



Faculté des lettres

Sciences du langage

Thèse de doctorat

présentée à la Faculté des lettres et des sciences humaines de l'Université
de Fribourg (Suisse)

Enseignement ciblé implicite ou explicite :

Étude comparative de l'apprentissage des représentations phonétiques et
phonologiques et du vocabulaire associé.

Approuvée par la faculté des lettres et des sciences humaines sur
proposition des professeur.e.s

Prof. Dr. Raphaël Berthele (premier rapporteur)
Prof. Dr. Jean-Pierre Chevrot (deuxième rapporteur)
Prof. Dr. Anita Thomas (troisième rapporteur)
Prof. Dr. Valérie Camos (membre du jury)
Prof. Dr. Claude Bourqui (président du jury)

Fribourg, le 5 octobre 2021
Le doyen Prof. Dr. Dominik Schöbi

Dherbey Chapuis, Nathalie
Nationalité française

2021

Remerciements

Avec passion

Je cultive ce que vous avez semé

Ces quelques mots ne sauraient dire

Tout le bonheur et la joie d'avoir accompli ce travail

Une brassée de merci chaleureux pour votre soutien sans faille

À Raphaël, fascinant, exigeant, généreux et toujours bienveillant. Pour ce moment de recherche merveilleux

À Cécilia, brillante, noble, positive et curieuse. Pour m'ouvrir ce chemin

À Sandra, captivante, drôle, inépuisable et joyeuse. Pour m'épargner d'un clic tant d'heures de travail

Aux élèves intéressés et curieux. Pour apprendre de bon cœur

Aux enseignantes ouvertes et dynamiques. Pour nos échanges didactiques, pédagogiques et sympathiques

Aux relecteurs courageux et bienveillants, Cecilia, Audrey, Lina et Jean. Pour votre lecture attentive

À l'Institut de Plurilinguisme, la Haute École Pédagogique de Fribourg et l'Université de Fribourg. Pour cet environnement de travail idéal

Au jury de soutenance. Pour cet échange vif, passionné et riche

Aux amis rencontrés en chemin,

À l'équipe Peapl, Thierry, Mireille et Lionel. Pour rêver à demain,

À Isa, Suzanne, Audrey, Giulia, Jan, Philippe, Zorana et tous mes collègues. Pour les moments chaleureux partagés

À Tom, Lucile et Albertine, le cœur de ma vie

Comme il faut tout un village pour l'éducation d'un enfant, il faut toute une équipe pour l'épanouissement d'une chercheuse.

Merci.

Table des matières

Introduction.....	5
1 Apprentissage des représentations phonologiques et du vocabulaire.....	11
1.1 Le modèle de mémoire de travail : Time Based Resource Sharing.....	12
1.2 Modèles de perception et d'apprentissage des représentations phonologiques en LE.....	19
1.3 Perception de l'information verbale et apprentissage des représentations phonologiques et lexicales	25
1.4 Enseignement des représentations phonétiques et phonologiques	31
1.5 Facteurs susceptibles de modifier les effets des séquences didactiques.....	35
1.5.1 Incidence des caractéristiques de l'apprenant sur l'effet du type d'enseignement.....	35
1.5.2 Incidence de la tâche sur l'effet du type d'enseignement	36
1.5.3 Incidence des caractéristiques de l'objet d'enseignement/apprentissage sur l'effet du type d'enseignement.....	47
1.6 Complexité et difficulté des phonèmes cibles	48
1.6.1. La voyelle nasale /ɔ̃/ et ses graphèmes <on ; om>.....	48
1.6.2 Le phonème consonantique /ʒ/ et ses graphèmes <j, g>	52
1.6.3 Complexité et difficulté comparée des deux phonèmes cibles	55
2 Définition des questions de recherche.....	59
2.1 Détermination des questions de recherche	59
2.2 Hypothèses et Prédiction.....	60
2.2.1 Question de recherche principale	60
2.2.2 Questions de recherche secondaire.....	60
3 Méthode expérimentale.....	63
3.1 Séquences didactiques enseignées	63
3.1.1 Prononciation	65
3.1.2 Discrimination	66
3.1.3 Correspondances phonème-graphème.....	67
3.1.4 Mémoire de travail.....	68
3.2 Tâches utilisées pour mesurer l'évolution des performances des compétences enseignées	69
3.2.1 Tâche de discrimination	69
3.2.2 Tâche de prononciation.....	72
3.2.3 Tâche de maintien en mémoire de travail	75
3.2.4 Tâche de mesure des relations phonème-graphème	80
3.2.5 Tâche de mesure du vocabulaire	82
3.3 Protocole expérimental.....	84

3.3.1 Participants.....	85
3.3.2 Procédure expérimentale.....	86
3.3.3 Pilote de l'expérimentation.....	88
4 Résultats	89
4.1 Résultats pour la tâche de discrimination	90
4.1.1 Analyse descriptive et observatoire	90
4.1.2 Analyse statistique	92
4.2 Résultats pour la tâche de prononciation	96
4.2.1 Analyse descriptive et observatoire	96
4.2.2 Analyse statistique	98
4.3 Résultats pour la tâche de maintien en mémoire de travail.....	104
4.3.1 Analyse descriptive et observatoire	104
4.3.2 Analyse statistique	110
4.4 Résultats pour la tâche de correspondance phonème-graphème.....	117
4.4.1 Analyse descriptive et observatoire	117
4.4.2 Analyse statistique	122
4.5 Résultats pour la tâche de vocabulaire réceptif.....	126
4.5.1 Analyse descriptive et observatoire	127
4.5.2 Analyse statistique	130
5 Discussion	135
5.1 Comparaison des effets des deux didactiques.....	135
5.1.1 Effets comparés des deux didactiques par compétence.....	136
5.1.2 Effets comparés des deux didactiques par phonème	143
5.1.3 Synthèse de la comparaison des deux didactiques.....	147
5.2 Recommandations pour l'enseignement du FLE en contexte scolaire	154
5.3. Implications théoriques.....	159
5.4. Analyse de la méthode utilisée	167
5.4.1 Difficultés, biais et limites de cette étude.....	167
5.4.2 Intérêts de cette étude.....	171
6 Conclusions.....	174
Annexes	179
Abréviations	179
Financement.....	204
Bibliographie	180

Liste des figures

Figure 1 : Représentations phonologiques en mémoire de travail	26
Figure 2 : Schéma de la progression didactique en fonction des compétences et des représentations.....	46
Figure 3 : Illustration des traits phonologiques qui séparent /ʒ/ des phonèmes les plus proches dans les trois répertoires phonologiques (français, allemand et suisse allemand).	66
Figure 4 : Planification des tests	87
Figure 5 : Discrimination moyenne par séquence didactique et par phonème.....	90
Figure 6 : Interaction entre la méthode didactique, le phonème en jeu et le temps	91
Figure 7 : Variances de la probabilité d'une discrimination réussie	92
Figure 8 : Représentation graphique des effets fixes du modèle.....	94
Figure 9 : Estimations des effets fixes du modèle relatif aux performances en discrimination	95
Figure 10 : Prédications par le modèle de la probabilité d'une discrimination réussie en fonction de l'interaction du phonème et de la séquence didactique sur les deux post-tests	96
Figure 11 : Progrès de la prononciation en fonction de la séquence didactique et du phonème	97
Figure 12 : Distribution de l'évaluation de la prononciation en fonction de la séquence didactique, du phonème et du test (pré-test ou post -test immédiat).....	97
Figure 13 : Interaction entre la séquence didactique, le phonème et le niveau au pré-test en fonction de la moyenne centrée (=ct0).....	98
Figure 14 : Ensemble des effets fixes de la modélisation du progrès de la prononciation	102
Figure 15 : Prédications du modèle de l'effet des deux séquences didactiques en fonction du phonème (autres prédicteurs maintenus constants).....	103
Figure 16 : Estimations par le modèle des effets fixes.....	104
Figure 17 : Progrès du nombre de phonèmes répétés en fonction de la série et de la séquence didactique (= Group).....	105
Figure 18 : Probabilité d'un progrès par séquence didactique pour l'ensemble des séries....	107
Figure 19 : Interactions entre les séquences didactiques et la série sur le nombre total de phonèmes répétés en fonction de l'inclusion ou non des phonèmes cibles	108
Figure 20 : Évolution du nombre moyen de phonèmes cibles et « parallèles » répétés en fonction des séries et des séquences didactiques	109
Figure 21 : Représentation graphique des effets fixes du modèle.....	112
Figure 22 : Estimations des effets fixes du modèle.....	114
Figure 23 : Progrès du nombre de phonèmes répétés entre le pré-test et le post-test immédiat	116

Figure 24 : Représentation de l'évolution du risque d'erreur par phonème cible et en fonction du test	118
Figure 25 : Réduction du risque d'erreur en fonction du phonème et de la séquence didactique	119
Figure 26 : Interactions observées sur la réduction du risque d'erreur : entre la séquence didactique et le phonème, entre le phonème et le type d'erreur, et entre la séquence didactique et le type d'erreurs.....	121
Figure 27 : Estimations des effets des prédicteurs lorsque les autres prédicteurs sont fixés à leur valeur moyenne	124
Figure 28 : Comparaison du risque et du type d'erreurs entre les deux phonèmes étudiés au post-test à 3 mois.....	126
Figure 29 : Représentation du taux de vocabulaire réceptif en fonction du test, du phonème et de la séquence didactique.....	127
Figure 30 : Évolution du taux moyen de vocabulaire réceptif en fonction du temps, du phonème et de la séquence didactique	129
Figure 31 : Influence du taux de pseudo-mots cochés par erreur (=FA) sur le taux de vocabulaire réceptif en fonction de la séquence didactique (=Group).....	130
Figure 32 : Estimation des effets fixes sur le vocabulaire réceptif	132
Figure 33 : Représentation graphique des effets fixes du modèle.....	133

Liste des tableaux

Tableau 1 : Séquence didactique et type d'enseignement des formes phonologiques	32
Tableau 2 : Complexité : Bilan des modifications de la grammaire phonologique nécessaires pour l'apprentissage des phonèmes /ʃ/-on et /ʒ/-j du FLE	55
Tableau 3 : Difficulté : Bilan des difficultés relatives à l'apprentissage des phonèmes /ʃ/-on et /ʒ/-j du FLE	56
Tableau 4 : Comparaison des actions pédagogiques engagées dans chacune des deux séquences didactiques	64
Tableau 5 : Pseudo-mots créés pour explorer le phonème /ʒ/	71
Tableau 6 : Pseudo-mots créés pour explorer le phonème / ʃ /	71
Tableau 7 : Comparaison des représentations graphémiques dans les deux langues d'un phonème commun	81
Tableau 8 : Pourcentage observé moyen de discrimination réussie par groupe, par test et par phonème	91
Tableau 9 : Effets fixes transformés en odds ratios et en probabilités d'une discrimination réussie.....	93
Tableau 10 : Probabilités estimées d'une discrimination réussie et leur intervalle de confiance	95
Tableau 11 : Moyenne du progrès entre pré- et post-test par séquence didactique et par phonème	98
Tableau 12 : Estimations des effets fixes du modèle sur le progrès de la prononciation.....	100
Tableau 13 : Estimations calculées à partir du modèle robuste	101
Tableau 14: Bilan de la comparaison des deux séquences didactiques en fonction du niveau au pré-test.....	103
Tableau 15 : Moyennes du progrès en pourcentage de phonèmes répétés correctement par séquence didactique et par série	106
Tableau 16 : Moyenne du progrès du nombre de phonèmes cibles correctement restitués par séquence didactique (max=8).....	109
Tableau 17 : Estimations des effets fixes sur le progrès lors de la répétition de pseudo-mots	111
Tableau 18 : Évolution du nombre de phonèmes répétés par série en %.....	113
Tableau 19 : Estimations des effets fixes sur l'évolution de la répétition des voyelles nasales	115
Tableau 20 : Observation de la réduction de la probabilité de faire une erreur	117
Tableau 21 : Probabilité observée du risque d'erreur par groupe, par test et par phonème...	118

Tableau 22 : Probabilité observée du risque d'erreur par groupe, par type d'erreur et par phonème	120
Tableau 23 : Effets fixes estimés sur le risque d'erreur au post-test à 3 mois	123
Tableau 24 : Estimations du modèle en termes de probabilités de faire une erreur au post-test à 3 mois et comparaison avec les valeurs observées à T0.....	125
Tableau 25 : Taux de vocabulaire réceptif observés	128
Tableau 26 : Effets fixes estimés sur l'évolution de l'identification d'un lexème comme connu	131
Tableau 27 : Estimations du taux de vocabulaire réceptif	134

Introduction

L'enseignement des représentations phonétiques et phonologiques en classe de français langue étrangère à l'école est limité à l'enseignement de la prononciation, quand il existe. En règle générale, l'enseignant.e demande aux apprenants de répéter quelques mots ou certains phonèmes isolés. Cependant les représentations phonétiques et phonologiques sont indispensables au fonctionnement de la langue. Elles sont mobilisées par l'ensemble des compétences langagières : parler, interagir et comprendre, et également lire, écrire et apprendre.

