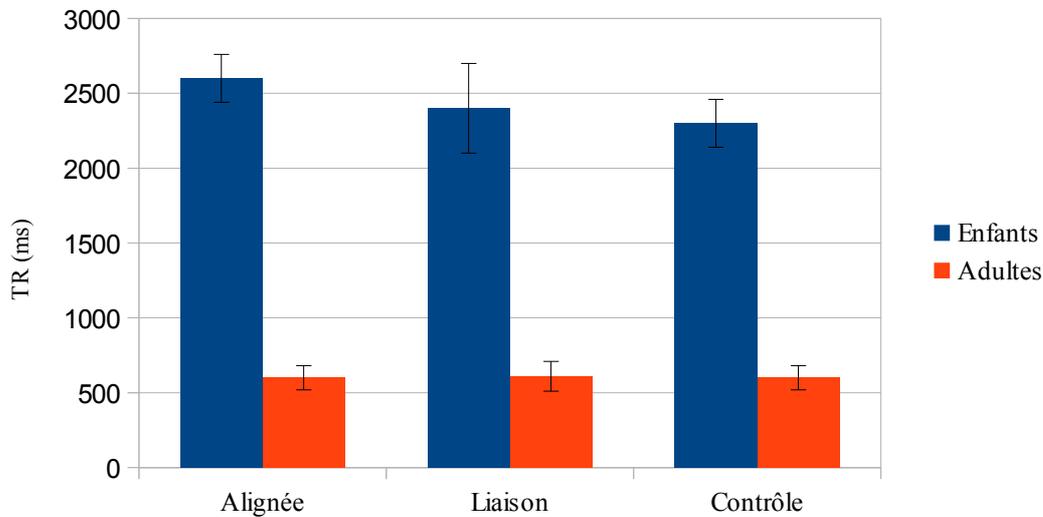
Avant d'analyser les résultats, nous avons écarté les données de 10 enfants (37 %) et 5 adultes (18 %) qui ont enregistré des TR trop courts (trois fois inférieur à l'écart-type) ou trop longs (trois fois supérieur à l'écart-type). Comme nous l'avons déjà fait pour AMOa, nous avons conduit une ANOVA à mesures répétées sur les variables dépendantes TR et scores de réponses correctes utilisant les sujets et les items comme variables aléatoires. Nous avons évalué le rôle des deux facteurs : l'âge et la condition.

Pour les scores de réponses correctes nous observons un effet principal du facteur « âge » ($F_1(1,46) = 177,100$, $p < .001$, $F_2(1,26) = 16,54$, $p < .05$), du facteur « condition » ($F_1(2,92) = 88,849$, $p < .001$, $F_2(2,52) = 3,27$, $p < .05$) et une interaction significative entre ces deux facteurs ($F_1(2,92) = 13,071$, $p < .001$, $F_2(2, 52) = 13,27$, $p < .05$). Comme nous pouvons le voir dans le graphique 24, l'âge est un facteur qui influence la production des réponses correctes : les enfants ont des scores de réponses correctes plus bas que les adultes. De plus, chez les enfants, la condition d'amorçage est un facteur qui influence la certitude des réponses : les trois conditions ont des scores différents et la condition « liaison » est celle qui cause le plus de problèmes.



Graphique 24. Réponses correctes à AMOb dans les trois conditions chez les enfants et les adultes

Pour les TR, l'analyse statistique de la variance (ANOVA) révèle qu'il y a un effet principal du facteur « âge » ($F_1(1,46) = 2,555$, $p < .001$, $F_2(1,26) = 623,678$, $p < .001$) mais pas d'effet pour le facteur « condition » ($F_1(2,92) = 2,305$, $p = .108$, $F_2(2,52) = 2,431$, $p = .117$) et pas d'interaction significative entre l'âge et la condition ($F_1(2,92) = 1,702$, $p = .107$, $F_2(2, 52) = 2,365$, $p = .105$). L'âge est un facteur qui influence les TR car les enfants ont, en général, des TR plus longs que les adultes (Graphique 25).



Graphique 25. TR dans AMOb pour les trois conditions chez les enfants et chez les adultes

Étant donné le grand nombre d'erreurs des enfants dans cette tâche, nous pourrions penser qu'ils ont été influencés par l'ordre de présentation des tâches. AMOa et AMOb sont deux tâches différentes qui demandent aux participants deux réponses différentes : dans AMOa, ils doivent juger si le mot cible était dans la phrase amorce ; dans AMOb, ils doivent juger l'orthographe de la cible sans être influencés par l'amorce. La présentation de AMOb après AMOa peut être source de certains problèmes. Les participants peuvent avoir pensé que dans AMOb, comme dans AMOa, il fallait encore prendre en considération la phrase amorce. Comme nous l'avons dit dans la section 10.5.2, nous avons présenté à la moitié des adultes et à la moitié des enfants les deux tâches dans un certain ordre. Nous avons inversé l'ordre de présentation pour l'autre moitié des participants. La procédure expérimentale que nous avons élaborée permet d'exclure l'influence de l'ordre de présentation des tâches sur les résultats.

10.5.4.4 Discussion des résultats de AMOb

Chez les enfants, dire que le mot TELEPHANT est mal orthographié après avoir entendu la suite *un petit[t]éléphant* est plus difficile que de le dire après la séquence acoustique *un grand arôme* ou *un gentil éléphant*. Ces résultats peuvent être

dus au fait que la forme phonologique [te.le.fã] de l'amorce en liaison active la forme lexicale TELEPHANT et donc peuvent signifier que refuser cette forme incorrecte est plus difficile. Nous pouvons penser que les enfants utilisent encore un découpage syllabique pour déterminer les frontières et également pour repérer les représentations écrites des Mots2. Une forme visuelle, où les frontières syllabiques correspondent aux frontières orthographiques, est jugée comme correcte même si elle est mal orthographiée.

Contrairement aux enfants, chez les adultes, nous ne remarquons pas d'influence de l'amorce. Les adultes ne sont pas influencés par l'amorce acoustique quand il leur est demandé de juger l'orthographe d'un mot à l'écrit. Cela implique qu'ils n'utilisent pas la phonologie dans AMOb mais seulement leurs compétences de langue écrite.

Ces données conduisent vers l'interprétation suivante : chez les enfants à ce stade de l'acquisition, l'accès aux mots écrits est plutôt guidé par l'unité syllabique de l'oral. En effet, elles montrent qu'à la différence des adultes, les enfants ne sont pas encore capables de se servir seulement des connaissances orthographiques pour accomplir la tâche et ils utilisent encore la phonologie dont ils ont besoin pour accéder au lexique et pour le jugement orthographique des mots à l'écrit.

10.6 Le rapport entre les résultats des tâches au CE1 et la fréquence des activités de lecture partagée

Dans cette section, nous mettons en rapport la fréquence des lectures partagées entre enfants et adultes avec les scores de productions écrites correctes (PROD.ECR^{CE1}), de lectures correctes (LECT^{CE1}) et de dénominations orales correctes (DENA^{CE1}). L'hypothèse initiale est que les enfants qui ont à la maison plus de relations avec la langue écrite ont aussi de meilleures performances en lecture, en écriture et en production orale, car ils connaissent mieux les frontières des mots à l'écrit. Nous considérons la classification de la fréquence des lectures partagées que nous avons présentée dans la section 8.6 de 28 enfants sur 37.

Pour PROD.ECR^{CE1}, le calcul de la valeur r de Pearson montre qu'il n'y a pas de corrélation entre les scores de segmentations correctes à l'écrit (productions écrites à « frontières lexicales ») et la fréquence des lectures partagées ($r(27) = .02$, $p = .91$).

Nous avons aussi essayé d'établir une corrélation entre les lectures correctes (REAL.LECT^{CE1}) et la fréquence des lectures partagées. Le calcul statistique montre qu'il n'y a pas de relation entre les scores de lectures correctes et la fréquence de lectures partagées ($r(27) = .17$, $p = .38$).

Pour la tâche DENa^{CE1}, l'analyse statistique des productions orales correctes en relation avec la fréquence de lecture partagée montre qu'il n'y a pas de corrélation entre les deux ($r(27) = .24$, $p = .25$).

Pour conclure, nous pouvons dire qu'au CE1, la fréquence des moments de lecture partagée ne semble pas jouer de rôle dans la production orale, écrite et dans la lecture en présence des phénomènes consonantiques.

10.7 Discussion et conclusions

10.7.1 Acquisition de la langue orale et *literacy*

Contrairement aux résultats présentés dans la littérature (Basset, 2000 ; Dugua & Baclesse, 2014), nous observons encore des non-réalisations et des insertions de consonnes erronées chez les enfants au CE1. Comme au CP, les non-réalisations sont les plus fréquentes. Ces résultats sont en accord avec Chevrot *et al.* (2013) qui remarquent la stabilité de non-réalisations après l'âge de 5 ans.

En ce qui concerne la *literacy*, nous pouvons voir que les compétences de la langue écrite s'améliorent pendant le CE1 : les enfants identifient mieux les frontières des mots en présence de phénomènes de resyllabation à l'oral et plus de 95 % des séquences sont lues correctement.

Nous observons aussi qu'au CE1, les enfants utilisent de plus en plus une stratégie lexicale pour lire et écrire. Ils orthographient plus fréquemment les <s> du pluriel à la fin des Mots2, même quand ils n'ont pas de correspondance à l'oral, ou les <h> muets au début du Mot2. De plus, la fréquence a un effet sur la lecture : les items

les mieux lus sont aussi les plus fréquents. Cependant, la stratégie phonologique a encore sa place dans le processus d'écriture et de lecture. Les élèves du CE1 réussissent à mieux représenter les frontières des mots à l'écrit mais ont encore des difficultés à bien orthographier les items. Nous enregistrons des erreurs dues au passage de l'oral à l'écrit comme, par exemple, des productions écrites <otruche> pour *autruche* ou <elefan> pour *éléphant*. Pour la lecture, nous remarquons encore l'utilisation de l'unité syllabique dans PRA.LECT (*les-[z]affiches* [le/za.fi]), où la CL est produite comme si elle était la consonne lexicale initiale de Mot2.

En ce qui concerne le marquage du pluriel, orthographier le <s> du pluriel à la fin des Mots1 et Mots2 n'implique pas la production correcte des liaisons à l'oral. Nos résultats confirment les conclusions de Mesquida *et al.* (2010) pour le CE1.

Dans les hypothèses présentées au chapitre 6, nous assumons que les productions des consonnes erronées diminuent en nombre à cause de l'acquisition de la *literacy* (H1). Or, nos données infirment cette hypothèse : nous n'observons pas de lien entre les compétences de la langue écrite (lectures et productions écrites) et les insertions d'une consonne erronée à l'oral. Une autre hypothèse présentée (H2) affirme que les non-réalisations à l'oral seront favorisées par l'acquisition de la langue écrite. Au CE1, nous observons exactement le contraire : à des compétences meilleures d'écriture et de lecture correspond un nombre plus bas de non-réalisations. Contrairement à ce que nous avons pensé au début de l'étude, l'apprentissage de la *literacy* permet de résoudre les erreurs de non-réalisations à l'oral. Apprendre l'orthographe aide à bien fixer les frontières des mots impliqué dans les phénomènes de resyllabation. Les activités de lecture aident à enchaîner les mots et à resyllabifier sur Mot2 la consonne qui appartient graphiquement à Mot1.

Même si les activités de lecture et d'écriture influencent les productions orales, nous ne pouvons pas établir de rapport entre elles au CE1. Plus encore qu'au CP, la lecture et l'écriture se différencient. Au CE1, mieux lire ne veut pas dire mieux écrire et vice-versa.

Pour comprendre les évolutions possibles après le CE1 dans l'acquisition de phénomènes de resyllabation, nous proposons dans la section suivante une comparaison entre nos résultats et les données de Dugua et Baclesse (2014).

10.7.2 Après le CE1 : une comparaison entre nos résultats et les données de Dugua et Baclesse (2014)

Dugua et Baclesse (2014) envisagent de donner un panorama sur la lecture de séquences en contexte de liaison chez les enfants de CE1 et CM1 (âge moyen : 9;2). Au moyen de trois tâches différentes, les deux chercheurs vérifient les compétences de liaisons obligatoires et facultatives chez 110 enfants francophones. Les 3 tâches proposées sont :

1. Tâche de dénomination guidée.

L'expérimentateur propose une carte et il la nomme en produisant la séquence au singulier *un[n]ours* et l'enfant doit produire la séquence au pluriel *des[z]ours* (car dans la carte il y a deux animaux).

2. Tâche de jugement normatif.

L'enfant doit juger des paires de séquences (sur deux cartes) : une où la liaison est réalisée, l'autre où elle n'est pas réalisée.

3. Tâche de lecture à voix haute.

L'enfant doit lire des phrases à voix haute.

Les expériences de Dugua et Baclesse testent la liaison obligatoire-catégorique et la liaison facultative. Nous présentons seulement les résultats pour la liaison obligatoire qui sont plus relatifs à notre recherche.

Les résultats pour la tâche de dénomination guidée montrent que les enfants réalisent bien la liaison dans 98 % des cas. Les erreurs de consonnes erronées (*un[z]ours* pour *un[n]ours*) arrivent dans seulement 0,15 % des cas et les non-réalisations subissent une chute : 1,85 % des cas où la liaison n'est pas réalisée. D'après ces données, nous pouvons comprendre que les enfants à cet âge maîtrisent la liaison obligatoire dans le contexte déterminant-nom en dénomination.

En ce qui concerne le jugement normatif, les enfants choisissent l'alternative où la liaison obligatoire est réalisée (94,6 % des choix de l'alternative où la liaison obligatoire est réalisée vs. 5,4 % des choix de l'alternative où la liaison obligatoire n'est pas réalisée).

Dans la tâche de lecture à voix haute, 96,5 % des liaisons obligatoires sont bien réalisées. En revanche, Dugua et Baclesse enregistrent encore des erreurs d'omission (3,1 %) et de consonne de liaison erronée (0,4 %). L'explication qu'ils donnent est que « du fait de sa situation à la frontière de deux mots, même si la liaison est systématique à l'oral, elle nécessite un traitement particulier en lecture chez des enfants » (p.136).

Nous allons d'abord comparer ces données de dénomination guidée à nos résultats au CE1 dans la tâche DENa^{CE1} (Tableau 91).

	%	
	CE1 (Ruvoletto)	Dugua et Baclesse (2014)
Omissions	3	1,8
Consonnes erronées	0,3	0,5

Tableau 91. Comparaison des résultats pour la dénomination guidée entre Ruvoletto (2016) et Dugua & Baclesse (2014)

Les données de Dugua et Baclesse et nos données sont en accord : elles montrent qu'à ce stade il y a plus d'erreurs d'omissions que de consonnes erronées. En revanche, les résultats que nous avons obtenus dans DEN^{CE1} sont plus mauvais que ceux que Dugua et Baclesse présentent dans leur article. Que ce soit sur les erreurs d'omissions ou sur les erreurs de consonne erronée, nous avons des pourcentages deux fois plus élevés que ceux calculés par Dugua et Baclesse (3,0 % vs. 1,85 % et 0,3 % vs. 0,15 %). Nous pouvons donc penser qu'entre le CE1 et le CM1, la compétence des enfants change, la maîtrise de la liaison augmente et les erreurs diminuent.

Pour la tâche de lecture, la différence entre nos résultats dans LECT^{CE1} et les résultats de Dugua et Baclesse devient très importante (Tableau 92). Dans nos lectures de séquences en liaison obligatoire [n] et [z], le pourcentage d'erreurs est dix fois plus élevé que celui calculé par Dugua et Baclesse. Dans nos résultats, la liaison en lecture n'est pas encore maîtrisée au CE1, car encore plus de la moitié des séquences en liaison (68,9 %) ne sont pas bien lues chez les 37 enfants testés.

	%	
	CE1 (Ruvoletto)	Dugua et Baclesse (2014)
Omissions	31,1	0
Consonnes erronées	3,1	0,4

Tableau 92. Comparaison des résultats pour la lecture entre Ruvoletto (2016) et Dugua & Baclesse (2014)

Une explication à ces résultats très différents peut être donnée par le fait que les performances en lecture s'améliorent beaucoup plus entre le CE1 et le CM1 qu'entre le CP et le CE1. Comme nous l'avons vu dans la section 10.3.2.2, il n'y a pas de différence significative entre les erreurs de lecture à la fin du CP et pendant le CE1. Nous pouvons donc prévoir après le CE1 une chute des erreurs d'omissions en lecture à voix haute tandis que les compétences de lecture et d'écriture s'améliorent.

Chapitre 11

L'évolution entre CP et CE1 : des données aux conclusions

11.1 Introduction

Dans un premier temps, nous discutons les problèmes méthodologiques des tâches que nous avons utilisées pour tester les enfants aux trois moments.

Dans un deuxième temps, nous proposons une comparaison des résultats présentés au CP et au CE1. Nous cherchons à évaluer les évolutions impliquées par l'année passée à l'école mais aussi par le développement cognitif et langagier des enfants entre 6 et 7 ans. Nous présenterons d'abord les résultats pour les tâches PERC (PERC^{CP1} et PERC^{CP2})⁵⁶. Ensuite nous présenterons une comparaison pour la tâche orale DEN (DEN^{CP1}, DEN^{CP2} et DEN^{CE1}). Dans cette comparaison nous mettrons en relation les résultats entre CP (octobre-mai) et CE1. Nous nous concentrons ensuite sur les tâches liées à la langue écrite à la fin du CP et au CE1, PROD.ECR et LECT (PROD.ECR^{CP2}, PROD.ECR^{CE1} et LECT^{CP2}, LECT^{CE1})⁵⁷. Nous chercherons également à voir comment les relations entre les tâches orales et les tâches liées à la langue écrite évoluent au cours des deux années de l'école primaire en analysant les corrélations à la fin du CP et au CE1. Il faut souligner que pour DEN, PROD.ECR et LECT nous comparons les résultats des 37 enfants testés aux trois moments⁵⁸.

56 Nous nous rappelons que pour PERC^{CP2} nous avons testé un échantillon de seulement 12 enfants.

57 Pour les tâches liées à la langue écrite nous avons seulement deux moments à comparer et à présenter : fin CP et CE1.

58 Nous écartons les résultats de 6 enfants qui n'étaient plus présents au CE1.

11.2 Des problématiques méthodologiques dans les tâches proposées

11.2.1 La tâche de perception d'erreurs

Une grosse limite de la tâche de perception d'erreurs et de reformulation est le choix du matériel. Dans cette tâche, nous avons trois conditions : non-réalisation, consonne erronée et contrôle. Nous avons vu que, pour l'élision, deux conditions proposent le même type d'erreurs : *la armoire*, pour la condition de non-réalisation ; *le artichaut*, pour la condition de consonne erronée. Normalement, pour cette dernière condition, une erreur du type *le lartichaut* [lə.laʁ.ti.fo] aurait pu être plus pertinente. Le choix de garder l'erreur *le artichaut* dans la condition de consonne erronée était la conséquence de recherches conduites dans la base de données CHILDES. Dans cette base de données, nous avons cherché à repérer des exemples d'erreurs du type *le lartichaut*, sans trouver de véritables occurrences. Cependant, dans les tâches suivantes, nous avons remarqué des productions orales de ce type chez les enfants.

Donc dans la tâche de perception d'erreurs et reformulation, il aurait mieux fallu introduire l'erreur d'élision *le lartichaut* dans la condition « consonne erronée », plutôt que de proposer deux erreurs du même type dans deux conditions différentes.

11.2.2 La tâche de dénomination guidée

Une des limites des tâches orales proposées est la quantité d'items tests utilisés. Nous avons seulement 9 items tests à produire. Pour améliorer nos tâches orales, nous aurions pu ajouter des items tests. Cela nous aurait permis de mieux gérer les statistiques. De plus, dans une des deux tâches de dénomination guidée, pour l'enchaînement, nous demandons la production d'une séquence adjectif + nom, tandis que pour les autres phénomènes nous demandons de produire une séquence article + nom. Il aurait fallu créer des contextes syntaxiques plus homogènes, où nous aurions pu faire produire des enchaînements dans des contextes article + nom.

Un autre problème de la tâche de dénomination guidée est que les contextes phonologiques testés n'étaient pas homogènes. Pour choisir des mots de fréquence moyenne et représentables sur des cartes, nous avons négligé les contextes

phonologiques. Par exemple, pour la liaison, la voyelle qui suit la CL que les enfants doivent produire n'est pas toujours la même ([a] dans *les[z]arrosoirs* [le.za.ʁɔ.zwaʁ] vs. [u] dans *les[z]ours* [le.zuʁs]). Pour réussir à choisir des items tests qui avaient les mêmes caractéristiques phonologiques, il aurait fallu utiliser des non-mots qu'on aurait fait apprendre en association à des images dans une phase d'entraînement précédente à la mise en place de la tâche.

Une tâche de dénomination guidée avec un nombre plus élevé d'items à produire choisis dans un répertoire de non-mots, insérés dans des contextes plus homogènes du point de vue syntaxique pourrait être utilisée dans de futures recherches.

11.2.3 Les tâches liées à la langue écrite

À notre avis, les tâches liées à la langue écrite que nous avons proposées ne posent pas de gros problèmes méthodologiques. Une amélioration possible serait l'introduction de *fillers* dans les tâches à la fin du CP. Nous avons choisi de ne pas introduire les *fillers* à ce stade car nous pensions qu'il y avait déjà assez de séquences à écrire ou à lire et nous ne voulions pas trop charger les enfants. Au moment du test, nous nous sommes aperçu que, même au CP, les enfants pouvaient écrire et lire un nombre plus élevé de séquences. Introduire des *fillers* n'aurait donc pas causé de problèmes.

11.2.4 La tâche d'amorçage

Dans l'article de Gaskell *et al.* (2002), les chercheurs laissent affiché le mot écrit pendant 500 ms ; dans notre tâche nous laissons le mot écrit affiché jusqu'au moment de la réponse. Les TR que nous avons observés dépendent donc également des temps de lecture. Pour exclure l'influence des temps de lecture, nous proposons ici une nouvelle version de la tâche où la lecture de la cible (ELEPHANT ou TELEPHANT) est faite avant la présentation de l'input acoustique. La nouvelle tâche d'amorçage est construite comme suit : le mot écrit (ELEPHANT ou TELEPHANT) est affiché sur l'écran et dès que les enfants l'ont lu, ils doivent presser une touche pour passer à l'écoute de l'input acoustique (la phrase *un petit éléphant*) et dire si le mot écrit qu'ils ont lu était dans la phrase. Vu que la lecture du mot précède la présentation de l'input

acoustique, les TR ne sont pas influencés par les temps de lecture. En revanche cette tâche risque d'être influencée par les performances de la mémoire à court terme. Étant donné que les enfants doivent se rappeler d'un mot qu'ils ont lu au tout début du stimulus, la mémoire à court terme peut jouer un rôle important sur les TR. Il faudrait encore réfléchir pour construire une tâche d'amorçage permettant de vérifier nos hypothèses et qui tienne compte des capacités cognitives des enfants à tester.

De plus, dans AMOb les enfants ont des problèmes à refuser TELEPHANT car l'input acoustique *un petit[t]éléphant* a permis l'activation phonologique de TELEPHANT. Il est par conséquent plus difficile de refuser un mot qui est congruent avec la structure phonologique. Pour vérifier si la congruence entre la forme acoustique et la forme orthographique détermine les résultats de cette tâche, nous pourrions créer une nouvelle tâche d'amorçage où l'input acoustique correspondrait à l'orthographe mais sans l'intervention des phénomènes de resyllabation, comme par exemple une amorce acoustique du type *Jean bombarde* [ʒã.bõ.baʁd] et une cible visuelle *jambon*. Si les enfants ont des problèmes à dire que *jambon* n'est pas dans la phrase [ʒã.bõ õ.baʁd], cela veut dire que la congruence entre représentation phonologique et représentation orthographique joue un rôle dans la reconnaissance des mots écrits et dans AMOb.

11.3 Les trois étapes de l'étude longitudinale : la comparaison des données

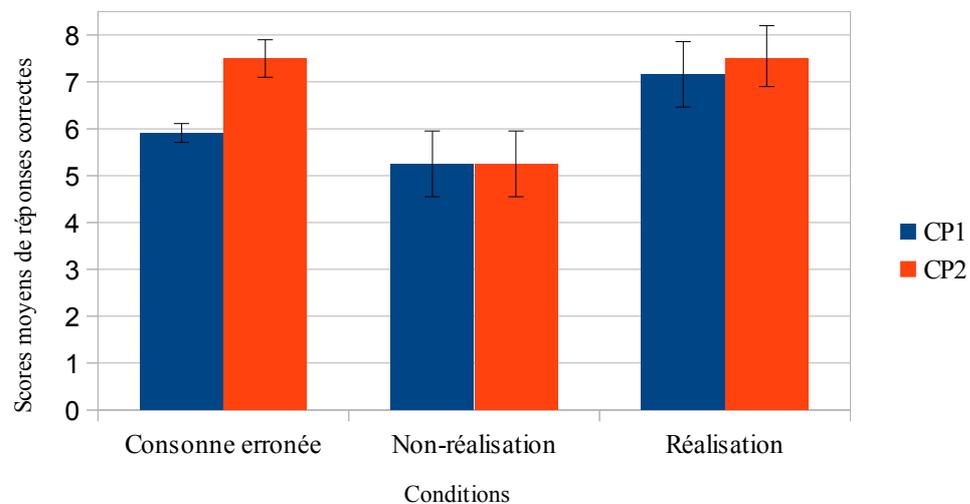
11.3.1 La tâche de perception d'erreurs : PERC^{CP1} et PERC^{CP2}

Nous avons établi une comparaison des résultats pour l'échantillon de 12 enfants testés au début du CP par PERC^{CP1} et à la fin par PERC^{CP2}. Nous avons considéré les scores de réponses correctes (les identifications d'insertion de consonne erronée dans CONDcons.err, de non-réalisation dans CONDnon-réal ainsi que les non-identifications d'erreurs dans la condition de réalisation CONDréal) pour les trois conditions⁵⁹. Au moyen d'un t-test, nous pouvons établir une différence statistiquement significative et une amélioration entre les scores de PERC^{CP1} et de PERC^{CP2} seulement pour la condition CONDcons.err ($t_1(12) = 0,96$, $p < .05$, $t_2(7) = 1,20$, $p < .05$). Cette différence et cette amélioration ne sont pas enregistrées pour CONDréal ($t_1(11) = 1,86$,

⁵⁹ Nous invitons à la lecture des chapitres 7 et 8 pour une description des résultats, du répertoire et de la procédure.

$p = 1,00$, $t_2(7) = 1,04$, $p = .32$) ni pour CONDnon-réal ($t_1(11) = 1,07$, $p = 0,30$, $t_2(7) = .80$, $p = .44$), comme il est bien représenté dans le graphique 26 (le score maximal est 8).

Les enfants à la fin du CP réussissent à mieux repérer l'insertion d'une consonne erronée entre Mot1 et Mot2. Cependant, au CP et au CE1, ils ont les mêmes difficultés à signaler les non-réalisations et à les considérer comme des fautes.



Graphique 26. Comparaison des réponses correctes pour PERC : CP vs. CE1

11.3.2 La tâche orale de dénomination guidée (DENa) entre CP et CE1

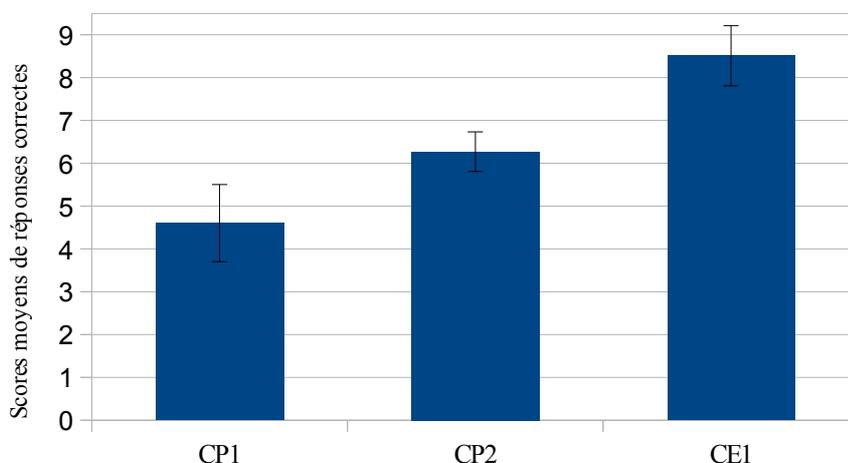
Nous proposons d'abord une comparaison entre les résultats au début du CP (DENa^{CP1}), à la fin du CP (DENa^{CP2}) et au CE1 (DENa^{CE1})⁶⁰.

L'analyse de la variance (ANOVA) pour les scores de réponses correctes (donc des productions correctes en dénomination) utilisant les sujets et les items comme variables aléatoires, montre un effet du facteur « temps » ($F_1(2, 72) = 36,023$, $p < .001$, $F_2(1, 2) = 6,02$, $p < .05$). L'analyse des contrastes montre une différence statistiquement significative pour les scores de réponses correctes entre DENa^{CP1} et DENa^{CP2} ($F_1(1,$

⁶⁰ Nous invitons à la lecture des chapitres 8, 9 et 10 pour une description des résultats, du répertoire et de la procédure.

36) = 12,831, $p < .001$, $F_2(1, 2) = 3,03$, $p < .05$) DENa^{CP2} et DENa^{CE1} ($F_1(1, 36) = 19,760$, $p < .001$, $F_2(1, 2) = 2,33$, $p < .05$). Cela veut dire que les performances des enfants changent au cours de l'année scolaire.

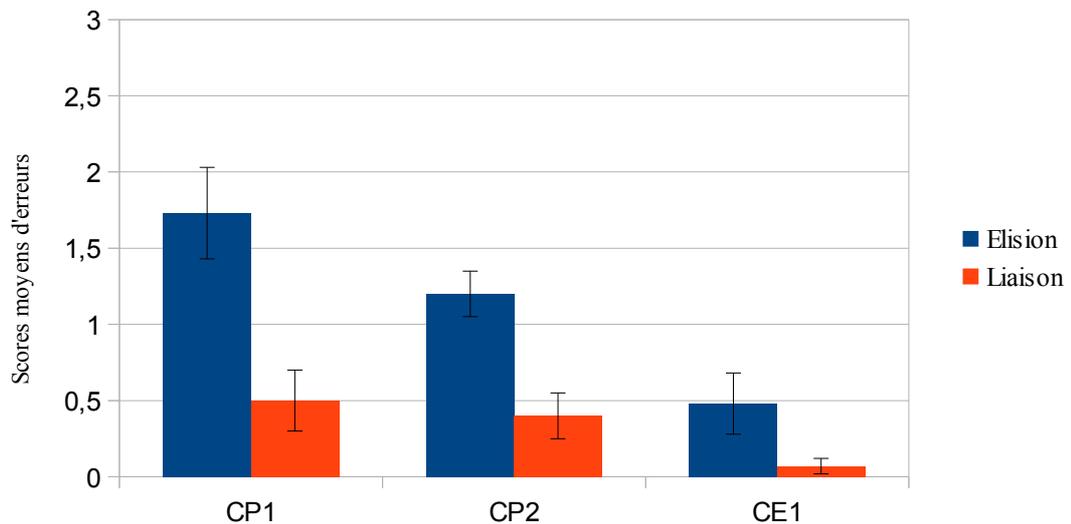
Dans le graphique 27, nous présentons les scores de réponses correctes pour la dénomination guidée aux 3 moments de recueil de données. Les résultats nous montrent une augmentation des productions correctes entre DENa^{CP1}, DENa^{CP2} et DENa^{CE1}. Il y a une très nette amélioration des productions orales entre CP et CE1.



Graphique 27. Comparaison des réponses correctes DENa dans les trois moments

Analysons plus spécifiquement les scores d'erreurs de segmentation (non-réalisation *les[Ø]ours* pour *le(s)[z]ours* et insertion d'une consonne erronée *le(s)[n]ours* pour *le(s)[z]ours*). L'analyse de la variance (ANOVA) pour les scores d'erreurs de segmentation utilisant les sujets et les items comme variables aléatoires ne révèle pas d'effet du facteur « temps » pour l'enchaînement ($F_1(2, 72) = .55$, $p = .57$, $F_2(2,4) = 24,40$, $p = .06$). Cette effet est observé pour l'élision ($F_1(2, 72) = 10,952$, $p < .001$, $F_2(2,4) = 32,09$, $p < .05$) et pour la liaison ($F_1(2, 72) = 5,68$, $p < .05$, $F_2(2,4) = 9,32$, $p < .05$). Pour l'élision, l'analyse des contrastes fait observer une différence significative sur les scores d'erreurs de segmentation aux trois moments (DENa^{CP1} vs. DENa^{CP2} $F_1(1, 36) = 4,24$, $p < .05$, $F_2(1,2) = 34,71$, $p < .05$; DENa^{CP2} vs. DENa^{CE1} $F_1(1, 36) = 5,47$, $p < .05$, $F_2(1,2) = 23,40$, $p < .05$). Pour la liaison, l'analyse de contrastes révèle une différence statistiquement significative sur les scores d'erreurs de segmentation entre CP et CE1 (DENa^{CP1} vs. DENa^{CP2} $F_1(1, 36) = 2,40$, $p = .19$,

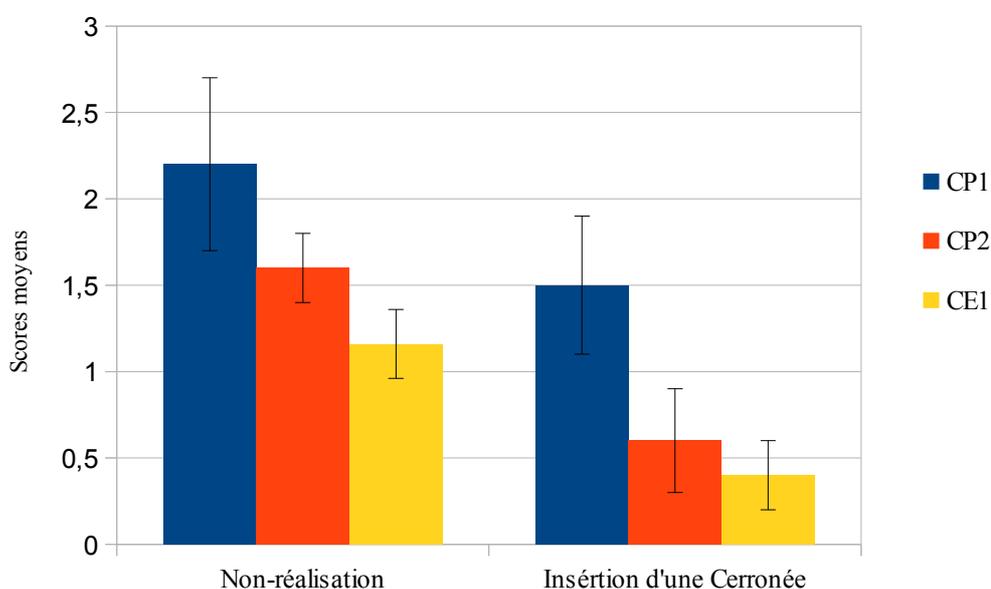
$F_2(1,2) = 2,81$, $p = .35$; $DENa^{CP2}$ vs. $DENa^{CE1}$ $F_1(1, 36) = 3,17$, $p < .05$, $F_2(1,2) = 12,34$, $p < .05$). Les erreurs de segmentation des séquences en élision diminuent entre $DENa^{CP1}$, $DENa^{CP2}$ et $DENa^{CE1}$ et les erreurs en liaison diminuent entre $DENa^{CP2}$ et $DENa^{CE1}$ (Graphique 28, où le score maximal est 3).