La didactique communicative a placé au centre de ses objectifs le processus communicatif. Parler, interagir et comprendre demande à l'apprenant¹ de disposer de représentations phonétiques et phonologiques de la langue étrangère pour pouvoir utiliser ses compétences de discrimination et de compréhension orale, et ses compétences de prononciation et de production orale pour pouvoir être compris. L'enseignement des représentations phonétiques et phonologiques peut concerner chaque élément de la représentation lexicale.

Phonème, cluster, morphème et représentation globale sont des représentations phonologiques identifiables dont les formes sonores constituent la base des représentations phonologiques lexicales. D'ailleurs, plus les formes sonores des représentations phonologiques sont proches de celles de la langue native, plus le lexème est facile à apprendre (Ellis, N. & Beaton, 1993). Cibler l'enseignement sur les représentations phonologiques pourrait faciliter l'apprentissage du vocabulaire en langues étrangères.

La présente étude a étudié les effets comparés de deux didactiques ciblées, l'une explicite et l'autre implicite, visant à enseigner les représentations phonétiques et phonologiques et le vocabulaire qui les inclut. Selon Norris et Ortega (2000), l'enseignement est considéré comme explicite lorsque l'apprenant doit consciemment faire attention à la forme particulière ciblée par l'enseignement. Au contraire, l'enseignement est considéré comme implicite lorsque la forme ciblée est mise en évidence pour être remarquée par l'apprenant sans qu'il en ait conscience et que l'objectif de l'enseignement est communicatif (Ellis, R., 2009). L'enseignement implicite vise une acquisition basée sur l'analyse des régularités, l'établissement de comparaisons, et l'organisation des représentations en fonction de leurs différences contrastives, alors que l'enseignement explicite vise un apprentissage attentionnel susceptible d'influencer le fonctionnement procédural pour construire une représentation ciblée. Deux séquences didactiques parallèles, l'une implicite et l'autre explicite, ont été

¹ J'ai renoncé à l'écriture épïcène pour le lexème *apprenant* qui revient fréquemment car cette écriture induit pour moi une charge cognitive importante. J'ai néanmoins gardé l'écriture épïcène pour *enseignant.e* pour essayer de représenter enseignant et enseignante et qui est un terme moins fréquent dans le texte.

élaborées pour permettre l'apprentissage des représentations phonétiques et phonologiques et du vocabulaire associé.

Les représentations du son, support de l'information verbale, sont qualifiées de phonologiques ou de phonétiques selon les faits linguistiques qu'elles codent (Best & Tyler, 2007). Une représentation phonologique représente une information abstraite permettant de différencier deux lexèmes (e.g., *ballon* = /balɔ̃/ ; *ballot* = /balɔ/). Une représentation phonologique est par exemple un phonème comme le phonème /ɔ̃/-on en français. Les représentations phonologiques d'un lexème regroupent l'ensemble de ses phonèmes, mais aussi celle des associations fréquentes de phonèmes comme par exemple celles de ses unités morphologiques, de ses syllabes ou sa représentation phonologique globale. Les représentations phonologiques permettent de séparer le sens de deux lexèmes. Au contraire, les représentations phonétiques permettent de produire un certain nombre de réalisations acoustiques allophoniques. Ces variations allophoniques sont produites sans indiquer une différence lexicale (e.g., la consonne *m* est produite avec plus ou moins de nasalisation en fonction des locuteurs sans que cette variation change l'identification lexicale). Les représentations phonétiques sont liées aux processus articulatoires et correspondent à un ensemble de réalisations acoustiques (Reetz & Jongman, 2020). Une représentation phonétique correspond par exemple à la place de l'articulation ou à la manière d'articuler un son verbal telle que la réalisation de la nasalisation par abaissement du voile du palais lors de la prononciation d'une voyelle nasale. Plusieurs représentations phonétiques contribuent à une représentation phonémique (i.e., phonologique). Les représentations phonétiques sont activées de façon automatique et systématique. Les représentations phonétiques et phonologiques sont essentielles au fonctionnement des processus impliqués dans la communication.

Un enseignement uniquement communicatif ne permet pas toujours le développement de la langue étrangère. L'enseignement ciblé et systématique, en particulier des difficultés phonologiques connues entre la langue première et la langue étrangère, est complémentaire (Detey, Fontan, & Pellegrini, 2016 ; Suzuki, Nakata, & DeKeyser, 2019). Certains phonèmes du français langue étrangère sont connus pour être des difficultés persistantes même chez des apprenants avancés (Racine, Detey, Bühler, Schwab..., & Kawaguchi, 2010; Ruellot, 2014). Les deux phonèmes choisis pour cette étude, les phonèmes /ɔ̃/-on et /ʒ/-j, sont des difficultés connues pour les apprenants, notamment germanophones. Ces phonèmes sont difficilement identifiés et souvent mal prononcés. Ces représentations difficiles à acquérir en instruction guidée ont été enseignées explicitement et implicitement dans la présente étude.

La présente étude, qui inclut 127 participants germanophones (âge moyen = 12;6 ans), s'est déroulée en contexte scolaire, pendant les heures d'enseignement dédiées au français langue étrangère. De nombreuses études ont montré la supériorité de l'enseignement explicite des représentations phonologiques en langue étrangère en contexte de laboratoire et avec des apprenants adultes, mais peu d'études ont ciblé le contexte scolaire et des apprenants non

adultes (Dekeyser, 2003; Norris & Ortega, 2000). La présente étude s'est proposée de compléter l'état actuel de la recherche en ciblant le contexte scolaire et des apprenants adolescents.

Dans la méta-analyse de Norris et Ortega (2000), seulement trois études ont été conduites en contexte scolaire et elles semblent indiquer une plus grande efficacité de l'enseignement explicite. Deux points méritent cependant d'être explorés. Premièrement, enseignement implicite et explicite sont comparés sans que l'enseignement implicite soit ciblé sur la forme visée. Deuxièmement, les études incluses dans cette méta-analyse sont centrées sur une à deux compétences, principalement des compétences grammaticales. Les compétences mobilisant les représentations phonétiques et phonologiques sont peu étudiées en contexte scolaire. La comparaison d'un enseignement implicite et d'un enseignement explicite, tous deux ciblés, semblent fort peu explorée, tout comme les représentations phonétiques et phonologiques en contexte scolaire avec des apprenants adolescents.

La présente étude a comparé les effets d'un enseignement implicite ciblé des représentations phonétiques et phonologiques dans une séquence didactique communicative actionnelle, et d'un enseignement explicite de ces mêmes représentations. L'étude longitudinale a suivi l'évolution des performances de cinq compétences grâce à un pré-test, un post-test et un test à 3 mois. Les compétences observées ont été identifiées comme essentielles pour l'apprentissage des représentations phonétiques et phonologiques dans le cadre théorique choisi pour cette étude.

L'apprentissage des représentations phonétiques et phonologiques en langue étrangère est couramment modélisé dans le cadre des théories basées sur la mémorisation d'exemplaires en la langue étrangère (e.g., PAM : Best & Tyler, 2007; SLM : Flege, Schirru & McKay, 2003; Flege, 1995). Ces modèles visent à prédire les difficultés qui seront rencontrées par un apprenant pour acquérir les représentations phonétiques et phonologiques de la langue étrangère. La présente étude a visé, quant à elle, à évaluer quelles compétences peuvent être développées, et par quelle méthode didactique les apprenants pourraient au mieux augmenter leurs performances lors de l'exercice de ces compétences. Les performances d'une compétence sont conditionnées par les difficultés cognitives induites par son fonctionnement.

Le modèle d'acquisition des représentations phonétiques et phonologiques, retenu pour la présente étude, a l'avantage de proposer un fonctionnement cognitif de l'apprentissage. Le modèle PRIMIR de perception et d'apprentissage des représentations phonétiques et phonologiques (Curtin, Hufnagle, Mulak & Escudero, 2017; Curtin & Werker, 2007; Werker & Curtin, 2005) est également un modèle basé sur l'acquisition d'exemplaires, mais il propose en plus une modélisation cognitive de cette acquisition. Il ne permet pas cependant de définir ce qu'est une charge cognitive, ni comment les représentations multimodales peuvent conduire à des traces en mémoire à long terme. Dans la présente étude, le modèle de mémoire de travail TBRS (Barrouillet & Camos, 2015) a renforcé et complété le modèle de perception PRIMIR. Le modèle TBRS propose un fonctionnement cognitif de la mémoire de travail, et une définition de la charge cognitive et de l'apprentissage. Les compétences clés impliquées dans

l'apprentissage des représentations phonétiques et phonologiques ont été identifiées par une synthèse des deux modèles.

Les cinq compétences observées dans cette étude ont montré un progrès significatif dans les deux séquences didactiques explicite et implicite. L'enseignement explicite a semblé plus à même de limiter les interférences cross-linguistiques. Il a conduit à des résultats supérieurs dans l'acquisition des relations phonème-graphème pour les deux phonèmes ciblés, pour toutes les compétences utilisant le phonème le plus difficile (= /ʒ/-j) et pour les élèves les plus faibles. Cependant en considérant la moyenne des deux phonèmes, aucune différence significative n'a été observée entre les deux séquences didactiques pour l'acquisition du vocabulaire réceptif et pour toutes les compétences observées (à l'exception des correspondances phonème-graphème). Ces résultats sont importants pour l'enseignement des langues étrangères en contexte scolaire.

Trois recommandations principales ont été émises. Premièrement, les effets positifs de la didactique implicite ciblée, incluse dans une didactique communicative actionnelle, ont ouvert la possibilité de pallier le manque de formation et de confiance des enseignants pour enseigner ces représentations. L'intégration de la didactique implicite ciblée peut être prévue dès la conception des méthodes et ainsi faciliter le travail des enseignants. Deuxièmement, l'absence de différence dans l'acquisition du vocabulaire réceptif entre une séquence didactique centrée sur la communication et une séquence explicite qui a exclu le recours à la valeur sémantique, a montré combien les représentations phonétiques et phonologiques sont importantes pour l'acquisition du vocabulaire en langue étrangère. Ce résultat souligne l'incohérence de l'abandon de leur enseignement. L'enseignement des représentations phonétiques et phonologiques est un élément clé pour l'apprentissage de la langue étrangère. Troisièmement, les compétences de discrimination et de correspondances phonème-graphème ont pu progresser après avoir été enseignées, ce qui pose la question de la raison de leur absence des méthodes d'enseignement.

Dans une première partie, le cadre théorique est établi afin de définir dans une seconde partie les questions de recherche et les hypothèses. Dans une troisième partie la méthode est décrite avant d'exposer les résultats dans une quatrième partie. Dans une cinquième et dernière partie, ces résultats sont discutés en comparaison des hypothèses émises et des résultats déjà présents dans la littérature.

1 Apprentissage des représentations phonologiques et du vocabulaire

Le lexique mental d'un adulte moyen, sachant lire, est structuré par l'association de multiples représentations. Les représentations sémantiques d'un lexème, telles que les valeurs sémantiques, pragmatiques, stylistiques et leurs collocations, sont liées à ses multiples représentations formelles telles que sa classe grammaticale, ses flexions morpho-phonologiques, ses représentations orthographiques et ses représentations phonétiques et phonologiques. L'ensemble de ces représentations compose la représentation lexicale. Chacune de ses représentations peut être activée et transmettre l'activation à une partie ou à l'ensemble des représentations du lexème. Cette activation se transmet à toutes les représentations associées quelle que soit la langue à laquelle elles appartiennent (Hulstijn, 2013). Bien que liées entre elles, toutes les représentations composant une représentation lexicale ne s'apprennent pas avec la même facilité.

L'association lexicale la plus complexe à acquérir est la relation arbitraire entre une représentation sémantique principale et ses représentations phonologiques et phonétiques (Hulstijn, 2013). Cette association est d'autant plus difficile à acquérir en langue étrangère (=LE) que les représentations phonologiques LE sont souvent erronément assimilées à celle de la langue première (=L1). Le système phonologique L1 sert de base au fonctionnement et à la construction du système LE (e.g., Flege, 1995; Weber & Cutler, 2004). Les caractéristiques spécifiques des représentations phonologiques LE peuvent être remarquées, et ainsi avoir une chance d'être apprises, seulement si ses particularités sont mises en évidence, mais aussi que suffisamment d'attention soit disponible pour permettre l'acquisition d'une procédure spécifique en mémoire de travail.

La mémoire de travail est considérée comme le centre de traitement obligatoire des inputs sensoriels (Baars, 1997; Baddeley, 2000; Cowan, 2008; Wen, Borges Mota, & McNeill, 2015). Le flux de parole est perçu, puis interprété grâce aux routines et procédures de la mémoire de travail. L'attention est la ressource qui permet à la mémoire de travail de traiter toutes les informations verbales entrantes comme sortantes (Baddeley & Hitch, 1974; Miller et al., 1960). La définition exacte de la mémoire de travail dépend du modèle choisi. Dans cette étude, le modèle retenu est le Time Based Resource Sharing (=TBRS ; Barrouillet, Bernardin, Portrat, Vergauwe & Camos, 2007; Barrouillet & Camos, 2012, 2015). Le cadre théorique proposé par le TBRS m'a permis de construire les séquences didactiques et leurs objectifs en organisant la tension entre charge cognitive, interférences cross-linguistiques et compétences à développer.

1.1 Le modèle de mémoire de travail : Time Based Resource Sharing

La mémoire de travail assume de nombreuses fonctions dont deux sont essentielles pour la présente étude : le maintien temporaire de l'information verbale d'origine sensorielle et son traitement. L'information verbale orale est en premier stockée en boucle phonologique avant d'être traitée par le système procédural pour être interprétée. L'information enregistrée en boucle phonologique est cependant très fragile et peu pérenne. Si elle n'est pas renouvelée, elle est partiellement ou totalement perdue avant même d'être traitée.

Le maintien de l'information et son traitement dépendent d'une même ressource attentionnelle. Ces deux fonctions doivent pouvoir bénéficier de ressources attentionnelles pour être accomplies. Or, une seule fonction à la fois peut être dans le focus de l'attention. Par conséquent, une seule fonction à la fois, maintien ou traitement, peut être réalisée. Les deux fonctions sont cependant opérées en parallèle grâce à une alternance très rapide de la distribution de l'attention entre elles. L'alternance très rapide du focus de l'attention entre maintien et traitement permet le déroulement parallèle des deux fonctions mais le fonctionnement de chacune peut limiter les performances de l'autre.

Le maintien de l'information peut ne pas être réalisé lorsque les ressources attentionnelles disponibles sont accaparées par le traitement de l'information. Le maintien de l'information verbale permet de lutter contre sa détérioration sous l'influence des interférences ou par son érosion au fil du temps. Seule sa rénovation régulière permet de limiter la perte d'information. Or, lorsque l'attention est focalisée sur le traitement de l'information, elle n'est pas disponible pour permettre la rénovation.

La rénovation de l'information verbale est assurée par deux mécanismes en fonction du buffer qui la contient (Camos, Lagner & Barrouillet, 2009; Camos, Mora & Barrouillet, 2013; Camos, Johnson, Loaiza, Portrat, ..., & Vergauwe 2018; Mora & Camos, 2013; Oftinger & Camos, 2015). Le buffer est un espace fonctionnel de stockage. Premièrement, la boucle phonologique est un buffer périphérique sensoriel qui encode et stocke l'information sensorielle, et l'information qu'il contient est renouvelée par la répétition subvocale. Deuxièmement, le buffer épisodique stocke les représentations élaborées par le système procédural, et ces représentations sont renouvelées par le rafraîchissement attentionnel.

La répétition subvocale assure la rénovation du matériel verbal contenu en boucle phonologique grâce au support des ordres moteurs articulatoires déjà inscrits en mémoire à long terme (e.g., Baddeley, 2003; Hulme, Roodenrys, Schweickert, Brown..., & Stuart, 1997; Smith & Jonides, 1999). La suite ordonnée des représentations phonétiques enregistrées en boucle phonologique est associée à une planification des ordres moteurs articulatoires. Chaque représentation phonétique est associée à l'ordre moteur qui permet de la produire. L'ordre moteur activé

réactive alors en retour la représentation phonétique contenue en boucle phonologique. La représentation phonétique réactivée est ensuite à nouveau disponible pour être traitée. La répétition subvocale est inconsciente et involontaire pour l'individu mais elle est importante lors de la restitution d'une liste de mots.

La quantité de mots entendus qu'il est possible de restituer est linéairement et positivement corrélée à la vitesse d'articulation de ces mots. Plus on articule vite, plus la liste de mots restitués est longue. Pour tenter de comprendre si les stades de développement de la capacité de la boucle phonologique en L1 étaient dus à une augmentation des capacités de la boucle phonologique ou à la vitesse de programmation des ordres moteurs articulatoires, Hulme, Thomson, Muir et Lawrence (1984) ont demandé à des enfants de 4, 7, et 10 ans et à des adultes de répéter des mono-, des bi- et des tri- et quadrisyllabiques. Les auteurs ont ainsi mesuré leur vitesse d'articulation. Puis ils ont demandé aux participants de répéter des listes de mots. En représentant graphiquement, la moyenne du nombre de mots répétés par rapport à la vitesse d'articulation de chaque mot pour chaque participant, les auteurs ont constaté la relation linéaire qui relie les deux variables. L'efficacité de la programmation des ordres moteurs articulatoires est donc fondamentale pour les performances de la rénovation subvocale et par conséquent pour la capacité à maintenir l'information verbale. Il est probable qu'en LE la non maîtrise des ordres moteurs articulatoires correspondant aux représentations phonétiques spécifiques de la LE puisse limiter les capacités de maintien de l'information verbale LE en mémoire de travail. La rénovation subvocale nécessite une quantité d'attention négligeable en L1 au contraire du rafraîchissement attentionnel (Baddeley, 2012).

Au niveau du buffer épisodique, la source de rénovation est le rafraîchissement attentionnel. Il rénove les représentations successivement élaborées par le système procédural et stockées dans le buffer épisodique, en créant des liens bidirectionnels entre ces nouvelles représentations et des représentations existantes en mémoire à long terme (Mora & Camos, 2015; Rose, Buchsbaum & Craik, 2014). Cependant, le rafraîchissement attentionnel ne peut être réalisé que lorsque le fonctionnement du système procédural est interrompu et qu'il est dans le focus de l'attention.

Une charge cognitive importante entraîne inévitablement une perte d'information verbale (e.g., Barrouillet, Bernardin & Camos, 2004; Barrouillet et al., 2007). La mobilisation de l'attention par le système procédural empêche le déroulement du rafraîchissement attentionnel. La lenteur intrinsèque du système procédural limite la quantité d'information qui peut être traitée avant d'être définitivement perdue (Barrouillet & Camos, 2015, p.119). Des routines (i.e., procédures automatisées) sont efficaces alors que de nouvelles procédures, comme c'est le cas pour le traitement des représentations phonétiques et phonologiques LE, prennent plus de temps pour être accomplies (Gathercole, Dunning, Holmes & Norris, 2019). Il semble possible que les nouvelles procédures à établir pour la spécificité des représentations LE prennent plus de temps pour être accomplies. Le temps pendant lequel le focus de l'attention est accaparé par le traitement procédural de l'information, est une période pendant laquelle la

renovation attentionnelle ne peut pas fonctionner. Ce temps de traitement représente la charge cognitive. Les représentations spécifiques de la LE ont de grandes chances d'induire une charge cognitive lors de leur traitement par le système procédural. D'une manière réciproque, comme le système procédural et le rafraîchissement attentionnel partagent le focus de l'attention, il est aussi possible que l'allongement du temps de rénovation attentionnelle limite le fonctionnement du système procédural.

Les performances de la mémoire de travail peuvent être limitées par la durée du processus de rafraîchissement attentionnel. Le système procédural ne peut fonctionner qu'en présence d'attention disponible, c'est-à-dire en dehors des périodes de rafraîchissement attentionnel. Plus le temps nécessaire à l'accomplissement du rafraîchissement attentionnel est long, moins les procédures peuvent être accomplies avant que l'information ne soit perdue. Case, Kurland, & Goldberg (1982) ont proposé une tâche de répétition à des étudiants et à des enfants, mais les listes étaient composées de pseudo-mots pour les étudiants et de mots pour les enfants. Les pseudo-mots incluaient uniquement des représentations phonétiques et phonologiques L1. La moyenne du nombre de pseudo-mots répétés par les étudiants a été identique à celle du nombre de mots répétés par les enfants, alors que leur capacité de maintien en mémoire de travail sont inférieures à celle des adultes. Les auteurs ont conclu que la vitesse de traitement des pseudo-mots, et en particulier le temps total de traitement de chaque représentation phonologique, était la cause principale expliquant le nombre de mots qu'un sujet peut restituer. La prononciation et l'appariement entre représentations phonétiques et articulatoires n'ont vraisemblablement pas constitué une difficulté puisque toutes les représentations phonétiques et phonologiques étaient connues des participants. Il est possible de faire l'hypothèse que le rafraîchissement attentionnel a demandé plus d'attention puisque la rénovation des pseudo-mots devait concerner chaque élément phonétique et phonologique. Au contraire lors de la répétition d'une liste mots, les participants peuvent utiliser une représentation lexicale globale et bénéficier du support lexical. Par conséquent, le temps de rafraîchissement attentionnel des mots LE inconnus devrait être allongé pour deux raisons. Premièrement, comme pour l'expérience de Case et al. (1982), le rafraîchissement attentionnel devrait être allongé par le traitement unitaire des éléments phonétiques et phonologiques qui composent le mot LE inconnu. Deuxièmement, le rafraîchissement attentionnel sera privé du soutien de la mémoire à long terme ou sera obligé d'utiliser une représentation L1 peu compatible avec l'information LE, car l'information verbale inclura des formes phonétiques et phonologiques LE non inscrites en mémoire à long terme.

Les interférences cross-linguistiques entre la LE et la L1 induisent une charge cognitive. L'appariement des représentations phonétiques LE comme celles associées au phonème /ɔ̃/-on à la représentation articulatoire L1 la plus proche comme celles du phonème /o/ provoque peut-être une surcharge cognitive quand une représentation naissante LE peu activée mais congruente est en compétition avec une représentation L1 fortement activée mais peu congruente. Miwa, Dijkstra, Bolger, et Baayen (2014) ont observé les effets de représentations

phonologiques communes entre japonais et anglais dans une tâche de décision lexicale en mesurant les temps de pose du regard des participants sur les stimuli et le temps de réponse (= étude d'*eye-tracking*). Une suite de lettre de l'alphabet latin a été visuellement présentée aux participants bilingues japonais et anglais. Ils devaient décider le plus vite possible et le plus justement si cette suite de lettres correspondait ou non à un mot d'anglais. La similarité phonologique entre candidats lexicaux potentiels des deux langues correspondant à la suite de lettres présentée a été mesurée par une évaluation sur une échelle de Lickert en 7 points par des locuteurs japonais. Dès la première fixation oculaire du participant, un effet de la similarité phonologique est observable. Cet effet est successivement précoce et inhibiteur, puis tardif et facilitateur. Les auteurs attribuent l'effet précoce inhibiteur à une co-activation des deux systèmes phonologiques qu'ils qualifient de « bruit ». Cette co-activation à partir d'un stimulus visuel spécifique, pourtant lié à un seul des deux systèmes phonologiques, indique que l'activation des représentations phonologiques est précoce, non-spécifique d'une langue, involontaire et est présente même en l'absence de stimulus sonore. Cette co-activation des deux systèmes phonologiques apparaît peut-être dès que les représentations phonologiques LE commencent à se construire. La compétition entre les deux systèmes phonologiques est une interférence cross-linguistique qui induit une charge cognitive.