Graphique 28. Comparaison des erreurs en liaison et élision dans les trois moments

Nous avons conduit une analyse des séquences tests pour les deux types d'erreurs de segmentation que nous avons remarqués dans la tâche : les erreurs de non-réalisation (*les[Ø]ours* [le.Øuʁs] pour *les[z]ours* [le.zuʁs], *la petite-orange* [la.pə.tit.o.vɑ̃ʒ] pour *la petite[t]orange* [la.pə.ti.to.vɑ̃ʒ], *le-oreiller* [lə.o.vɛ.jɛ] pour *l'oreiller* [lə.vɛ.jɛ]) et les erreurs d'insertion d'une consonne erronée (*les[n]ours* [le.nuʁs] pour *les[z]ours* [le.zuʁs], *le [n]oreiller* [lə.nɔ.vɛ.jɛ] pour *l'oreiller* [lə.vɛ.jɛ], et *la petite[z]orange* [la.pə.ti.zo.vɑ̃ʒ] pour *la petite orange* [la.pə.ti.to.vɑ̃ʒ]). L'analyse de la variance (ANOVA) pour les scores de non-réalisation utilisant les sujets et les items comme variables aléatoires montre un effet du facteur « temps » ($F_1(2,72) = 3,134$, $p < .05$, $F_2(2, 16) = 6,889$, $p < .001$). L'analyse des contrastes nous révèle qu'il n'y a pas de différence statistiquement significative entre $DENa^{CP1}$ et $DENa^{CP2}$ ($DENa^{CP1}$ vs. $DENa^{CP2}$ $F_1(1,36) = 2,57$, $p = .11$, $F_2(1, 8) = 1,49$, $p = .20$) mais qu'elle a lieu entre ces deux dernières et $DENa^{CE1}$ ($DENa^{CP1}$ et $DENa^{CP2}$ vs. $DENa^{CE1}$ $F_1(1,8) = 4,31$, $p < .05$, $F_2(1, 8) = 10,88$ $p < .05$). L'analyse de la variance (ANOVA) pour les scores d'insertion d'une consonne erronée utilisant les sujets et les items comme variables aléatoires montre également un effet du facteur « temps » ($F_1(2,72) = 5,663$,

$p < .001$, $F_2(2, 16) = 8,577$, $p < .001$). L'analyse des contrastes permet d'observer une différence statistiquement significative entre $DENa^{CP1}$ et $DENa^{CP2}$ et une différence statistiquement non significative entre $DENa^{CP2}$ et $DENa^{CE1}$ ($DENa^{CP2}$ vs. $DENa^{CE1}$ $F_1(1,36) = 0,36$, $p = .54$, $F_2(1,8) = 0,64$, $p = .80$; $DENa^{CP2}$ et $DENa^{CE1}$ vs. $DENa^{CP1}$ $F_1(1,36) = 5,536$, $p < .05$, $F_2(1, 8) = 10,46$, $p < .05$). Les erreurs d'insertion d'une consonne erronée diminuent pendant le CP mais restent stable entre CP et CE1, tandis que les erreurs de non-réalisation restent stables pendant le CP et diminuent entre CP et CE1 (Graphique 29⁶¹).



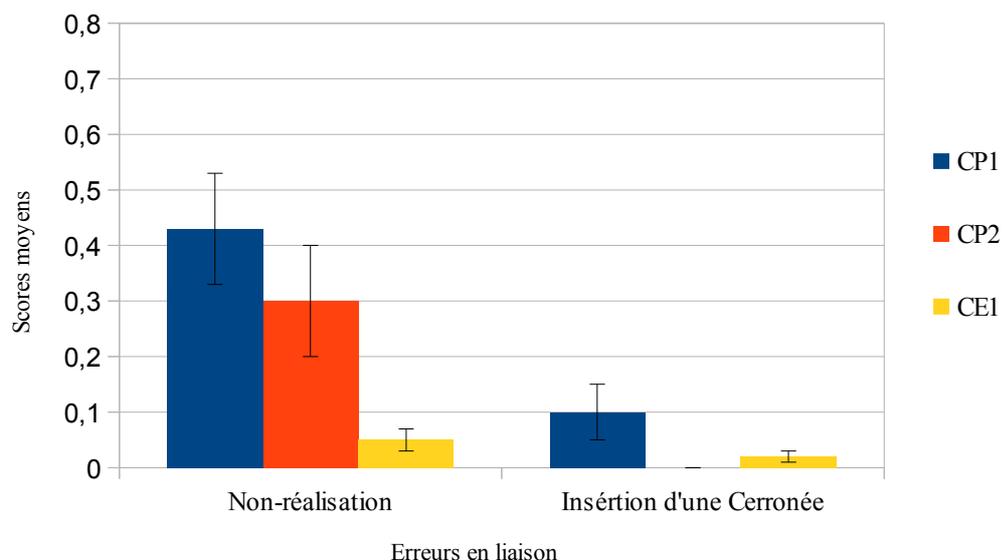
Graphique 29. Non-réalisations et insertions d'une consonne erronée entre CP et CE1

Nous enregistrons également une différence statistiquement significative entre le nombre d'erreurs de non-réalisation et le nombre d'erreurs d'insertion d'une consonne erronée dans $DENa^{CP1}$ ($t_1(41) = 3,528$, $p < .001$; $t_2(8) = 2,34$, $p < .05$), $DENa^{CP2}$ ($t_1(42) = 1,442$, $p < .001$, $t_2(8) = 2,14$, $p < .05$) et $DENa^{CE1}$ ($t_1(36) = 3,19$, $p < .05$, $t_2(8) = 3,10$, $p < .05$). Nous remarquons que les erreurs de non-réalisation sont les plus nombreuses à chacun des trois moments (Graphique 29). Puisque les non-réalisations sont les plus nombreuses au CP et au CE1, nous les analysons plus en détail dans les prochains paragraphes.

Il faut souligner que dans le chapitre 9, nous avons vérifié au moyen d'une autre tâche de dénomination guidée (DENb) quelles erreurs étaient conséquentes à la procédure adoptée et quelles erreurs caractérisent les productions des enfants

61 Le score maximal est 9.

indépendamment de cette procédure. Nous avons vu que les erreurs d'élision et d'enchaînement étaient générées par la tâche et que les seules erreurs qui n'étaient pas en rapport avec la procédure étaient les erreurs de liaison. Les données sur la liaison représentent mieux les productions des enfants car elles ne sont pas biaisées par la tâche proposée. Nous proposons donc une analyse statistique des non-réalisations et des insertions d'une consonne erronée, pour la liaison, au CP et au CE1. L'analyse de la variance (ANOVA) pour les scores de non-réalisations utilisant les sujets et les items comme variables aléatoires montre un effet du facteur « temps » ($F_1(2,72) = 4,090, p < .05, F_2(2, 4) = 7,86, p < .05$). L'analyse des contrastes révèle qu'il n'y a pas de différence statistiquement significative entre $DENa^{CP1}$ et $DENa^{CP2}$ ($DENa^{CP1}$ vs. $DENa^{CP2}$ $F_1(1,36) = 0,66, p = .41, F_2(1, 2) = 0,19, p = .70$), mais qu'il y a une différence statistiquement significative entre ces deux derniers et $DENa^{CE1}$ ($DENa^{CP2}$ et $DENa^{CP2}$ vs. $DENa^{CE1}$ $F_1(1,36) = 14,831, p < .001, F_2(1,2) = 7,34 p < .05$). Dans le graphique 30, nous présentons les scores moyens de non-réalisation et d'insertion de consonne erronée pour la liaison. Nous pouvons observer que, pour les 37 enfants testés aux trois moments, les erreurs d'insertion se réduisent pendant le CP et le CE1⁶². Les erreurs de non-réalisation restent stables pendant le CP et diminuent drastiquement entre CP et CE1.



Graphique 30. Non-réalisations et insertions d'une consonne erronée pour la liaison

62 Nous ne proposons pas une analyse de la variance avec ANOVA car à la fin du CP (CP2), nous enregistrons 0 erreur d'insertion d'une consonne erronée). Cela implique l'absence de la variance et donc de la possibilité de l'analyser.

11.3.2.1 L'analyse phonétique du traitement des non-réalisations en production entre CP et CE1

Nous avons analysé phonétiquement les productions orales des enfants pour comprendre comment l'espace sonore entre Mot1 et Mot2 est traité quand les séquences produites en dénomination guidée présentent des non-réalisations ([le.Øuʁs], [la.pə.tit.o.ʁɑ̃ʒ], [lə.o.ʁε.je]). Nous avons déjà remarqué trois types de productions : insertion d'un coup de glotte, production d'une véritable pause ou production d'un continuum (§ 8.3.3.1). Nous analyserons en détail ces trois types de productions pour les phénomènes de resyllabation testés aux trois moments de recueil des données (CP1, CP2 et CE1).

La non-réalisation de la liaison dans DENa^{CP1}, DENa^{CP2} et DENa^{CE1}

Pour la liaison, la non-réalisation du phénomène implique la non-réalisation de la CL au début du Mot2 (les[Ø]ours [le.Øuʁs]). Dans le tableau 93, nous montrons les trois types de productions qui émergent dans notre analyse phonétique entre CP et CE1. En regardant le tableau, nous pouvons résumer l'évolution de ces productions en trois étapes.

	%		
	CP1	CP2	CE1
Coup de glotte (ex. [le.ʁuʁs])	22,2	62	40
Pause (ex. [le/uʁs])	11,10	8	0
Continuum (ex. [le.uʁs])	66,7	30	60

Tableau 93. Analyses phonétiques des non-réalisations de liaison entre CP et CE1

1. CP1 : continuité entre Mot1 et Mot2

Au début du CP, les enfants produisent plus de continuums que de pauses ou de coups de glotte. Il ne produisent pas la CL mais ils cherchent à donner une continuité à la chaîne Mot1 + Mot2.

2. CP2 : Mot1 et Mot2 comme deux formes séparées

À la fin du CP, les enfants commencent à maîtriser l'écriture et ils semblent avoir compris qu'il faut bien marquer la séparation entre Mot1 et Mot2. Ils produisent un coup de glotte qui prend la place de la CL non réalisée.

3. CE1 : enchaîner Mot1 et Mot2

Au CE1, les enfants préfèrent encore produire des continuums ou des coups de glotte qui prennent la place de la CL. La production de coups de glotte montre une continuité avec l'étape précédente. La production de continuums peut être le résultat de la volonté de reproduire la chaîne orale et l'enchaînement entre mots en liaison. L'exercice de la lecture qui porte à enchaîner des mots séparés, peut causer la multiplication des réalisations de continuums.

L'analyse phonétique fait émerger une évolution des non-réalisations. Au début du CP, les enfants ne sont pas encore tout-à-fait capables de produire séparément Mot1 et Mot2, car les frontières entre les deux ne sont pas encore claires. À la fin de l'année scolaire, ils commencent à comprendre cette séparation : ils ne réalisent pas la CL, mais séparent Mot1 et Mot2 au moyen d'un coup de glotte. À la fin du CE1, influencés aussi par la lecture, ils essaient d'enchaîner les deux mots en liaison même si la CL n'est pas produite.

La non-réalisation de l'élision dans DENa^{CP1}, DENa^{CP2} et DENa^{CE1}

Pour l'élision, dans les non-réalisations, le schwa ne chute pas et la position d'attaque syllabique, normalement occupée par une consonne, reste vide (ex. [lə.o.ʁε.je]). En regardant le tableau 94, nous pouvons voir que, pendant le CP, nous enregistrons l'ajout phonétique d'un coup de glotte entre Mot1 et Mot2. L'ajout de ce coup de glotte permet de remplir l'attaque vide au début du Mot2 et de reproduire une séquence syllabique du type CV (et d'empêcher la chute du noyau vocalique).

Au CE1, ces productions de coups de glotte subissent une chute en faveur de la production d'un continuum. Nous pouvons penser que les productions de continuums essaient de reproduire l'unité sonore des séquences de Mot1 et Mot2 en élision, où il n'y a pas de véritable pause entre les deux mots à l'oral ni d'espaces blancs à l'écrit.

	%		
	CP1	CP2	CE1
Coup de glotte (ex. [lə.ʔo.ʁε.jε])	72,4	41,90	8,4
Pause (ex. [lə/o.ʁε.jε])	24,1	35,5	25
Continuum (ex. [lə.o.ʁε.jε])	3,5	22,5	66,40

Tableau 94. Analyses phonétiques des non-réalisations d'élision entre CP et CE1

La non-réalisation d'enchaînement dans $DENa^{CP1}$, $DENa^{CP2}$ et $DENa^{CE1}$

Dans les non-réalisations, pour les séquences en enchaînement, la consonne d'enchaînement est prononcée mais elle n'est pas resyllabifiée sur Mot2 (ex. *la petite-orange* [la.pə.tit.o.ʁãʒ] pour *la petite [t] orange* [la.pə.ti.to.ʁãʒ]).

Comme nous l'avons déjà vu dans les chapitres précédents, nous remarquons pour l'enchaînement seulement deux types de productions phonétiques : ajout d'un coup de glotte et ajout d'une pause. Le tableau 95 montre que la production de coups de glotte reste basse tandis que la production de pauses augmente du CP au CE1. L'augmentation des pauses peut montrer que les enfants sont de plus en plus influencés par la *literacy* et essaient de reproduire l'espace blanc entre Mot1 et Mot2 à l'écrit.

	%		
	CP1	CP2	CE1
Coup de glotte (ex. [la.pə.tit.ʔo.ʁãʒ])	71,4	42,85	30
Pause (ex. [la.pə.tit/o.ʁãʒ])	18,6	57,15	70

Tableau 95. Analyses phonétiques des non-réalisations d'enchaînement entre CP et CE1

Les idées exposées dans ce paragraphe tiennent compte de l'influence des activités de *literacy* sur les productions orales. Nous analyserons ces influences et leurs évolutions dans la section 11.3.5 de ce chapitre.

11.3.3 L'analyse de la fréquence dans les productions orales pour DENa^{CP1}, DENa^{CP2} et DENa^{CE1}

Nous avons essayé de vérifier si la fréquence des items dans le lexique des enfants influence leurs productions orales quand ces items sont insérés dans des phénomènes de resyllabation. Selon le modèle lexical-constructionniste, la fréquence joue un rôle fondamental dans le traitement des frontières de mots. Comme nous l'avons déjà dit, si le Mot2 à produire est plus fréquent au pluriel, les enfants peuvent le trouver plus facilement dans des séquences comme *les[z]ours* ([le.zuʁs]) et ils peuvent par conséquent penser que ce mot commence par [z]. Cela peut donc causer des erreurs du type *un[z]ours*. Au CP et au CE1, nous avons cherché à établir une corrélation entre la fréquence SFI au pluriel des items à produire et les erreurs d'insertion d'une consonne erronée [z] au début du nom dans DENa^{CP1}, DENa^{CP2} et DENa^{CE1}. Dans le tableau 96, nous présentons les indices de corrélation.

	CP1	CP2	CE1
<i>r</i> (8)	.21 (p = .60)	-.21 (p = .54)	.63 (p = .06)

Tableau 96. Indice de corrélation entre les erreurs de production et la fréquence entre CP et CE1

A aucun des trois moments, il n'y a de corrélation entre les erreurs d'insertion d'une consonne erronée [z] et la fréquence au pluriel des items à produire. À la lumière de ce résultat, nous ne pouvons pas dire, qu'à ce stade, la fréquence intervient dans la production des mots en liaison et dans l'identification de leurs frontières lexicales.

11.3.4 Les tâches liées à la langue écrite entre CP et CE1

Comparons maintenant les résultats de la tâche d'écriture PROD.ECR et de lecture LECT aux deux moments (CP2 et CE1) pour comprendre si les compétences sur la langue écrite subissent une amélioration.

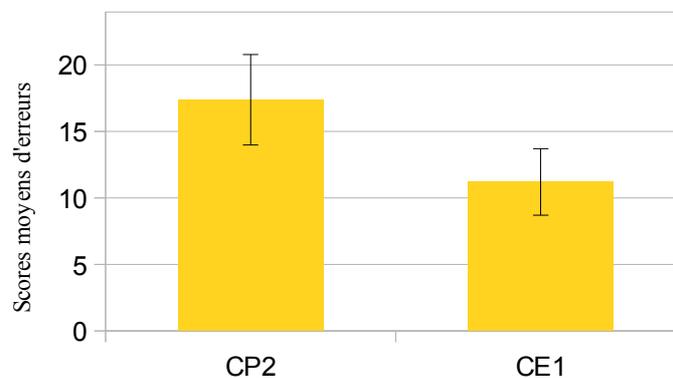
11.3.4.1 La tâche d'écriture : PROD.ECR^{CP2} et PROD.ECR^{CE1}

Nous comparons les résultats de la tâche d'écriture proposée au CP (PROD.ECR^{CP2}) et au CE1 (PROD.ECR^{CE1}). Nous présentons d'abord les résultats généraux qui considèrent toutes les erreurs dans les productions écrites enregistrées à l'intérieur du mot (<une otruche> pour *une autruche*, <les affiche> pour *les affiches*), ou sur la frontière entre Mot1 et Mot2 (<les zoranges> pour *les oranges*). Ensuite, nous analysons les réalisations à l'écrit des phénomènes de resyllabation, en comparant les productions écrites à « frontières lexicales » et les productions écrites à « frontières syllabique » entre CP et CE1. Enfin, nous traitons les questions du marquage du pluriel, du graphème *h* et de la familiarité et de la fréquence des items à écrire.

11.3.4.1.1 Les erreurs dans les productions écrites

Nous considérons d'abord les erreurs totales, où les formes lexicales sont mal orthographiées (ex. <une otruche> pour *une autruche*, <les affiche> pour *les affiches*) et également où une mauvaise segmentation est présente à l'écrit (productions écrites à « frontières syllabiques », <les zoranges> pour *les oranges*). Cette analyse permet d'évaluer le niveau global des enfants et leur évolution en ce qui concerne l'activité d'écriture entre CP et CE1.

Le résultat du t-test nous montre qu'il y a une différence statistiquement significative entre le nombre d'erreurs dans les productions écrites dans PROD.ECR^{CP2} et PROD.ECR^{CE1} ($t_1(36) = 7,642$, $p < .001$; $t_2(19) = 9,365$, $p < .001$) et qu'il y a une amélioration au CE1 car les erreurs diminuent (Graphique 31).



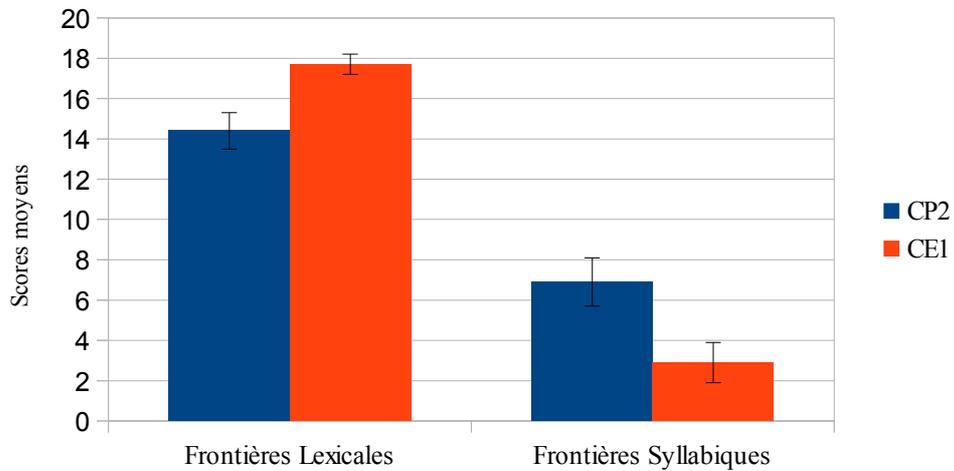
Graphique 31. Erreurs en PROD.ECR du CP au CE1

Nous pouvons donc dire que les enfants réussissent à écrire mieux au CE1. Ces résultats ne sont pas surprenants : nous nous attendions à des progrès consécutifs au passage à un niveau scolaire supérieur.

11.3.4.1.2 Les productions écrites qui respectent les frontières lexicales ou les frontières syllabiques

Nous comparons maintenant les données sur les productions écrites qui respectent les frontières lexicales (<les affiches>, <l'éléphant>, <une ambulance>) ou les frontières syllabiques (<les zoranges>, <lélephant>, <une nambulance>) dans les tâches PROD.ECR^{CP2} et PROD.ECR^{CE1}. Les enfants qui montrent des productions écrites qui respectent les frontières lexicales sont capables de bien orthographier les séquences caractérisées par les phénomènes de resyllabation. En revanche, les enfants qui respectent les frontières syllabiques orales des mots pour écrire, orthographient la consonne qui resyllabifie sur Mot2 comme l'initiale de ce mot.

Nous enregistrons une différence statiquement significative aux deux moments (CP2 vs. CE1), pour les productions écrites à « frontières syllabiques » ($t_1(36) = 4,40$, $p < .05$, $t_2(19) = 2,77$, $p < .05$), comme pour les productions écrites « à frontières lexicales » ($t_1(36) = 4,53$, $p < .05$, $t_2(19) = 2,70$, $p < .05$). Comme on peut le voir dans le graphique 32, nous remarquons une réduction des productions écrites « à frontières syllabiques » et une augmentation des productions « à frontières lexicales » au CE1.



Graphique 32. Productions écrites : frontières syllabiques et frontières lexicales entre CP et CE1

Nous pouvons conclure que les enfants au CE1 maîtrisent mieux les frontières entre les mots à l'écrit en présence des phénomènes de resyllabation. Ils réussissent à bien orthographier les séquences caractérisées par des phénomènes de resyllabation en proposant des productions écrites correctes où les frontières lexicales des mots sont bien identifiées et représentées.

11.3.4.1.3 Le marquage du pluriel à l'écrit

Dans cette section, nous comparons les productions écrites de la liaison pour évaluer s'il y a une amélioration dans la pratique du marquage graphique du pluriel entre CP et CE1. Pour vérifier le marquage du pluriel, nous prenons en considération les séquences en liaison au pluriel et nous tenons compte du marquage du Mot2, car le marquage du Mot1 est déjà acquis à partir du CP (voir chapitre 8). Nous pouvons voir dans le tableau 97 qu'il y a une augmentation du marquage graphique du pluriel au CE1 comparé au CP2. Si nous regardons les pourcentages totaux, l'augmentation est forte (de 25,01 % au CP2 nous passons à 43,24 % au CE1).

	% marquage graphique du pluriel	
	CP2	CE1
Les oranges	8,10	35,13
Des ailes	16,21	40,50
Les ours	60,1	67,56
Les affiches	8,01	29,72
TOTAL	25,01	43,24

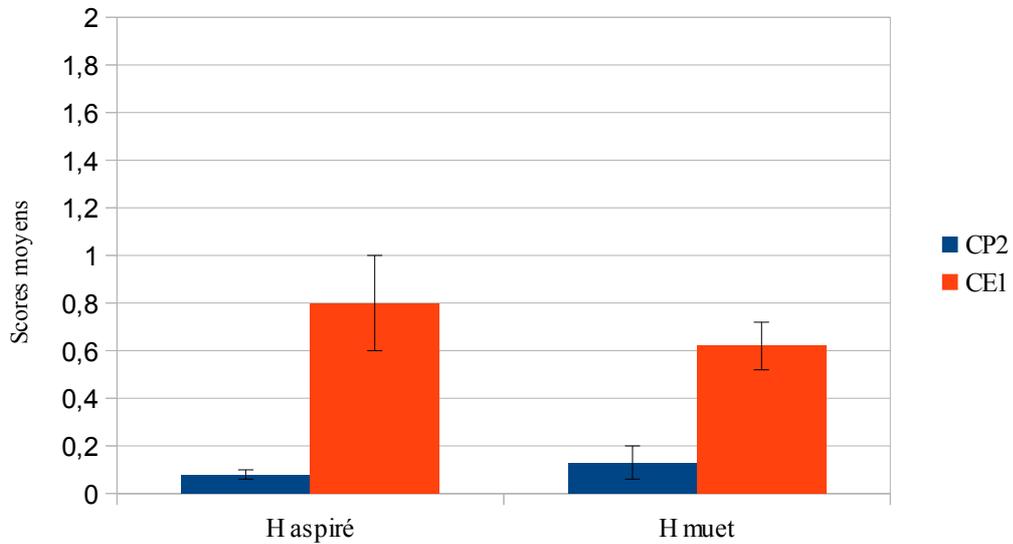
Tableau 97. Marquage du pluriel entre CP et CE1

Nous pouvons voir que le marquage de *ours* ne subit pas de changements entre CP2 et CE1. En revanche nous observons une augmentation des pourcentages du marquage pour les autres items. Il faut souligner que pour le mot *ours*, le <s> orthographique à la fin de la forme plurielle et singulière est réalisé oralement. Cette réalisation n'a pas lieu pour les autres items. L'augmentation du nombre de marques du pluriel pour les autres items montre donc que les enfants, petit à petit, sont en train d'apprendre que pour un mot au pluriel il est nécessaire de mettre un <s> à la fin de sa forme orthographique, même si ce <s> n'est pas réalisé oralement. Cela implique qu'au CE1 ils sont en train de construire un lexique orthographique qu'ils peuvent utiliser pour écrire les mots sans avoir toujours besoin de passer par les correspondances graphème-phonème.

11.3.4.1.4 H aspiré et h muet à l'écrit

Nous avons également cherché à comprendre si les enfants orthographient le *h* au début du Mot2 plus fréquemment au CE1 dans PERC^{CE1} qu'à la fin du CP dans PERC^{CP2}. Nous comparons les scores de *h* orthographié pour les 4 contextes (2 contextes en *h* aspiré et 2 contextes en *h* muet). Nous observons une différence statistiquement significative dans le cas de *h* aspiré ($t_1(36) = 5,020$, $p < .001^{63}$) et dans le cas de *h* muet ($t_1(36) = 3,682$, $p < .001$) entre PROD.ECR^{CP2} et PROD.ECR^{CE1}. Les *h* initiaux (aspirés ou muets) sont orthographiés plus fréquemment au CE1 (Graphique 33).

63 Le nombre d'items à comparer pour *h* aspiré et *h* muet est très bas (seulement 2 items). Pour cette raison, nous allons seulement considérer le résultat du t_1 .



Graphique 33. Scores moyens de productions écrites de <h> au début du Mot2

Nous pouvons conclure que les enfants orthographient mieux les mots qui commencent par *h* au CE1. Comme nous l'avons déjà dit pour le marquage graphique du pluriel, au CE1 les enfants renforcent leur lexique orthographique et réussissent à mieux orthographier les mots, sans devoir toujours faire référence à la correspondance graphème-phonème. Cependant, comme nous l'avons déjà vu au CP et au CE1, les enfants ne font pas de différence entre les deux contextes.

11.3.4.1.5 La familiarité des contextes et la fréquence des items à écrire

Nous avons également évalué l'influence de la familiarité d'un contexte à l'écrit en cherchant à voir si un contexte connu (une séquence de Mot1 + Mot2 apprise à l'école) est mieux orthographié qu'un contexte inconnu (une séquence de Mot1 + Mot2 jamais traitée à l'école). Les frontières des mots sont-elle mieux identifiées si le contexte est connu à l'écrit ? La comparaison entre les contextes connus et les contextes inconnus dans PROD.ECR^{CP2} et PROD.ECR^{CE1} montre qu'au CP seulement les frontières des mots sont mieux identifiées et orthographiées grâce à la familiarité.

Nous nous posons dans ce paragraphe une deuxième question : la fréquence des items à écrire influence-t-elle la tâche d'écriture ?

Dans les chapitres 9 et 10, nous avons calculé deux corrélations :

1. Entre productions écrites globalement correctes (<les affiches>) et fréquence.
2. Entre productions écrites à « frontières syllabiques » (<les saffiches>) et fréquence.

Nous avons considéré deux indices de fréquence : la fréquence absolue des items (nombre total d'occurrences) dans notre base de données au CP et la fréquence SFI donnée par MANULEX (au CP et au CE1, en accord avec le niveau scolaire). Comme nous l'avons vu dans les chapitres précédents, les productions écrites ne sont pas en rapport avec la fréquence des Mots2 à écrire. À la lumière de nos résultats, nous pouvons conclure que la fréquence des items dans le lexique des enfants n'aide ni à orthographier les items, ni à identifier, à l'écrit, les frontières de mots perturbées par les phénomènes de resyllabation à l'oral.

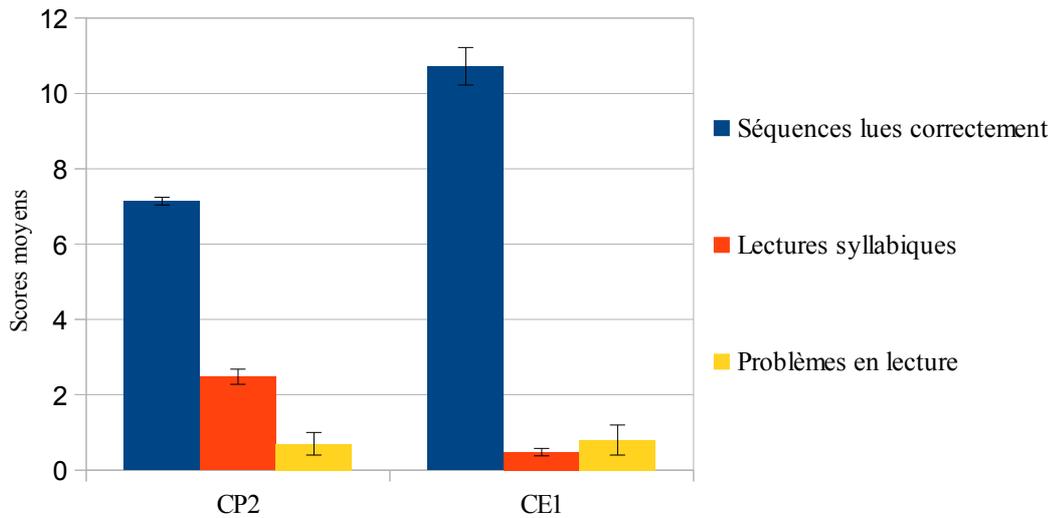
11.3.4.2 La tâche de lecture : LECT^{CP2} et LECT^{CE1}

Pour évaluer si les performances des enfants en lecture s'améliorent, nous proposons d'abord une analyse générale des lectures dans LECT^{CP2} et LECT^{CE1} et ensuite une analyse plus spécifique, centrée sur les phénomènes de resyllabation.

11.3.4.2.1 Les lectures correctes et les lectures syllabiques

Nous proposons d'abord une comparaison des trois catégories de lectures utilisées dans la classification générale au CP et au CE1 (§ 9.5.2.2, § 10.3.2.2) : séquences lues correctement (la lecture des items est bien réalisée), lectures syllabiques (ex. [le/le/fã] pour *l'éléphant*, ou [le/za/fif] pour *les affiches*) et problèmes de lecture (lectures erronées/inventées). Nous observons une différence statistiquement significative pour les séquences lues correctement ($t_1(36) = 6,159$, $p < .001$; $t_2(19) = 4,56$, $p < .05$) et pour les lectures syllabiques ($t_1(36) = 4,332$ $p < .001$; $t_2(19) = 5,123$ $p < .001$). Les lectures syllabiques diminuent au CE1 et les lectures associées à un bon niveau augmentent (Graphique 34⁶⁴). En revanche, aucune différence statistiquement significative n'est enregistrée pour les problèmes de lecture qui restent stables entre les deux moments ($t_1(36) = 2,70$, $p = .10$; $t_2(19) = 1,10$, $p = .41$).

64 Le score maximal est 20.

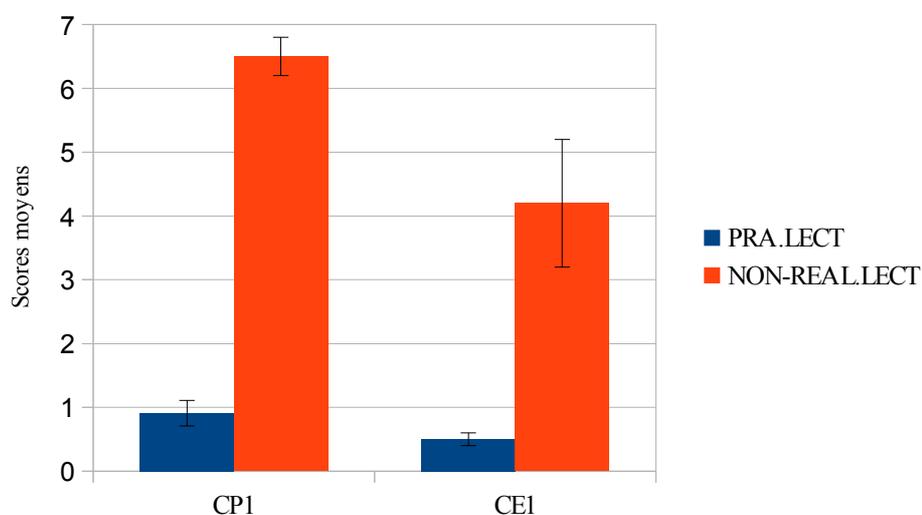


Graphique 34. Niveaux de lecture des enfants entre CP et CE1

11.3.4.2.2 Les erreurs de segmentation en lecture

Nous considérons la classification faite dans les chapitres 9 et 10. Nous écartons les données pour les contextes *h* muet et *h* aspiré que nous allons traiter dans une section dédiée. Nous comparons les deux types d'erreurs de segmentation en lecture : les non-réalisations du phénomène de resyllabation (NON-REAL.LECT, ex. *le(s)[Ø]affiches*) et les productions à attaque remplie (PRA.LECT, ex. *le(s)-pause-[z]affiches*). Le t-test montre qu'il y a une différence statistiquement significative entre le nombre de PRA.LECT^{CP2} et de PRA.LECT^{CE1} ($t_1(36) = 2,37$, $p < .05$, $t_2(15) = 2,25$, $p < .05$) et également entre le nombre de NON-REAL.LECT^{CP2} et de NON-REAL.LECT^{CE1} ($t_1(36) = 3,871$, $p < .001$, $t_2(15) = 4,92$, $p < .05$). Comme on peut le voir dans le graphique 35, le nombre de PRA.LECT et de NON-REAL.LECT diminue au CE1, même si à ce moment le nombre de ce dernier type de lecture (NON-REAL.LECT) est encore le plus élevé. Dans la prochaine section, nous proposons une analyse phonétique de NON-REAL.LECT qui permet de comparer le traitement de l'espace sonore entre Mot1 et Mot2.