Les interférences cross-linguistiques demandent de l'attention pour être résolues comme tout conflit entre les représentations co-activées en mémoire de travail. Dans une revue de littérature, Kim, Shin et Hur (2016) arguent que la résolution de conflit pour incongruence en mémoire de travail repose autant sur le fonctionnement des processus de maintien de l'information que sur les contraintes exercées par les représentations présentes en mémoire à long terme sur le système procédural. Par conséquent, il est possible de supposer que les interférences cross-linguistiques accroissent le temps de traitement de l'information verbale soit par une augmentation du temps de rafraîchissement, soit par un ralentissement du fonctionnement du système procédural, soit par les deux.

La mémoire de travail verbale a des capacités limitées non seulement par la concurrence pour l'attention entre traitement et maintien de l'information, mais aussi par la capacité de stockage des buffers. Les différents buffers ont une capacité de stockage limitée, environ équivalente à l'information verbale qu'il est possible d'articuler en 2s pour la boucle phonologique et à quatre représentations complexes pour le buffer épisodique (Baddeley, 2012; Cowan, 2015). Malgré ces capacités limitées, la mémoire de travail gère la distribution de l'attention pour arriver à la fois à maintenir et à traiter l'information.

Le système procédural de la mémoire de travail, quand il est dans le focus attentionnel, est responsable du traitement de l'information. Le système procédural en association avec le buffer épisodique forme la boucle exécutive qui élabore les représentations multimodales successives, uniques et épisodiques. Pour former ces représentations multimodales, il associe les représentations recodant l'information d'origine sensorielle et celles de la mémoire à long terme. Ces nouvelles représentations sont ponctuellement stockées dans le buffer épisodique.

Puis le système procédural analyse cycliquement les représentations multimodales nouvellement créées. En fonction du but de la tâche à accomplir ou de l'information contenue en mémoire à long terme, il les modifie par l'activation du schéma procédural le plus adapté à l'objectif de la tâche (Anderson, Bothell, Byrne, Douglass..., & Qin., 2004). Il gère ainsi l'activation des représentations présentes dans les buffers et en mémoire à long terme dans le but de lier celles qui sont le plus congruentes (étape de *binding*) pour construire la représentation suivante (Abdallah, 2010). Les représentations d'origine sensorielles peuvent ainsi être associées à des représentations en mémoire à long terme dont elles n'ont pas déclenché directement l'activation.

Des éléments inadéquats ou de nouveaux éléments peuvent être associés à une représentation élaborée alors que l'information sensorielle qui l'a induite ou son recodage ne contenait pas ces éléments. Par exemple, l'assimilation d'un phonème LE à un phonème L1 peut être expliquée par une activation des éléments de représentations partagés par les deux phonèmes et un abandon de ceux qui diffèrent au fil des mises à jour de la représentation élaborée. Une voyelle nasale du français langue étrangère (= FLE) peut être ainsi assimilée à une voyelle orale L1 dont elle partage la place d'articulation. D'autres observations de cet effet *top-down* peuvent être expliquées par ce mécanisme (Baddeley, 2012). Il s'agit par exemple de l'influence de la fréquence lexicale ou sub-lexicale (Adlof & Patten, 2017; Gathercole, Frankish, Pickering & Peaker, 1999; Masoura & Gathercole, 2005; Vitevitch, Luce, Charles-Luce & Kemmerer, 1997), de la lexicalité (Roodenrys, Hulme, Alban, Ellis, A. W., & Brown, 1994) ou de la concrétude d'un lexème (Majerus & van der Linden, 2003; Tse & Altarriba, 2009). L'effet *top-down* est le résultat de l'association de représentations lexicales activées en mémoire à long terme à la représentation multimodale en cours d'élaboration.

Les nouveaux savoirs, c'est-à-dire les nouveaux états mentaux, semble être produits par le rafraîchissement attentionnel des représentations multimodales élaborées. Le rafraîchissement attentionnel consiste en une rénovation attentionnelle des représentations construites par le fonctionnement de la boucle exécutive entre le système procédural et la mémoire à long terme. Les traces de l'activation des représentations en mémoire de travail sont des états mentaux passifs qui constituent la mémoire à long terme. Pour l'information verbale, le rafraîchissement attentionnel inscrit des états mentaux associant par exemple les représentations phonologiques et les représentations orthographiques (Cowan, 2015; Linck, Osthus, Koeth & Bunting, 2014; Meschyan & Hernandez, 2002; Wang, Karipidis, Pleisch, Fraga-González & Brem, 2020). Les traces en mémoire à long terme (= états mentaux) ne sont pas dégradées par le temps lorsqu'elles sortent du focus de l'attention.

Les états mentaux persistent après la sortie du focus de l'attention mais sont perpétuellement modifiés par les nouvelles traces des élaborations en boucle exécutive, ce qui permet l'apprentissage. Par exemple, l'apprentissage de nouveaux phonèmes LE peut s'appuyer en début d'apprentissage sur les états mentaux L1, en particulier sur les représentations phonotactiques L1 (Dehaene-Lambertz et al., 2000; Dupoux, Pallier, Kakehi & Mehler, 2001)

et sur les représentations phonologiques L1 (Fourtassi, Schatz, Varadarajan & Dupoux, 2014; Jacquemot, Pallier, LeBihan, Dehaene & Dupoux, 2003). Les représentations LE s'individualiseraient alors progressivement à partir des représentations L1 comme décrit par les modèles d'acquisition phonémique L2 basés sur la théorie des exemplaires (e.g PAM : Best & Tyler, 2007 ; SLM : Flege et al., 2003). Le rafraîchissement attentionnel permet l'apprentissage de la LE mais le contraint également.

Lors de l'usage de la LE, la construction des représentations multimodales active systématiquement les états mentaux L1, ce qui peut expliquer certaines interférences cross-linguistiques. L'absence d'une représentation phonétique LE en mémoire à long terme empêche la prise en compte d'une caractéristique phonétique LE pourtant présente dans l'information d'origine sensorielle, comme par exemple lors de surdit  phonologique. Les apprenants d'une LE semblent incapables d'entendre certains marqueurs phonétiques sp cifiques de la LE (e.g., stress) lorsqu'ils sont absents de la L1 (Dupoux, Pallier, Sebasti n-Gall s & Mehler, 1997, Dupoux, Sebasti n-Gall s, Navarret, & Peperkamp, 2008 ; Jacquemot et al., 2003). La valorisation ou non d'une caractéristique LE de la repr sentation phon tique d'origine sensorielle, par le syst me proc dural, peut  tre consid r e comme un ensemble de r gles implicites appel es « grammaire phonologique ». Il est possible que la seule fa on de contraindre le syst me proc dural   valoriser une caract ristique LE soit d'attirer l'attention sur cette caract ristique par un enseignement cibl .

L'enseignement cibl  a prouv  son efficacit  dans de nombreuses  tudes, y compris pour permettre   l'apprenant avanc  de d passer ses difficult s (e.g., Norris & Ortega, 2000; Rousse-Malpat & Verspoor, 2012; Spada, Jessop, Tomita, Suzuki & Valeo, 2014). L'enseignement cibl  utilise une part de l'attention disponible au d triment des fonctions de maintien et de traitement. Le maintien peut n anmoins toujours  tre assur  par la r novation subvocale mais cette r novation ne participe pas   l'apprentissage.

La r novation subvocale en boucle phonologique ne laisse pas de traces directes en m moire   long terme et par cons quent ne peut pas permettre un apprentissage (Barrouillet & Camos, 2015, p.134 ; Loaiza & McCabe, 2012; Mora & Camos, 2013, 2015). La r novation subvocale assure le maintien temporaire des informations verbales contenues en boucle phonologique qui sont sous forme de repr sentations phon tiques transitoires. Ces repr sentations transitoires ne peuvent pas participer   la construction d'une repr sentation multimodale (Baddeley, 2012; Cowan, 2015; Ricker, 2015). Les interf rences en boucle phonologique ne laissent ainsi pas de traces en m moire   long terme ce qui peut expliquer qu'elles peuvent perdurer m me apr s de longues p riodes d'enseignement en LE ou d'immersion en langue seconde. Les repr sentations contenues dans les buffers p riph riques sont plus fragiles que les repr sentations  labor es.

La r novation subvocale ne peut pas modifier un  tat mental existant en m moire   long terme comme par exemple une repr sentation phonologique. En boucle phonologique, les premi res repr sentations phon tiques sont rapidement inactives et ne peuvent pas  tre li es aux

représentations déjà inscrites en mémoire à long terme (Ricker, 2015). Les représentations phonétiques sont recodées en représentations phonétiques élaborées qui, elles, sont intégrables dans une représentation multimodale par la boucle exécutive. Un temps de traitement minimal ($\approx 250\text{ms}$, Shapiro et al., 1994, cité dans Barrouillet & Camos, 2015, p.124) est nécessaire pour construire une représentation élaborée susceptible de rester activée dans le buffer épisodique et donc de pouvoir permettre un apprentissage. La rénovation subvocale peut néanmoins induire des modifications de l'information.

La rénovation subvocale peut modifier la structure de la future représentation multimodale en mémoire à long terme lors du recodage de l'information verbale en boucle phonologique. La rénovation subvocale lie les représentations phonétiques aux ordres moteurs articulatoires déjà présents en mémoire à long terme (= mémoire articulatoire). Puis l'association des deux représentations activées vient réactiver à son tour ou recoder la représentation phonétique en boucle phonologique. Le recodage est inhérent au fonctionnement du maintien de l'information mais il peut induire des modifications de l'information d'origine sensorielle.

La rénovation subvocale peut recoder une information verbale présentée visuellement, mais aussi lorsqu'elle est présente seulement sous forme de représentations phonétiques dans la boucle phonologique ; ce qui peut expliquer l'influence de l'orthographe L1 sur la perception et la prononciation LE (Bassetti, Escudero & Hayes-Harb, 2015). L'influence de la L1 s'observe que le support écrit soit présent ou juste activé en mémoire à long terme. Lorsque les deux langues partagent le même système d'écriture, les interférences induites par l'information sub-lexicale, et en particulier par des correspondances phono-graphémiques différentes entre L1 et LE, peuvent limiter l'apprentissage phonétique LE. Le recodage des caractéristiques LE de l'information phonétique d'origine auditive est perturbé par le recodage suivant le code phono-graphémique L1 des informations visuelles présentes ou activées. La rénovation subvocale peut recoder et maintenir une information visuelle si cette information peut être nommée verbalement (Baddeley, 2003). Les représentations phonétiques spécifiques de la LE peuvent donc être perdues dès le recodage en boucle phonologique sous l'influence du code phono-graphémique de la L1. Ce possible recodage limite fortement une association congruente des représentations phonétiques élaborées LE dans une représentation multimodale lexicale LE. Sans association dans une représentation multimodale, les représentations spécifiques de la LE sont vraisemblablement très difficiles à apprendre. L'absence de traces en mémoire à long terme de ce recodage limite probablement la possibilité du système procédural d'adapter les procédures aux spécificités de la LE.

Le traitement et le maintien de l'information a été décrit seulement pour le matériel verbal car cette étude est centrée sur les représentations phonologiques en LE. Le modèle TBRS m'a permis de définir la charge cognitive, et d'identifier un système procédural dont les routines de fonctionnement permettent d'élaborer les représentations multimodales et un système complexe de maintien de l'information verbale. Les possibles effets d'interférences cross-linguistiques ont été situés dans ce modèle, mais les relations entre les systèmes phonologiques L1 et LE

n'ont pas été explicitées. Les modèles de perception en LE complètent la compréhension du traitement de l'information verbale en LE.

1.2 Modèles de perception et d'apprentissage des représentations phonologiques en LE.

La perception de l'information verbale est cruciale pour l'apprentissage du vocabulaire car elle permet l'établissement des liens entre sons de la parole et valeur sémantique (e.g., Werker & Yeung, 2005). Les modèles de perception de la parole en LE restent discutés.