Il faut souligner que pour la liaison, au CE1, nous avons dû introduire une troisième catégorie : NONEnch.LECT. Dans ces productions, la CL est réalisée à la fin du Mot1 et, après une pause, les enfants lisent le Mot2 qui est traité comme un mot à voyelle initiale (ex *les[z]affiches* [lez/a.fiʃ] pour *les[z]affiches* [le.za.fiʃ]). L'apparition de ce type de production peut avoir une explication orthographique. La CL est représentée graphiquement sur Mot1. Les enfants respectent les frontières lexicales des mots et produisent la CL à la fin du Mot1.



Graphique 35. PRA et NON-REAL en lecture au CP et au CE1

11.3.4.2.3 L'analyse phonétique des NON-REAL.LECT^{CP2} et NON-REAL.LECT^{CE1}

Nous proposons une analyse phonétique des NON-REAL.LECT^{CP2} et NON-REAL.LECT^{CE1}, en considérant les trois catégories déjà bien décrites dans la section 9.5.2.2.3.1 : insertion d'un coup de glotte, continuum et pause. Nous traitons seulement la liaison et l'enchaînement (Tableau 98), car pour l'élision nous n'enregistrons pas de nombreuses non-réalisations.

Pour la liaison, nous remarquons la diminution d'insertion de coups de glotte et l'augmentation des insertions de véritables pauses entre Mot1 et Mot2. Nous observons que la même tendance est enregistrée pour l'enchaînement.

	% pour les séquences en liaison		% pour les séquences en enchaînement	
	CP2	CE1	CP2	CE1
Continuum	5,8	18	//	//
Coup de glotte	50	32	81,5	52,94
Pause	30	61	18,94	47,9

Tableau 98. Analyses phonétiques des lectures entre CP et CE1 pour la liaison et l'enchaînement

Nous pouvons donc conclure que, dès que les enfants apprennent à mieux lire, ils insèrent moins de coups de glotte entre Mot1 et Mot2. Ils préfèrent ne pas réaliser de sons entre les deux mots mais les lire séparément. Il y a deux explications possibles à cette diminution d'insertion des coups de glotte. Ces deux explications font référence à deux statuts différents de l'occlusive glottale :

1. Le coup de glotte représente une hésitation.

Les enfants au CP ont plus de problèmes dans la tâche de lecture. L'insertion d'une occlusive glottale entre Mot1 et Mot2 peut marquer une hésitation dans la lecture des suites Mot1 + Mot2.

Au CE1 la lecture devient plus fluente et rapide. Il y a moins d'hésitation et par conséquent moins de coups de glotte.

Dans cette explication, l'occlusive glottale n'aurait donc aucune valeur phonologique mais seulement une valeur phonétique.

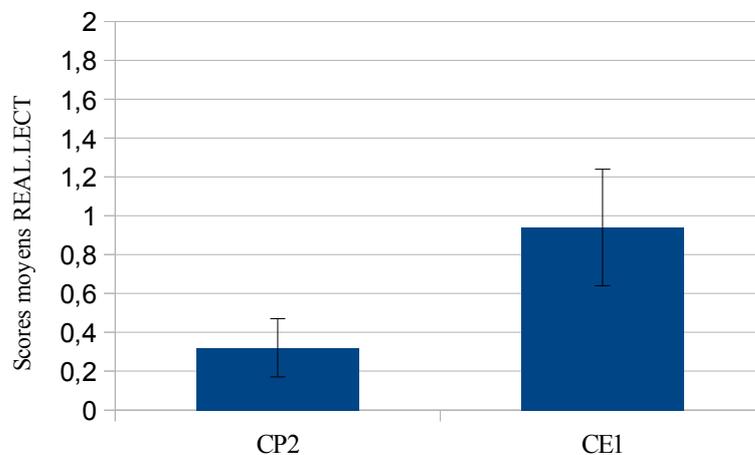
2. Le coup de glotte est l'attaque d'une syllabe CV.

Les enfants qui commencent à lire utilisent l'unité syllabique pour la lecture. Un Mot2 qui commence par une voyelle, comme dans nos séquences, est un mot à attaque vide caractérisé par un noyau vocalique seul. Les enfants au CP qui utilisent la syllabe pour lire, pour remplir l'attaque vide d'un mot qui commence par une voyelle et régulariser la structure syllabique V en CV, remplissent l'attaque vide avec une occlusive glottale. Plus la lecture progresse, plus les enfants sont capables de remplir l'attaque avec la consonne qui resyllabifie sur Mot2 et moins ils proposent d'insertions de coups de glotte. Cette analyse de l'occlusive glottale donne à cette production consonantique une valeur phonologique.

11.3.4.2.4 H aspiré et h muet en lecture

Comme nous l'avons déjà dit plusieurs fois dans les paragraphes précédents, dans les séquences où le Mot2 commence par *h* muet, les phénomènes de resyllabation ont lieu. En revanche, dans les séquences où Mot2 commence par *h* aspiré, les phénomènes de resyllabation ne sont pas réalisés. Nous nous attendons donc à une réalisation des phénomènes de resyllabation en lecture pour les séquences où Mot2 commençant par un *h* muet et une non-réalisation en lecture pour les séquences où Mot2 commençant par *h* aspiré (*un[Ø]hibou* [ɛ̃Øi.bu]).

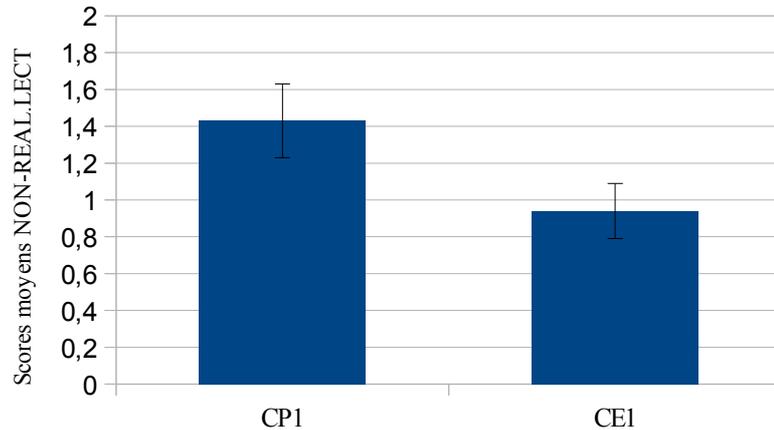
Nous comparons d'abord les résultats des réalisations pour les contextes à *h* muet initial en LECT^{CP2} et en LECT^{CE1} (*un[n]hippopotame* [ɛ̃.ni.po.po.tam]). Le t-test montre qu'il y a une différence statistiquement significative entre le nombre de REAL.LECT^{CP2} et REAL.LECT^{CE1} ($t_1(36) = 4,234$, $p < .001$, $t_2(1) = 2,15$, $p < .05$). Si nous regardons le graphique 36, nous pouvons dire qu'au CE1 les réalisations des phénomènes de resyllabation sont plus fréquentes pour ce contexte.



Graphique 36. Réalisations des phénomènes consonantiques en lecture entre CP et CE1

Nous avons comparé le nombre de NON-REAL.LECT^{CP2} et NON-REAL.LECT^{CE1} pour les séquences où le Mot2 commence par *h* aspiré (*un[Ø]hibou* [ɛ̃Øi.bu]). Il y a une différence statistiquement significative entre les non-réalisations à la fin du CP et au CE1 ($t_1(36) = 2,63$, $p < .05$, $t_2(1) = 3,55$, $p < .05$). Les non-réalisations diminuent au CE1 (Graphique 37). Les données nous suggèrent que les enfants ne

traitent pas différemment le contexte déterminant-nom à *h* aspiré initial et les autres contextes, car ils essaient de produire en lecture les phénomènes de resyllabation comme si Mot2 commençait par un *h* muet ou par une voyelle.



Graphique 37. Non-réalisations des phénomènes de resyllabation en lecture entre CP et CE1

11.3.4.2.4.1 L'analyse phonétique de NON-REAL.LECT^{CP2} et NON-REAL.LECT^{CE1} dans les contextes à *h* initial

Nous avons analysé phonétiquement les NON-REAL.LECT dans les contextes à *h* aspiré et *h* muet initiaux, pour arriver à bien décrire l'espace sonore entre Mot1 et Mot2 dans le cas où les phénomènes de resyllabation ne sont pas réalisés (*un[Ø]hibou* [ẽØi.bu], *un[Ø]hippopotame* [ẽØi.po.po.tam]).

Nous avons classifié les NON-REAL.LECT dans ces contextes selon les trois types de productions phonétiques que nous avons déjà vus (coup de glotte, pause, continuum). Les résultats sont bien montrés dans le tableau 99.

	% pour les séquences à <i>h</i> aspiré initial		% pour les séquences à <i>h</i> muet initial	
	CP2	CE1	CP2	CE1
Continuum	4,07	15,2	0	3,2
Coup de glotte	53,06	30,4	43,24	12,9
Pause	42,87	15,20	56,76	83,9

Tableau 99. Analyses phonétiques des lectures en contexte à *h* initial

Comme pour les autres phénomènes de resyllabation, le pourcentage de coups de glotte baisse entre CP et CE1 mais les productions de pauses augmentent. Pour l'explication, nous vous invitons à la lecture de la section 11.3.4.2.3.

11.3.4.2.5 La familiarité des contextes et la fréquence des items à lire

Nous avons essayé de voir si la familiarité avec un contexte à l'écrit influence la lecture et la réalisation correcte des phénomènes de resyllabation. La comparaison des résultats entre les contextes connus et les contextes inconnus (voir § 7.2.2 pour une classification) montre qu'au CP et au CE1, cette facilitation n'existe pas. La variable « familiarité » n'influence la lecture ni aux premières étapes de l'acquisition de la lecture, ni à un stade plus avancé.

Nous avons évalué le rôle de la fréquence en lecture en vérifiant deux types de corrélations :

1. Lectures correctes, où la lecture est globalement correcte et les phénomènes de resyllabation sont bien réalisés (ex. *les[z]ours* [le.zuʁs]) – fréquence
2. Erreurs en lecture (ERR.LECT, ex. *le(s)[Ø]ours* [leØuʁs], *le(s)[n]ours* [le.nuʁs], *les-[z]ours* [le/zuʁs]) – fréquence.

Nous prenons en considération les résultats de $LECT^{CP2}$ et $LECT^{CE1}$ et deux types de fréquences : la fréquence absolue (nombre d'occurrences) calculée dans notre base de données au moyen des cahiers des enfants (Fréq.) et la fréquence SFI dans MANULEX (au CP et au CE1 en accord avec le niveau scolaire). Les résultats sont résumés dans le tableau 100.

Corrélation	Indice	CP	Pos./Neg.	CE1	Pos./Neg.
Lectures correctes – fréquence	Fréq.	$r(19) = .45,$ $p < .05$	Positive	$r(19) = .45,$ $p < .05$	Positive
	SFI	$r(19) = .61,$ $p < .05$	Positive	$r(19) = .45,$ $p < .05$	Positive
ERR.LECT – fréquence	Fréq.	$r(19) = -.37,$ $p = .09$	//	$r(19) = -.60,$ $p = .80$	//
	SFI	$r(19) = -.53,$ $p < .05$	Négative	$r(19) = -.51,$ $p < .05$	Négative

Tableau 100. Corrélations entre la lecture et la fréquence au CP et au CE1

Nous pouvons voir que la fréquence (soit la fréquence absolue, soit le SFI) est en rapport avec les lectures correctes au CP et au CE1. Ce rapport est positif : les lectures et les réalisations des phénomènes de resyllabation sont meilleures quand les items à lire sont plus fréquents. De plus, le SFI influence les ERR.LECT au CP et au CE1. Le rapport entre ERR.LECT et la fréquence SFI est négatif. Cela veut dire que les Mots2 les plus fréquents sont associés à un nombre inférieur d'erreurs dans la réalisation des phénomènes de resyllabation en lecture.

Nous résumons les résultats des tâches d'écriture et de lecture au CP et au CE1.

- Amélioration des productions écrites au CE1.

Les enfants orthographient mieux les mots et réussissent à mieux identifier les frontières lexicales à l'écrit. Ils marquent plus fréquemment les graphèmes qui n'ont pas de réalisation sonore comme par exemple les <h> muets au début et les <s> à la fin de Mot2. Au CE1, les enfants utilisent de plus en plus les stratégies lexicales pour orthographier les mots. De plus, nous pouvons dire que les enfants commencent à insérer dans la représentation orthographique des traits morpho-syntaxiques.

- Amélioration des lectures au CE1.

Les enfants lisent mieux les mots et réussissent à mieux réaliser les phénomènes de resyllabation en enchaînant les items. Cependant, ils ne traitent pas différemment les contextes *h* muet et *h* aspiré. La différence de traitement présente à l'oral dans la littérature, n'est pas reproduite en lecture.

- Diminution des coups de glotte en lecture.

L'analyse phonétique des NON-REAL.LECT montre une diminution générale du nombre d'insertions de coups de glotte entre Mot1 et Mot2. Cela peut être expliqué par une amélioration dans la lecture qui implique moins d'hésitations ou un remplissage de l'attaque du Mot2 sans que le bon segment soit réalisé.

- Rôle de la familiarité et de la fréquence.

Pour l'écriture, la familiarité avec le contexte Mot1 + Mot2 influence les productions écrites au CP. Les contextes connus sont mieux orthographiés que les contextes inconnus.

La fréquence des items a un effet seulement pour la lecture. La fréquence du Mot2 à écrire n'influence pas les productions écrites des enfants : un Mot2 plus fréquent n'implique pas des productions écrites plus justes. Pour la lecture, au CP comme au CE1, la fréquence a une influence positive sur le nombre de lectures correctes et également sur la réalisation des phénomènes de resyllabation. Plus le Mot2 est fréquent, plus il est facile à lire et à enchaîner avec les autres.

11.3.5 Comparaison des corrélations

Dans cette section, nous cherchons à mettre en rapport les indices de corrélation que nous avons calculés pour la fin du CP et pour le CE1 (§ 9.6 et § 10.4). Cela nous permettra de comprendre l'évolution et les enjeux qui existent entre la *literacy* et la phonologie en ce qui concerne cette tranche d'âge et les phénomènes de resyllabation au centre de notre étude.

11.3.5.1 Rapport entre la *literacy* et les productions orales

Nous traiterons d'abord les corrélations qui mettent en rapport les productions écrites et orales. Ensuite, nous nous concentrerons sur les rapports entre la lecture et la production orale pour enfin établir la relation entre l'écriture et la lecture. Pour réaliser ces comparaisons, nous proposerons l'analyse de l'indice r de Pearson au CP et au CE1.

11.3.5.1.1 Les productions écrites et les productions orales

Dans ces premières comparaisons, nous prenons en considération les productions écrites correctes (« frontières lexicales » dans PERC, ex. <les affiches>) respectant les frontières lexicales, hors des erreurs d'orthographe à l'intérieur du mot (ex. <une otruche> pour *une autruche*). Ces productions sont liées au niveau de connaissance des frontières de mots à l'écrit chez les enfants. Nous avons établi trois relations avec cette variable :

1. Les productions écrites correctes dans PROD.ECR (« frontières lexicales » <les affiches>) et les non-réalisations à l'oral dans DENa (*le(s)[Ø]affiches*) ;
2. Les productions écrites correctes dans PROD.ECR (« frontières lexicales » <les affiches>) et les productions correctes à l'oral avec la réalisation des phénomènes de resyllabation DENa (*le(s)[z]affiches*) ;
3. Les productions écrites correctes (« frontières lexicales » <les ours>) dans PROD.ECR et les insertions d'une consonne erronée dans DENa (*le(s)[n]ours* [le.nuʁs])

Le tableau 101 présente l'indice de corrélation pour ces deux types de rapport dans les deux classes.

	CP2	CE1
Frontières lexicales - Non-réalisations à l'oral	r(41) = .16	r(36) = -.546* *
Frontières lexicales - Productions correctes à l'oral	r(41) = .440**	r(36) = .43*
Frontières lexicales - Insertions d'une consonne erronée à l'oral	r(41) = .40	r(36) = .21

**p<.001, *p<.05

Tableau 101. Corrélations écrit-oral entre CP et CE1

Comme on peut le voir dans le tableau, le rapport entre les productions écrites correctes et les productions orales correctes reste stable entre CP et CE1. En revanche, le rapport entre les productions écrites à « frontières lexicales » et les non-réalisations à l'oral s'établit seulement au CE1 et il a une valeur négative. Pendant le CE1, les frontières entre mots dans des séquences resyllabifiées deviennent plus claires et stables à l'écrit. Une meilleure connaissance du découpage en mot à l'écrit influence

l'exactitude des productions orales et corrèle à une baisse de non-réalisations à l'oral. Les erreurs d'insertion d'une consonne erronée ne sont pas en rapport avec les productions écrites.

11.3.5.1.2 Le marquage du pluriel à l'écrit et la réalisation correcte de la liaison en [z]

Pour les séquences en liaison, nous pouvons penser que la connaissance du <s> à la fin des mots, pour marquer orthographiquement le pluriel, peut influencer la production correcte de liaison en [z] à l'oral. Au CP et au CE1, nous avons établi une corrélation entre le marquage du Mot2 au pluriel en PERC et les réalisations correctes des liaisons en [z] dans les productions orales pour DENa. Si nous regardons les résultats dans le tableau 102, nous pouvons voir qu'il n'existe pas de corrélations entre ces deux variables au CP et au CE1.

	CP	CE1
Marquage à l'écrit-réalisation de la liaison en [z]	$r(41) = .11$	$r(36) = .16$

Tableau 102. Rapport entre oral et écrit pour le marquage du pluriel

Nos données montrent que le marquage du pluriel n'est pas lié à la réalisation de la liaison entre déterminant et nom à ce stade de l'acquisition.

11.3.5.1.3 La lecture et la production orale

L'autre composante de la *literacy* que nous avons testée est la lecture. Comment évolue le rapport entre la lecture et la production orale au cours de notre étude longitudinale ? Nous considérons ici trois rapports que nous avons déjà présentés dans les chapitres 9 et 10 :

1. Les non-réalisations à l'oral dans DENa (ex. *le(s)[Ø]ours* [leØuʁs]) et les non-réalisations en LECT (ex. *le(s)[Ø]ours* [leØuʁs]), où les phénomènes de resyllabation ne sont pas réalisés.

2. Les réalisations des phénomènes consonantiques en LECT (REAL.LECT, ex. *le(s)[z]ours* [le.zuʁs]) et productions orales correctes dans DENa (ex. *le(s)[z]ours* [le.zuʁs]), où les phénomènes de resyllabation sont bien réalisés.
3. Les insertions d'une consonne erronée dans DENa (ex. *le(s)[n]ours* [le.nuʁs]) et les lectures correctes dans LECT (REAL.LECT, ex. *les[z]ours* [le.zuʁs]).

	CP	CE1
Non-réalisations à l'oral - NON-REAL.LECT	r(40) = 0,20	r(36) = 0,02
Productions orales correctes - REAL.LECT	r(40) = .47*	r(36) = .39*
Insertion d'une consonne erronée à l'oral - REAL.LECT	r(40) = .37	r(36) = .018

*p<.05

Tableau 103. Rapports productions orales et lectures au CP et au CE1

Dans le tableau 103, nous pouvons voir qu'il y a une corrélation positive entre les REAL.LECT et les productions orales correctes au CP et au CE1. En revanche nous ne pouvons établir de corrélations entre les non-réalisations à l'oral et en lecture et les insertions d'une consonne erronée et les lectures correctes ni au CP, ni au CE1. Les enfants qui réussissent à bien enchaîner deux mots en lecture et à réaliser les phénomènes de resyllabation sont aussi capables de bien produire les deux mots enchaînés à l'oral, en dénomination guidée. De plus les lectures correctes ne provoquent pas la diminution des insertions d'une consonne erronée à l'oral. Les influences réciproques entre lecture et dénomination guidée ne changent pas selon le niveau scolaire.

11.3.5.1.4 Le rapport entre la lecture et l'écriture au CP et au CE1

Nous vérifions l'évolution des trois corrélations qui mettent en rapport l'écriture et la lecture à la fin du CP (§ 9.5.3) et au CE1 (§ 10.3.3) :

1. Les non-réalisations en lecture (NON-REAL.LECT, ex. *le(s)[Ø]ours* [leØuʁs]) et les productions écrites correctes à « frontières lexicales » où les frontières des mots sont bien identifiées (ex. <les ours>) ;

2. Les productions écrites à « frontières syllabiques » (ex. <les zaffiches>) et les lectures syllabiques (ex. [le/za/fiʃ]) ;
3. Les productions écrites correctes où les frontières sont bien identifiées (à « frontières lexicales », ex. <les ours>) et les lectures correctes où les phénomènes de resyllabation sont bien réalisés (REAL.LECT, ex. les[z]ours [le.zuʁs]).

	CP	CE1
NON-REAL.LECT – frontières lexicales	r(40) = -.21	r(36) = -.21
Frontières syllabiques - lectures syllabiques	r(40) = .96*	r(36) = .07
Frontières lexicales – REAL.LECT	r(40) = .39*	r(36) = .03

*p<.05

Tableau 104. Rapport entre l'écriture et la lecture au CP et au CE1

Comme nous pouvons le voir dans le tableau 104, nous avons des corrélations statistiquement significatives seulement au CP. À ce stade de l'acquisition de la *literacy*, l'écriture et la lecture s'influencent réciproquement : les enfants qui arrivent à mieux identifier les frontières des mots à l'écrit peuvent aussi bien les enchaîner pour les lire ; les enfants qui utilisent encore la syllabe pour lire les séquences des mots, l'utilisent aussi pour écrire les séquences resyllabifiées au moyen de productions écrites qui respectent les frontières syllabiques (<les sours>). Cependant, ce rapport disparaît au CE1. À ce moment de la scolarisation, la lecture et l'écriture doivent être traitées comme deux activités différentes qui ne s'influencent pas réciproquement.

11.3.6 Les autres variables analysées

11.3.6.1 La variable sociolinguistique

Nous avons essayé de vérifier si le niveau scolaire des parents et leur milieu socioculturel influencent les productions orales de liaisons chez les enfants pendant le CP et le CE1. Les résultats montrent que ces deux variables influencent seulement les

résultats au début du CP (CP1). Au début du CP, les enfants avec des parents « cadre » ont de meilleurs résultats en dénomination guidée. Cependant, cette tendance n'est plus observée dès la fin du CP.

11.3.6.2 La variable « lecture partagée »

Nous avons cherché à évaluer si la fréquence de lectures partagées entre les enfants et les parents peut être mise en rapport avec les résultats des tâches orales et écrites au CP (CP1 et CP2) et au CE1. Que ce soit pour les tâches orales ou pour les tâches liées à la langue écrite, à chaque moment de recueil des données, il n'y a pas de corrélation entre les résultats et cette variable. En ce qui concerne les phénomènes de resyllabation, la fréquence de lecture partagée ne semble pas influencer la bonne segmentation en mots à l'écrit ou la production correcte en dénomination ou en lecture.

À partir de ces résultats de l'étude longitudinale, dans le chapitre suivant, nous discuterons les questions de recherche et nous chercherons à valider les hypothèses présentées dans le chapitre 6. Nous proposons aussi une section où nous présenterons des perspectives futures pour continuer notre recherche.

Partie 4.

Discussion et conclusions

Chapitre 12

Discussion des résultats et vérification des hypothèses

12.1 Les données sur la langue orale après 6 ans

Les recherches sur l'acquisition de la liaison à l'oral chez les enfants francophones entre 2 et 5 ans, présentées dans le chapitre 4, enregistrent des productions erronées qui ont été appelées omissions ou non-réalisations (*u(n)[Ø]avion* [ẽ.a.vjõ] pour *un[n]avion* [ẽ.na.vjõ]) et insertions d'une consonne erronée ou remplacements (*de(s)[n]éléphants* [de.ne.le.fã] pour *des[z]éléphants* [de.ze.le.fã]) (Chevrot & Fayol, 2001 ; Wauquier & Braud, 2005 ; Wauquier, 2009 ; Nardy & Dugua, 2011 ; Wauquier & Shoemaker, 2013 ; Chevrot *et al.*, 2009, 2001, 2005, 2013 ; Dugua, 2002). Les études développementales (longitudinales ou transversales) observent une diminution de ces erreurs entre 4 et 5 ans (Basset, 2000 ; Dugua & Chabanal, 2006). Après l'âge de 5 ans, seules les erreurs de non-réalisations subsistent à l'oral (Chevrot *et al.* 2013 ; Nardy & Dugua, 2011).

Dans cette thèse, nous avons cherché à décrire l'acquisition de la liaison et des deux autres phénomènes de resyllabation (élision et enchaînement) chez un groupe d'enfants francophones après l'âge de 5 ans. Dans ce but, nous avons proposé une étude longitudinale d'une durée de 2 ans, pendant laquelle nous avons testé le même groupe d'enfants à trois moments différents : au début du CP (CP1), à la fin du CP (CP2) et au CE1. Pour recueillir des données sur la langue orale, nous avons proposé deux tâches orales :

- une tâche de perception d'erreurs et reformulation (PERC) ;
- une tâche de dénomination guidée (DENa).

Ces deux tâches ont permis de tester la production et la compréhension à l'oral des trois phénomènes de resyllabation au centre de l'étude. Contrairement aux

observations présentées par les recherches précédentes, nos résultats montrent que les enfants ont encore des difficultés à traiter les frontières de mots à l'oral après 5 ans. Au début et à la fin du CP, il est encore très difficile pour eux de juger comme erronée une séquence où le phénomène de resyllabation n'est pas produit (pour la liaison *les[Ø]ours* [leØuʁs], pour l'élision *la-armoire* [la.aʁ.mwaʁ], pour l'enchaînement *une-ambulance* [yn.ɑ̃.by.lãs]), indépendamment du contexte syntaxique. De plus, au CP et au CE1 les productions orales des enfants sont encore caractérisées par des insertions de consonne erronée entre Mot1 et Mot2 (pour la liaison *un[z]ours*, pour l'élision *la[n]oie* et pour l'enchaînement *la petite[z]orange*), mais surtout par de nombreuses non-réalisations (pour la liaison *les[Ø]ours*, pour l'élision *le-orage* [lə.o.ʁaʒ], pour l'enchaînement *la petite-orange* [la.pə.tit.o.ʁãʒ]).

La comparaison des non-réalisations et des insertions de consonne erronée dans les productions orales aux trois moments permet d'identifier deux patterns développementaux :

1. le nombre de non-réalisations reste stable pendant le CP mais diminue au CE1.
2. le nombre d'insertions d'une consonne erronée diminue drastiquement pendant le CP. Le peu d'erreurs qui restent subsistent au CE1.

Nous avons vu que la liaison en DENa est le seul phénomène que les enfants réalisent sans être influencés par la procédure de la tâche (§ 9.4.3). Pour la liaison nous observons le même pattern développemental pour les non-réalisations (stabilité au CP, diminution au CE1). Pour les insertions de consonnes erronées nous notons une diminution en nombre au CP et au CE1. À la fin de l'année scolaire ces erreurs sont très rares. Nous avons cherché à à établir un lien entre ces deux patterns développementaux et l'acquisition de la *literacy* dans la section 12.8.

Aux trois moments (CP1, CP2 et CE1), les non-réalisations restent les plus fréquemment produites à l'oral et non-identifiées pendant la tâche de jugement. Ces données sont en accord avec les données sur la liaison présentées en littérature (Chevrot *et al.*, 2013 ; Nardy & Dugua, 2001) qui montrent une subsistance des non-réalisations à cet âge.

12.2 La variable sociolinguistique

Nous avons analysé au CP les productions orales de liaisons obligatoires/catégoriques des enfants en rapport avec le niveau socioculturel des parents⁶⁵. Les résultats montrent qu'au début du CP, « les enfants de cadres » produisent moins d'erreurs que les autres. Ces données confirment les résultats de Nardy & Dugua (2001) qui remarquent une diminution des erreurs sur la liaison adjectif-nom après 5 ans chez « les enfants de cadres ». Chez les enfants que nous avons testés, cette différence disparaît à la fin du CP. Pendant les premières années de l'école primaire, les enfants passent beaucoup de temps en classe (environ 40 heures par semaine), où ils sont exposés au même input (guidé par la maîtresse). Nous pouvons penser que l'input oral qui est perçu pendant l'année scolaire est plus homogène et formel que celui perçu à la maison avec leurs parents. Par conséquent ils sont moins exposés à la variante sociolinguistique de la langue parlée chez eux.

De plus, la variante sociolinguistique influence les productions orales des phénomènes consonantiques dans les contextes catégoriques seulement avant l'acquisition de l'écriture et de la lecture. La disparition de l'influence de cette variable, sur les productions, peut être liée à l'apprentissage de la *literacy*.

Les recherches qui font émerger l'importance de la composante sociolinguistique sur les productions orales décrites dans le chapitre 8 font référence au phénomène de liaison non-catégorique ou au contexte adjectif-nom. En revanche, nos résultats concernent la liaison catégorique en contexte déterminant-nom. Nous pouvons penser que dans la langue adulte, il y aurait une variation impliquée par le niveau socioculturel qui influencerait l'input de l'adulte et à son tour les productions des enfants. Cependant, à notre connaissance, aucune étude ne montre de différences dans les taux de réalisations de la liaison en contexte déterminant-nom influencées par cette variable. D'autres études doivent être conduites pour approfondir la compréhension que nous avons de ces phénomènes sociolinguistiques en acquisition.

65 Nous invitons à la lecture des chapitre 8 et 9 pour comprendre les buts descriptifs de cette partie de la recherche

12.3 Les données sur la langue écrite

Pour recueillir des données sur la langue écrite, nous avons testé les enfants à la fin du CP et au CE1 au moyen de deux tâches liées à la langue écrite :

- une tâche d'écriture de séquences de mots en liaison, en élision et en enchaînement (PROD.ECR)
- une tâche de lecture à voix haute de séquences de mots en liaison, en élision et en enchaînement (LECT).

Ces données nous ont fourni des informations sur les stratégies de lecture et d'écriture utilisées dans les premières années de l'école primaire et nous ont permis d'enrichir le panorama sur l'apprentissage du français écrit.

Nous remarquons d'abord que les connaissances de *literacy* s'améliorent entre le CP et le CE1. À l'écrit les enfants réussissent à mieux orthographier les mots et à identifier plus facilement les frontières des mots, même en présence de phénomènes de resyllabation. En lecture, les enfants produisent moins de lectures syllabiques (ex. *le-le-phant*) et ont moins de difficultés à enchaîner Mot1 et Mot2.

Nos données nous permettent de remarquer, surtout à la fin du CP, l'utilisation d'une stratégie phonologique pour lire et écrire. Les enfants utilisent la langue orale et ses unités pour établir les correspondances graphème-phonème et orthographier les mots. Pour cette raison, les mots les plus inconsistants sont les plus difficiles à écrire. Ces résultats sont en accord avec les études concernant la *literacy* et l'importance de la stratégie phonologique au début de l'acquisition du français écrit (Sprenger-Charolles & Serniclaes, 2003 ; Ziegler & Goswami, 2005). Comme montré par Sprenger-Charolles *et al.* (1997), nous observons également que la stratégie phonologique est déjà utilisée en parallèle avec la stratégie lexicale au CP. Cette dernière est de plus en plus utilisée au CE1 où les enfants commencent à orthographier fréquemment les graphèmes non-réalisés à l'oral (le <s> à la fin de <oranges>, le <h> de <hippopotame>). Nous remarquons également un effet de la fréquence pour la lecture au CP et au CE1 qui est symptomatique de la mise en œuvre d'une stratégie lexicale. Cependant, cet effet n'est pas enregistré pour l'écriture à aucun des deux moments. Pour passer de l'oral à l'écrit les enfants adoptent encore une stratégie phonologique.

En ce qui concerne la représentation des frontières des mots dans des séquences caractérisées par des phénomènes de resyllabation à l'oral, la littérature observe de nombreuses productions écrites qui respectent les frontières syllabiques, du type <un

nage> pour *un ange* (Soum, 1997 ; Sabio, 2000 ; Soum-Favaro, 2014). En revanche, nous enregistrons plus de productions écrites correctes (<les ours>, <une ambulance>, <l'éléphant>) que de productions écrites respectant les frontières syllabiques (<les zours>, <une nambulance>, <le léléphant>), au CP comme au CE1.

Au CP, nous remarquons beaucoup de non-réalisations des phénomènes de resyllabation en lecture (NON-REAL.LECT, ex. *les[Ø]affiches*, pour *les[z]affiches*). Elles diminuent au CE1 en faveur des réalisations (REAL.LECT, ex. *les[z]affiches*), mais restent tout de même nombreuses. Nous allons les analyser phonétiquement dans la section 12.4.

Nous avons vu qu'en France, les enseignants utilisent deux méthodes pour enseigner la langue écrite : la méthode globale et la méthode alphabétique. Les résultats que nous avons présentés ne sont pas influencés par la méthode car, comme nous l'avons dit dans le chapitre 7, les maîtresses adoptent une méthode mixte qui combine les deux méthodes.

12.4 L'analyse phonétique des non-réalisations en lecture

Nous avons analysé phonétiquement l'espace sonore entre Mot1 et Mot2 quand les enfants ne produisent pas les phénomènes consonantiques attendus. L'analyse phonétique est centrée sur les non-réalisations des liaisons (*le(s)[Ø]affiches* [le.a.fi]) et des enchaînements (*une-ambulance* [yn.ã.by.lãs]) car nous n'avons pas de non-réalisations en élision. Nous remarquons 3 types de productions possibles communes à ces deux séquences : l'insertion d'un coup de glotte ([le.ʔa.fi], [yn.ʔã.by.lãs]) et la production d'une véritable pause entre les deux ([le/a.fi], [yn/ã.by.lãs]).