La relation d'identification directe entre fait acoustique et ordre moteur articuloire a justifié les modèles défendant un système préétabli de règles qui régiraient une hiérarchie symétrique entre les représentations du système sensoriel et moteur et les représentations phonologiques (e.g., Clements, 2003, 2006; Clements & Hallé, 2010; Liberman & Mattingly, 1985). Les modèles émergentistes, basés également sur la correspondance entre le fait acoustique et l'ordre moteur articuloire, ont proposé une hiérarchie des représentations qui se créeraient au fil de l'acquisition et qui par conséquent se construirait en LE par confrontations successives au système de la L1 déjà en place pour en modifier les règles (e.g., Brown, 2000; Goad & White, 2008; Hancin-Bhatt, 1994; Ohala, 2005). Ces modèles, règles préétablies ou émergées, ont été critiqués en particulier pour leur difficulté à expliquer comment les importantes variations de réalisations acoustiques n'entravent pas l'intercompréhension et aussi pour leur incapacité à expliquer comment les caractéristiques du lexique (e.g., fréquence, similarité phonologique, voisins orthographiques, cognats...) peuvent influencer la perception. Les modèles plus actuels proposent la constitution progressive d'une structure de référence composée d'exemplaires des catégories phonétiques et phonologiques de la langue concernée.

Les actuels modèles d'acquisition phonémiques en L2 prédisent la difficulté d'acquisition d'une représentation phonémique en fonction de sa relation de similarité avec les exemplaires de la L1 (e.g., PAM : Best & Tyler, 2007 ; SLM : Flege et al., 2003 ; L2LP : Escudero, 2005; van Leussen & Escudero, 2015). Ces modèles impliquent que l'établissement d'une nouvelle catégorie phonémique précéderait obligatoirement l'établissement d'un contraste lexical. Cependant ces modèles ne permettent pas d'expliquer pourquoi des apprenants avancés de LE sont capables de différencier une paire minimale lexicale et ne sont pas capables d'identifier les catégories phonémiques en jeu mieux que des apprenants débutants (Cutler, Weber & Otake, 2006; Escudero, Hayes-Harb & Mitterer, 2008; Hayes-Harb & Masuda, 2008; Levy, 2009; Weber & Cutler, 2004). Ils ne peuvent pas non plus expliquer comment la charge cognitive peut influencer les performances de perception en LE (Piske, Flege, McKay & Meador, 2002). Ces modèles sont importants pour anticiper les difficultés cross-linguistiques d'apprentissage des représentations phonologiques en LE mais ils n'informent pas sur le lien acquisitionnel entre vocabulaire et représentations phonologiques au contraire du modèle PRIMIR.

Le modèle PRIMIR (Curtin & Werker, 2007; Werker & Curtin, 2005; Werker & Yeung, 2005; Yeung & Werker, 2013) est un modèle connexionniste qui peut s'inscrire dans le cadre théorique de la mémoire de travail et qui repose sur le stockage d'exemplaires. Le modèle PRIMIR (= *Processing Rich Information from Multidimensional Interactive representations*) est un cadre théorique initialement élaboré pour expliquer les relations réciproques entre la construction du lexique et la perception de la parole chez l'enfant. Ce cadre théorique a été étendu aux enfants bilingues (Curtin, Byers-Heinlein & Werker, 2011) et est directement transposable à l'apprentissage d'une LE. Ce modèle inclut des mécanismes d'apprentissage pour modifier les représentations existantes et créer de nouvelles représentations, des espaces de représentations pour conserver l'information, et des filtres dynamiques qui contraignent le traitement de l'information sensorielle.

Les mécanismes d'apprentissage organisent l'information en mémoire à long terme par des analyses statistiques des régularités, comme celles supposées être déployées par le système procédural en boucle exécutive. L'analyse statistique des régularités permet la formation de groupes ou catégories qui partagent des caractéristiques similaires (e.g White, Peperkamp, Kirk & Morgan, 2008). Ainsi une représentation phonémique serait une catégorie qui regroupe plusieurs catégories de représentations phonétiques. La représentation du phonème /ɔ̃/-on regrouperait les représentations phonétiques de la nasalisation, de place de la langue et de la forme des lèvres permettant d'identifier ce phonème. Plusieurs autres modèles s'appuient sur la même hypothèse notamment pour l'acquisition lexicale (e.g., Li, Farkas & McWhinney, 2004; Li & Farkas, 2002). Une représentation lexicale regrouperait plusieurs catégories de représentations : phonologiques, sémantiques, orthographiques... Les représentations en mémoire à long terme seraient conservées dans des espaces multidimensionnels.

Dans le modèle PRIMIR, l'organisation des représentations en mémoire à long terme est non hiérarchique et les représentations sont regroupées dans trois espaces multidimensionnels liés par de puissantes interactions. Premièrement, l'espace général perceptuel est organisé progressivement par les mécanismes d'apprentissage sous l'influence des représentations de l'information verbale qui ont été produites par le système perceptuel humain. Cet espace conserve des représentations phonétiques et leurs informations relatives (articulatoires, visuelles, identité du locuteur...). Deuxièmement, l'espace des formes verbales conserve des exemplaires des séquences de sons extraites de la parole sans qu'elles soient préalablement associées à une valeur sémantique. L'association statistique d'un son et d'un sens se développerait avec la compréhension orale. L'augmentation du nombre de liens entre forme phonologique et valeur sémantique induirait la création du troisième espace multidimensionnel : l'espace phonémique.

Les représentations phonémiques émergeraient de l'analyse statistique et de la généralisation des liens existants entre les informations son-sens conservées dans l'espace des formes verbales et celles des représentations phonétiques de l'espace général perceptuel. Les phonèmes seraient ensuite utilisés pour réduire la variation phonétique liée au contexte et au locuteur. Les

phonèmes pourraient ainsi guider, si nécessaire, l'acquisition de nouvelles formes lors de la perception de l'information verbale.

Dans ce modèle, les phonèmes sont représentés en fonction de leur distribution phonotactique. Les représentations d'un même phonème sont différentes en fonction de sa position dans le lexème, ce qui a été confirmé dans d'autres études notamment en étudiant les effets de la fréquence de la position phonotactique sur la perception (e.g., Di Liberto, Wong, Melnik & de Cheveigné., 2019; Järvikivi, Pyykkönen, & Niemi, 2009; Peperkamp, Pettinato & Dupoux, 2002). Ces représentations différentes du même phonème permettent peut-être une organisation plus précise des représentations lexicales, notamment dans la structure morphologique : le phonème /ɔ̃/-on par exemple est un marqueur flexionnel verbal uniquement en position finale. Dans ce modèle, l'espace phonémique est placé au centre de l'organisation des différents espaces multidimensionnels. L'organisation en espaces multidimensionnels étroitement connectés permet au modèle PRIMIR d'expliquer les relations existantes entre les représentations, leur acquisition et leur traitement.

Acquisition et traitement des représentations sont contraints par les filtres dynamiques. Trois filtres dynamiques influencent l'acquisition et le traitement des représentations : les biais perceptuels (décrits aussi par les modèles exemplaires-base précédemment évoqués), le niveau développemental de l'apprenant et la charge cognitive imposée par la tâche. Premièrement, les biais perceptuels orientent l'attention vers les caractéristiques phonétiques permettant d'établir les contrastes en L1, ce qui limite l'acquisition aux représentations langagières plausibles et évite d'enregistrer des exemplaires atypiques. Les biais perceptuels de la L1 sont la base des modèles de perception en LE reposant sur l'idée d'une collection d'exemplaires. Deuxièmement, la charge cognitive et le but de la tâche orientent les priorités de traitement. Dans une étude de 2020, Atkinson, Allen, Baddeley, Hitch et Waterman montrent que les participants sont capables de prioriser une partie de l'information verbale dans une tâche de répétition de numéros en fonction de l'importance accordée ou non au stimulus (voir également Atkinson, Waterman, & Allen, 2019). Avec une tâche concurrente induisant une charge cognitive, ils sont toujours capables de prioriser mais leurs performances chutent. Dans le modèle TBRS, la charge cognitive induite par la seconde tâche limite l'attention disponible ce qui explique la diminution des performances, et le maintien des consignes de la tâche (i.e., priorisation) induit des modifications des routines du système procédural pour répondre aux demandes de la tâche. Troisièmement, le niveau développemental impose des limites à l'apprenant. Ces trois filtres dynamiques ont été décrits par différents modèles d'acquisition de la LE et sont compatibles avec le modèle TBRS. Dans le modèle PRIMIR, les liens entre les espaces perceptuels sont en permanence activés.

Le système PRIMIR est à tout moment interactif entre les espaces de représentations ce qui facilite le traitement, la modification et la conservation de l'information. L'interaction continue explique l'activation simultanée de toutes les langues du locuteur dans toute tâche verbale (voir également par exemple Flege, 1995; Flege et al., 2003; Hulstijn, 2013). Cependant cette co-

activation des langues du locuteur pourrait empêcher l'acquisition de nouvelles formes en privilégiant les biais perceptuels de la forme dominante (i.e., la L1) si le modèle n'incluait pas un système de régulation.

Le modèle PRIMIR inclut en plus du système général d'apprentissage statistique un mécanisme permettant la comparaison et le contraste de l'information verbale. Les comparaisons permettent le regroupement de l'information en catégories et les contrastes permettent de séparer les informations en fonction des dimensions qui les différencient. Ce mécanisme organise les relations entre les représentations au sein, et entre les espaces multidimensionnels. Il permet d'identifier les similarités et les différences entre représentations indépendantes et d'établir des liens entre elles (e.g le phonème *on-/ɔ̃/* en position finale est commun aux lexèmes *maison, ballon, pont...*). Ce système permet par exemple au bilingue de construire un double jeu d'analyses distributionnelles pour établir les catégories phonétiques pertinentes pour chaque langue (Curtin et al., 2011). Un input qui bloquerait ce mécanisme pourrait empêcher l'apprentissage.

Les difficultés observées pour l'acquisition de certaines représentations en LE, comme les relations phonèmes-graphèmes des voyelles nasales (Dherbey-Chapuis & Berthele, 2020), peuvent être expliquées par un blocage de ce mécanisme. Le graphème <*on*> est commun à la L1 (ici allemand) et au FLE ; il est statistiquement lié à des représentations phonologiques différentes en fonction de la langue : par exemple pour une position finale, en FLE */ɔ̃/-on* (e.g., *maison*) et en allemand est */ɔn/-on* (e.g., *von*). La construction des représentations phonétiques du FLE est bloquée par les biais perceptuels qui valorisent le plus souvent l'association de la L1 : <*on*> et */ɔn/*. Les liens phonétique-lexique-phonème deviennent alors très difficiles à construire (Darcy, Daidone, & Kojima, 2013). La construction des liens est bloquée entre les espaces multidimensionnels et l'apprentissage de la LE est limité.

Les mécanismes d'apprentissage (i.e. analyse statistique et le système comparaison-contraste) peuvent être ceux du système procédural de la boucle exécutive dans le modèle de mémoire de travail TBRS. Ils participent de la même façon dans les deux modèles à la perception et à l'apprentissage. Les mécanismes d'apprentissage et les filtres dynamiques contraignent le traitement online de l'information verbale entrante en se référant aux informations déjà conservées dans les différents espaces multidimensionnels. Tout comme dans le modèle TBRS, les informations, déjà présentes en mémoire à long terme (e.g., phonétiques, phonologiques, orthographiques, sémantiques...) et qui sont activées par une relation de similarité, participent à l'élaboration de la nouvelle représentation de l'information en y étant associées par des liens dynamiques. Ce système de comparaison-contraste participe à la fois à l'apprentissage et à la perception.

Lors de la perception, le traitement précoce de l'information verbale active toutes les représentations phonétiques et phonémiques dans le modèle PRIMIR (Curtin et al., 2011; Curtin & Werker, 2007; Norris, McQueen, & Cutler, 2003; Werker & Curtin, 2005; Werker & Logan,

1985). De même dans le modèle TBRS, les représentations phonétiques sont les premières représentations de l'information verbale. Elles sont conservées dans la boucle phonologique (Baddeley, 2003; Baddeley, Papagno & Vallar, 1988; Barrouillet & Camos, 2015; Cowan, 2008) et elles précèdent l'accès au sens. Après l'écoute d'un stimulus, l'activation des représentations phonétiques et phonémiques mesurée par le potentiel évoqué (= ERP, *Event-Related potential*) apparaît en environ 50ms, celles des représentations lexicales en environ 250-300ms et l'intégration lexico-sémantiques en environ 400ms (e.g., Camos et al., 2018; Hickok, 2012; Mesgarani, Cheung, Johnson & Chang, 2014; Stemmer & Whitaker, 2008). Les élaborations successives des représentations d'un input verbal suivent un ordre de succession temporelle.