En général à la fin du CP, les enfants produisent plutôt des coups de glotte entre Mot1 et Mot2. La production de coups de glotte diminue entre le CP et le CE1 mais la production de pauses augmente. L'augmentation de la production de pauses est clairement motivée par le repérage de l'espace blanc séparant les mots à l'écrit. La diminution des réalisations de coups de glotte peut avoir deux explications qui font référence à deux natures possibles de l'occlusive glottale :

1. L'occlusive glottale est une production phonétique et représente une hésitation dans la tâche de lecture.

2. L'occlusive glottale a une valeur phonologique et elle est associée à une position sous-jacente qui fait que l'attaque du Mot2 est remplie et que le début du mot montre une structure syllabique CV, qui guide les enfants dans le processus de lecture.

Dans les deux cas, la chute des productions d'occlusives glottales peut être expliquée par une approche plus lexicale que phonétique-phonologique de la lecture. Les formes orthographiques des Mots1 et Mots2 mieux acquises au CE1, guident la lecture : l'espace blanc (donc l'absence de graphèmes) entre les deux est traduit à l'oral au moyen d'une pause (absence de sons).

12.5 La lecture et l'écriture : quel rapport ?

Dans les chapitres 9 et 10, nous avons essayé d'établir plusieurs relations entre les productions écrites et la lecture. Les compétences de lecture et d'écriture semblent entrer en relation seulement au CP. À ce moment, nous observons une corrélation très significative entre les productions écrites à « frontières syllabiques » (<les zaffiches>) et les lectures syllabiques (*le-[z]a-fi-che*). Les enfants qui ont le plus de difficultés à déterminer les frontières des mots à l'écrit ont également des difficultés à lire correctement les séquences. Comme ils utilisent la syllabe pour lire, ils l'utilisent aussi pour écrire et identifier les frontières des mots. De plus, pendant la première année de l'école primaire, nous observons aussi que les enfants qui réussissent plus facilement à représenter les frontières des mots à l'écrit réussissent aussi à réaliser le plus de lectures correctes. Le fait que les niveaux de lecture et d'écriture soient corrélés ne doit pas faire penser que ces deux activités sont équivalentes. Comme Cossu & Marshall (1985) l'ont souligné, elles font référence à deux processus cognitifs distincts. Cette différence est visible en analysant le rôle de la fréquence qui a un effet sur la lecture exclusivement (§ 9.5.2.2.4) et de la familiarité des contextes qui, au contraire, a un impact sur les productions écrites au CP uniquement.

12.6 Les noms à *h* initial

Dans la tâche d'écriture, réussir à orthographier les *h* initiaux au début du Mot2 est plus facile au CE1 qu'au CP. Les performances en lecture montrent qu'au CE1 les réalisations augmentent pour les séquences à *h* muet initial. Pour les séquences à *h* aspiré initial le nombre de non-réalisations diminue au CE1.

Nous observons qu'au CE1 les enfants réalisent plus de resyllabations en présence des Mots2 à *h* aspiré et à *h* muet initiaux. Conformément à la tendance générale qui implique une augmentation des réalisations des phénomènes de resyllabation en lecture, les enfants enchaînent Mot1 et Mot2 indistinctement.

Au chapitre 9, nous avons essayé d'établir un rapport entre les productions écrites et les lectures pour les contextes *h* muet et *h* aspiré. Nous faisons l'hypothèse que les enfants qui orthographient une consonne <h> à l'écrit sont aussi les enfants qui produisent une consonne glottale au début du Mot2 en cas de non-réalisation du phénomène de resyllabation. Les résultats nous montrent qu'il n'y a pas de corrélation entre l'ajout d'un coup de glotte en lecture et l'orthographe correcte du <h> au début du nom.

À partir de nos données, nous avons l'impression que les enfants ne font pas vraiment la différence entre les contextes à *h* muet et à *h* aspiré initiaux dans les deux premières années de l'école primaire. Ces données sont confirmées par les résultats de notre tâche de perception (PERC). Nous avons vu que, dans la plupart des cas, les enfants considèrent comme fautive une phrase caractérisée par la non-réalisation de la liaison en présence de *h* aspiré. Ils corrigent en réalisant la liaison ce qui pour eux est une erreur.

Nous pouvons conclure qu'au CP et au CE1, il n'y a pas de véritable différence dans le traitement de ces deux contextes, en ce qui concerne la réalisation phonologique et en ce qui concerne l'écrit et la lecture.

12.7 Le rôle de la lecture partagée

Nous avons analysé le rapport entre les compétences en langue écrite ou orale des enfants et la fréquence des activités de lecture partagée avec leurs parents. Comme nous l'avons vu dans le chapitre 5, les enfants qui partagent avec leurs parents plus de

moments de lecture montrent de meilleurs résultats en écriture et en lecture et dans des tâches métalinguistiques (Ecalte & Magnan, 2008 ; Goswami, 2001 ; Bertelson *et al.*, 1985 ; Casalis & Louis-Alexandre, 2000 ; Ecalte, 2000 ; Bradley & Bryant, 1978 ; Goswami & Bryant, 1990 ; Lecocq, 1986 ; Lundberg, 1982 ; Lundberg *et al.*, 1980 ; Morais *et al.*, 1979 ; Perfetti, 1985 ; Shankweiler, 1999 ; Stanovich, 1986 ; Stanovich & Siegel, 1994). Pourtant, nos données montrent que nous ne pouvons pas établir de corrélation entre les niveaux de lecture et d'écriture des enfants et la variable lecture partagée au CP et au CE1 en ce qui concerne les phénomènes de resyllabation. La fréquence de la lecture partagée n'influence pas les productions orales des enfants non plus.

L'absence d'un effet de la fréquence de lecture partagée peut être liée au fait que les enfants sont testés après l'entrée à l'école, c'est-à-dire au CP. Les heures d'apprentissage explicite en classe et à la maison deviennent les moments prédominants dans l'acquisition de la langue et empêchent d'observer un effet de la fréquence de la lecture partagée.

De plus, nous pouvons penser que lire et écrire correctement les phénomènes de resyllabation est un processus compliqué qui a besoin d'un apprentissage explicite de la langue écrite et d'une réflexion métalinguistique consciente. L'absence d'un effet de fréquence de la lecture partagée sur la lecture et l'écriture serait donc expliquée par la nature de ces phénomènes qui impliquent une incongruence entre la production orale et la représentation orthographique. Nous ne nions pas l'effet de la fréquence de la lecture partagée sur la *literacy* et sur les compétences orales, en général⁶⁶, mais nous pouvons dire que pour le cas particulier du phénomène de resyllabation, cet effet n'est pas effectif.

Dans nos données, nous ne remarquons pas de différence de lecture partagée selon les milieux socioculturels et le niveau scolaire des parents (§ 8.5). Il faut souligner que les données récoltées sont limitées aux moments de lecture partagée entre l'enfant et les parents et ne prennent pas en compte les moments de lecture avec l'entourage familial ou les professionnels qui gardent l'enfant. De plus, les données sur la lecture partagée ont été récoltées au moyen d'un questionnaire complété par les parents. Ces données peuvent manquer de fiabilité, comme cela est bien expliqué par Pierrehumbert (2003 : 178) : « Toutefois l'inconvénient évident du questionnaire

66 Dans le chapitre 7 nous avons vu que la fréquence des lectures partagées est liée à les compétences de base d'écriture et de lecture testées au moyen du livret de passation à la rentrée.

rempli par les parents est la subjectivité, ou, en termes plus techniques la « désirabilité sociale ». Selon la conception qu'il a de son rôle, son degré de confiance envers les chercheurs, ses connaissances dans l'objet d'étude, pour ne pas parler de sa propre histoire, le parent peut en effet vouloir restituer une image « désirable »[...]». Des parents qui passent plus de temps avec leurs enfants et qui partagent plus de moments de lecture avec eux peuvent être vus socialement comme plus « désirables ». Pour cette raison, la « désirabilité » peut avoir influencé les réponses au questionnaire sur la fréquence de la lecture partagée. Par conséquent, nous ne pouvons pas faire confiance à 100 % à ces données et aux résultats obtenus dans les analyses qui les utilisent.

12.8 Discussion des questions de recherche et vérification des hypothèses

Nous discutons les questions de recherche et les hypothèses présentées au chapitre 6.

- **Q1 : quel modèle de l'acquisition de la liaison décrit le mieux le découpage en mots après 5 ans ?**

Nous avons vu que pour décrire l'acquisition des frontières de mots en français, deux modèles étaient proposés (§ 4.2.3.1) : le modèle lexical-constructionniste, qui souligne l'importance de la fréquence de mots dans le lexique pour l'acquisition de leurs frontières (Chevrot *et al.*, 2013) ; le modèle morpho-phonologique, qui met en rapport l'acquisition des frontières des mots avec des compétences morpho-phonologiques (Wauquier et Braud, 2005). Dans cette thèse, nous essayons de bien comprendre lequel de ces deux modèles est le plus adapté pour décrire l'acquisition de la liaison et, en général, des frontières de mots en français, après l'âge de 5 ans.

Comme nous l'avons déjà dit, la fréquence a une importance notable dans le modèle lexical-constructionniste (Bybee, 2005). Dans cette étude, nous avons mis en rapport les productions orales des enfants de 6 à 7 ans avec la fréquence des mots dans leur lexique en utilisant l'indice de fréquence SFI dans MANULEX. Nous avons cherché à voir s'il y a une relation entre un Mot2 plus fréquent au pluriel et l'insertion de consonne erronée [z] au début de ce mot. À aucun des trois moments (CP1, CP2 et CE1), une relation entre la fréquence des items et la production de ces erreurs n'a été observée. Ces données sont en

accord avec les résultats de Dugua *et al.* (2009), qui révèlent qu'après 6 ans, la fréquence des items à produire n'influence plus les productions. Quel modèle adopter alors pour décrire les productions des enfants après 5 ans ? Pour comprendre si nous pouvons utiliser le modèle morpho-phonologique, nous cherchons à valider l'existence qu'il propose d'une position sous-jacente entre Mot1 et Mot2. Comme nous l'avons vu à la fin du chapitre 8, l'analyse phonétique des productions d'élosion et d'enchaînement révèle une position prosodique sous-jacente dans la représentation de Mot1 + Mot2 en liaison et valide le double-flottement de la CL.

À la lumière de ces résultats, nous pouvons pencher pour une acquisition des frontières de mots centrée sur un modèle représentationnel. Cependant, le modèle morpho-phonologique jusqu'à maintenant ne propose pas une représentation pour les non-réalisations qui subsistent dans la phase de l'acquisition que nous avons étudiée. De plus, il ne considère pas une variable très importante qui influence l'acquisition des mots en français : la *literacy*. Nous avons donc décidé de développer le modèle morpho-phonologique dans la section 12.9, pour essayer de donner une représentation des non-réalisations et d'intégrer les connaissances de la langue écrite. Avant de passer à la modélisation, nous cherchons à mieux comprendre les influences de la lecture et de l'écriture sur les productions orales, en développant la question de recherche Q2.

◆ **Q2 : quelle est l'influence de l'apprentissage de l'orthographe sur la production orale et sur le découpage en mots ?**

Quand les enfants apprennent à écrire, les séquences entendues à l'oral prennent forme sur papier. À l'écrit, les frontières des mots sont bien délimitées par des symboles graphiques comme des espaces blancs (forme orthographique *les ours* [le.zuʁs]) ou des apostrophes (*l'éléphant*). De plus, dans les séquences en liaison, la CL a une véritable représentation à l'écrit en tant que graphème attaché au mot de gauche (Mot1). Nous avons vu que l'apprentissage de l'écriture d'une langue influence les compétences orales de cette langue (chapitre 2). L'apprentissage de l'orthographe chez les enfants français au début de l'école primaire peut donc être vu comme un facteur qui facilite la

segmentation en mots à l'oral, car elle permet de fixer sur papier les frontières lexicales des mots.

Pour vérifier si connaître les formes écrites a un effet sur la langue orale, nous proposons une comparaison entre les résultats des tâches orales et des tâches liées à la langue écrite (lecture et écriture) au CP et au CE1. Nous nous attendions à vérifier deux hypothèses.

- **Hypothèse 1 (H1) : la disparition des erreurs de remplacement ou d'insertion de consonnes erronées**

Dans H1, nous proposons que la disparition des insertions de consonnes erronées à l'oral est liée à l'apprentissage de la *literacy*. Nos données, tous phénomènes confondus, montrent que les erreurs de consonne erronée diminuent pendant le CP mais restent stables au CE1. Pour la liaison, les erreurs d'insertion d'une consonne erronée diminuent significativement pendant le CP et deviennent très rares au CE1.

Nous avons cherché à évaluer si l'apprentissage de la lecture et de l'écriture pouvait être lié à ce type de productions. Cependant, à aucun des deux moments, nous ne pouvons établir de lien entre l'insertion d'une consonne erronée et la *literacy*. Les activités de lecture et d'écriture ne favorisent pas la diminution ou la stabilité de ces types de productions. Une autre explication qui ne concerne pas la *literacy* doit donc être apportée.

Les données récoltées ne permettent pas de donner une explication précise, mais permettent d'affirmer que les insertions de consonne erronée disparaissent plus facilement que les non-réalisations et, étant donné le nombre très réduit de ces productions qui restent stables au CE1, que celles qui restent deviennent des résidus sédimentés dans la langue orale pour certains enfants.

- **Hypothèse 2 (H2) : la stabilisation des non-réalisations**

L'hypothèse 2 (H2) suppose la stabilité des non-réalisations. Pour cette hypothèse, l'apprentissage des formes écrites des mots aurait une influence négative sur les productions orales : apprendre Mot1 + Mot2 comme deux formes séparées à l'écrit impliquerait de les produire séparément, sans réaliser les phénomènes de resyllabation à l'oral. Par conséquent, nous nous

attendions à observer une stabilité des non-réalisations pendant les premières années de l'école.

Nos résultats montrent que les non-réalisations restent stables pendant le CP mais baissent au CE1 pour tous les phénomènes. Cette tendance a été enregistrée également pour la liaison.

À la fin du CP (CP2), quand les non-réalisations à l'oral restent stables, elles ne sont pas en rapport avec l'identification des frontières des mots à l'écrit (§ 9.6). En revanche, cette relation est mise en place quand elles baissent, au CE1. À ce moment, les lectures justes (où Mot1 et Mot2 sont correctement enchaînés) et les productions écrites correctes (où les frontières des Mots1 et Mots2 sont bien identifiées, ex. <les ours>) sont associées à un nombre inférieur de non-réalisations. La diminution des non-réalisations est donc liée à de meilleures compétences en lecture et en écriture qui auraient une influence exclusivement positive sur les productions orales.

De plus, au CP comme au CE1, nous remarquons qu'identifier les frontières des mots à l'écrit et les enchaîner correctement en lecture implique des productions orales plus justes. Ces données permettent d'infirmer l'hypothèse H2 : l'apprentissage de la *literacy* n'a pas d'effet défavorable sur la production orale.

Comme nous l'avons déjà dit à la fin du chapitre 9, nous ne pouvons pas établir avec certitude la direction de la relation entre la *literacy* et la production orale. Nos données ne permettent pas de dire si ce sont les compétences de la langue écrite qui influencent l'oral et la phonologie ou vice-versa. Cependant, elles permettent de souligner le rôle fondamental de la *literacy* dans l'acquisition des phénomènes de resyllabation. Nous chercherons à intégrer cette composante au modèle de l'acquisition de la liaison ci-après.

○ **Hypothèse 3 (H3) : une relation entre les connaissances morphologiques à l'écrit et les productions orales**

Dans H3, nous faisons l'hypothèse que l'apprentissage du marquage du pluriel à l'écrit (le <s> à la fin du Mot1 et Mot2) implique des productions correctes de liaison au pluriel (liaison en [z]). Les données récoltées au CP

et au CE1 sont en accord avec Mesquida *et al.* (2010) : il n'y a pas de relation entre l'amélioration du marquage du pluriel et les progrès dus à l'apprentissage de l'orthographe. Nous infirmons donc notre hypothèse 3.

12.9 Un modèle pour l'acquisition de la liaison après 6 ans

Le modèle morpho-phonologique explique l'acquisition de la liaison suivant 3 étapes. L'acquisition se termine au stade 3, quand les enfants acquièrent le double flottement. Comme nous l'avons vu dans le chapitre 4, ce modèle à étapes permet de très bien expliquer les erreurs de consonne erronée (étape 2 du modèle). Cependant, il ne s'intéresse pas à la représentation des non-réalisations. Notre recherche a pris en compte ces non-réalisations et a montré qu'elles doivent être reconsidérées par un modèle qui explique l'acquisition de la liaison.

Sur la base des analyses phonétiques, nous chercherons à représenter l'acquisition de la liaison du CP au CE1 en proposant une étape intermédiaire à intégrer avant l'acquisition du double-flottement.

Notre point de départ est l'étape 2 du modèle morpho-phonologique présenté par Wauquier et Braud (2005) en figure 41. La représentation proposée explique très bien l'apparition des erreurs d'insertion d'une consonne erronée et le fait que Mot1 et Mot2 sont déjà interprétés comme deux formes séparées.

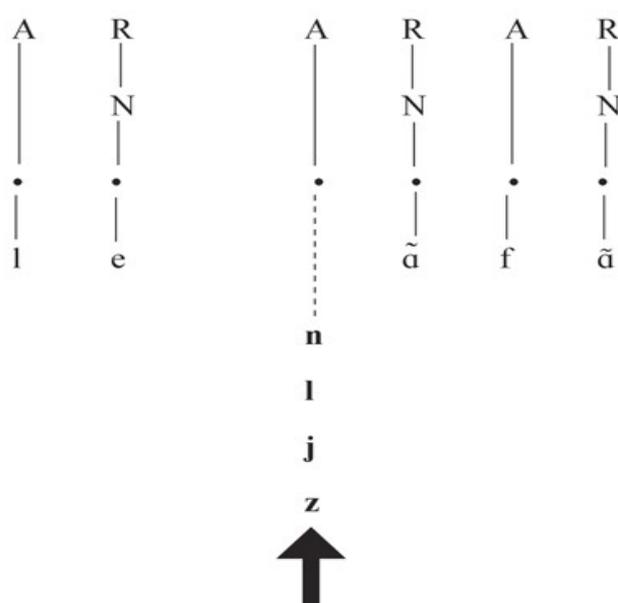


Figure 41. Stade 2 dans le modèle de l'acquisition de la liaison proposé par Wauquier et Braud (2005)

Pour réussir à donner une représentation phonologique des non-réalisations de liaisons, nous considérons l'analyse phonétique que nous avons conduit au CP, où nous observons :

- un continuum sonore entre Mot1 et Mot2 : la CL n'est pas réalisée et Mot2 est produit après Mot1.
- L'insertion d'un coup de glotte entre Mot1 et Mot2.

Nous représentons ces productions en figure 42. Nous appelons cette étape « 3bis » et nous l'insérons avant l'acquisition du double-flottement. La consonne de liaison occupe la position d'attaque mais le contenu segmental n'est pas lié à la position syllabique. La CL n'est donc pas réalisée. Parfois une occlusive glottale occupe la place d'attaque du Mot2 et réalise la structure syllabique CV au début du nom. Parfois cette occlusive n'est pas réalisée et la production de Mot1 et Mot2 est perçue comme un continuum sonore.

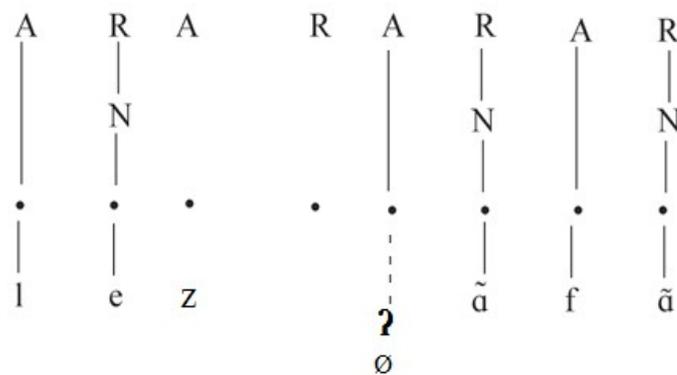


Figure 42. Étape 3 bis

Au CE1, nous remarquons encore des productions de coups de glotte entre Mot1 et Mot2 en parallèle à des productions de continuum. Nous pouvons penser que les enfants jusqu'au CE1 fluctuent encore entre les deux réalisations possibles prévues pour 3bis. L'étape finale, suite à cette nouvelle étape intermédiaire est l'acquisition du double-flottement (Figure 43).

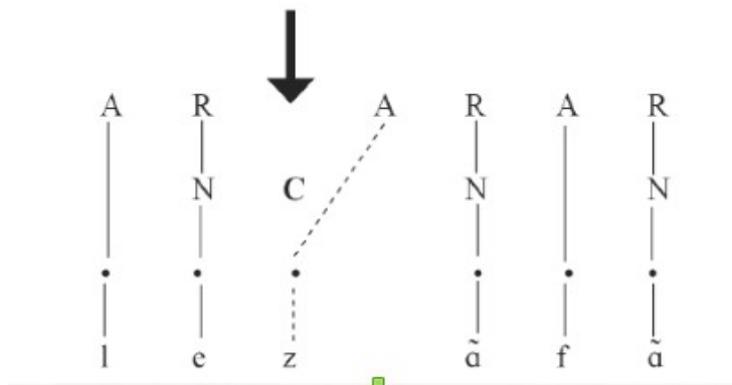


Figure 43. Dernière étape, acquisition du double flottement

Une représentation de l'étape 3bis a déjà été donnée par Scheer *et al.* (2015). Comme on peut le voir dans la figure 44, dans son étape 3 du schéma (*stage 3*) les informations segmentales sont séparées des informations syllabiques.

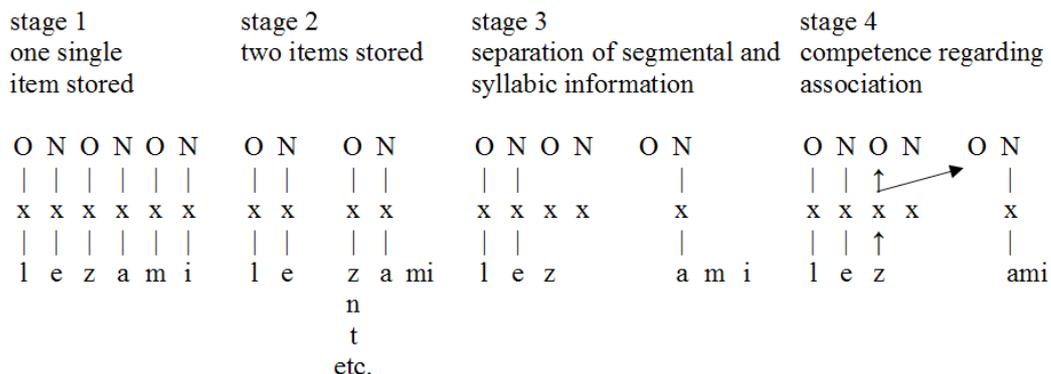


Figure 44. Nouvelle version du modèle phonologique (Scheer *et al.*, 2015)

L'introduction de cette étape permet d'expliquer les observations de Splendido (2014) qui enregistre dans sa thèse des liaisons non-enchaînées chez des enfants bilingues franco-suédois. De plus, elle peut bien expliquer nos non-réalisations, où nous enregistrons un continuum entre Mot1 et Mot2 et où la CL n'est pas réalisée. Cependant elle ne tient pas compte de la possibilité de l'insertion d'un coup de glotte entre Mot1 et Mot2. Pour cette raison nous avons introduit l'étape 3bis. Le coup de glotte apparaît pour remplir la position rendue vide à l'initiale du Mot2 devant la voyelle à la suite du détachement du déterminant.

Comme nous l'avons vu dans le chapitre 2, la littérature montre que l'apprentissage de la lecture et de l'écriture peut impliquer des changements

phonologiques. De plus, nous avons vu que le traitement correct des frontières des mots en lecture et en écriture apporte des productions orales correctes (§ 9.6, § 9.8.3, § 10.4).

Dans le modèle morpho-phonologique où nous avons représenté l'étape intermédiaire, un des moteurs qui permettrait le passage de l'étape 3bis au double-flottement serait donc l'apprentissage de la *literacy*. Pour le modèle présenté par Wauquier (2009), les enfants encodent la consonne flottante dans la représentation lexicale sous-jacente par *bootstrapping* morpho-phonologique. Cela implique qu'ils maîtrisent l'alternance morpho-phonologique des formes du déterminant. Nous pouvons donc penser que ce n'est pas seulement le *bootstrapping* morphologique qui permet l'encodage de la CL dans les représentations lexicales sous-jacentes, mais aussi l'acquisition des formes écrites. Dans cette explication, le déterminant est acquis comme un mot qui porte une consonne flottante qui est associée à un graphème orthographié toujours à la fin de ce mot (le <s> dans <les>). Cette acquisition est confirmée par l'apparition au CE1 des lectures NONEnch.LECT en liaison (ex. *les[z]affiches* [lez/a.fiʃ] pour *les[z]affiches* [le.za.fiʃ]). Si le Mot2 commence par une voyelle à l'écrit comme à l'oral, la CL resyllabifie sur le mot suivant. Ce processus est favorisé par la lecture qui entraîne la resyllabation sur Mot2 d'une consonne qui ne lui appartient pas visuellement.

Le modèle morpho-phonologique considère la liaison comme un phénomène exemplaire d'opacité phonologique (Klein & Brandão de Carvalho, 2010) car toutes les formes variables (formes de surfaces) sont conditionnées par une seule forme sous-jacente. Le locuteur ne peut pas partir des formes de surfaces pour faire ses généralisations, mais doit connaître la forme sous-jacente car « devant un mot liaisonnant, seule la forme sous-jacente explique pourquoi on fait la liaison, par exemple, dans *peti* [t] *arbre*, mais pas dans *joli arbre* » (Klein & Brandão de Carvalho, 2010 : 46). Les activités de *literacy* permettent de bien encoder la CL, de l'attacher correctement au Mot1 mais aussi de la resyllabifier sur Mot2, comme prévu par la représentation du double-flottement. La *literacy* peut donc être vue comme une clé pour accéder à la forme sous-jacente à partir des formes de surfaces et résoudre la question de l'opacité.

Le modèle morpho-phonologique permet facilement d'intégrer la connaissance de la lecture et de l'écriture et d'expliquer assez clairement nos résultats. Ce n'est pas le cas pour le modèle lexical-constructionniste. Il propose une représentation des liaisons

à partir d'un découpage de surface aboutissant à /un + nX/ où la consonne est encodée sur le Mot2. Ses prédictions sont donc des formes orthographiques du type <les sours> et des lectures comme *les/[z]ours*. Nos résultats montrent qu'au CP et au CE1 les enfants produisent plutôt des formes attendues (<les ours>) et des réalisations de la liaison *les[z]ours*. Nous observons des lectures du type PRA.LECT (*les/[z]ours*) où le nom est lu comme un mot à consonne initiale même si dans sa représentation écrite il commence par une voyelle. Cependant le nombre de ces lectures est inférieur au nombre de lectures correctes. De plus, les PRA.LECT peuvent être expliquées par l'utilisation de l'unité syllabique en lecture qui, comme nous l'avons vu dans le chapitre 5, reste une des unités primaires pour accéder au français écrit.

En conclusion, nous pouvons dire que le modèle morpho-phonologique permet de bien expliquer l'acquisition de la liaison après 5 ans car il permet non seulement de représenter les non-réalisations mais aussi d'intégrer la *literacy* et d'expliquer les données sur la langue écrite.

12.10 L'élision comme un phénomène facile à maîtriser

Nous avons vu que pour la tâche DENa, les productions en élision sont provoquées par l'input en liaison donné par l'expérimentateur. Les nombreuses erreurs sont donc le fait d'une difficulté à segmenter les suites Mot1 + Mot2 en liaison. Néanmoins les erreurs d'élision diminuent en nombre entre CP et CE1. Cela implique que les enfants réussissent de plus en plus à segmenter les séquences en liaison, à identifier les frontières des mots, grâce aussi à l'acquisition de la *literacy*.

Dans les tâches liées à la langue écrite, les enfants écrivent correctement les mots en élision. Surtout au CE1, les erreurs où la consonne qui resyllabifie sur Mot2 est écrite comme si elle était son initiale sont peu nombreuses et les erreurs en lecture le sont encore moins.

Ces données nous suggèrent que l'élision est un phénomène plus facile à acquérir que la liaison. Ce phénomène implique d'abord un répertoire plus réduit de segments impliqués ([l] surtout mais aussi [s]) comparé à la liaison. De plus, contrairement à la liaison qui prévoit l'apparition d'un segment latent et de sa

resyllabation, l'élision implique la chute d'une voyelle pour éviter un hiatus. Il est clairement plus facile d'effacer une voyelle et produire une élision que de produire un segment latent en liaison.

En ce qui concerne le rapport avec la *literacy*, le [l] est orthographié par le graphème <l> correspondant et il est représenté orthographiquement au début de ce mot où il est prononcé. Le seul élément qui sépare les deux mots est l'apostrophe. Nous pouvons penser qu'il est donc plus facile de le lire au début du Mot1 à cause de l'absence d'espace blanc entre déterminant et nom et de la resyllabation d'une consonne qui appartient orthographiquement au déterminant, comme pour la liaison.

12.11 L'enchaînement : des résultats vagues

Comme pour l'élision, les productions d'enchaînements dans DENa sont causées par la procédure de la tâche et par l'input en liaison donné par l'expérimentateur.

Les résultats de DENb montrent que, dans la plupart des cas, l'enchaînement est produit entre adjectif et nom. Les non-réalisations ou les insertions d'une mauvaise consonne de liaison sont globalement peu nombreuses. Cette tendance est aussi enregistrée pour le contexte déterminant-nom dans les tâches liées à la langue écrite. Au CP et au CE1, nous avons plus de réalisations d'enchaînement en lecture que de non-réalisations. De plus, quand en lecture les enfants ne réalisent pas l'enchaînement, ils insèrent parfois un coup de glotte, surtout au CP. Comme nous l'avons déjà dit dans la section 10.3.2.2.2, nous pouvons donner à cette occlusive glottale une valeur phonologique : elle peut prendre la place de la CF dans les non-réalisations d'enchaînement.

Ces données permettent de remettre en question l'idée de variabilité associée à ce phénomène par Encrevé (1983) et de proposer que l'enchaînement soit catégorique pour les contextes déterminant-nom et adjectif-nom. Cependant, surtout au CE1, nous n'enregistrons pas de difficultés à identifier les frontières écrites du Mot2 en présence d'enchaînement. Nous pouvons penser que ces résultats en écriture sont impliqués par le caractère variable de ce phénomène et par le fait que, dans l'input oral de ces séquences, ils peuvent fréquemment entendre des Mots1 + Mots2 non-enchaînés.

Dans la tâche de dénomination guidée, la plupart des non-réalisations d'enchaînement sont enregistrées dans la séquence *un minuscule ogre*, entre adjectif et nom. Nous pensons que la longueur de la séquence test à produire a influencé la non-réalisation de ce phénomène. Comme il l'a déjà été vu pour la liaison (Fougeron *et al.*, 2001), nous observons un taux de réalisations d'enchaînement plus bas dans une séquence plus longue.

À la lumière de ces résultats, des travaux supplémentaires doivent être conduits pour mieux comprendre la nature de ce phénomène. De futures recherches basées sur des enregistrements du parlé spontané chez les enfants et sur les productions d'enchaînement pourront nous confirmer sa nature catégorique ou variable.

◆ **Q3 : Quel processus cognitif derrière le découpage en mots ?**

Au moyen de deux tâches d'amorçage chez notre groupe d'enfants et chez un groupe d'adultes, nous avons cherché à comprendre le processus cognitif sous-tendant le découpage en mots en français. Nous cherchons à voir si repérer un Mot2 dans une séquence caractérisée par la liaison cause des coûts dans les processus cognitifs et à comprendre si la *literacy* joue un rôle dans ce traitement. Nous cherchons à vérifier les hypothèses (H4) et (H5) du chapitre 6.

○ **Hypothèse 4 (H4) : un coût pour le traitement cognitif de l'input en liaison chez les enfants**

Nos résultats chez les adultes sont en accord avec la littérature (Wauquier-Gravelines, 1996 ; Gaskell *et al.*, 2002). Ils montrent qu'il n'y a pas de coûts dans le traitement cognitif des séquences en liaison chez les adultes (chapitre 4). En revanche, chez les enfants du CE1, nous observons que la resyllabation est responsable d'un coût du traitement cognitif. Dans AMOa, identifier le mot ELEPHANT dans la séquence en liaison *un petit éléphant* demande plus de temps qu'identifier le même mot dans la séquence *un gentil éléphant*, où les frontières syllabiques sont alignées aux frontières lexicales. Pour les enfants, un nom inséré dans un input aligné est plus facile à identifier qu'un mot inséré dans un input resyllabifié.

Les résultats de cette tâche d'amorçage chez les enfants, différents des résultats des adultes, peuvent être expliqués par le niveau de *literacy* en confirmant H4. En accord avec les études présentées dans le chapitre 4,

quand les adultes entendent un input acoustique, ils activent la représentation orthographique des mots dans la séquence. Qu'ils entendent *un petit éléphant* ou *un gentil éléphant*, ils activent tout de suite la représentation orthographique du mot <éléphant>. Cette activation permet de ne pas passer par le découpage de la séquence orale pour juger si le mot ELEPHANT est présent dans la séquence et cause des TR équivalents pour les deux conditions.

Chez les enfants les mots écrits à voyelle initiale ne sont pas encore stockés de manière stable dans leur lexique orthographique. À ce stade, ils utilisent encore la phonologie et l'unité syllabique pour repérer les frontières des mots. Pour cette raison, identifier ELEPHANT dans un input resyllabifié implique des TR plus longs que de le repérer dans un input où les frontières syllabiques sont alignées aux frontières lexicales. Dans cette tâche, le passage par la phonologie cause un ralentissement dans le traitement cognitif des mots.

- **Hypothèse 5 (H5) : le rôle de la syllabe entre écrit et oral chez les enfants**

Nos données, nous montrent qu'une des composantes majeures qui influence le découpage en mots chez les enfants du CE1 est la syllabe. Dans la tâche AMOb, nous avons vu que les enfants ont des difficultés à dire que le mot TELEPHANT n'est pas orthographié correctement après avoir entendu la séquence *un petit[t]éléphant*. Ces résultats montrent que les enfants favorisent les frontières syllabiques dans le découpage des séquences en liaison et posent la frontière du nom sur la frontière de la syllabe. L'attaque syllabique est plutôt choisie comme le point où le mot doit commencer et une représentation orthographique qui respecte cette structure syllabique est considérée comme correcte. Ces résultats sont en accord avec différentes études qui ont souligné l'importance de la syllabe dans le découpage de la chaîne orale (Babineau & Shi, 2011 ; Nazzi *et al.*, 2006, 2008).

De plus, les données confirment l'importance de cette unité dans les premières pratiques de lecture et d'écriture, comme plusieurs études l'ont confirmé (Ziegler & Goswami, 2005 ; Goswami, 1993, 1999 ; Goswami

& Bryant, 1990 ; Goswami & East, 2000 ; Kandel & Valdois, 2006 ; Colé *et al.*, 1999 ; Bastien-Toniazzo *et al.*, 1999 ; Doignon & Zagar, 2006 ; Chetail & Mathey, 2012). Les enfants constituent la forme orthographique à partir de la forme phonologique et ont des difficultés à refuser une forme orthographique non-correcte (TELEPHANT) mais dont la structure syllabique respecte un découpage en syllabes CV (*un petit[t]éléphant*). Au contraire, les adultes ne passent plus par la phonologie pour juger l'exactitude des mots écrits et ils font référence seulement à la représentation écrite. Pour cette raison, ils n'ont pas de problèmes à refuser un mot comme TELEPHANT même après avoir entendu une séquence phonologique comme *petit[t]éléphant*.

Les résultats dans la tâche AMOb confirment l'importance de la phonologie et de l'unité syllabique à ce stade de l'acquisition de la langue écrite et remettent à nouveau en jeu la variable *literacy* pour l'interprétation de ces résultats d'amorçage.

12.12 Perspectives futures

12.12.1 Amélioration de la base de données

Comme nous l'avons dit, la base de données MANULEX décrit le lexique orthographique des enfants à partir de l'école primaire. Cette base de données est très bien réalisée mais ne permet pas d'avoir des informations explicites sur la langue orale des enfants. Une base de données structurée et normée comme MANULEX mais qui décrive la langue orale des enfants serait une perspective envisageable. De plus, comme nous l'avons déjà montré, MANULEX donne la fréquence des mots isolés. Cela ne permet pas de bien comprendre la fréquence des co-occurrences ni de bien évaluer leur rôle dans l'acquisition des phénomènes de resyllabation. Une des possibilités pour résoudre ces problèmes est de réussir à créer une base de données pour l'oral et pour l'écrit qui donne la fréquence des couples de mots chez les enfants.

12.12.2 Les données phonétiques chez les plus petits

Nous avons analysé les productions orales des enfants au moyen du logiciel PRAAT. Ces analyses nous ont aidée à mieux comprendre la possible représentation sous-jacente que les enfants utilisent pour produire à l'oral.

À partir de nos analyses, nous avons vu que la CL (ex. le [z] dans *les[z]ours* [le.zuʁs]) a une longueur différente de la consonne lexicale (ex. le [z] dans *les zèbres* [le.zɛbʁ]). Cette différence de longueur fait penser à un traitement différent de la CL et de la consonne lexicale, ce qui implique plutôt la validation du modèle morpho-phonologique. Une des perspectives de recherches futures possibles est donc l'analyse acoustique des consonnes lexicales et des CL chez les enfants entre 2 et 5 ans, dans une situation qui permet d'enregistrer des productions spontanées. La production spontanée peut être guidée par l'utilisation de cartes ou de jouets qui aident les enfants à produire certains items tests. Si la longueur de la CL est différente de la longueur de la consonne lexicale chez les plus petits, nous serons obligés de leur attribuer un statut différent dès le début de l'acquisition.

Une autre analyse intéressante à conduire sur ce type de données serait de voir si la longueur de la consonne erronée insérée entre Mot1 et Mot2 (le [n] dans *les[n]ours* [le.nuʁs]), fréquemment produite chez les enfants plus petits, est plus proche de la longueur d'une CL (le [n] dans *un[n]ours* [ɛ̃.nuʁs]) ou d'une consonne lexicale (le [n] dans *les navires* [le.na.viʁ]) pour vérifier son statut.

Dans la production spontanée, il faudrait aussi analyser les non-réalisations. Les insertions de coups de glotte entre Mot1 et Mot2 que nous avons enregistrées peuvent être la conséquence de la tâche et avoir une valeur seulement phonétique en tant qu'hésitation. Retrouver les coups de glotte dans les données de productions spontanées pourrait confirmer nos données et conclusions.

12.12.3 La tâche d'amorçage : tester d'autres groupes de participants et formaliser un modèle de la lecture

Nous avons comparé les résultats de notre tâche d'amorçage chez les enfants avec les résultats trouvés par Gaskell *et al.* (2002) chez les adultes. Nous avons expliqué la différence entre les résultats de Gaskell *et al.* (2002) et les nôtres par le

niveau de *literacy*. Pour bien vérifier si c'est vraiment la *literacy* qui joue un rôle dans cette tâche, il faudrait envisager de tester des enfants avec un niveau de *literacy* différent, comme par exemple des enfants au CP, dont les connaissances orthographiques sont moins développées. Il faudrait aussi tester des enfants plus grands, par exemple à la fin de l'école primaire, pour vérifier si ces résultats sont comparables aux résultats chez les adultes.

Une autre façon pour évaluer l'influence de la *literacy* sur cette tâche est de tester les adultes illettrés qui apprennent le français comme langue seconde au moyen de l'oral, sans avoir de contact avec l'orthographe. Normalement, on s'attend à ce que ces adultes aient les mêmes résultats que les enfants au CE1, car ils n'ont pas de vrai lexique orthographique de la langue française mais seulement des compétences phonologiques basées sur l'oral.

Tester d'autres groupes de participants nous permettrait non seulement de vérifier nos résultats et conclusions mais aussi de comprendre plus en profondeur le processus cognitif derrière le découpage en mots en présence de la liaison.

Ces nouvelles données peuvent nous permettre de reconsidérer aussi la question des modèles de lecture. Dans le chapitre 5, nous avons présenté plusieurs modèles de lecture. Même si dans certains paragraphes de notre texte nous faisons référence à l'existence d'un lexique orthographique, dans cette thèse nous ne choisissons pas un modèle de lecture. Nous pourrions approfondir cette question dans de futures recherches, faire un choix entre les différents modèles proposés dans la littérature et penser aussi à une possible implémentation du modèle choisi qui rende compte de nos données sur la lecture.

12.12.4 D'autres recherches qui mettent en rapport la langue orale et la *literacy*

Dans cette thèse, nous avons essayé de comprendre le rapport entre phonologie et *literacy*. Nous avons réussi à établir un rapport et une continuité entre la lecture et l'écriture et les productions orales qui nous ont permis d'apporter des améliorations aux modèles expliquant l'acquisition de la liaison. Cette perspective peut être utilisée pour expliquer d'autres phénomènes du français qui causent une incongruence entre écrit et oral comme par exemple les consonnes muettes ou les digrammes. Ces arguments peuvent être également exploités pour d'autres langues opaques comme l'anglais.

Nous pensons aussi que l'étude des relations entre la phonologie et la *literacy* que nous avons conduite pour des enfants à développement typique peut inspirer d'autres recherches sur les troubles du langage et de l'apprentissage.

Références bibliographiques

- Abercrombie, D. (1967). *Elements of General Phonetics*. Edinburgh : Edinburgh University Press.
- Adams, M., J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge: MIT Press.
- Ågren, J. (1973). *Enquête sur quelques liaisons facultatives dans le français de conversation radiophonique*. Uppsala : Acta Universitatis Upsaliensis.
- Alegria, J. & Leybaert, J. (1991). Mécanismes d'identification des mots chez le sourd. In R. Kolinsky, J. Morais & J. Segui (Eds.), *La reconnaissance des mots dans les différentes modalités sensorielles*. Paris : PUF, 231-252.
- Ashby, W., J. (1981). French liaison as a sociolinguistic Phenomenon. In William W. Cressy and Donna Jo Napoli (ed.), *Linguistic Symposium on romance languages 9*. Washington D.C. : Georgetown University Press, 46-57.
- Babineau, M. & Shi, R. (2011). Processing of French liaisons in toddlers. In *BUCLD 35: Proceedings of the 35th annual Boston University conference on language development*. Boston, MA: Cascadilla Press, 1, 25-37.
- Baddeley, A., D. (1986). *Working memory*. Oxford: Oxford University Press.
- Badian, N., A. (2001). Phonological and orthographic processing: Their roles in reading prediction. *Annals of Dyslexia*, 51, 179–202.
- Bassano, D. (2010). L'acquisition du déterminant nominal en français : une construction progressive et interactive de la grammaire. *Cognitextes*, 5. URL : <http://cognitextes.revues.org/315>.
- Basset, B. (2000). *La liaison à 3, 7 et 11 ans: description et acquisition*. Unpublished Master dissertation. University of Grenoble 3.
- Bastien-Toniazzo, M., Magnan, A., L. & Bouchafa, H. (1999), Nature des représentations du langage écrit aux débuts de l'apprentissage de la lecture: un modèle interprétatif, *Journal International de Psychologie*, 34 (1), 43-58.
- Béchade, H., D. (1992). *Phonétique et morphologie du français moderne et contemporain*. Paris: Presses universitaires de France.
- Bertelson, P., Morais, J., Alegria, J. & Content, A. (1985). Phonetic analysis capacity and learning to read. *Nature*, 313, 73-74.
- Bybee, J. (2001a). *Phonology and language use*. Cambridge: Cambridge University Press.
- Bybee, J. (2005). La liaison : effets de fréquence et constructions. *Langages*, 158, 24-37.
- Blanche-Benveniste, C. & Chervel, A. (1968). *L'Orthographe*. Paris : Maspéro. (Nouvelle édition : 1978, Edition augmentée).
- Blanche-Benveniste, C. & Jeanjean, C. (1987). *Le français parlé: transcription et édition*. Paris: Didier érudition.

- Boersma, P. (2005). *Some listener-oriented accounts of h-aspiré in French*. Mr. University of Amsterdam.
- Booij, G. & De Jong, D. (1987). The domain of liaison: theories and data. *Linguistics*, 25, 1005-1025.
- Bordal, G. & Lyche, C. (2008). *La liaison en terre africaine*. Journées PFC décembre 2008: Phonologie du français contemporain: variation, interfaces, cognition.
- Boula de Mareüil, P., Adda Decker, M. & Gendner, V. (2003). *Liaisons in French : a corpus-based study using morpho-syntactic information*. In ICPHS International Congress of Phonetic Science, Barcelone, 1329-1332.
- Bradley, D., C., Sánchez-Casas, R., M. & Gracia-Albea, J., E. (1993). The status of the syllable in The Perception of Spanish and English. *Language & Cognitive Processes*, 8, 197-233.
- Bradley, L., L., & Bryant, P. (1978). Difficulties in auditory organization as a possible cause of reading backwards. *Nature*, 271, 746-747.
- Braud, V. (2003). *Acquisition de la prosodie chez des enfants francophones : les phénomènes de tronctions*. Unpublished doctoral dissertation, University of Nantes.
- Brouillet, T., & Syssau, A. (2005). Étude du lien entre l'évaluation de la valence et les réponses verbales dans une tâche de décision lexicale: Les réponses oui et non sont-elles des réponses affectives?. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 59(4), 255-261.
- Bürki, A. & Frauenfelder, U., H. (2012). Producing and recognizing words with two pronunciation variants: Evidence from novel schwa words. *Quarterly Journal of Experimental Psychology*, 65, 796-824.
- Calfee, R., C., Chapman, R. & Venezky, R. (1972). How a child needs to think to learn and read. In L. Gregg (éd.), *Cognition and learning an memory*. New York : John Wiley, 139-182.
- Caramazza, A., Miceli, G. & Villa, G. (1986). The role of the (output) phonological buffer in reading, writing, and repetition. *Cognitive Neuropsychology*, 3, 37-76.
- Casalis, S., & Louis-Alexandre, M. (2000). Morphological analysis, phonological analysis and learning to read French : A longitudinal study. *Reading and Writing : An Interdisciplinary Journal*, 12, 303-335.
- Castro-Caldas, A., Petersson, K., M., Reis, A., Stone-Elander, S. & Ingvar, M. (1998). The illiterate brain. Learning to read and write during childhood influences the functional organization of the adult brain. *Brain*, 121 (6), 1053–1063.
- Catach, N. (1986). *L'orthographe française*. Paris : Nathan .
- Chabanal, D. & Embarki, M. (2002). *L'acquisition d'un marqueur socio-stylistique : L'exemple de la liaison facultative*. XXIVèmes Journées d'Étude sur la Parole, Nancy, 24-27 juin 2002.
- Chabanal, D. & Liégeois, L. (2014). Production de liaisons dans l'input parental. Dans C. Soum-Favaro, A. Coquillon & J.-P. Chevrot (eds.). *La liaison : approches contemporaines*. Bern: Peter Lang, 263-282.

- Chetail, F., & Mathey, S., (2009), Syllabic Priming in Lexical Decision and Naming Tasks: The Syllable Congruency Effect Re-examined in French. *Canadian Journal of Experimental Psychology*, 63 (1), 40-48.
- Chetail, F., & Mathey, S., (2012). Effect of Syllable Congruency in Sixth Graders in the Lexical Decision Task With Masked Priming. *Scientific Studies of Reading*, 00 (0), 1-13.
- Cheung, H., Chen, H.-C., Yip Lai, C., Wong, O., C. & Hills, M. (2001). The development of phonological awareness: Effects of spoken language experience and literacy skills in Chinese children receiving English instruction. *Journal of Educational Psychology*, 102(3), 712-728.
- Chevrot, J.-P. (1996). Prononciation et situations de lecture : données sociolinguistiques, approches expérimentales, modèles psycholinguistiques. *Lidil* 13, 127-158.
- Chevrot, J.-P. (1998). Acquisition phonologique tardive et apprentissage de la lecture-écriture, *Travaux de l'Institut de Phonétique de Strasbourg*, 28, 19-37.
- Chevrot, J.-P. & Malderez, I. (1999). L'effet Buben : de la linguistique diachronique à l'approche cognitive (et retour). *Langue française, l'orthographe et ses scripteurs*, 124, 104-125. doi:10.3406/lfr.1999.6309.
- Chevrot, J.-P., Beaud, L. & Varga, R. (2000). L'apprentissage des unités phonologiques variables : l'exemple du /R/ post-consonantique final en français. *Linx*, 42, 89-100.
- Chevrot, J.-P. & Fayol, M. (2001). *Acquisition of French Liaison and Related Child Errors*. Selected paper of the 8th congress for the study of child language, San Sebastian, julliet 1999.
- Chevrot, J.-P., Chabanal, D. & Dugua, C. (2007). Pour un modèle de l'acquisition des liaisons basé sur l'usage : trois études de cas. *Journal of French Language Studies*, 17, 103-128.
- Chevrot, J.-P., Dugua, C. & Fayol, M. (2009). Liaison, word segmentation and construction in French: a usage-based account. *Journal of Child Language*, 36 (3), 557-596.
- Chevrot, J.-P., Nardy, A. & Barbu, S. (2011). Developmental dynamics of SES-related differences in children's production of obligatory and variable phonological alternations. *Language Sciences*, 33 (1), 180-191.
- Chevrot, J.-P., Dugua, C., Harnois-Delpiano, M., Siccardi, A. & Spinelli, E. (2013). Liaison acquisition: debates, critical issues, future research. *Language Sciences*, 39, 83-94.
- Chevrot, J.-P., Siccardi, A., Parisse, C., & Spinelli, E (à paraître). Multiword sequences and phonological variants in the lexicon: What can we learn from the acquisition of prenominal liaison in French?.
- Chilounaki, K. & Bryant, P. (2007). How Children Learn About Morphological Spelling Rules. *Child Development*, 78 (4), 1360-137.
- Christophe, A. (1993) *The role of prosody in segmenting speech into words*. PhD in Cognitive Psychology, EHESS, PARIS.

- Clements, G., N. & Keyser, S., J. (1981). *A Three-tiered Theory of the Syllable*, Occasional Paper no. 19, Center for Cognitive Science, M.I.T.
- Clements, G., N. & Keyser, S., J. (1983). *CV Phonology: a Generative Theory of the Syllable* (Linguistic Inquiry Monograph 9). MIT Press : Cambridge, Ma.
- Colé, P. (1996). *Phonology processing at the end of Primary school*, oral paper at the XXVI International Congress of Psychology, Montréal, Canada, 16-21 August.
- Colé, P., Magnan, A. & Grainger, J. (1999). Syllable-sized units in visual words recognition: evidence from skilled and beginning readers. *Applied Psycholinguistics*, 20, 507-532.
- Colé, P., & Sprenger-Charolles, L. (1999). Syllabic processing and visual word recognition in eleven year-old French dyslexics, below average readers and normal readers / Traitement syllabique au cours de la reconnaissance de mots écrits chez des enfants dyslexiques, lecteurs en retard et normo-lecteurs de 11 ans. *Revue de Neuropsychologie*, 9 (4), 323-360.
- Colé, P., Royer, C., Leuwers, C. & Casalis, S., (2004). Les connaissances morphologiques dérivationnelles et l'apprentissage de la lecture chez l'apprenti-lecteur français du CP au CE2. *L'année psychologique*, 104 (4), 701-750.
- Coltheart, M., Curtis, B., Atkins, P. & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed processing approaches. *Psychological Review*, 100, 589-608.
- Content, A. & Leybaert, J. (1992). L'acquisition de la lecture: influence des méthodes d'apprentissage. In P. Lecocq (Ed.), *La lecture. Processus, apprentissage, troubles*. Lille : Presses Universitaires de Lille, 181-211.
- Coquillon, A., Durand J., Lyche C. & Eychenne J. (2010). French Liaison: from global results to local varieties. *Colloque international PHONLEX*, Université de Toulouse II-Le Mirail (8-10 septembre 2010).
- Cossu, G. & Marshall, J., C. (1985). Dissociation between reading and written spelling in two Italian children. Dyslexia without dysgraphia? *Neuropsychologia*, 23, 697-700.
- Cossu, G, Gugliotta, M. & Marshall, J., C. (1995). Acquisition of reading and written spelling in a transparent orthography: Two non parallel processes? *Reading and Writing*, 7 (1), 9-22.
- Côté, M.-H. (2005). Le statut lexical des consonnes de liaison. *Langages*, 158, 66-78.
- Côté, M.-H. (2008a). Empty elements in schwa, liaison and h-aspiré: The French Holy Trinity revisited. In J. M. Hartmann, V. Hegedüs & H. van Riemsdijk (éd.), *Sounds of silence: Empty Elements in Syntax and Phonology*. Amsterdam : Elsevier, 61-103.
- Côté, M.-H. (2014). Liaison et assibilation en français laurentien. Dans C. Soum-Favaro, A. Coquillon & J.-P. Chevrot (eds.). *La liaison : approches contemporaines*. Bern: Peter Lang, 8-26.

- Crain-Thoreson, C. & Dale, P., S. (1992). Do early talkers become early readers? Linguistic PROFICIENCY, preschool language, and emergent literacy. *Developmental Psychology*, 28, 421-429.
- Creuzet, V. & Pasa, L. (2009). L'écriture inventée : que signifie écrire une lettre pour une syllabe ?. *Spirale*, 44, 31-47.
- Cutler, A. & Carter, D., M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*, 2, 133-142.
- Cutler, A. & Norris, D. (1998). The role of strong syllables in segmentation for lexical access. *J. Exp. Hum. Percept. Perform*, 14, 113-121.
- Dawson, C., Klinger, L., F., Panagiotides, H., Hill, D. & Spieker, S., (1992). Frontal lobe activity behavior of infants of mothers with depressive symptoms. *Child Development*, 63, 725-737.
- Delattre, P. (1947). La liaison en français, tendances et classification. *The French Review*, 21 (2), 148-157.
- Delattre, P. (1951). *Principes de phonétique française à l'usage des étudiants anglo-américains*. Middlebury : Middlebury College. Delattre, P. (1951). *Principes de phonétique française à l'usage des étudiants anglo-américains*. Middlebury : Middlebury College.
- Delattre, P. (1965). *Comparing the phonetic features of English, French, German and Spanish*. Heidelberg: Julius Gross Verlag.
- Delattre, P. (1966). *Studies in French and comparative Linguistics*. The Hague: Mouton.
- Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Nunes Filho, G., Jobert, A., Dehaene-Lambertz, G., Kolinsky, R., Morais, J. & Cohen, L. (2010). How learning to read changes the cortical networks for vision and language. *Science*, 330, 1359-1364.
- Dejean de La Bâtie, B. (1993). *Word boundary ambiguity in spoken French*. Unpublished Doctoral Dissertation. Monash University, Victoria, Australia.
- Dejean de la Bâtie, B. & Bradley, D. (1995). Resolving word boundaries in spoken French: native and non-native strategies. *Journal of Applied Psycholinguistics*, 16, 59-81.
- Dell, F. (1973a). *Règles et les sons*. Paris : Hermann.
- Dell, F. (1973b). E muet: fiction graphique ou réalité linguistique? In Stephen R. Anderson & Paul Kiparsky (éd.) *A festschrift for Morris Halle*. New York: Holt, Rinehart & Winston, 26-50.
- Demuth, K., & Fee, E., J. (1995). *Minimal Words in Early Phonological Development*. Ms., Brown University and Dalhousie University.
- Demuth, K. (2001). Prosodic constraints on morphological development. In J. Weissenborn & B. Höhle (eds.), *Approaches to Bootstrapping: Phonological, Syntactic and Neurophysiological Aspects of Early Language Acquisition*. Amsterdam: John Benjamins. Language Acquisition and Language Disorders Series, 24, 3-21.

- Desrosières, A., & Thévenot, L. (1988). Les catégories socio-professionnelles. Editions La Découverte : Paris.
- Di Cristo A. (1998). Intonation in French. Dans D. Hirst, & A. Di Cristo, (eds), *Intonation Systems: A Survey of Twenty Languages*. Cambridge University Press : Cambridge, 195-218.
- Dijkstra, T., Roelofs, A. & Fieuws, S. (1995). Orthographic effects on phoneme monitoring. *Journal of Experimental Psychology*, 49 (2), 264-271.
- DiPietro, J., A. (2000). Baby and the brain: advances in child development. *Annu Rev Public Health*, 21, 455-71.
- Doignon, N. & Zagar, D. (2006). Les enfants en cours d'apprentissage perçoivent-ils la syllabe à l'écrit ?. *Revue canadienne de psychologie expérimentale*, 60 (4), 258-274.
- Dubois, J. (1965). *Grammaire structurale du Français: Nom et Prenom*. Paris: Larousse.
- Dugua, C. (2006). *Liaison, segmentation lexicale et schémas syntaxiques entre 2 et 6 ans. Un modèle développemental basé sur l'usage*. Thèse de doctorat, Université Stendhal, Grenoble III.
- Dugua C., Chabanal, D. (2006). Acquisition de la liaison chez l'enfant francophone, formes lexicales des mots². Poster lors du colloque International : *JEP (Journées d'Etude sur la Parole. Dinard, France, (CD –rom)*.
- Dugua, C., Chevrot, J.-P. & Fayol, M. (2006). Liaison, segmentation des mots et schémas syntaxiques entre 2 et 6 ans : un scénario développemental. *Entretiens de Bichat : Orthophonie*, 29-30 septembre 2006, Paris : Expansion, formation et éditions, 230-244.
- Dugua, C. & Spinelli, E. (2009). La liaison : Effets de la fréquence et du rapport à l'écrit sur son acquisition et son usage. *Les cahiers de l'infantile*, 7, 17-41.
- Dugua, C., Spinelli, E., Chevrot, J.-P., Fayol, M. (2009). Usage-based account of the acquisition of liaison: Evidence from sensitivity to the singular/plural orientation of nouns. *Journal of Experimental Child Psychology*, 102, 342-350.
- Dugua, C., & Baclesse, M. (2014). Incidence d'effets de fréquence sur l'usage de la liaison en lecture à haute voix et dans des jugements normatifs chez des enfants de CE2-CM1. Dans C. Soum-Favaro, A. Coquillon & J.-P. Chevrot (eds.). *La liaison : approches contemporaines*. Bern: Peter Lang, 117-139.
- Durand, J., Laks, B., Calderone, B. & Tchobanov, A. (2011). Que savons-nous de la liaison aujourd'hui ?. *Langue Française*, 169, 1, 103-126.
- Durand, J. & Lyche, C. (2008). French liaison in the light of corpus data. *Journal of French Language Studies*, 18 (1), 33-66.
- Durand, M. (1936), *Le genre grammatical en français parlé, à Paris et dans la région parisienne*. Bibliothèque du français moderne : Paris.
- Dumay, N., Frauenfelder, U., H. & Content, A. (2002). The role of the syllable in lexical segmentation in French: word-spotting data. *Brain Lang.*, 81, 144-161.

- Durgunoğlu, A., Y. & Öney, B. (2002). Phonological awareness in literacy development: It's not only for children. *Scientific Studies of Reading*, 6, 245–266.
- Ecalte, J. & Magnan, A., (2008), Relations between print exposure and literacy skills: New evidence from Grande 1-5. *British Journal of Development Psychology*, 26, 525-544.
- Ecalte, J. (2000). Prédiction de réussite scolaire en lecture-écriture au cycle II. *Revue européenne de psychologie appliquée*, 50, 81-85.
- Eychenne, J., Lyche, C., Durand, J. & Coquillon, A. (2014). Quelles données pour la liaison aujourd'hui : la question des corpus, dans C. Soum, A. Coquillon & J.-P. Chevrot (eds) *La liaison : approches contemporaines*, Bern: Peter Lang, 33-60.
- Ehri, L., C. & Wilce, L., S. (1985). Movement into reading: Is the first stage of printed word learning visual or phonetic?. *Reading Research Quarterly*, 20, 163-179.
- Ehri, L., C. & Wilce, L., S. (1982). The salience of silent letters in children's memory and spellings. *Memory and Cognition*, 104, 155-166.
- Ehri, L., C. & Wilce, L., S. (1980). The influence of orthography on readers' conceptualization of the phonemic structure of words. *Applied Psycholinguistics*, 1, 371-385.
- Encrevé, P. (1983). La liaison sans enchaînement. *Actes de la recherche en sciences sociales*, 46, 39-66.
- Encrevé, P. (1988). *La liaison avec et sans enchaînement : phonologie tridimensionnelle et usage du français*. Paris: Seuil.
- Ferreiro, E. & Teberosky, A. (1979). *Los sistemas de escritura en el desarrollo del niño*. México: Siglo XXI.
- Ferreiro, E., Pontecorvo, C. & Zucchermaglio, C. (1996). Pizza or Piza? How Children Interpret the Doubling of Letters in Writing. *Children's Early Text Construction*, New Jersey: Lawrence Erlbaum Associated Publisher, 145-164.
- Ferreiro, E. (2000). *L'écriture avant la lettre*. Paris : Hachette.
- Ferguson, C. & Farwell, C. (1975). Words and Sounds in Early Language Acquisition. *Language*, 51, 419-439.
- Fletcher, J. (1991). Rhythm and final lengthening in French. *Journal of Phonetics*, 19, 193-212.
- Fougeron, C., Goldman, J.-Ph. & Dart, A. (2001). *Influence de facteurs stylistiques, syntaxiques et lexicaux sur la réalisation de la liaison en français*. TALN, 2001, Tours.
- Fougeron, C., Goldman, J., P. & Frauenfelder, U., H. (2001). Liaison and schwa deletion in French : an effect of lexical frequency and competition. *Proceedings of Eurospeech Conference*, 639-642.
- Fougeron, C. & Delais-Roussarie, E. (2004). Liaisons et enchaînements « Fais_en à Fez_en parlant », *Actes des Journées d'Etudes sur la Parole*, 221-224.
- Francescato, G. (1970). *Il linguaggio infantile. Strutturazione e apprendimento*. Torino : Einaudi.

- Frauenfelder, U., H., Segui, J. & Dijkstra, A. (1990). Lexical effects in phonemic processing: Facilitatory or inhibitory?. *Journal of Experimental Psychology: Human Perception and Performance*, 16 (1), 77-91.
- Friederichi, A., D. & Thierry, G. (2008). *Early language development : bridging brain and behavior*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Frith, U. (1985). Beneath the surface of developmental dyslexia. In K.E. Patterson, J.C. Marshall & M. Coltheart (Ed.). *Surface dyslexia: Cognitive and neuropsychological studies of phonological reading*. London: Erlbaum, 301-330.
- Frith, U. (1986). A developmental framework for developmental dyslexia. *Annals of Dyslexia*, 36, 69-81.
- Gaatonne, D. (1978a). Forme sous-jacente unique ou liste d'allomorphes? (À propos des consonnes de liaison en français). *Linguistics*, 214, 33-54.
- Gabriel, C., & Meisenburg, T. (2004). *Silent onsets? An optimality-theoretic approach to French h aspiré words*, *Phonetik und Phonologie* 1, Potsdam 2004, at OCP2.
- Ganapole, S., J. (1987). The development of word consciousness prior to first grade. *Journal of Reading Behavior*, 19, 415-436.
- Gaskell, G., Spinelli, E. & Meunier, F. (2002). Perception of Resyllabification in French. *Memory and Cognition*, 30, 798-810.
- Girard, F. & Lyche C. (1997). *Phonétique et phonologie du français*. 3rd ed. Oslo: Universtitetsforlaget.
- Glushko, R., J. (1981). Principles for pronouncing print: The psychology of phonography. In Lesgold, A., M. & Perfetti C., A. (Eds), *Interactive processes in reading*. Hillsdale, N.J.: Lawrence Erlbaum, 61-83.
- Goswami, U. (2001). Early Phonological Development in the Acquisition of Literacy. *Handbook of Early Literacy Research*, Neuman, S., B., & Dickinson, D., K. (eds). The Guilford Press : New York, 111-125.
- Goswami, U. (1993). Toward an interactive analogy model of reading development : Decoding vowel graphemes in beginning reading. *Journal of Experimental Child Psychology*, 56, 443-475.
- Goswami, U. & Bryant, P. (1990). *Phonological skills and learning to read*. Hillsdale : Erlbaum.
- Goswami, U., Gombert, J., E. & de Barrera, L., F. (1998). Children's orthographic representations and linguistic transparency: Non-sense word reading in English, French, and Spanish. *Applied Psycholinguistics*, 19, 19-52.
- Goswami, U. (1999). Causal connections in beginning reading : The importance of rhyme. *Journal of Research in Reading*, 22, 217-240.
- Goswami, U. & East, M. (2000). Rhyme and analogy in beginning reading : Conceptual and methodological issues. *Applied Psycholinguistics*, 21, 63-93.
- Goswami, U., Ziegler, J. C. & Richardson, U. (2005). The effects of spelling consistency on phonological awareness: A comparison of English and German. *Journal of Experimental Child Psychology*, 92 (4), 345-365.

- Gougenheim, G. (1938). *Système grammatical de la langue française*. Paris, d'Artrey
- Grammont, M. (1914). *Traité pratique de prononciation française*. Paris : Delagrave.
- Green, J., N. & Hintze, M., -A. (1988). A reconsideration of liaison and enchainement. *Occasional Papers*. University of Essex: Department of Languages and Linguistics, 136-168.
- Guasti, M., De Lange J., Gavarro, A. & C. Caprin (2008). Article omission across child languages. *Language Acquisition*, 15, 89-119.
- Halle, M. & Vergnaud, J.,-R. (1980). Three Dimensional Phonology. *Journal of Linguistic Research*, 1 (1), 103-118.
- Hallé, P., Chéreau, C., & Segui, J. (2000). Where is the /b/ in “absurde” [apsyrd]? It is in French listeners' minds. *Journal of Memory and Language*, 43, 618-639.
- Hay, J. & Foulkes, P. (to appear) The evolution of medial (-t) in real and remembered time. To appear in *Language*.
- Harris, M. & Coltheart, M. (1986). *Language processing in children and adults: An introduction*. London: Routledge and Kegan.
- Hohne, E., A. & Jusczyk, P., W. (1994). Two-month-old infants' sensitivity to allophonic differences. *Perception and Psychophysics*, 56, 613-623.
- Humphreys, G., W. & Evett, L., J. (1985). Are there independent lexical and nonlexical routes in word processing? An evaluation of the dual route theory of reading. *The Behavioral and Brain Sciences*, 8, 689-739.
- Hutzler, F., Ziegler J., C. , Perry, C., Wimmer, H & Zorzi, M. (2004). Do current connectionist learning models account for reading development in different languages?. *Cognition*, 91, 272-296.
- Jack, A. I., & Shallice, T. (2001). Introspective physicalism as an approach to the science of consciousness. *Cognition*, 79, 161–196.
- Jacob, V., & D'Odorico, L. (2001). *Aspetti prosodici e lessicali dell'input linguistico materno*. Relazione al Congresso Nazionale dell'Associazione Italiana di Psicologia – Selezione di Psicologia dello sviluppo, Palermo.
- Jakimik J., Cole R., A. & Rudnicky A., I. (1985). Sound and spelling in spoken word recognition. *Journal of Memory & Language*, 24, 165–178.
- Jusczyk, P. W., Houston, D. M. & Newsome, M., (1999). The beginning of word segmentation in English-learning infants. *Cogn. Psychol.*, 39, 159-207.
- Justice, L., M. & Ezell, H., K. (2001). Written language awareness in preschool children from low-income households: A descriptive analysis. *Communication Disorders Quarterly*, 22, 123-134.
- Kaye, J., D., & Lowenstamm, J. (1982). Compensatory Lengthening in Tiberian Hebrew, Theoretical Implications. *Studies in Compensatory Lengthening*. Leo Wetzels & Engin Sezer (eds.).
- Kaye, J. , D. & Lowenstamm, J., (1984). De la syllabicité. In F.Dell, D.Hirst & J.-R. Vergnaud (eds.), *Formes sonores du lanagage: structure de representation en phonologie*. Paris : Hermann, 123-159.