Trois conséquences de l'ordre temporel d'activation et de construction des différentes représentations sont importantes : 1) Il est possible de cibler l'étude d'un type de représentation en fonction de la durée du silence après un stimulus oral dans une tâche (e.g., Werker & Logan, 1985). Un silence très court donne peu de chances à la représentation lexicale d'être activée et les représentations phonétiques et phonologiques seront principalement impliquées ; 2) La différence de temps de latence entre l'activation des représentations phonétiques et phonémiques et celle des représentations lexicales pourrait indiquer une acquisition implicite possible pour certaines compétences comme la discrimination, car ces compétences utilisent des représentations sollicitées au début du traitement de l'information verbale, et un apprentissage plutôt explicite du vocabulaire qui demande la construction de représentations multimodales très élaborées utilisant des représentations activées plus tardivement comme la valeur sémantique (Ullman, 2013) ; 3) l'activation lexicale influence la construction des représentations phonétiques et phonologiques après le début de leur construction. Cet effet prédictif (= *top-down*) indique un lien étroit entre l'apprentissage du lexique et l'acquisition des représentations phonétiques et phonologiques et inversement.

La valeur sémantique lexicale est en interaction notamment lors de l'apprentissage avec toutes les représentations phonétiques et phonologiques qui lui sont associées. Les représentations lexicales en mémoire à long terme peuvent avoir été encodées au niveau phonétique de façon globale (=encodage lexical global). Les représentations lexicales sont alors liées à un ensemble de représentations phonétiques (e.g., pour la L1 : Swingley, 2003; Vihman & Keren-Portnoy, 2019 ; pour la LE/L2 : Cutler et al., 2006; Darcy et al., 2013; Escudero et al., 2008). Les représentations lexicales peuvent également être encodées par leurs représentations phonologiques syllabiques ou phonémiques (e.g., pour une revue Laganaro, 2019 ; Dupoux et al., 2001; Ota, Hartsuiker & Haywood, 2009). En LE, les représentations phonétiques d'un contraste phonémique ne précèdent pas obligatoirement l'encodage du contraste phonémique au niveau lexical.

Les effets *top-down* peuvent être induits par toutes les valeurs (e.g., phonétiques, phonologiques, orthographiques, phonotactiques...) incluses dans la représentation lexicale. La valeur sémantique de la représentation lexicale peut suppléer et corriger une représentation

phonologique erronément associée. L'information verbale est toujours accompagnée d'une valeur sémantique portée par plusieurs éléments qui sont indépendamment traités en mémoire de travail. L'activation de la valeur sémantique (ou orthographique) de la représentation lexicale peut permettre d'inclure aux représentations multimodales de nouvelles représentations phonologiques dans la boucle exécutive. Ainsi l'absence de modification des routines permettant de détecter un contraste phonémique LE peut être contrebalancée par des représentations phonétiques liées de façon fiable à une valeur sémantique ou orthographique. L'enseignement des représentations phonologiques peut donc s'appuyer sur la valeur sémantique et sur un encodage global des représentations phonétiques dont l'apprentissage est implicite.

Les représentations phonétiques ont moins de chances d'être liées aux représentations lexicales adéquates lorsqu'elles n'incluent pas une caractéristique LE de l'information verbale au moment de leur sortie de la boucle phonologique et de leur intégration dans la boucle exécutive (Darcy et al., 2012; Dupoux et al., 2008; Ota et al., 2009). Darcy et al. (2012) ont montré que des apprenants de FLE d'un niveau intermédiaire n'encodent pas différemment, dans les représentations phonologiques globales lexicales, des contrastes phonémiques qu'ils étaient pourtant capables de catégoriser aussi justement que des apprenants avancés. Les apprenants avancés étaient, eux, capables d'encoder au niveau lexical le contraste phonémique comme des locuteurs L1. L'association entre les représentations phonétiques et phonologiques globales lexicales peut être favorisée par les processus attentionnels.

L'association en boucle exécutive d'une représentation phonémique à des représentations phonétiques peut induire secondairement une modification des routines et procédures du système procédural. Cet apprentissage, d'abord attentionnel, devrait pouvoir être porté par un enseignement explicite de l'association des représentations phonémiques et phonétiques (Ellis, N., 2005, 2015). L'association entre représentation phonémique et représentation lexicale peut être renforcée par l'acquisition de la représentation orthographique phonémique qui participera à l'établissement des liens sémantiques, orthographiques, phonémiques et phonétiques.

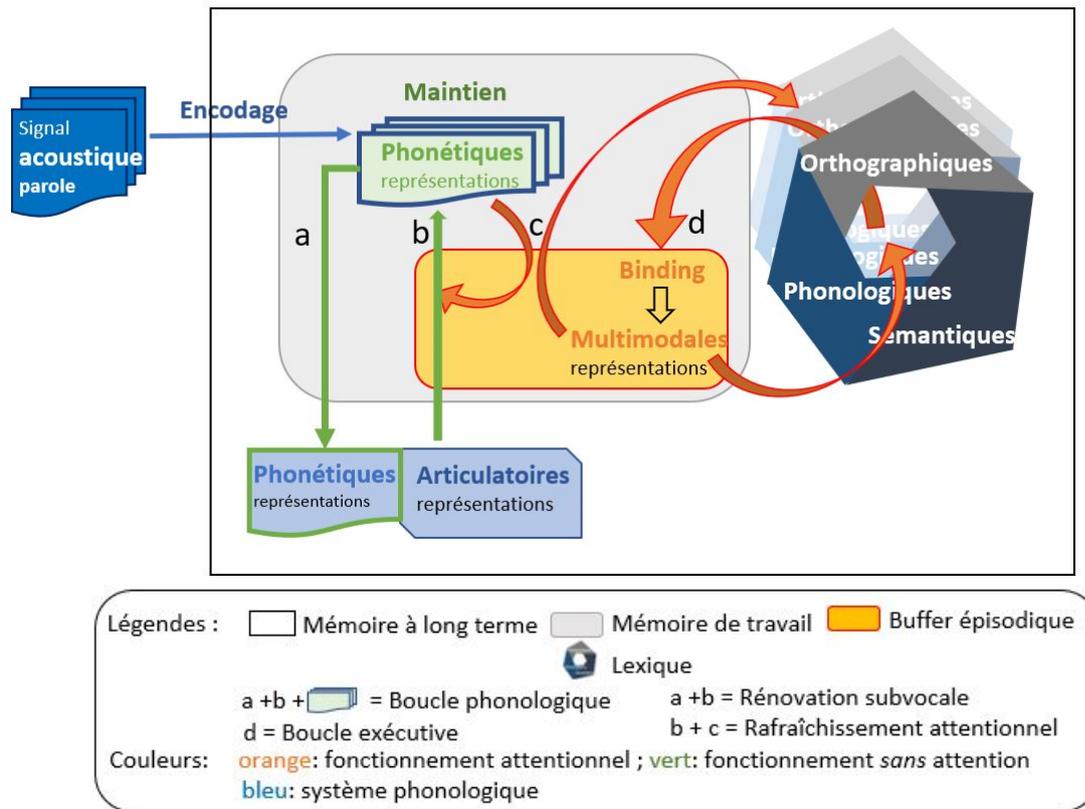
L'apprentissage des représentations phonétiques et phonologiques liées à un lexème peut commencer par celui des représentations phonétiques du mot entier ou par celui des représentations phonémiques (Pallier et al., 2001). Ces représentations sont mobilisées à des étapes différentes de la perception. L'ordre de mobilisation des représentations et des compétences en jeu lors de la perception peut être instructif pour la construction de la séquence didactique.

1.3 Perception de l'information verbale et apprentissage des représentations phonologiques et lexicales

Dans le modèle PRIMIR, les routines procédurales analysent simultanément tous les niveaux de représentations phonétiques et phonologiques : prosodiques, segmentation du flux de la parole, identification de la représentation phonologique globale d'un mot, extraction de syllabes, de représentations sublexicales, de phonèmes ou de représentations phonétiques (voir également la revue de Bürki, 2018). En parallèle de l'extraction, les différentes informations extraites sont confrontées entre elles et sont sélectionnées en fonction des différentes contraintes *bottom-up* et *top-down*. Cependant, dans une démarche schématique, les représentations verbales en mémoire de travail vont être présentées dans les paragraphes suivants de manière séquentielle en décrivant le parcours théorique effectué par une information verbale, depuis sa perception jusqu'à son inscription en mémoire à long terme.

Le traitement de l'information verbale en mémoire de travail commence en boucle phonologique. La boucle phonologique est le buffer périphérique sensoriel de la mémoire de travail et correspond à l'espace perceptuel du modèle PRIMIR. La perception du matériel verbal, après les premiers codages sensoriels au niveau de l'oreille et de leur transmission neurologique, se matérialise par l'inscription en boucle phonologique des premières représentations phonétiques : c'est l'encodage sous formes de représentations phonétiques en boucle phonologique (Figure 1).

Figure 1 : Représentations phonologiques en mémoire de travail



La rénovation subvocale est un des deux mécanismes de rénovation de l'information verbale en boucle phonologique. Les représentations phonétiques de la boucle phonologique sont rapidement détruites si elles ne sont pas rénovées (Murakami, Kell, Restle, Ugawa & Ziemann, 2015; Ricker, 2015). La rénovation subvocale se met en place en identifiant les représentations phonétiques d'origine sensorielle à des représentations phonétiques qui partagent certaines caractéristiques et qui sont déjà présentes en mémoire à long terme (Bouchard, Mesgarani, Johnson & Chang, 2013; Mesgarani et al., 2014). Ces représentations phonétiques activées en mémoire à long terme sont associées aux représentations articulatoires qui permettent de les produire. La rénovation subvocale est constituée de la planification des ordres moteurs articulatoires correspondant à la suite ordonnée des représentations phonétiques contenues en boucle phonologique. Les ordres moteurs articulatoires correspondent à l'objectif auditif défini par les représentations phonétiques d'origine sensorielle (Hickok, 2012; Hickok, Rogalsky, Chen, Herskovits..., & Hillis, 2014). Cette étape est préattentionnelle (Flege, 2016), ce qui signifie que la personne qui entend le stimulus n'a pas l'intention explicite d'analyser le stimulus et qu'elle peut y accorder de l'attention sans action volontaire de sa part.

La discrimination préattentionnelle analyse l'adéquation de l'appariement entre les représentations articulatoires et les représentations phonétiques (figure 1) réalisé par la répétition subvocale (Zatorre & Gandour, 2008). Cette étape correspond au fonctionnement du système comparaison-contraste du modèle PRIMIR. Les représentations articulatoires et

phonétiques peuvent se correspondre et le processus de perception suivra son cours *bottom-up*, ou ne pas se correspondre. Dans le cas où elles ne sont pas congruentes, comme lors de l'apprentissage d'une LE par un débutant, la non-correspondance des représentations articulatoires et phonétiques peut être détectée par le système comparaison-contraste. La capacité de détecter l'adéquation de la correspondance entre représentations phonétique et articulatoire, c'est-à-dire l'existence et la disponibilité des représentations phonétiques LE et de leur liens avec les représentations articulatoires, est mesurée par une tâche discrimination AX (Curtin & Werker, 2007).

La détection d'une représentation phonétique inhabituelle, réalisée en perception et dans une tâche de discrimination, est indépendante des langues. Cette discrimination préattentionnelle témoigne de l'absence de représentation similaire en mémoire à long terme. En neurologie, l'inadéquation entre les représentations phonétique et articulatoire est observable et mesurable même en l'absence d'attention volontaire par la mesure des différences de potentiels évoqués (=ERP ; Näätänen & Winkler, 1999; Schröger, 1997; Stemmer & Whitaker, 2008). La discrimination est préattentionnelle, mais peut accaparer l'attention car la résolution du conflit identifié demande de l'attention.