- Kandel, S., & Valdois, S. (2006). Syllables as functional units in the copying task : a viuo-orthographic and grapho-motor approach. *Language and Cognitive Process*, 21 (4), 432-452.
- Kandel, S., Alvarez, C. & Vallée, N., (2006). Syllables as processing units in handwriting production. *Journal of Experimental Psychology*, 312 (1), 18-31.
- Keeney, T. J, & Wolfe, J. (1972). The acquisition of agreement in English. *Journal of Verbal Learning and Verbal Behavior*, 11, 698-705.
- Klausenburger, J. (1974). Rule inversion, opacity, conspiracy: French liaison and elision. *Lingua*, 34 (2-3), 167-179.
- Landerl, K., Firth, U. & Wimmer, H. (1996). Intrusion of orthographic knowledge on phoneme awareness: strong in normal readers, weak in dyslexic readers. *Applied Psycholinguistics*, 17, 1-14.
- Landry, S., & Smith, K., E. (2006). The Influence of Parenting on Literacy Skills. *Handbook of Early Literacy*, Reaserch Volume 2, Neuman, S., B. & Dickinson, D., K. (eds). The Guilford Press : New York, 135-148.
- Laks, B. (2005). La liaison et l'illusion. *Langage*, 158,101- 125.
- Lavine, L. (1977). Differentiation of letter like forms in prereading children. *Developmental Psychology*, 13, 89-94.
- Largy, P., Cousin, M.-P., & Fayol, M. (2004). Acquérir le pluriel des noms : existe-t-il un effet de fréquence du nom?. *Lidil*, 30, 39-54.
- Lecocq., P. (1986). Sensibilité à la similarité phonétique chez les enfants dyslexiques et les bons lecteurs. *L'Année psychologique*, 86, 201-221.
- Lecocq, P. (1991). *Apprentissage de la lecture et dyslexie*. Bruxelles : Mardaga.
- Lété, B., Sprenger-Charolles, L., & Colé, P. (2004). MANULEX: A grade-level lexical database from French elementary school readers. *Behavior Research Methods*, 36 (1), 156-166.
- Lewkowitz, N. (1980). Phonemic awareness training : What to teach and how to teach it. *Journal of Educational Psychology*, 72, 686-700.
- Liberman, I. Y., Shankweiler, D., Fischer, F. W. & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology*, 18, 201-212.
- Lowenstamm, J. (199). CV as the only syllable type. In J. Durand et B. Laks (éds.), *Current Trends in Phonology, Models and Methods*, pp. 419-443. Salford : European Studies Research Institute.
- Lundberg, I. (2006). *The Child's Route Into Literacy : a Double-Track journey* dans McKeogh, A., Phillips, L., M., Timmons, V., Lupart, J. L. eds *Understanding Literacy Development : a Global view*, Publisher : London.
- Lundberg, I., Olofsson, A. & Wall, S. (1980). Reading and spelling skills in the first school years predicted from phonemic awareness skills in kindergarten. *Scandinavian Journal of Psychology*, 21, 159-173.

- Lundberg, I. (1982). Linguistic awareness as related to dyslexia. In P. Berman & J. Zotterman (éd.), *Dyslexia : Neuronal, cognitive and linguistic aspects*. New York : Pergamon, 141-153.
- MacWhinney, B. & Snow, C. (1984). The Child Language Data Exchange System (CHILDES).
- Malécot, A. (1975). French liaison as a function of grammatical, phonetic and paralinguistic variables. *Phonetica*, 32, 161-179.
- Mallet, G. (2008). *La liaison en français : descriptions et analyses dans le corpus PFC*. Thèse de doctorat, Université Paris Ouest-Nanterre-La Défense.
- Marsh, G., Friedman, M., Welch, V. & Desberg, P. (1981), «A cognitive-developmental theory of reading acquisition», in T.G. Waller & G.E. McKinnon (eds). *Reading research : advances in theory and practice*, 3, New York : Academic Press, 199-221.
- Martin, N. (2005). *Réseaux sociaux et variations phonétiques*. Mémoire de master 1, Université Stendhal, Grenoble.
- Mathey, S., Zagar, D., Doignon, N. & Seigneuric, A. (2006). The nature of syllabic neighbourhood in French. *Acta Psychologica*, 123, 372-393.
- Mathôt, S., Schreij, D. & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, 44(2), 314-324.
- Mattys, S., L. & Jusczyk, P., W. (2001). Phonotactic cues for segmentation of fluent speech by infants. *Cognition*, 78(2), 91-121.
- McClelland, J., L. & Rumelhart D., E. (1981). An Interactive Activation Model of Context Effects in Letter Perception : Part 1. An account of basic findings. *Phonological Review*, 88 (5), 375-407.
- Mehler, J., Dommergues, J. Y., Frauenfelder, U. & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior*, 20, 298-305.
- Menn, L., & Matthei, E. (1992). The “two-lexicon” model of child phonology: Looking back, looking ahead. in Ferguson, Menn, & Stoel-Gammon (eds.), *Phonological Development: Models, Research, Implications*, 211-247.
- Mesquida, C., Largy, P., Favaro-Soum, C., & Gunnarsson, C., (2010), Lorsque qu'une liaison gêne puis facilite la production de pluriel des noms chez les enfants. *Synergies Pays Scandinaves*, 5, 113-121.
- Metsala, J., L., & Walley, A., C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: precursors to phonemic awareness and early reading ability. In J.L. Metsala & L.C. Ehri (Eds.), *Word recognition in beginning literacy*. Mahwah, NJ: Lawrence Erlbaum Associates, 89-120.
- Morais, J., Cary, L., Alegria, J. & Bertelson, P., (1979). Does awareness of speech as a sequence of phones arise spontaneously?. *Cognition*, 7, 323-331.

- Morais, J., Alegria, J. & Content, A. (1987). The relationships between segmental analysis and alphabetic literacy : An interactive view. *Cahiers de psychologie cognitive*, 7, 415-438.
- Morgan, T., A. (1989-1990). Orthographic Constraints on Sound Change : The Non-Neutralization of Graphemes in Castilian. *The Lacus Forum*, 16, 303-312.
- Morin, Y.-C. & Kaye, J. (1982). The syntactic bases for French liaison. *Journal of Linguistics*, 18, 291-330.
- Morin, Y.-C. (1992). Un cas méconnu de la déclinaison de l'adjectif français : les formes de liaison de l'adjectif antéposé. Dans A. Clas (Ed.), *Le mot, les mots, les bons mots. Word, words, witty words. Hommage à Igor A. Mel'cuk*. Montréal : Presses de l'Université de Montréal, 233-250.
- Morin, Y.-C. (2003). Remarks on prenominal liaison consonants in French. *Living on Edge -28 papers in honour of Jonathan Kaye, Stefan Ploch (eds.)*, Berlin/New York : Mouton de Gruyter, 291-330.
- Morin, Y.-C. (2005a). La liaison relève-t-elle d'une tendance à éviter les hiatus. Réflexions sur son évolution historique. *Langages*, 158, 8-23.
- Morin, Y.-C. (2005b). Liaison et enchaînement dans les vers aux XVI^e et XVII^e siècles, in Jean-Michel Gouvard, *De la langue au style*, Presses Universitaires de Lyon, 299-318.
- Morton, J. & Patterson, K. (1980). A new attempt at an interpretation, or an attempt at a new interpretation. In: Coltheart, M., Patterson, K. & Marshall, J. (Eds.), *Deep Dyslexia*. London: Routledge & Kegan Paul, 91-118.
- Morton, J. (1989). An information-processing account of reading acquisition. In A.M. Galaburda (Ed.), *From reading to neurones*. Cambridge, MA: Bradford Book, MIT Press, 43-66.
- Nardy, A. (2008). *Acquisition des variables sociolinguistiques entre 2 et 6 ans: facteurs sociologiques et influences des interactions au sein du réseau social*. Doctorat. Université Grenoble 3 (téléchargeable à : <http://tel.archives-ouvertes.fr/tel-00466276/fr/>).
- Nardy, A. & Dugua, C., (2011). Le rôle de l'usage sur le développement des constructions nominales chez les enfants pré-lecteurs. *Travaux de linguistique*, 1, 129-148.
- Nardy, A., Chevrot, J.-P., & Chauvin, C. (2014). La liaison facultative dans les formes récitées du folklore enfantin. Dans C. Soum-Favaro, A. Coquillon & J.-P. Chevrot (eds.). *La liaison : approches contemporaines*. Bern: Peter Lang, 239-262
- Nazzi, T., Iakimova, G., Bertocini, J., Fredonie, S. & Alcantara, C. (2006). Early segmentation of fluent speech by infants acquiring French: Emerging evidence for crosslinguistic differences. *Journal of Memory and Language*, 54 (3), 283-299.

- Nazzi, T. (2008). Segmentation précoce de la parole continue en mots : évaluation inter-linguistique de l'hypothèse d'initialisation rythmique. *L'Année Psychologique*, 108, 309-342.
- New, B., Pallier, C., Ferrand L. & Matos, R. (2001) Une base de données lexicales du français contemporain sur internet: LEXIQUE. *L'Année Psychologique*, 10, 447-462.
- Nguyen, N., Wauquier-Gravelines, S., Lancia, L. & Tuller, B. (2007). Detection of Liaison consonants in speech processing in French, Experimental Data and theoretical implications. In P. Prieto, J. Mascaro, M.-J. Solé (Eds), *Segmental and Prosodic Issues in Romance Phonology*. Current Issues in Linguistic Theory n° 282, John Benjamins, 3-25,.
- Nyrop, K., R. (1925). *Manuel de phonétique du français parlé (4e ed)*. Copenhagen : Gyldendalske Boghandel Nordisk Forlag
- Nicolaci-Da Costa, A. & Harris, M. (1983). Redundancy of syntactic information: An aid to young children's comprehension of sentential number. *British Journal of Psychology*, 74, 343-352.
- Nishibayashi, L.-L., Goyet, L. & Nazzi, T. (2015). Early speech segmentation in French-learning infants: monosyllabic words versus embedded syllables. *Language & Speech*, 58(3), 334-350.
- Paap, K., R. & Noel, R., V. (1991). Dual-route models of print to sound : still a good horse race. *Psychological Research*, 53, 13-34.
- Pacton, S., Fayol, M. & Perruchet, P. (1996). *On the knowledge of double letters in French young spellers*. S.I.G.Writing. Barcelone, Octobre.
- Pacton, S., Fayol, M. & Perruchet, P. (1998). *How do young French spellers use different graphemes to transcribe one phoneme ? The case of the phoneme /O/*. S.I.G. Writing, Poitiers, juillet.
- Pagliano, C. (2003) *L'épenthèse consonantique en français. Ce que la syntaxe, la sémantique et la morphologie peuvent faire à la phonologie : parles-en de ta numérotation ? impossible*. Thèse de doctorat, Nice.
- Parisse, C., & Maillart, C. (2004). Le développement morphosyntaxique des enfants présentant des troubles de développement du langage : données francophones. *Enfance*, 56 (1), 20-35.
- Passy, P. (1917). *Les sons du français*. Paris: Firmin-Didot.
- Pattamadilok, C., Perre, L., Dufau, S. & Ziegler, J. C. (2009). On-line Orthographic Influences on Spoken Language in a Semantic Task. *Journal of Cognitive Neuroscience*, 21(1), 169-179.
- Peereman R., Dufour S., Burt J. S. (2009). Orthographic influences in spoken word recognition: the consistency effect in semantic and gender categorization tasks. *Psychon. Bull.*, 16, 363-368.

- Perfetti, L. (1985). *Reading ability*. New York : Oxford University Press.
- Perfetti, C. (1992) The representation problem in reading acquisition. In Laurence R., & Perfetti, C. (eds). *Learning to read : basic research and its implications*, Hillsdale : NJ Erlbaum.
- Perre, L., & Ziegler, J., C. (2008). On-line activation of orthography in spoken word recognition. *Brain Research*, 1188, 132-138.
- Perre, L., Bertrand, D. & Ziegler, J., C. (2011). Literacy affects spoken language in a non-linguistic task: An ERP study. *Frontiers in Psychology*, 2, 1-8.
- Perre, L., Pattamadilok, C., Montant, M. & Ziegler, J., C. (2009). Orthographic effects in spoken language: On-line activation or phonological restructuring? *Brain Research*, 1275, 73-80.
- Peters, A. (1985) *Language segmentation: Operating principles for the perception and analysis of language*. In: Slobin DI, editor.
- Petersson, M., K., Ingvar, M., & Reis, A. (2009). Language and Literacy from a Cognitive Neuroscience Perspective in *The Cambridge Handbook of Literacy*, Olson, R., D., & Torrence, N. (eds.), 152-181.
- Petersson, K., M., Silva, C., Castro-Caldas, A., Ingvar, M., & Reis, A. (2007). Literacy: a cultural influence left-right differences in the inferior parietal cortex. *European Journal of Neuroscience*, 26, 791-799.
- Pike, K., L. (1972) [1945]. The intonation of American English. In Bolinger, D. (ed.): *Intonation*. Harmondsworth: Penguin, 53-83.
- Pierrehumbert, B. (2003). *Le premier lien : théorie de l'attachement*. Odile Jacob Ed. : Paris.
- Plénat, M. (2008). *La liaison « obligatoire » avec et sans enchaînement*. In J. Durand, B. Habert & B. Laks (éd.), *Congrès Mondial de Linguistique Française – CMLF'08*, 1657-1667. Paris : Institut de Linguistique Française. [<http://www.linguistiquefrancaise.org>].
- Pulgram, E. (1970). *Syllable, word, nexus, cursus*. The Hague: Mouton.
- Purcell-Gares, V. (1996). Lexical and syntactic knowledge of written narrative held by well-read-to kindergarteners and second graders. *Research in the Teaching of English*, 22, 128-160.
- Ramus, F., Nespors, M. & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 73, 265-292.
- Ranson, D. (2008). La liaison variable dans un corpus du français méridional : L'importance relative de la fonction grammaticale, in J. Durand, B. Habert & B. Laks (eds)(2008) *Congrès mondial de linguistique française. Recueil des résumés et CD-ROM des actes*. Paris: Institut de Linguistique Française et EDP Sciences, 1657-1671.
- Rhodes, R. (2008). The Phonological History of Métchif. In L. Baronian & F. Martineau (Eds.), *Le français : d'un continent à l'autre. Mélanges offerts à Yves-Charles Morin*. Québec : Presses de l'Université Laval.

- Ruvoletto, S. (2011). *Fonologia autosegmentale e alfabetizzazione emergente: modello multilineare e geminate tra scrittura pre-convenzionale e formalizzata*. Master 2, Université de Padoue.
- Sabio, F. (2000). Les difficultés de la notion de mot : l'exemple des liaisons graphiques dans les textes d'enfants. *Linx*, 42, 1-10.
- Saffran, J., R., Newport, E. & Aslin, R. (1996b). Word segmentation : The rôle of distributional cues. *J. Me. Lang.*, 35, 606-621.
- Savin, H., B. & Bever T., G. (1970). The nonperceptual reality of the phoneme. *J. Verbal Learn. Verbal Behav.* 9, 295-302.
- Seidenberg, M., S. & Tanenhaus, M., K. (1979). Orthographic effects in rhyme and monitoring. *Journal of Experimental Psychology: Human Learning and Memory*, 5, 546-554.
- Seidenberg, M., S. & McClelland, J. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96, 523-568.
- Seymour, P., H., K. (1986). *A cognitive Analysis of Dyslexia*. London : Routledge & Kegan Paul.
- Selkirk, E. (1972). *The phrase phonology of English and French*. Thèse de doctorat, MIT.
- Selkirk, E. (1974). French Liaison and X notation. *Linguistic inquiry*, 5, 573-590.
- Schane, S., A. (1967). L'élision et la liaison en français. Adaptation du Chapitre 1 de *French Phonology and Morphology*. Cambridge, Massachusetts: M.I.T. Press.
- Schane, S. A. (1968). *French Phonology and Morphology*. Cambridge, Massachusetts: M.I.T. Press.
- Scheer, T. (1998). A unified model of Proper Government. Dans *The Linguistic Review* 15.1, 39-64.
- Scheer, T. (1999). A Theory of Consonantal Interaction. *Folia Linguistica*. 32, 201-237.
- Scheer, T. (2000). *De la Localite, de la Morphologic et de la Phonologic en Phonologic*. Habilitation thesis, University of Nice.
- Scheer, T. & Ségéral, P. (2001). Les séquences consonne + yod en gallo-roman. *Recherches Linguistiques de Vincennes*, 30, 87-120.
- Scheer, T., Encrevé, P. & Wauquier, S. (2015). *Autosegmental news from h aspiré and liaison without enchaînement*. 13th Conference of the French Phonology Network, Bordeaux, 29 juin – 2 juillet 2015.
- Share, D., L. (1995). Phonological recoding and self-teaching: *sine qua non* of reading acquisition. *Cognition*, 55 (2), 151-208.
- Shankweiler, D. (1999). Words to meanings. *Scientific Studies of Reading*, 3, 113-127.
- Shoemaker, E. (2010). The exploitation of fine phonetic detail in the processing of L2 French. In *Research in Second Language Processing and Parsing*, by VanPatten, Bill and Jill Jegerski (eds.). Amsterdam: John Benjamins.

- Soum, C. (1997). *L'apprentissage de l'écriture ; contraintes orthographiques : contraintes orales*. Thèse de doctorat, Université de Toulouse 2.
- Soum-Favaro, C., Gunnarsson, C. & Largy, P. (2014). La liaison à l'interface entre l'oral et l'écrit. In Soum-Favaro, C., Coquillon, A-L. & Chevrot, J-P. (éds.), *La liaison : approches contemporaines*. Peter Lang, 141-167.
- Splendido, F. (2014). *The initial development of Voice Onset Time in early L2 French*. Ph.D dissertation, University of Lund.
- Spinelli, E., McQueen, J. G. & Cutler, A. (2003). Processing resyllabified words in French. *Journal of Memory and Language*, 48, 233-254.
- Sprenger-Charolles, L., Siegel, S., L., & Béchennec, D. (1997). L'acquisition de la lecture et de l'écriture en français: étude longitudinale, dans Rieben, L., Fayol, M., Perfetti, A., C. (eds) *Des orthographes et leur acquisition*. Paris : delechaux et niestré, 359-384.
- Sprenger-Charolles, L., Colé, P., Lacert, P. & Serniclaes, W. (2000). On subtypes of developmental dyslexia : Evidence from processing time and accuracy scores. *Canadian Journal of Experimental Psychology*, 54(2), 87-104.
- Sprenger-Charolles, L., Siegel, L.S., Béchennec, D. & Serniclaes, W. (2003). Development of phonological and orthographic processing in reading aloud, in silent reading, and in spelling: A four-year longitudinal study. *Journal of Experimental Child Psychology*, 84(3), 194-217.
- Sprenger-Charolles, L. (2012). L'apprentissage de la lecture dans les écritures alphabétiques. *Information grammaticale*, 113, 15-21.
- Sprenger-Charolles, L. (2013). Place des compétences phonologiques dans l'apprentissage typique et atypique (dyslexie) de la lecture. Séminaires de Recherches en Phonétique et Phonologie (S-RPP). 01 Mars 2013, ILPGA, Paris.
- Snow, C., Burns, M. S. & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- Stanovich, K., E. (1988). Explaining the differences between dyslexic and the garden-variety poor reader: The phonological-core variable-difference model. *Journal of Learning Disabilities*, 21, 590-612.
- Stanovich, K., E. & Siegel, L., S. (1994). Phenotypic performance profile of children with reading disabilities: A regression-based test of the phonological-core variable-difference model. *Journal of Educational Psychology*, 86, 24-53.
- Sung, K.-P.(1989). H-aspiré in French. *Romance linguistics & literature review*, 2, 1-21.
- Tabossi, P., Collina, S., Mazetti, M., & Zoppello, M. (2000). Syllables in the processing of spoken Italian. *J. Exp. Psychol. Hum. Percept. Perform.*, 26, 758–775.
- Ternes, E. (1977). Konsonantische Anlautveränderungen in den keltischen und romanischen Sprachen. *Romanistisches Jahrbuch*, 28, 19-53.

- Thevenin, M.-G., Totereau, C., Jarousee & J.-P, Fayol, M. (1999). L'apprentissage/enseignement de la morphologie écrite du nombre en français. *Revue française de pédagogie*, 126, 39-52.
- Thompson, P. M., Giedd, J. N., Woods, R. P., MacDonald, D., Evans, A., C. & Toga, A. W. (2000). Growth patterns in the developing brain detected by using continuum mechanical tensor maps. *Nature*, 404, 190–193.
- Tolchinsky, L. (2003). *The cradle of culture and what children know about writing and numbers before being taught*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Tolchinsky, L. (2006). The emergence of writing. In C. A. MacArthur, S. Graham & J. Fitzgerald (Eds.), *Handbook of writing research*. New York: The Guilford Press, 83-95.
- Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press.
- Totereau, C., Thevenin, M.-G. & Fayol, M. (1997). Acquisition de la morphologie du nombre à l'écrit en français, 147-165, dans Rieben, L., Fayol, M., Perfetti, A., C. (eds) *Des orthographes et leur acquisition*. Paris : delechaux et niestré.
- Tranel, B . (1981), *Concreteness in Generative Phonology: Evidence from French*. University of California Press.
- Tranel, B. (1995). Current issues in French phonology : Liaison and position theories. In J. A. Goldsmith (Ed.), *The Handbook of Phonological Theory*, pp. 798-816. Cambridge : Blackwell.
- Treiman, R. (1985). Onsets and rimes as units of spoken syllables: Evidence from children. *Journal of Experimental Child Psychology*, 39,161-181.
- Treiman, R., Cohen, J., Mulqueeny, K., Kessler, B. & Schechtman, S. (2007). Young Children's Knowledge About Printed Names, *Child Development*, 78 (5), 1458-1471.
- Treiman, R. (1993). *Beginning to spell: A study of first-grade children*. New York: Oxford University Press.
- Ventura, P., Morais, J., Pattamadilok, C., & Kolinsky, R. (2004). The locus of the orthographic consistency effect in auditory word recognition. *Language and Cognitive Processes*, 19, 57 – 95.
- Vergnaud, J.-R. (1982). *On the foundation of phonology*. Communication au colloque de GLOW (Paris). (inédit).
- Vihman, M. (1996). *Phonological Development: The origins of language in the child*. Oxford : Blackwell.
- Vroomen, J., van Zon, M., & de Gelder, B. (1996). Cues to speech segmentation: Evidence from juncture misperceptions and word spotting. *Memory & Cognition*, 24, 744-755.
- Vroomen, J. & de Gelder, B. (1999). Lexical access of resyllabified words: evidence from phoneme monitoring. *Mem Cognit*, 27 (3), 413-421.

- Wauquier-Gravelines, S. (1996). *Organisation phonologique et traitement de la parole continue*. Thèse de doctorat en linguistique théorique et formelle, UFRL, Paris 7, ms non publié.
- Wauquier-Gravelines, S. & Braud, V., (2005). Proto-déterminant et acquisition de la liaison obligatoire en français. *Langages*, 158, 53-65.
- Wauquier-Gravelines, S. (2004). Les liaisons dangereuses. Psycholinguistique et phonologie : une interface complexe, in Ferrand, L. & Grainger, J. (eds.) *Psycholinguistique cognitive, Essais en l'honneur de Juan Segui*, Bruxelles : De Boeck Université, Collection Neurosciences et Cognition.
- Wauquier, S., (2009). Acquisition de la liaison en L1 et L2 : stratégies phonologiques ou lexicales ? *AILE Lia*, 2, 93-130.
- Wauquier S. (2010), Des ornithorynques et consonnes « doublement flottantes », pour une théorisation unifiée de la liaison. *Faire signe*, pour Pierre Encrevé.
- Wauquier S. (2010), Acquisition de la phonologie « du » français, des usages à la structure. In M. Barra (ed), *Le(s) français: formaliser la variation, Langue française*, 168 (4), 127-144.
- Wauquier, S. & Yamaguchi, N. (2013). Templates in French, in Vihman, M. & Keren-Portnoy, T., *The Emergence of Phonology : Whole word Approaches and Cross-linguistic Evidence*, Cambridge : Cambridge University Press, 317-342.
- Wauquier, S. & Shoemaker, E. (2013). Convergence and divergence in the acquisition of French liaison by native and non-native speakers: A review of existing data and avenues for future research. *Language, Interaction and Acquisition*, 4 (2), 161-189.
- Weiss, J., J., & Hagan, R. (1988). A key to literacy: Kindergarteners' awareness of the functions of print. *The Reading Teacher*, 41, 574-579.
- Whitehurst, G., J., & Lonigan, C., J. (1998). Child development and emergent literacy. *Chil. Dev.*, 69 (3), 848-72.
- Ziegler, J., C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review*, 5, 683-689.
- Ziegler, J., C, Muneaux, M., & Grainger, J. (2003). Neighborhood effects in auditory word recognition: Phonological competition and orthographic facilitation. *Journal of Memory & Language*, 48, 779-793.
- Ziegler, J., C., Ferrand, L. & Montant, M. (2004). Visual phonology: The effects of orthographic consistency on different auditory word recognition tasks. *Memory & Cognition*, 32, 732-74.
- Ziegler, J., C., & Goswami, U., C. (2005). Reading acquisition, developmental dyslexia and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, 131(1), 3-29.
- Ziegler, J., C., & Muneaux, M. (2007). Orthographic facilitation and phonological inhibition in spoken word recognition: a developmental study. *Psychonomic Bulletin & Review*, 14, 75-80.

Annexes

Annexe 1. Formulaire d'autorisation et questionnaire pour les parents



Titre de thèse : « Étude longitudinale des phénomènes consonantiques du français chez les enfants entre compétence phonologique et apprentissage de l'écriture »

Doctorante : Samantha Ruvoletto
(Université Paris 8 / Laboratoire « Structures Formelles du Langage » UMR 7023)

Directrice de thèse : Sophie Wauquier
(Laboratoire « Structures Formelles du Langage » UMR 7023)

Cette étude porte sur la manière dont le langage oral influence la langue écrite, et en particulier comme les enfants qu'ils commencent à apprendre l'écriture au CP vont considérer les phénomènes consonantiques, comme la liaison, que dans la langue française sont sujets d'un traitement différent entre l'écrit et l'orale et qu'ils sont toujours sources des erreurs d'orthographe.

Nous avons choisi de suivre des enfants à partir du CP jusqu'à CE1. Pendant les deux années d'école primaire nous allons proposer régulièrement des tâches linguistiques pour comprendre l'évolution des relations entre écriture et langue parlée selon le niveau scolaire.

Au fin d'améliorer l'analyse des données recueillies successive nous vous demandons de compléter le questionnaire parental ci-joint. Cela nous permettra d'établir un profil sociolinguistique plus précis de votre enfant.

Les tâches choisies ne posent pas de danger pour aucun des participants et seront administrées pendant l'horaire scolaire dans les bâtiments de l'école en accord avec la direction et les enseignants.

Nous étudions et utilisons les enregistrements audio et vidéo uniquement au sein de la communauté universitaire et pour les seuls buts concernant la recherche.

Nous vous invitons à compléter l'autorisation et le questionnaire parental et, en suite, leurs donner à l'enseignant.

☞

Je soussigné....., parent / tuteur légal de l'élève autorise :

(Mettre une croix dans la ou les cases correspondantes)

- la participation de mon fils / ma fille à l'activité de recherche de Mlle Samantha Ruvoletto pendant les années scolaires 2013-2014 et 2014-2015 sur les phénomènes consonantiques du français entre compétence phonologique et apprentissage de l'écriture.

- le traitement et l'utilisation anonymes des données et des enregistrements dans la situation spécifique de diffusion scientifique (conférences, articles scientifiques).

Toutes les données seront utilisées seulement pour les buts de la recherche et tous les droits à la vie privée seront respectés.

Date,..... Signature

QUESTIONNAIRE POUR LES PARENTS

Toutes les données sont traitées sous forme anonyme et utilisées seulement pour le projet de thèse. Aucune question ne concerne l'enfant. Les autres questions vous concernent et concernent les activités de lecture partagée avec votre enfant.

Le nom et le prénom demandés sont seulement indicatifs.
La date de naissance est utilisée pour calculer l'âge par mois et années.
Les autres informations ont valeur purement linguistique afin de contribuer à la recherche.

Enfant

Prénom :..... Nom :.....
.....

Date de naissance :.....
Nationalité :.....
Langue /s
parlée /
s :.....

Parents : Mère (ou tuteur légal)

Nationalité :.....

Langue/s parlée/s (écrire aussi la variété du français, par exemple « Français (Nord de la France) » :

Langue 1 :

Langue 2 :

Langue 3 :

Langue 4 :

.....

Donnez votre évaluation sur la compétence en cette/ces langue (entre **langue maternelle – bilingue ou trilingue – courant – moyen – notions – scolaire**) :

Langue 1 :

Langue 2 :

Langue 3 :

Langue 4 :

.....
Profession actuelle :
Diplôme ou niveau
d'étude :

Parents : Père (ou tuteur légal)

Nationalité :
Langue /s parlée /s (écrire aussi la variété du français, par exemple « Français (Nord de la France) »):.....
.....
Langue 1 :
Langue 2 :
Langue 3 :
Langue 4 :
.....

Donnez votre évaluation sur la compétence en cette/ces langue (entre **langue maternelle – bilingue ou trilingue – courant – moyen – notions – scolaire**) :
Langue 1 :
Langue 2 :
Langue 3 :
Langue 4 :
.....

Profession actuelle :
Diplômes ou niveau
d'étude :

Je vous en prie de répondre à ces 3 questions

1) Quelle langue utilisez-vous à la maison ?

.....
.....
.....

2) Quelle langue utilisez-vous pour parler avec votre enfant ?

.....
.....
.....

3) Est-ce que vous lisez quelques livres illustrés avec votre enfant ? Si oui, quel type de livre et avec quelle fréquence ? (par exemple « une fois par semaine, une fois par jour »). Donnez quelques titres des livres préférés.

.....
.....

LIVRET DE PASSATION

Maîtrise de la langue

JOUR 1

Exercice n°1

Dénombrer les syllabes d'un mot

Dire aux élèves :

- Ecoutez bien et regardez bien. Je vais vous dire le nom de chaque dessin. Ensuite, vous compterez les syllabes de ces mots et vous dessinerez le nombre de ronds correspondants.

Nous allons faire 2 exemples ensemble. Posez votre doigt sur le † : « lapin » et « escargot ».

- Ecoutez bien, maintenant vous le ferez seul. Posez votre doigt sur la ☹ :
 - mouton
 - éléphant
 - chat
 - caméléon
 - hérisson

Codage pour la correction :

- Item 1**
- 1 Le nombre de syllabes est exact pour tous les mots
 - 4 Erreur sur le mot de 4 syllabes (*caméléon*)
 - 9 Erreur sur les mots de 1, 2 ou 3 syllabes
 - 0 Absence de réponse

Exercice n°2**Localiser une syllabe dans un mot**

Dire aux élèves :

- Vous allez trouver où est une syllabe demandée dans un mot. Je vais vous dire le nom de chaque dessin. Ensuite, vous cocherez la case qui correspond à l'endroit où l'on entend la syllabe.
- Nous allons faire 2 exemples ensemble. Posez votre doigt sur le  :
 - « maman ». Si vous entendez {man} dans la première syllabe du mot, vous cocherez la première case ; si vous entendez {man} dans la deuxième syllabe, vous cocherez la deuxième case. *Répéter et vérifier que les élèves ont bien compris la consigne.*
 - « manteau ». Cochez la case qui correspond à la place de la syllabe {man}.
- Posez votre doigt sur la . Ecoutez bien ; vous cocherez à chaque mot la case qui correspond à la syllabe {ro}.
 - robot, robinet, Zorro, arrosoir
- Posez votre doigt sur le . Ecoutez bien ; vous cocherez à chaque mot la case qui correspond à la syllabe {mi}.
 - fourmi, momie, miroir, cheminée
- Posez votre doigt sur la . Ecoutez bien ; vous cocherez à chaque mot la case qui correspond à la syllabe {la}.
 - lavabo, koala, lapin, chocolat

Codage pour la correction :

- | | | |
|---------------|----------|--|
| | 1 | La localisation de la syllabe est exacte pour au moins 12 mots |
| Item 2 | 4 | La localisation de la syllabe est exacte pour au moins 9 mots |
| | 9 | La localisation de la syllabe est exacte pour moins de 9 mots |
| | 0 | Absence de réponse |

Exercice n°3**Isoler un phonème dans un mot***Utilisation du cache*

- Vous cherchez un son dans un mot. Je vous dirai le nom de chaque dessin. Ensuite, vous entourerez l'image si vous entendez le son demandé, vous barrerez l'image si vous n'entendez pas le son demandé.
- Nous allons faire un exemple ensemble. Posez votre cache sous la ligne de la C. Ecoutez bien ; entourez l'image si vous entendez le son {ch}, barrez l'image si vous n'entendez pas le son {ch}. Attention, on cherche le son {ch} : serpent, château, mouche, girafe, biche.

Répéter le son recherché régulièrement (au moins 2 fois pour une même ligne).

Dire aux élèves :

La consigne est la même pour le reste de l'exercice :

- Posez votre cache sous la ligne du Δ. Ecoutez bien ; entourez l'image si vous entendez le son {ou}, barrez l'image si vous n'entendez pas le son {ou} :
 - ours, téléphone, cheval, hibou, bouchon
- Posez votre cache sous la ligne du O. Ecoutez bien ; entourez l'image si vous entendez le son {i}, barrez l'image si vous n'entendez pas le son {i} :
 - pirate, champignon, ballon, hirondelle, tomate
- Posez votre cache sous la ligne du □. Ecoutez bien ; entourez l'image si vous entendez le son {m}, barrez l'image si vous n'entendez pas {m} :
 - banane, maison, nénuphar, pomme, moustique

Codage pour la correction :

1 point par dessin bien entouré ou barré (note globale sur 15)

1 oubli est considéré comme une erreur

Item 3	1	14 ou 15 points
	4	10, 11, 12 ou 13 points
	9	De 0 à 9 points
	0	Absence de réponse

Exercice n°4**Discriminer et mémoriser un mot**

Montrer l'étiquette- mot CAROTTE (voir annexe 1)

Dire aux élèves :

- Regardez bien ce mot. C'est le mot CAROTTE. Regardez-le bien, après il sera caché. Il faudra se rappeler de ce mot, il faudra le garder en mémoire.

Cacher l'étiquette.

- Le mot CAROTTE est écrit 2 fois. Entourez les 2 mots CAROTTE.

Codage pour la correction :

- 1 Les 2 mots CAROTTE sont entourés.
- 4 Un seul des 2 mots CAROTTE est entouré et aucun autre mot.
- Item 4** 9 Aucun mot CAROTTE n'est entouré ou le mot CAROTTE + 1 autre mot sont entourés.
- 0 Absence de réponse

Exercice n°5**Etablir des relations entre quantité d'oral et quantité d'écrit**

Nous allons faire 1 exemple ensemble :

Dire aux élèves :

- Posez votre doigt sur . Vous voyez 3 étiquettes. Je vais vous lire ce qui est écrit dans l'une des étiquettes. Ecoutez bien « chocolat ». Entourez l'étiquette que j'ai lue.
- Mettez votre doigt sur le . Ecoutez bien « Le chat attrape la souris. ». Entourez l'étiquette que j'ai lue.
- Mettez votre doigt sur la . Ecoutez bien « une grosse voiture bleue ». Entourez l'étiquette que j'ai lue.
- Mettez votre doigt sur . Ecoutez bien « les oiseaux ». Entourez l'étiquette que j'ai lue.

Codage pour la correction :

- Item 5
- 1 Les 3 réponses sont correctes
 - 4 2 réponses sont correctes (1 erreur ou 1 oubli)
 - 9 1 seule réponse est correcte ou aucune réponse n'est correcte
 - 0 Absence de réponse

Exercice n°6	Discrimination visuelle
--------------	-------------------------

Dire aux élèves :	<ul style="list-style-type: none">• Dans chaque cadre vous avez un modèle qui est au début de la ligne et qui est dans la grande case. <i>(l'enseignant peut montrer sur le document)</i>• Vous devez entourer tous les groupes de lettres qui sont tout à fait pareils au modèle. <i>Proposer un service de gommage</i>• Regardez bien le modèle ; entourez les groupes de lettres qui sont tout à fait pareils au modèle.
-------------------	---

Codage pour la correction :

- Item 6
- 1 7 bonnes réponses entourées (1 oubli toléré)
 - 4 2 oublis tolérés mais pas d'erreur
 - 9 1 erreur ou plus (mauvaise reconnaissance)
 - 0 Absence de réponse

Exercice n°7	Discriminer et mémoriser un mot
--------------	---------------------------------

Dire aux élèves :	<ul style="list-style-type: none">• Vous vous souvenez, je vous avais demandé de vous rappeler du mot CAROTTE <i>Ne pas remonter l'étiquette</i>• Voici une deuxième liste de mots. Entourez le mot CAROTTE, il y est aussi 2 fois.
-------------------	--

Codage pour la correction :

- 1 Les 2 mots CAROTTE sont entourés
4 Un seul des 2 mots CAROTTE est entouré et aucun autre mot
Item 7 9 Aucun mot CAROTTE n'est entouré
ou le mot CAROTTE + 1 autre mot sont entourés
0 Absence de réponse

Fin du jour 1

Lire l'histoire « Retour de vacances » dans la journée (voir annexe 2)

JOUR 2

Exercice n°8

Reconnaître des mots courants

Dire aux élèves :

- Vous voyez des mots écrits en écriture scripte. Vous devrez entourer les mots que je demande. S'assurer à chaque fois que les élèves ont pris le bon feutre.
- Maintenant prenez le feutre vert et entourez le mot « papa ».
- Maintenant prenez le feutre bleu et entourez le mot « lundi ».
- Maintenant prenez le feutre rouge et entourez le mot « anniversaire ».
- Maintenant prenez le feutre jaune et entourez le mot « le ».
- Maintenant prenez le feutre rose et entourez le mot « école ».
- Maintenant prenez le feutre noir et entourez le mot « une ».

Codage pour la correction :

- Item 8 1 De 4 à 6 mots entourés avec la bonne couleur

- 4 3 mots entourés avec la bonne couleur
- 9 Autres réponses
- 0 Absence de réponse

Exercice n°9	Reconnaître les types d'écrits rencontrés dans la vie quotidienne
---------------------	--

Dire aux élèves :	<ul style="list-style-type: none"> • Sur la page, il y a des séries d'images. • Posez votre doigt sur →. Prenez votre crayon à papier, entourez l'image qui correspond à « journal ». • Posez votre doigt sur le ✱. Prenez votre crayon à papier, entourez l'image qui correspond à « album ».
-------------------	---

Codage pour la correction :

- | | |
|---------------|---|
| Item 9 | <ul style="list-style-type: none"> 1 Les 2 types d'écrits sont identifiés 4 1 seul type d'écrit est identifié 9 Aucun type d'écrit n'est identifié 0 Absence de réponse |
|---------------|---|

Exercice n°10	Se repérer dans un livre (connaissance du vocabulaire technique)
----------------------	---

Dire aux élèves :	<ul style="list-style-type: none"> • Sur cette page vous avez 4 images qui représentent 4 pages d'un livre. Elles sont en désordre. Prenez votre feutre rouge, cherchez la page de <u>couverture</u>. Quand vous l'avez trouvée, faites une croix rouge dans la case sous l'image. • Maintenant prenez le feutre vert. Cherchez la <u>quatrième de couverture</u> ou le dos du livre. Quand vous l'avez trouvée, faites une croix verte dans la case sous l'image. • Maintenant prenez le feutre bleu, entourez le <u>titre</u> de l'histoire. • Prenez le feutre jaune, entourez le dessin du <u>héros</u> de
-------------------	--

cette histoire (c'est le personnage principal de l'histoire).
Si des enfants signalent qu'il y est plusieurs fois, dites qu'ils peuvent alors entourer plusieurs fois.

Codage pour la correction :

- Item 10**
- 1 Tous les repérages sont corrects
 - 4 3 repérages sont corrects
 - 9 Moins de 3 repérages sont corrects
 - 0 Absence de réponse

Exercice n°11

S'orienter dans l'espace de la page

Dire aux élèves :

- Cette fois-ci, on a photocopié une page dans un album.
- Entourez le numéro de la page de cet album.
- Mettez une croix où on doit commencer la lecture.
- Soulignez le dernier mot de la page de l'album.

Codage pour la correction :

- Item 11**
- 1 Les 3 réponses sont correctes
 - 4 2 réponses sont correctes
 - 9 Moins de 2 réponses sont correctes
 - 0 Absence de réponse

Exercice n°12

Connaître le langage technique

Dire aux élèves :

- Posez votre doigt sur ✎. Regardez bien tout ce qui est écrit sur cette ligne, entourez le chiffre.
- Posez votre doigt sur ✎, entourez la lettre

- Posez votre doigt sur le 🚲, entourez le point d'exclamation
- Posez votre doigt sur la 🏠, entourez le mot
- Posez votre doigt sur ☆, entourez le premier mot
- Posez votre doigt sur la 🎵, entourez la majuscule
- Posez votre doigt sur 👁, entourez la phrase
- Posez votre doigt sur le ↩, entourez le dernier mot
- Posez à nouveau votre doigt sur le ↩, entourez la première lettre du mot souligné

Codage pour la correction :

	1	Au moins 7 réponses correctes
Item 12	4	5 ou 6 réponses correctes
	9	Moins de 5 réponses correctes
	0	Absence de réponse

Fin du jour 2

Relire l'histoire « Retour de vacances » dans la journée

JOUR 3

Exercice n°13

Comprendre des consignes données de manière collective.

Consignes pour le maître :

Dire les consignes deux fois sans les reformuler.
S'assurer que tous les élèves ont terminé pour passer à la tâche suivante.

Dire aux élèves :

- Sur la feuille, il y a une image. On voit un garçon qui donne à boire à un cheval et une petite fille qui jardine. Vous allez bien écouter et, quand je vous le dirai, vous utiliserez les crayons de couleur pour faire ce que je vous demande.
- Écoutez bien :
 - Coloriez en vert les chaussettes de la petite fille, puis entourez l'arrosoir.
 - Barrez en bleu la main qui est posée sur le robinet et reliez le cheval à la barrière.

Codage pour la correction :

Items 13

- 1 Les 4 actions sont bien exécutées
- 4 3 actions sont bien exécutées
- 9 Moins de 3 actions sont bien exécutées
- 0 Absence de réponse

Exercice n°14	Comprendre le temps des verbes pour exprimer le passé, le présent, le futur
----------------------	--

Dire aux élèves :

Il y a 4 images, elles ne sont pas dans l'ordre. Écoutez bien.

- Entourez en jaune l'image qui montre « la petite fille qui décorera le bonhomme de neige ».
- Entourez en rouge l'image qui montre « la petite fille qui a décoré le bonhomme de neige ».
- Écoutez bien, entourez en bleu l'image qui montre « la petite fille qui décore le bonhomme de neige ».

Codage pour la correction :

(2 réponses sont possibles pour « décorera »)

Item 14

- 1 Les 3 temps des verbes sont identifiés
- 4 2 des temps sont compris
- 9 Autres réponses
- 0 Absence de réponse

Exercice n°15

Comprendre une histoire lue par l'enseignant, la raconter en restituant les enchaînements chronologiques.

Prévoir les images en Annexe 2bis

Les découper à l'avance et les distribuer, faces cachées, avant la relecture.

Dire aux élèves :

- Je vais vous relire l'histoire « Retour de vacances ». Quand j'aurai fini de lire, je vous demanderai de prendre seulement les images qui parlent de l'histoire. Il faudra les ranger (les coller) devant vous pour qu'elles racontent « Retour de vacances ».
- Vous collerez celles qui ne parlent pas de cette histoire dans la « boîte ».

Codage pour la correction :

Item 15

- 1 L'ordre chronologique est correct pour les images utilisées
- 4 Deux événements sont intervertis
- 9 La plupart des événements sont intervertis
- 0 Absence de réponse

Item 16

- 1 Les 2 images incorrectes sont dans la boîte (Poursuite de 2 chiens dans la rue et Léo sur un quai)
- 4 Pas de code 4 pour cet item
- 9 La sélection d'images dans la boîte n'est pas correcte
- 0 Absence de réponse

Exercice n°16

Comprendre un vocabulaire pertinent.

Dire aux élèves :

- Sur la page, il y a une série d'images. Vous écouterez chaque mot puis, quand je vous le dirai, vous utiliserez le crayon à papier pour entourer l'image qui correspond au mot.
- Posez votre doigt sur le . Ecoutez bien : « parasol ».
Prenez le crayon à papier pour entourer l'image qui correspond à « parasol ».
- Idem pour :
 - ☀ -> « indiquer »
 - ↔ -> « renverser »
 - 🤗 -> « joyeusement »

Codage pour la correction :

- Item 17**
- 1 Les 4 mots sont compris
 - 4 3 mots sont compris
 - 9 Moins de 3 mots sont compris
 - 0 Absence de réponse

JOUR 4

Exercice n°17

Comprendre un vocabulaire pertinent

Dire aux élèves :

- Sur cette page, vous voyez une silhouette de bonhomme. Dessinez ce que je vais vous demander : des cheveux frisés, un chapeau pointu. (*Laisser dessiner*).
- Dessinez un bouquet dans une main et terminez son visage.

Codage pour la correction :

- Item 18**
- 1 Les 7 éléments (cheveux, frisés, chapeau, pointu, bouquet, dans la main, visage avec au moins yeux nez et bouche) sont illustrés
 - 4 5 éléments au moins sont illustrés
 - 9 Moins de 5 éléments sont illustrés
 - 0 Absence de réponse

Exercice n°18**Comprendre le système alphabétique
de codage de l'écriture**

Dire aux élèves :

- Sur les lignes, je vais vous demander d'écrire des mots. C'est un peu difficile, mais vous devez écrire comme vous croyez, le mieux que vous pouvez. Si vous vous trompez, ce n'est pas grave.
- Mettez votre doigt sur le lit. Ecrivez à côté le mot « lit »
- Maintenant mettez votre doigt sur le marteau. Ecrivez à côté le mot « marteau »
- Maintenant mettez votre doigt sur la moto. Ecrivez à côté le mot « moto »
- Maintenant mettez votre doigt sur le dernier dessin. Ecrivez à côté la phrase « La moto est sur le lit »

Codage pour la correction :

- | | | |
|----------------|---|---|
| Item 19 | 1 | Correspondance grapho-phonétique d'au moins 4 lettres |
| | 4 | Correspondance grapho-phonétique de 2 ou 3 lettres |
| | 9 | Moins de 2 correspondances grapho-phonétique |
| | 0 | Absence de réponse |
| Item 20 | 1 | Présence de segmentation de la phrase : au moins 2 espaces |
| | 4 | Présence de segmentation de la phrase : 1 espace |
| | 9 | Pas de segmentation |
| | 0 | Absence de réponse |
| Item 21 | 1 | Permanence de l'écrit : les mots « moto » et « lit » sont écrits de la même façon que dans l'item 19 |
| | 4 | Permanence de l'écrit : un des 2 mots, « moto » ou « lit », est écrit de la même façon que dans l'item 19 |
| | 9 | Pas de permanence de l'écrit |
| | 0 | Absence de réponse |

Exercice n°19**Apprendre le geste de l'écriture**

Consignes de passation :

S'assurer que tous les élèves aient terminé pour passer à la tâche suivante.

Dire aux élèves :

- Vous écrivez votre prénom en écriture cursive (attachée) sur la ligne qui a un . [item 22]
- Vous recopiez la phrase « C'est le mois de septembre » en écriture cursive (attachée), comme sur le modèle, sur la ligne qui a une . [item 23]

Codage pour la correction :

- Item 22**
- 1 Toutes les lettres du prénom sont écrites et lisibles
 - 4 Des lettres sont oubliées
 - 9 Autre réponse
 - 0 Absence de réponse

- Item 23**
- 1 Tous les mots sont écrits en cursive et lisibles
 - 4 3 mots sont bien écrits
 - 9 Toute autre réponse
 - 0 Absence de réponse

BONUS

Dessiner un bonhomme

Dire aux élèves :

- Tournez la page. Sur la page blanche dessinez un joli bonhomme, le plus joli que vous pouvez.

Aucun codage pour cet exercice bonus.

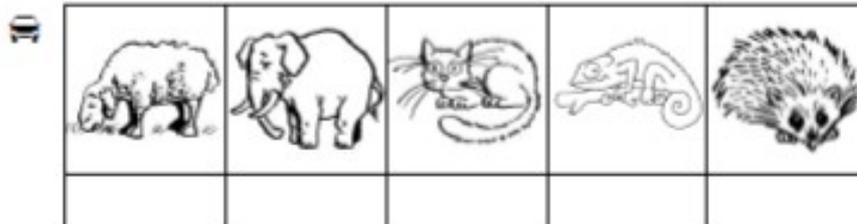
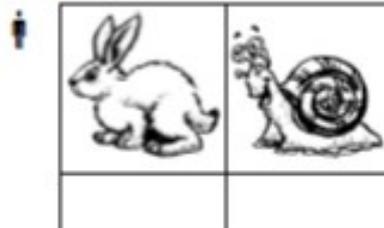
Celui-ci permet d'évaluer l'idée que l'enfant possède de son schéma corporel ainsi que de donner des informations sur l'image du corps de l'enfant.

Annexe 3. Matériel donné aux enfants dans le livret de passation

JOUR 1

Exercice 1 :

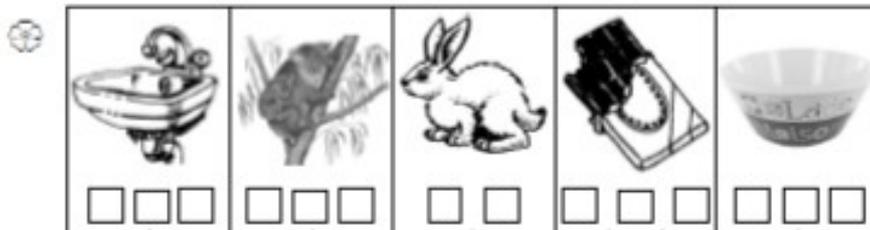
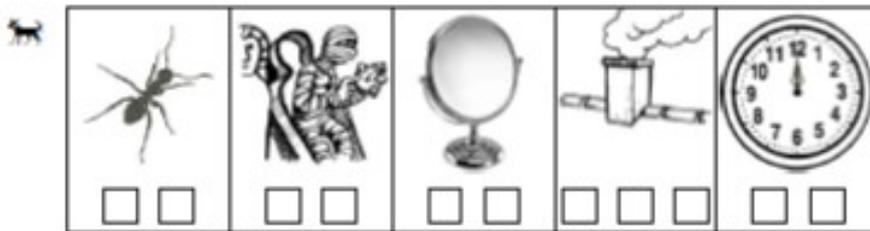
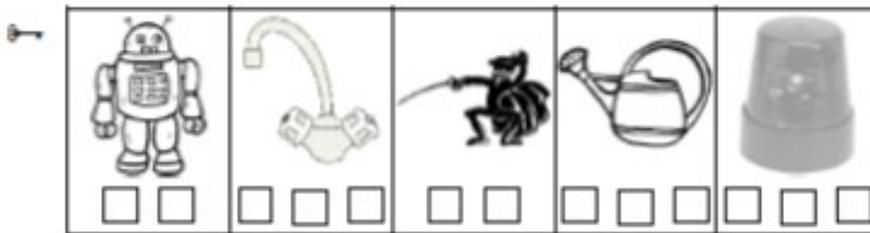
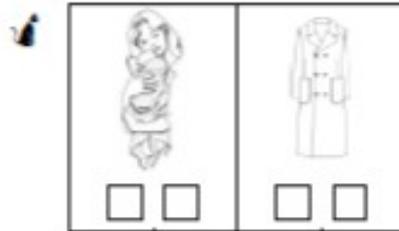
Exemple :



Item 1

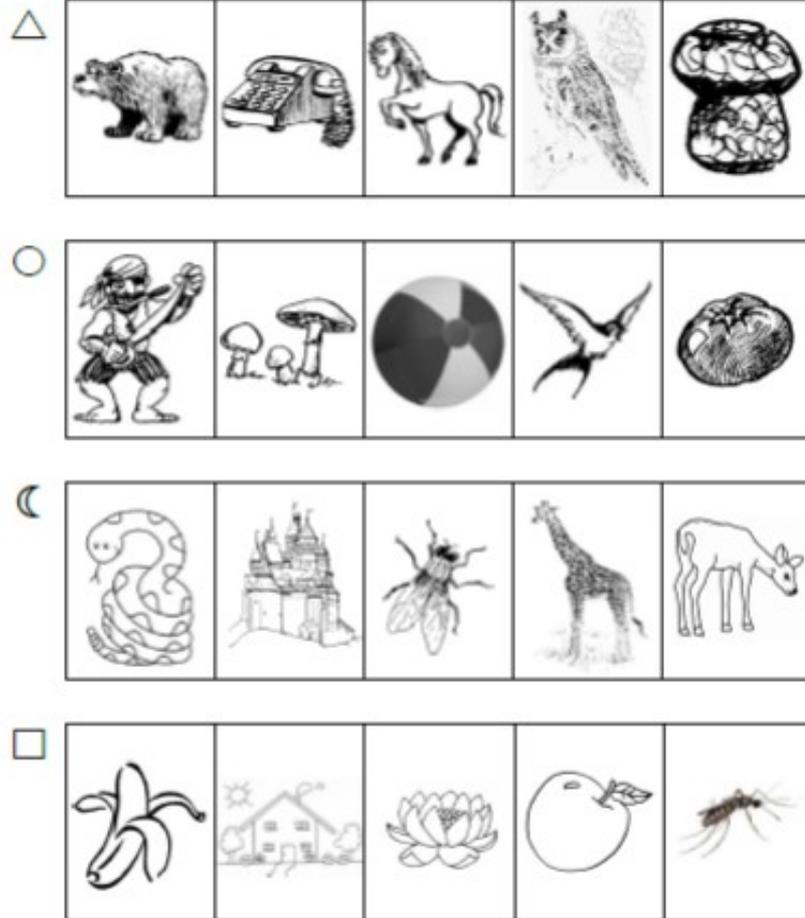
1	4	9	0
---	---	---	---

Exercice 2 :



Item 2
1 4 9 0

Exercise 3 :



Item 3

1	4	9	0
---	---	---	---

Exercice 4 :



Item 4

1	4	9	0
---	---	---	---

Exercice 5 :

Exemple :



chocolat

Lisa joue avec sa poupée.

Les escargots ont deux petites cornes, une jolie coquille jaune et ils avancent lentement, ils ne sont pas pressés.



Le facteur

Le chat attrape la souris.

Les trois petits cochons s'enfuient devant le loup.



une voiture

une voiture bleue.

une grosse voiture bleue.



les oiseaux

Les oiseaux volent.

Les oiseaux volent très haut.

Item 5

1	4	9	0
---	---	---	---

Exercise 6 :

ra	ro	ar	ra	na	ra	va
----	----	----	----	----	----	----

ham	him	lam	ham	ham
	hom	ham	ahm	ham

MNOP	MNOP	MNOP	MONP	PNOM	NMOP
	MNOP	MMOP	MNOP	NNOP	WNOP

Item 6

1	4	9	0
---	---	---	---

Exercice 7 :



Item 7
1 4 9 0

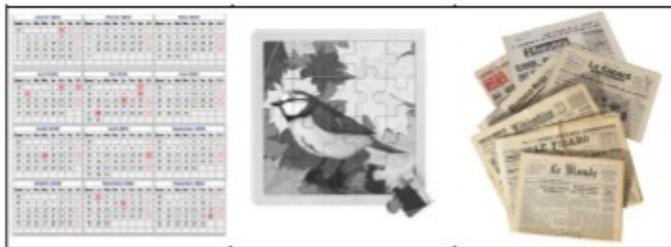
JOUR 2

Exercice 8 :

la	jeudi	les
école	un	ce
cantine	mercredi	papa
lait	anniversaire	une
lundi	samedi	gâteau
maman	le	dimanche

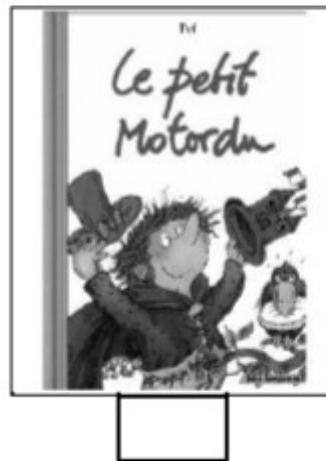
Exercice 9 :

Item 8
9 0



Item 9
1 4 9 0

Exercice 10 :



Exercice 11 :



Exercice 12 :

 a e 2 c ?

 3 6 b , +

 7 où m @ ! île

 123 ☀ soleil qpdbm

 dans école maison lit

 Arthur mange des galettes.

 34295 abcdef428
fille Le chat dort.

 Le chat dort sur le coussin.

Item 12
1 4 9 0

JOUR 3

Exercice 13 :



Item 13
1 4 9 0

Exercice 14 :



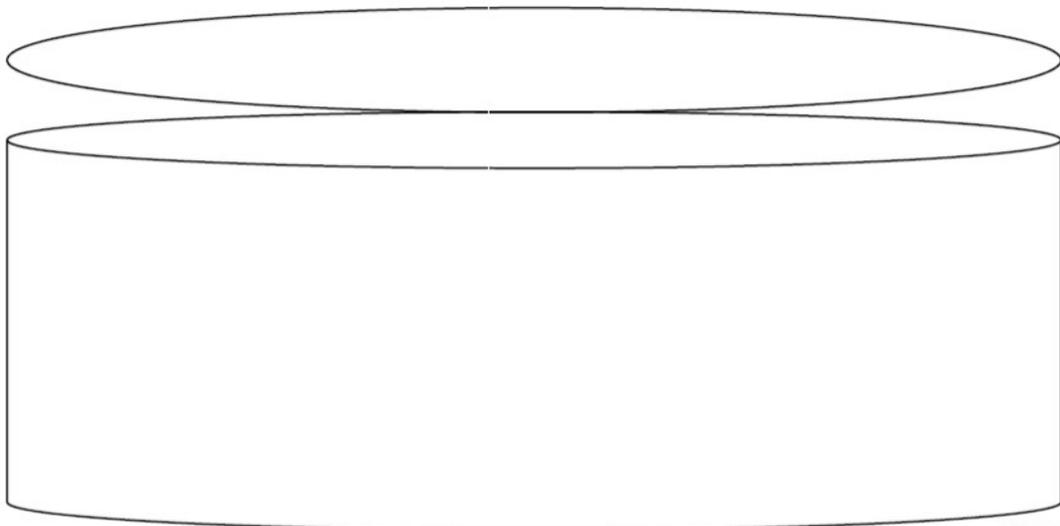
Item 14
1 4 9 0

Exercice 15 :



Item 15

1	4	9	0
---	---	---	---



Item 16

1	4	9	0
---	---	---	---

JOUR 4

Exercice 17 :

Sur cette feuille blanche, dessine le plus beau bonhomme que tu saches faire, un bonhomme frisé avec un chapeau pointu.

Exercise 18 :



.....



.....



.....



.....

Item 19 Item 20 Item 21
1 4 9 0 1 4 9 0 1 4 9 0

Exercice 19 :

Écris ton prénom :



Recopie la date écrite au tableau :



Item 22 Item 23
1 4 9 0 1 4 9 0

Annexe 4. Phrases *fillers* dans PREC

Phrases <i>fillers</i>	Bloc	Réponse attendue	Erreurs et corrections
Les petites filles jouent avec des animaux en peluches	1	Oui	-

J'ai un tapis rouge dans ma chambre	3	Oui	-
Fabienne et Thibaud sont allés au cinéma hier soir	6	Oui	-
La marmotte prennent des pissenlits pour se nourrir	1	Non	<i>prennent</i> (pour <i>prend</i>)
Moi ce que je préfère quand je vais au parc c'est faire du togobban	1	Non	<i>togobban</i> /tɔ.gɔ.bã/ pour <i>toboggan</i> /tɔ.bɔ.gã/
Les pompiers ont des caks /kaks/ qui brillent	1	Non	<i>caks</i> /kaks/ pour <i>casque</i> /kask/
Louis a peur des se(r)pent, surtout les gros	4	Non	<i>se(r)pent</i> /sɛ.pã/ pour <i>serpent</i> /sɛʁ.pã/
Les enfants aiment les glaces	5	Oui	-
Les phoques sautent et nagent dans l'eau	2	Oui	-
Anthony veut manger des biscuits	4	Oui	-
Michelle et Lucille ont mangé des glaces	5	Oui	-
Les glaçons est dans le verre	2	Non	<i>est</i> pour <i>sont</i>
Mon père regarde toujours la télésvion le soir	6	Non	<i>télésvion</i> /te.le.zi.vjõ/ pour <i>télévision</i> /te.le.vi.zjõ/
j'ai arrangé mes jouets dans le grand pla[z]card	2	Non	<i>pla[z]card</i> /plaz.kas/ pour <i>placard</i> /pla.kas/
Les ou(r)s bruns habitent au Canada et ils mangent du miel	3	Non	<i>ou(r)s</i> /us/ pour <i>ours</i> /uʁs/
Maman a cuisiné	6	Oui	-

des pâtes pour nous ce soir			
Maman tire les rideaux pour qu'il fasse nuit avant de dormir	2	Oui	-
Le pirate a trouvé un trésor enterré sous le sable	4	Oui	-
Sylvia et Nadine regardent la télé	4	Oui	-
Les papillons s'envolent tous ensemble	5	Oui	-
Mes parents font les courses au mazagin	3	Non	<i>mazagin /ma.za.gẽ/ pour magasin /ma.ga.zẽ/</i>
La marguerite a des pé[s]tales blanches	3	Non	<i>pé[s]tales /pe.stal/ pour pétales /pe.tal/</i>
Moi ce que je préfère à l'école c'est la réc(r)éation	5	Non	<i>réc(r)éation /ʁe.ke.a.sjõ/ pour recréation /ʁe.ke.a.sjõ/</i>
Je vais à l'école en vélo	6	Oui	-

Annexe 5. Phrases d'entraînement dans PERC.

Phrases d'entraînement	Réponses attendues	Erreurs
Les filles sont très beaux	Non	<i>beaux pour belles</i>
Le [n]éléphant est le plus gros animal de la Savane	Non	<i>néléphant /ne.le.fã/ pour éléphant /e.le.fã/</i>
Le chevaux courent dans le bois	Non	<i>le pour les</i>
J'ai vu des c(r)ocod[r]iles au zoo	Non	<i>cocodriles /kɔ.kɔ.dʁil/ pour crocodiles /kʁɔ.kɔ.dil/</i>
Louis s'est pronemé à la plage	Non	<i>pronemé /pʁɔ.nə.me/ pour promené /pʁɔ.mə.ne/</i>
Le chien mange ses croquettes	Oui	-
Papa m'a a[s]jeté un beau vélo	Non	<i>a[s]jeté /a.sə.te/ pour acheté /a.fə.te/</i>
Hier au cinéma mes amis ont vu un super dessin aminé	Non	<i>dessin aminé /dɛ.sɛ̃ a.mi.ne/ pour dessin animé /dɛ.sɛ̃ a.ni.me/</i>
Je mange des biscuits au chocolat	Oui	-
Maman aime bien les chats	Oui	-

Annexe 6. Entraînement dans DENa

Entraînement		
Singulier → Pluriel		
Contexte	Input : expérimentateur	Réponse attendue : enfant
det+nom	le bateau	les bateaux
det+nom	la fleur	les fleurs
det+nom	la maison	les maisons
adj+nom	un gros poisson	des gros poissons
adj+nom	un grand chien	des grands chiens
Pluriel → Singulier		
Contexte	Input : expérimentateur	Réponse attendue : enfant
det+nom	les stylos	le stylo
det+nom	les livres	le livre
det+nom	les cochons	le cochon
adj+nom	des petits chats	un petit chat

Annexe 7. Filler dans DENa

Contexte <i>filler</i>	Input <i>filler</i> : expérimentateur	Réponse attendue <i>filler</i> : enfant
Singulier → Pluriel		
dét+nom	La pomme	Les pommes
dét+nom	Le cahier	Les cahiers
dét+nom	La cerise	Les cerises
Pluriel → Singulier		
dét+nom	Les gâteaux	Le gâteau
dét+nom	Les zèbres	Le zèbre
dét+nom	Les lapins	Le lapin
dét+nom	Les noix	La noix
dét+nom	Les maisons	La maison
dét+nom	Les parapluies	Le parapluie

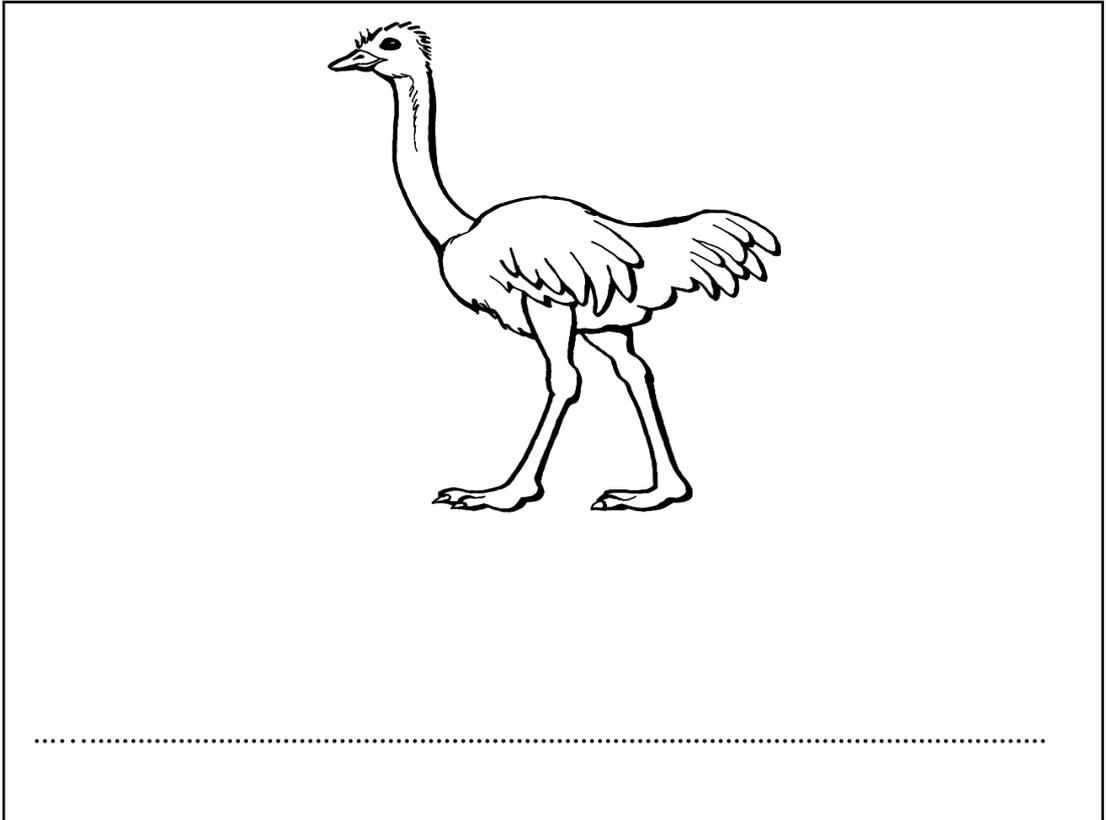
Annexe 8. *Fillers* dans DENb

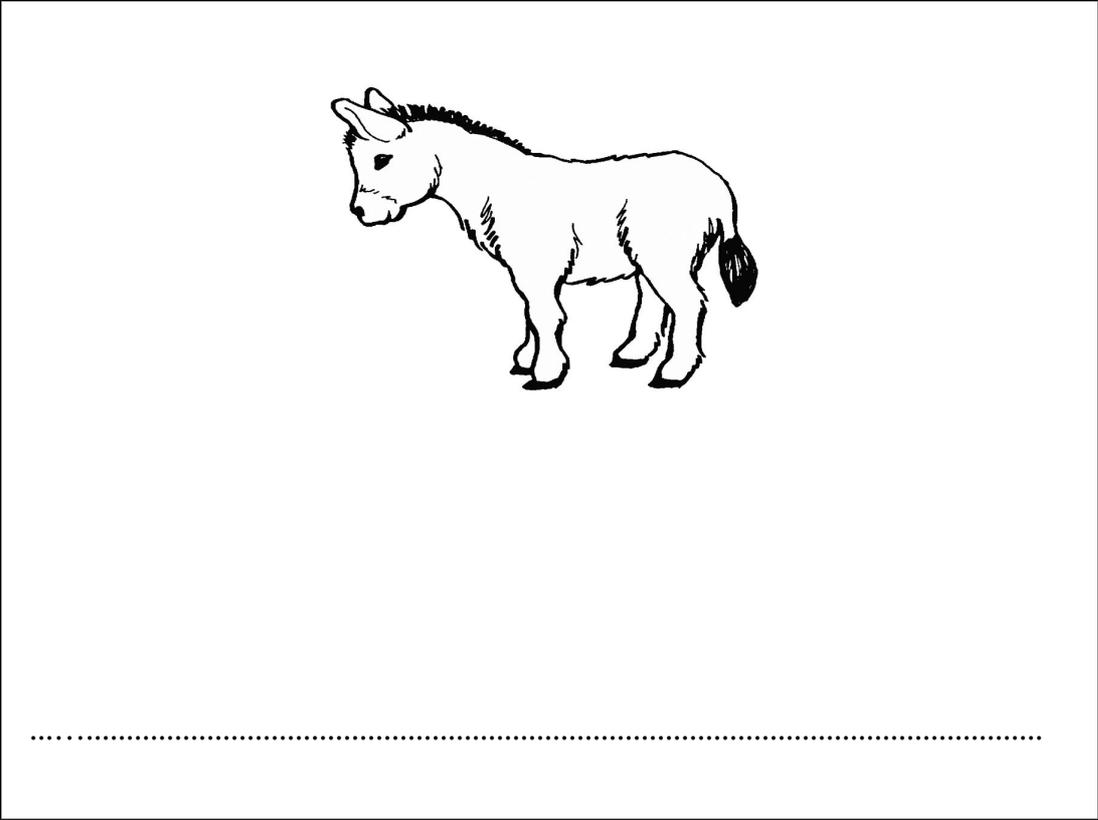
Bloc	Expérimentateur : input <i>filler</i>	Enfants : cible <i>filler</i>
1	Le raisin	Le vélo
	Le zèbre	Le tigre
	Les clés	Les nuages
2	La fraise	La cuillère
	Les chats	Les coquillages
	Les livres	Les chiens
3	La fenêtre	La chaise
	Les stylos	Les canards
	Le lion	Le fantôme