La détection d'une représentation phonétique atypique est corrélée à une charge cognitive de traitement supérieure et à une réduction de l'apprentissage du vocabulaire (Abdallah, 2010; Holt & Lotto, 2010; Masoura & Gathercole, 1999). En LE, la non adéquation entre les représentations phonétique et articulatoire entraîne un coût cognitif supérieur de l'encodage, soit par un ralentissement du recodage soit par une diminution de la durabilité du code phonétique en boucle phonologique (Piske et al., 2002). Un enseignement de la discrimination permettrait peut-être d'attirer l'attention des apprenants sur les caractéristiques phonétiques indispensables à une perception efficiente de la LE (caractéristiques spécifiques de la LE), et faciliter ainsi l'acquisition des représentations phonétiques et phonologiques correspondantes. La discrimination peut être un axe didactique de l'enseignement du vocabulaire en LE car elle peut aider à lever la difficulté de perception. Lorsque les représentations phonétiques d'origine sensorielle ne peuvent pas être liées à des représentations congruentes et déjà présentes en mémoire à long terme, la construction d'une représentation élaborée dépend de l'information contenue en mémoire à long terme.

Pour ne pas bloquer la perception lorsque les représentations sont atypiques, les régions du *Superior Temporal Gyrus* et du *Superior Temporal Sulcus* (ici regroupées par commodité sous le terme STS) assurent un appariement entre les représentations phonétique et articulatoire qu'il y ait parfaite adéquation ou non (Hickok, 2012; Hickok et al., 2014). Le STS permet la construction d'une représentation de l'association phonétique-articulatoire grâce à un processus attentionnel (Hickok, 2012). De même dans le modèle TBRS (voir 1.1), les conflits peuvent être résolus par des processus attentionnels en boucle phonologique pour permettre le déroulement de la rénovation (Heald & Nusbaum, 2014). Il est possible de penser que cette étape de perception correspond au début du rafraîchissement attentionnel en mémoire de travail

et du début de traitement de l'information en boucle exécutive, qui sont eux-aussi des processus attentionnels impliqués dans la perception.

L'appariement par le STS entre représentation phonétique et articulatoire est activé non seulement en perception, mais aussi en production, et lors du maintien des représentations phonémiques (Hickok, 2012). La représentation phonémique peut activer les représentations lexicales, phonétiques et phonologiques qui lui sont liées et donc faciliter leur apprentissage par l'établissement de liens entre les espaces multidimensionnels. Deux conséquences sont importantes pour cette étude : 1) l'enseignement des représentations phonémiques peut faciliter la perception en LE et donc l'acquisition du vocabulaire ; 2) une tâche de prononciation informe sur le fonctionnement perceptif et sur l'acquisition phonémique ; le progrès de la prononciation des phonèmes LE devrait accompagner l'acquisition du vocabulaire.

Le fonctionnement perceptif est largement influencé par les représentations multimodales présentes en mémoire à long terme. L'appariement imparfait phonétique-articulatoire peut être modifié par les autres représentations activées en mémoire à long-terme (Heald & Nusbaum, 2014), en particulier phonologiques (*top-down* processus). La représentation articulatoire initialement activée en boucle phonologique peut être remplacée en boucle exécutive par une autre représentation articulatoire plus cohérente avec l'ensemble des représentations activées et notamment par une représentation articulatoire codant une autre place articulatoire (Murakami et al., 2015). La modification de la part phonologique de la représentation multimodale peut aussi avoir des effets négatifs et expliquer par exemple les phénomènes de surdité phonologique ou d'assimilation phonémique (Hickok, 2012). Un enseignement de la prononciation permet peut-être : 1) de favoriser la création des représentations phonétiques LE, et de favoriser leur lien à des représentations articulatoires adéquates, 2) de favoriser la construction de représentations phonologiques LE en limitant la correction par les représentations lexicales et sublexicales (figure 1). L'enseignement de la prononciation peut cibler toutes les tailles de représentations phonologiques, et en particulier la représentation lexicale globale ou le phonème.

Le phonème est une représentation sublexicale importante dans mon cadre théorique car il est explicitement défini dans le modèle PRIMIR comme le lien dans les relations entre les espaces multidimensionnels et la forme cognitive la plus légère à manipuler pour le système procédural. Les phonèmes sont représentés en fonction de leur distribution (i.e., leur position dans le mot), et sont liés à leurs représentations phonétiques, articulatoires et lexicales (orthographiques et sémantiques). Cette organisation peut être considérée comme la base des « règles » phonotactiques d'une langue. En effet, l'analyse statistique de leur distribution effectuée par le système procédural permet d'établir des prévisions de leur distribution qui réduisent le coût cognitif de la perception. Plusieurs auteurs ont montré que les interférences cross-linguistiques phonotactiques gênaient la perception en LE (Cutler, 2015; Escudero, Simon & Mulak, 2014; Gaonac'h, 2000).

Lorsque les deux langues de l'apprenant partagent le même système d'écriture mais pas les mêmes correspondances phonème-graphème ou les mêmes règles phonotactiques, comme dans la présente étude, l'enseignement des correspondances phonème-graphème et des règles phonotactiques peut peut-être limiter les difficultés en perception et favoriser l'acquisition du vocabulaire LE (figure 1). En effet, l'acquisition des correspondances phonème-graphème et des règles phonotactiques en LE n'est pas rapide. Après plus de 300 heures d'enseignement, certains apprenants n'arrivent pas à reconnaître l'écriture selon le code LE de pseudo-mots entendus, en particulier lorsque le phonème n'existe pas dans le répertoire phonologique L1 ou lorsqu'il occupe une place illégale pour les règles phonotactiques L1 (Dherbey-Chapuis & Berthele, 2020). L'enseignement des correspondances phonème-graphème peut peut-être lever cet obstacle et ainsi favoriser la perception et l'acquisition du vocabulaire. Le niveau de maîtrise des correspondances phonème-graphème influence d'ailleurs le fonctionnement de la boucle phonologique en mémoire de travail (Szewczyk et al., 2018).

Une tâche simple de répétition de pseudo-mots mesure principalement les performances de la boucle phonologique (e.g., Gathercole, 2006). La boucle phonologique et la production de la parole s'appuient sur les mêmes mécanismes neurologiques pour planifier les ordres moteurs articulaires (Glanz, Derix, Kaur, Schulze-Bonhage..., & Ball, 2018; Hickok et al., 2014; Pulvermüller, Huss, Kherif, Moscoso del Prado..., & Shtyrov., 2006; Yeung & Werker, 2013); ce qui indique que : 1) une tâche de répétition de non-mots est pertinente pour mesurer les performances de la boucle phonologique et évaluer son fonctionnement (Coady & Evans, 2008; Metsala, 2011), 2) un enseignement de virelangues ou la mémorisation de phrases en LE peut entraîner la boucle phonologique et potentiellement augmenter ses performances, 3) comme le maintien en mémoire de travail mesuré par une tâche de répétition de pseudo-mots est corrélé positivement au niveau de vocabulaire (e.g., Abreu & Gathercole, 2012; Gathercole, Willis, Emslie & Baddeley, 1992; Masoura & Gathercole, 1999), on peut émettre l'hypothèse que l'apprentissage du vocabulaire puisse être favorisé par un meilleur fonctionnement de la boucle phonologique, 4) la répétition d'un mot en boucle pendant le recodage de l'information verbale bloque le fonctionnement de la boucle phonologique; c'est la condition de suppression articulaire pendant une tâche de répétition de pseudo-mots (Baddeley, 2012). Cette condition permet une évaluation des performances du rafraîchissement attentionnel et de la disponibilité des représentations sublexicales sur lesquelles il s'appuie.

En LE, le lien entre le maintien de l'information verbale et le vocabulaire est peut-être en partie exercé par la disponibilité et la qualité des représentations phonémiques et de leurs représentations associées (e.g., Curtin & Werker, 2007; Werker & Curtin, 2005). Le degré d'activation des représentations articulaires est corrélé positivement au maintien de l'information verbale en boucle phonologique lorsqu'il est mesuré par les performances d'une tâche de répétition de pseudo-mots (Buchsbaum, Baldo, Okada, Berman..., & Hickok, 2011; Herman, Houde, Vinogradov & Nagarajan, 2013; Hickok & Poeppel, 2007; Murakami et al., 2015). La répétition de pseudo-mots repose essentiellement sur l'association des

représentations phonémiques, articulatoires et phonétiques. L'amélioration des représentations articulatoires par un enseignement ciblé peut peut-être augmenter le maintien des représentations phonétiques et phonémiques associées. L'enseignement de la prononciation en LE devrait favoriser un meilleur maintien de l'information verbale LE. Cependant un résultat opposé peut aussi être observé en début d'acquisition phonémique.

Une représentation phonémique LE peu stable, comme c'est le cas en début d'acquisition, peut induire des performances inférieures comparativement à l'absence de représentation phonémique LE lors d'une tâche de répétition de pseudo-mots en LE. Perez et al. (Pérez, Carreiras, Gillon Dowens & Duñabeitia, 2015) observent que lorsque le stimulus est en LE, les représentations phonétiques sensorielles sont prépondérantes pour la perception, au détriment des processus mobilisant les connaissances en mémoire à long terme. La présence d'un phonème non familier semble renforcer l'importance des processus *bottom-up* basés sur les représentations phonétiques d'origine auditive ce qui induit une charge cognitive supplémentaire (Ernestus & Mak, 2004; Hisagi, Shafer, Strange & Sussman, 2015; Strange, 2011). Toute charge cognitive entraîne une perte d'information en boucle phonologique (voir 1.1). Par conséquent, des représentations phonémiques en cours d'acquisition (i.e., liées à des représentations phonétiques mal définies et peu spécifiques) peuvent limiter le maintien de l'information verbale LE par la charge cognitive supplémentaire qu'elles induisent. Une rénovation limitée entraîne une perte d'information, information qui ne sera pas incluse au sein de la boucle exécutive.

Les représentations phonologiques activées en boucle exécutive représentent la base des représentations multimodales qui permettent l'apprentissage d'un lexème inconnu (Kaushanskaya & Marian, 2008; Masoura & Gathercole, 2005). Les représentations phonologiques activent et sont activées par les autres représentations lexicales et sublexicales pour constituer les représentations multimodales (figure 1), et ce indépendamment de la langue auxquelles les représentations appartiennent (Friesen, Jared, & Haigh, 2014). Le vocabulaire de la L1 peut aider l'encodage de la LE uniquement si les représentations phonologiques sont communes à la L1 et la LE (Kaushanskaya & Marian, 2008). Les représentations phonologiques LE sont probablement des représentations importantes pour l'apprentissage du vocabulaire LE.

Pour observer les liens entre l'acquisition du vocabulaire réceptif et celle des représentations phonologiques LE, cette étude propose d'observer les effets d'une séquence didactique visant à enseigner les représentations phonétiques et la compétence de discrimination, les représentations phonémiques et articulatoires et la compétence de prononciation, les correspondances phono-graphémiques et la compétence orthographique correspondante, le maintien en mémoire de travail et les représentations lexicales.

L'enseignement explicite vise à modifier les associations de représentations dans les représentations multimodales, et l'enseignement implicite vise à induire des modifications des procédures utilisées par le système procédural. La forme de l'enseignement explicite ou

implicite peut influencer l'acquisition des représentations phonétiques, phonologiques et du vocabulaire.

1.4 Enseignement des représentations phonétiques et phonologiques

L'objectif d'un enseignement est d'induire un changement des états mentaux de l'apprenant, c'est-à-dire de provoquer un apprentissage (voir 1.2). Apprendre une LE, selon Ellis et Wulff (Ellis, N. & Wulff, 2012, p.12), demande à l'apprenant de modifier les biais attentionnels, qu'il a construits par l'usage de manière implicite lors de l'acquisition et de l'utilisation de sa L1, pour pouvoir accorder de l'attention aux particularités de la LE. En LE, l'apprenant doit cibler son attention sur les particularités de l'objet à apprendre, en particulier lorsque cet objet est peu saillant au niveau perceptif ou redondant dans l'input (Ellis, N., 2015), pour que les caractéristiques de cet objet puissent être remarquées et apprises lors de son traitement cognitif. Dans mon cadre théorique, l'objectif est de contraindre les processus procéduraux de la boucle exécutive (TBRS) ou le système de comparaison-contraste (PRIMIR) à prendre en compte les particularités de la LE (e.g., considérer comme contrastif la nasalisation d'une voyelle ou le caractère voisé d'une fricative). Contraindre l'attention de l'apprenant à se centrer sur une forme de structure linguistique peut être réalisé par un enseignement ciblé.