Annexe 9. Ordre de présentation de blocs pour PROD.ECR

Groupes d'enfants	Classe	Ordre de présentation des blocs
Groupe 1, 10 enfants	CP B	1, 2, 3
Groupe 2, 12 enfants	CP B	2, 3, 1
Groupe 3, 11 enfants	CP A	3, 2, 1
Groupe 4, 10 enfants	CP A	1, 3, 2

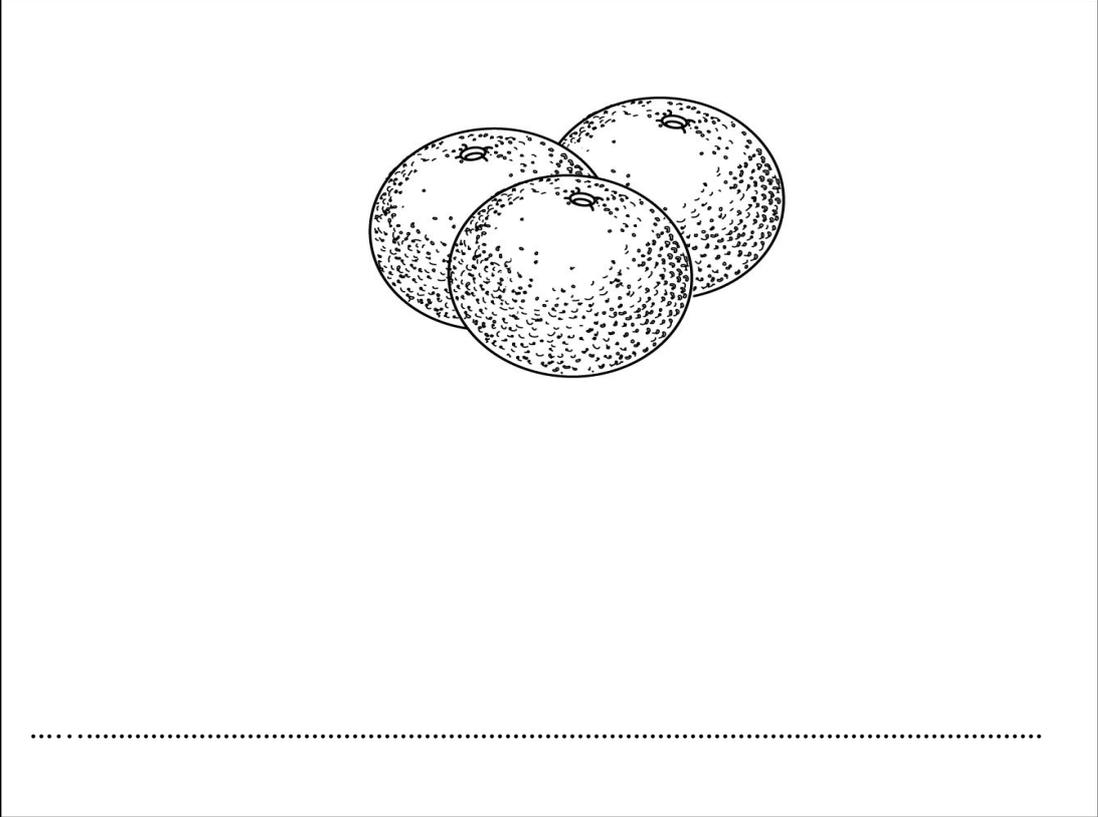
Annexe 10. Livret donné aux enfants pour PROD.ECR

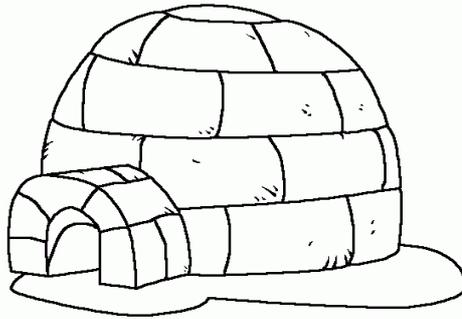




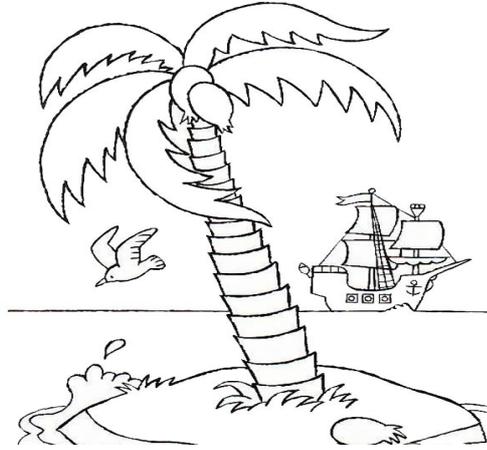


.....

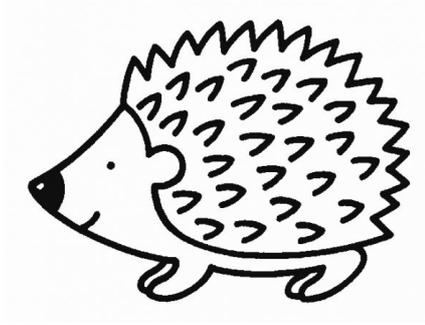




.....



.....



.....

AVIS AUX DESSINATEURS
VENEZ
NOMBREUX
FILER UN COUP DE MAIN
AU DINDON



REPLÉZ DE SERIEUSEMENT SANS DE 24 JANVIER
DANS LE CADRE DE COTIER INVISIBLE

PICCOLETTI
CIRCUS

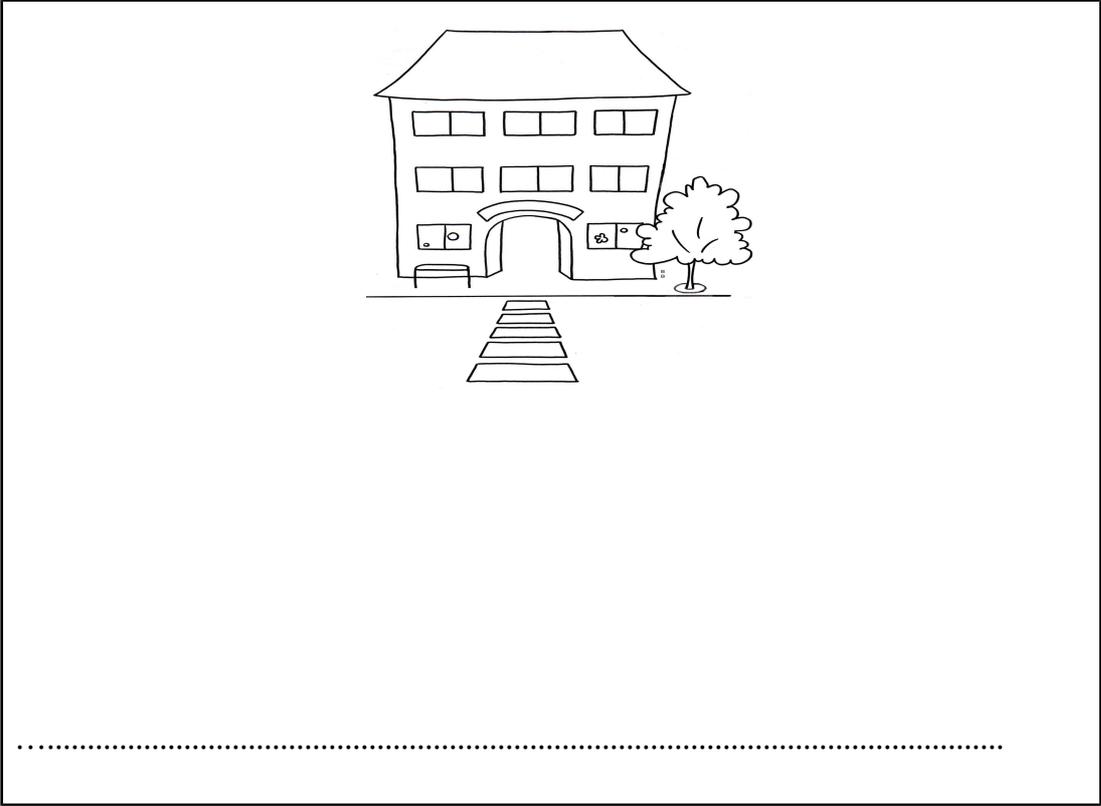


CE SOIR !

coloriages-enfants.com

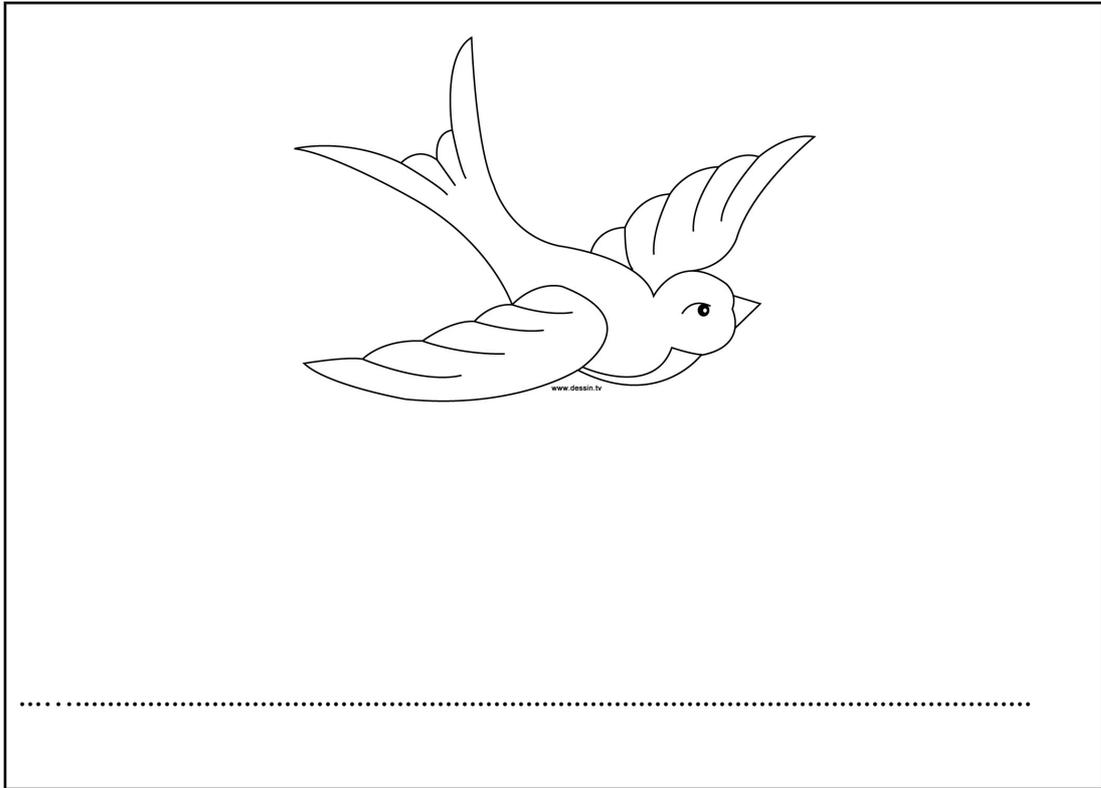


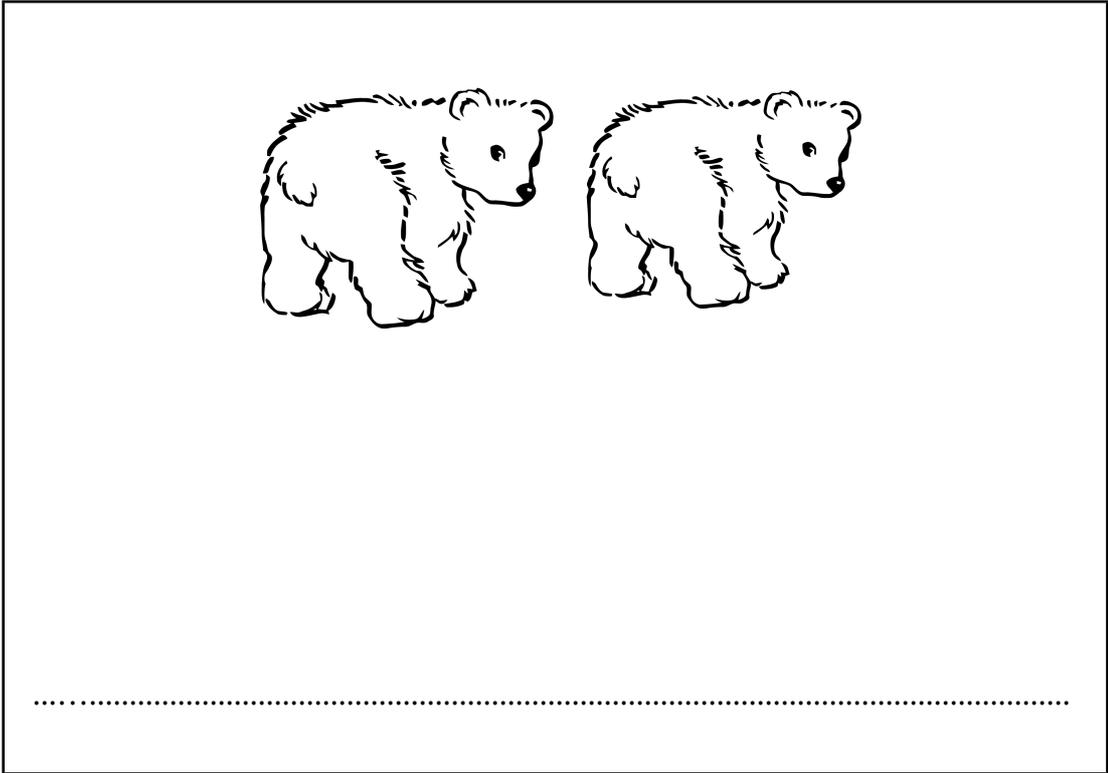
.....

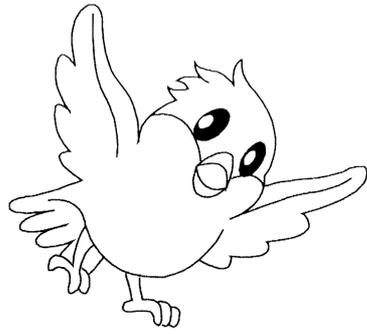


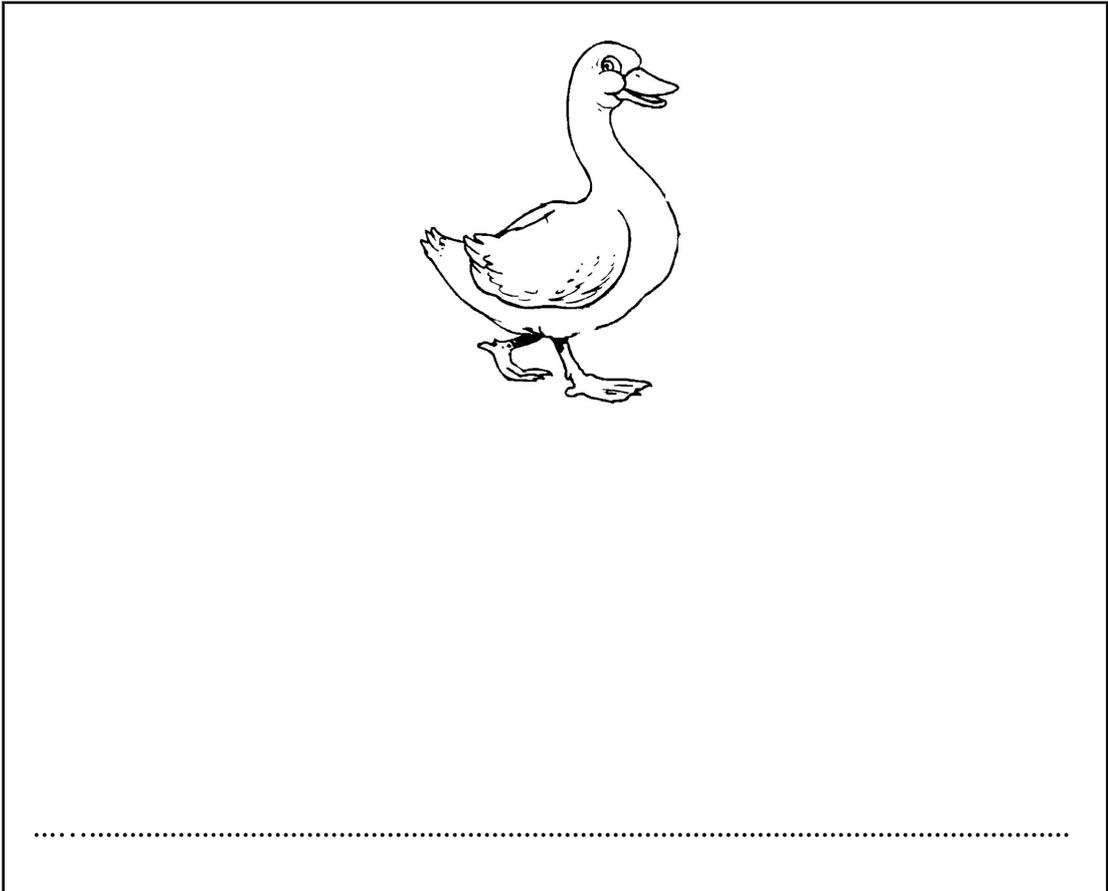


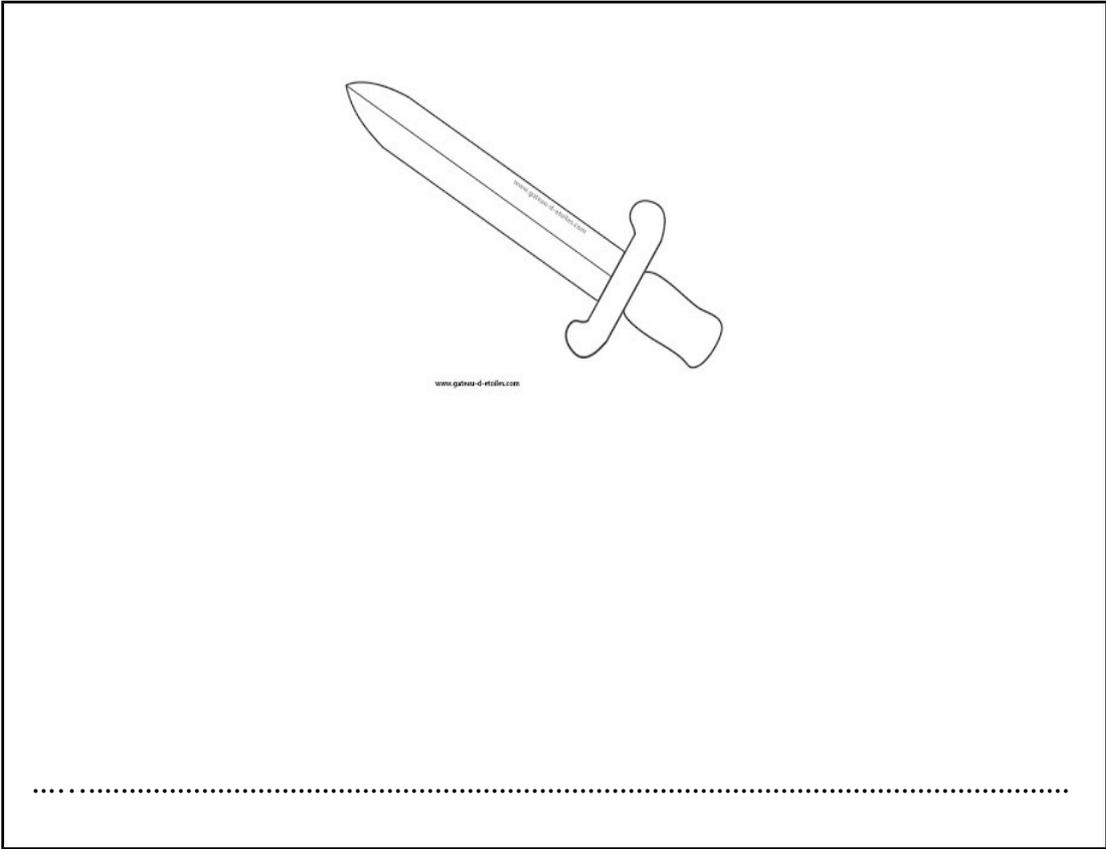
.....

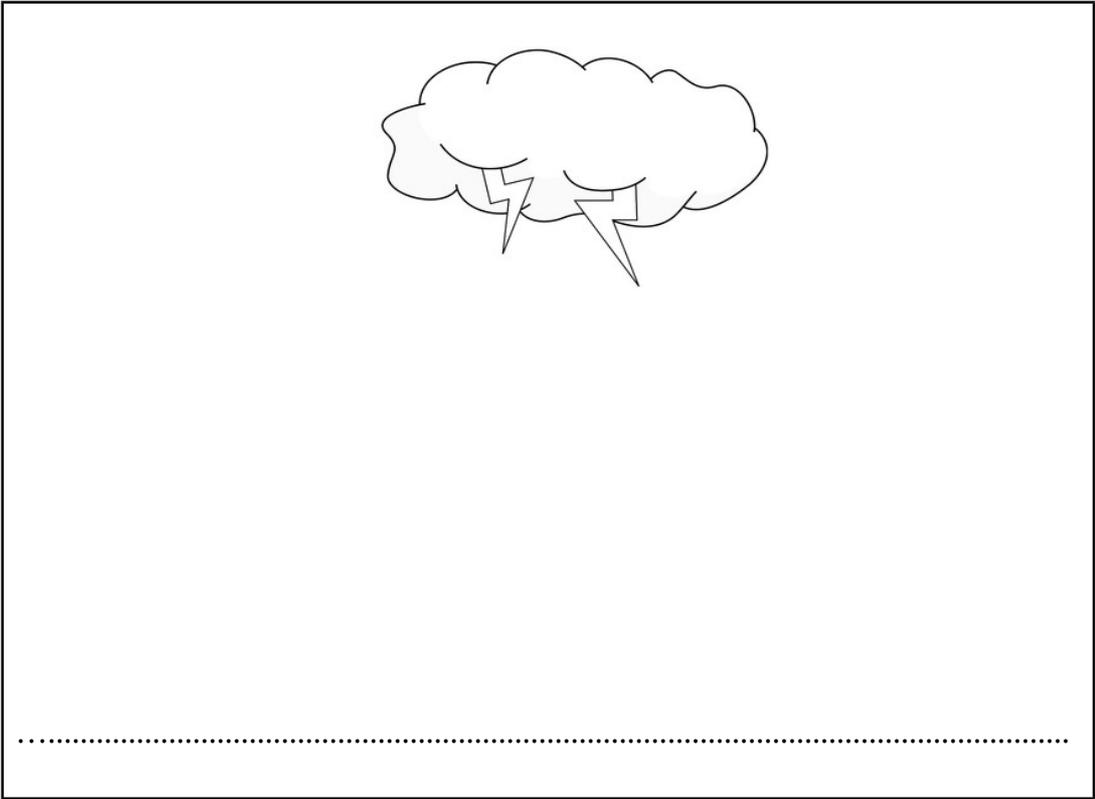


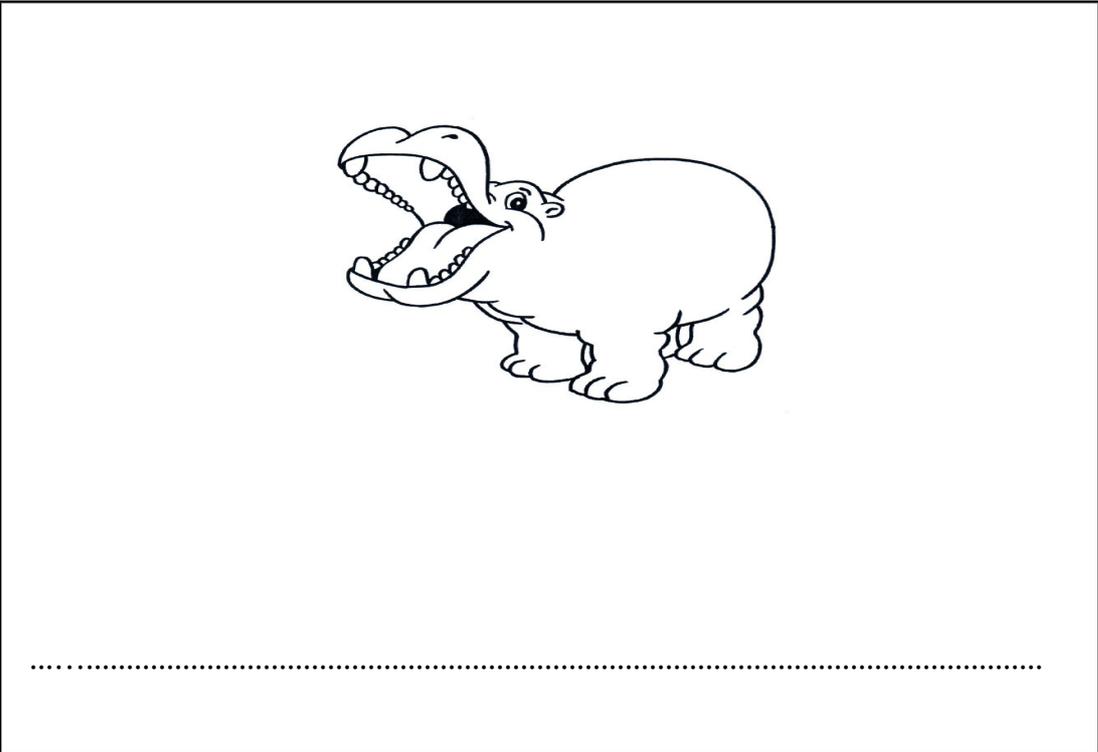


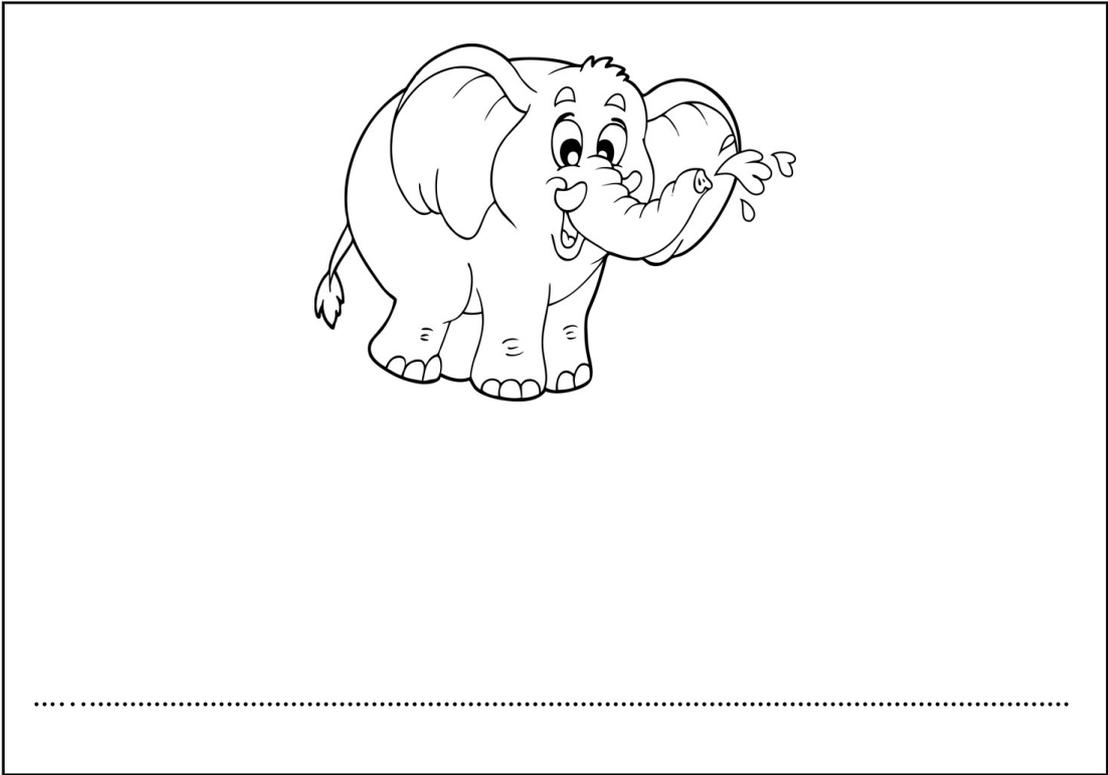












Annexe 11. Ordre de présentation des blocs pour la tâche d'écriture PROD.ECR au CE1

Groupes d'enfants	Ordre de présentation des blocs
Groupe 1, 13 enfants	1, 2, 3
Groupe 2, 13 enfants	2, 3, 1
Groupe 3, 11 enfants	3, 2, 1

Annexe 12. Questionnaire pour les participants à la tâche d'amorçage

QUESTIONNAIRE POUR LES PARTICIPANTS

Toutes les données sont traitées sous forme anonyme et utilisées seulement pour le projet de thèse.

Le nom et le prénom demandés sont seulement indicatifs.

La date de naissance est utilisée pour calculer l'âge par mois et années.

Les autres informations ont valeur purement linguistique afin de contribuer à la recherche.

Prénom : Nom :
.....

Date de naissance :

Nationalité :

Langue /s parlée /s (écrire aussi la variété du français, par exemple « Français (Nord de la France) ») :

Langue 1 :

Langue 2 :

Langue 3 :

Langue 4 :

.....

Donnez votre évaluation sur la compétence en cette/ces langue (entre **langue maternelle – bilingue ou trilingue – courant – moyen – notions – scolaire**) :

Langue 1 :

Langue 2 :

Langue 3 :

Langue 4 :

.....

Profession actuelle :

Diplôme ou niveau

d'étude :

Je vous en prie de répondre à ces 3 questions

1) Est-ce que vous êtes scolarisé en France ? (Mettez une croix dans la case correspondante à

votre réponse) OUI NON

Si OUI, à partir de quel niveau ?

(Mettez une croix dans la case correspondante à votre réponse. Plusieurs réponses sont possibles)

Maternelle

Primaire

Secondaire

2) Quelle langue utilisez-vous à la maison ?

.....
.....

3) Est-ce que pendant la journée vous lisez ou écrivez de textes en **français** ? Pour quelle raison et avec quelle fréquence (« une fois par semaine, une fois par jour ») ? (par exemple « Je suis étudiante donc plusieurs fois par jours je suis obligée de lire et d'écrire »)

.....
.....
.....
.....

Autorisation pour la participation à l'activité de recherche

Je soussigné.....j'autorise la participation à l'activité de recherche de Mlle Samantha Ruvoletto sur les phénomènes consonantiques du français entre compétence phonologique et apprentissage de l'écriture et le traitement et l'utilisation anonymes des données dans la situation spécifique de diffusion scientifique (conférences, articles scientifiques).

Toutes les données seront utilisées seulement pour les buts de la recherche et tous les droits à la vie privée seront respectés.

Date,.....

Signature

.....

Annexe 13. Fillers pour la tâche AMOa

Cible - <i>Filler</i>	Amorçage - <i>filler</i>		
TRAIN	un petit train	un joli train	un grand tonnerre
TAPIS	un petit tapis	un joli tapis	un grand timbre
LION	un petit lion	un joli lion	un grand bruit
LOUP	un petit loup	un joli loup	un grand cheval
SERPENT	un gros serpent	un mini serpent	un vieux bâtiment
SAC	un gros sac	un mini sac	un vieux jouet
ZEBRE	un gros zèbre	un mini zèbre	un curieux tableau
ZOO	un grand zoo	un mini zoo	un curieux détail
NOEL	un bon Noël	un vrai Noël	un fin tissu
NOMBRE	un bon nombre	un vrai nombre	un fin papier
PERE	un bon père	un vrai père	un excellent repas
POLICIER	un bon policier	Un vrai policier	un excellent dessert
HIBOU	un gros hibou	un mini hibou	un affreux château
HERISSON	un bon hérisson	un vrai hérisson	un banal courrier

Annexe 14. Fillers pour la tâche AMOb

Cible - Filler	Amorçage - filler		
RAIN	un petit train	un joli train	un grand tonnerre
TAPIS	un petit tapis	un joli tapis	un grand timbre
LION	un petit lion	un joli lion	un grand bruit
LOUP	un petit loup	un joli loup	un grand cheval
SERPENT	un gros serpent	un mini serpent	un vieux bâtiment
AC	un gros sac	un mini sac	un vieux jouet
EBRE	un gros zèbre	un mini zèbre	un curieux tableau
ZOO	un grand zoo	un mini zoo	un curieux détail
OEL	un bon Noël	un vrai Noël	un fin tissu
NOMBRE	un bon nombre	un vrai nombre	un fin papier
PERE	un bon père	un vrai père	un excellent repas
OLICIER	un bon policier	Un vrai policier	un excellent dessert
HIBOU	un gros hibou	un mini hibou	un affreux château
HERISSON	un bon hérisson	un vrai hérisson	un banal courrier