Selon Long (1998), l'enseignement ciblé a trois formes possibles. Premièrement, basé sur l'idée que si l'input est riche et a du sens, un focus sur le sens devrait permettre un apprentissage procédural (*incidental learning*). Plusieurs études ont cependant montré que l'apprentissage était systématiquement inférieur lors d'un enseignement *incidental* en comparaison d'un enseignement explicite (e.g., méta-analyse de Norris & Ortega, 2000 ; revue de Saito, 2012). Deuxièmement, basé sur l'idée que la complexité de certaines formes linguistiques impose une approche isolée et organisée (i.e., un focus sur des formes) pour qu'elles puissent être apprises, un focus sur la forme devrait permettre de créer des représentations multimodales qui l'incluent. Cette forme d'enseignement ciblé ne présente pas le sens comme le premier but de l'activité, mais organise la mobilisation de compétences utilisant la forme ciblée. Troisièmement, basé sur une approche mixte à la fois guidée par le sens et prenant également en charge une forme ciblée, le travail bref et ponctuel d'une forme qui pose problème peut permettre de l'apprendre et de l'utiliser en contexte communicatif. Spada (1997) élargit cette troisième approche en introduisant la notion de conscience de l'apprenant de l'objet à apprendre. La conscience est volontaire et diffère de l'attention qui réfère au fonctionnement cognitif. Les formes d'enseignement retenues pour cette étude sont ciblées (i.e., autres qu'*incidental*), mais elles diffèrent sur l'aspect implicite versus explicite.

La définition des deux séquences didactiques de cette étude a suivi la classification proposée par Norris et Ortega (2000) en fonction de leur nature implicite ou explicite, et de leur cible

sémantique ou sur la forme. Premièrement, la séquence didactique communicative actionnelle (= séquence didactique C) est considérée comme ciblée sur la forme et le sens (= « FonF ») car 1) elle développe des tâches dans lesquelles le sens est la priorité de l'apprenant, 2) le matériel pédagogique est intégralement authentique, 3) la démarche didactique vise à permettre à l'apprenant de remarquer les particularités de la LE sans pour autant nommer, expliquer ou travailler explicitement les formes ciblées. Cette séquence didactique est considérée comme une *integrated practice* (Kim & Godfroid, 2019), c'est à dire un enseignement implicite des formes phonétiques et phonologiques, qui sont cependant didactiquement présentées pour être remarquées, et un enseignement explicite du vocabulaire dont la valeur sémantique est l'objectif de toutes les tâches proposées à l'apprenant (tableau 1).

Tableau 1 : Séquence didactique et type d'enseignement des formes phonologiques

Séquence didactique	Discrimination	Prononciation	Maintien en mémoire de travail	Correspondances phonème graphème	Vocabulaire réceptif
Séquence communicative (= C)	I	I	I	I	Sémantique= E Forme phonologique = I
Séquence explicite (= E)	I	Méthode verbo-tonale : I/E Méthode articulatoire : E	I	Métalinguistique = E	Sémantique= / Forme phonologique = E

Légendes : I = implicite ; E = explicite

Deuxièmement, la séquence didactique explicite (= séquence didactique E) est considérée comme ciblée sur les formes (= « FonFs ») car 1) elle n'utilise jamais la valeur sémantique du matériel verbal enseigné, 2) toute l'attention de l'apprenant est centrée sur la forme ciblée. Les règles de grammaire phonologiques et phonotactiques et la prononciation (par la méthode articulatoire) sont explicitement travaillées, par contre, la discrimination, la prononciation (par la méthode verbo-tonale) et le maintien de l'information verbale en mémoire de travail sont enseignés implicitement. Pour le niveau métalinguistique, la conscience phonologique est travaillée avec des jeux d'assemblage et de segmentation phonémique sans recours à la valeur sémantique. L'enseignement de la conscience phonologique est explicitement orienté vers la forme phonologique. La forme phonologique globale du vocabulaire est explicitement travaillée lors de l'enseignement du lien son-orthographe et de la prononciation du mot mais sans jamais utiliser le sens du lexème associé à cette forme phonologique. La nature de

l'enseignement explicite versus implicite est aux mains de l'enseignant.e mais l'apprentissage est le fruit d'une coopération des savoirs implicite-explicite (Ellis, N., 2005 ; voir 1.1 et 1.2).

Dans mon cadre théorique, l'apprentissage est le fruit d'un processus attentionnel. La construction des représentations multimodales est le résultat des processus procéduraux implicites mais leur inscription en mémoire à long terme est le résultat de la boucle attentionnelle.

L'enseignement explicite attire l'attention de l'apprenant sur la structure de la forme ciblée et lui fixe comme but son usage. La construction explicite et attentionnelle de nouvelles représentations LE permettrait l'inscription en mémoire à long terme d'exemplaires de référence (= traces des représentations multimodales élaborées par le système procédural). Les nouveaux exemplaires induiraient, par des processus de généralisation, la mise à jour implicite des processus procéduraux (Ellis, N., 2015). Par exemple, l'enseignement explicite de la nasalisation obligatoire du phonème /ɔ̃/-on induirait le lien entre représentations phonétiques de la nasalité et la représentation sémantique de *maison, bonbon, tonton* puis ces représentations phonétiques guideraient le recodage dans la boucle phonologique et permettrait le maintien de l'information phonétique « nasalité ».

L'enseignement implicite viserait à faciliter la reconnaissance par le système procédural des spécificités de la LE. Le traitement implicite des spécificités de la LE par le système procédural pourrait être directement influencé par une mise en évidence des contrastes existants en LE et pas en L1. La comparaison des représentations phonétiques LE aux représentations existantes en mémoire à long terme s'appuierait sur les résultats de l'analyse contrastive et induirait la création de nouvelles catégories LE quand celle-ci diffèrent suffisamment des représentations L1 existantes (voir 1.2 : Modèle PRIMIR et autres modèles basés sur l'acquisition d'exemplaires). Par exemple, l'enseignement implicite de la nasalisation obligatoire du phonème /ɔ̃/-on pourrait être facilité par une juxtaposition dans l'input du phonème /ɔ̃/-on et de ses concurrents communs à la LE et la L1 (i.e., phonèmes /u/-ou et /o/-<eau, au, o...>) dans des structures syllabiques simples comme par exemple « Tu me donnes un bonbon ? » (= /ty mɑ dɔ̃n œ̃ bɔ̃bɔ̃/ ?). Les effets de l'enseignement, explicite versus implicite, ont été comparés dans de nombreuses études.

L'enseignement explicite et ciblé sur la forme semble plus efficace en LE qu'un enseignement implicite pour favoriser l'acquisition des représentations implicites et explicites (e.g., revue de Dekeyser, 2003 ; revue de Ellis, R., 2015 ; méta-analyse de Norris & Ortega, 2000 ; méta-analyse de Spada & Tomita, 2010). Les enseignements ciblés, explicites ou implicites, sont rarement comparés dans des études visant à mesurer les effets d'un enseignement des représentations phonologiques en salle de classe contrairement aux études concernant la grammaire ou le vocabulaire. Par exemple, l'acquisition incidente de la grammaire semble être moins efficace qu'un apprentissage explicite. Denhovska, Serratrice et Payne (2016) ont comparé les performances de deux groupes d'adultes anglophones dans

l'apprentissage/acquisition de l'accord de l'adjectif en fonction du cas grammatical et du genre en russe LE. Les cent participants ont reçu soit un enseignement explicite qui incluait des exemples, des traductions, des indications métalinguistiques et des images illustrant le sens des phrases utilisées comme exemples, soit un enseignement visant un apprentissage incidentiel à partir des mêmes images et des mêmes phrases regroupées sous formes de petites séquences narratives. Immédiatement après la phase d'enseignement, les participants ont dû compléter les marques flexionnelles finales des adjectifs dans un texte à trous. Le groupe explicite a obtenu des résultats significativement très supérieurs à ceux du groupe incidentiel ($p < .001$). Cependant les études sur les effets du *feed-back* semblent contester la supériorité de l'approche explicite.

Les *feed-back* explicites et implicites semblent tout autant favoriser l'apprentissage de la LE quand ils peuvent être individualisés (Saito, 2021). Le caractère implicite versus explicite du *feed-back* est un continuum (Lyster, Saito & Sato, 2013). Pour une comparaison plus aisée entre les deux séquences didactiques de cette étude, les *feed-back* sont donnés dans les deux séquences didactiques comparées sous forme de *recast* qui est la forme de *feed-back* la plus largement utilisée dans un contexte scolaire (e.g., Lyster et al. 2013) et la plus efficace (Saito, 2021). Par exemple pour la prononciation, le *feed-back* explicite, dans la séquence didactique E, fournit des explications articulatoires et métalinguistiques tout en transmettant la forme correcte (= *form recast*). Au contraire, le *feed-back* implicite, dans la séquence didactique C, consiste à répéter et/ou reformuler partiellement ou complètement la production de l'apprenant sans lui indiquer quelle partie de sa production nécessite une amélioration (forme correcte donnée = *conversational recast*) (e.g., Li, 2010; Lyster et al., 2013; Sheen & Ellis, R., 2011). Les *feed-back* implicites et explicites semblent avoir le même effet positif sur l'apprentissage de formes grammaticales (e.g., Kang, 2009; Shin & Christianson, 2012) ou sur les formes phonétiques et articulatoires (e.g., Lyster, 2018). Le *feed-back* implicite semble favoriser un apprentissage plus durable que l'explicite, et le *feed-back* explicite semble être plus performant à court terme que l'implicite (e.g., Li, 2010). La méta-analyse de Lee, Jang et Plonsky (2015) montre que d'une façon générale l'instruction explicite des représentations phonétiques et articulatoires, dont le *feed-back* est un des outils, favorise le développement de la prononciation de la LE en comparaison des apprenants qui ne reçoivent aucune activité centrée sur la prononciation. Cependant, d'autres facteurs semblent influencer l'effet des *feed-back* correctifs (e.g., Saito, 2021) et de l'enseignement.

1.5 Facteurs susceptibles de modifier les effets des séquences didactiques.

Les effets de la séquence didactique et du type de *feed-back* donné sont influencés par d'autres facteurs liés à l'apprenant, la compétence développée et à l'objet ciblé.

1.5.1 Incidence des caractéristiques de l'apprenant sur l'effet du type d'enseignement

Certaines caractéristiques de l'apprenant influencent les effets comparés de l'enseignement explicite et implicite, notamment son âge et son niveau de langue (e.g., DeKeyser, 2003). De nombreuses études ont montré la supériorité de l'enseignement explicite des représentations phonologiques en LE en contexte de laboratoire avec des apprenants adultes, mais peu d'études ont ciblées le contexte scolaire et des apprenants non adultes (DeKeyser, 2003 ; Norris & Ortega, 2000). Dans la méta-analyse de Norris et Ortega (2000), seulement trois études ont été conduites en contexte scolaire et elles semblent indiquer une plus grande efficacité de l'enseignement explicite.

Une des caractéristiques de l'enseignement explicite est de favoriser le développement de la conscience métalinguistique (DeKeyser, 2005), ce qui aurait pour conséquence d'augmenter l'habileté à remarquer les particularités de la LE (Lichtman, 2013). Un apprenant adulte diffère d'un apprenant plus jeune notamment par des capacités métacognitives supérieures. Un adulte est plus réceptif à l'enseignement explicite qu'à l'enseignement implicite. Au contraire, les enfants semblent apprendre plus facilement implicitement (DeKeyser, 2003). L'approche didactique explicite n'est d'ailleurs pas forcément plus efficace en contexte scolaire dans toutes les études.

Dans une étude en contexte scolaire, Papachristou (2011) mesure l'effet comparé d'un enseignement explicite et de *feed-back* correctifs implicites sous forme de *recast* sur la prononciation des voyelles de l'anglais par 15 apprenants grecs âgés de 16 ans. L'instruction (pour un total de 7 heures) semble avoir peu d'effets et aucune différence significative n'est observable entre les deux groupes expérimentaux et le groupe contrôle qui n'a reçu aucun enseignement de la prononciation. Le même résultat a été observé lors de l'apprentissage de séquences grammaticales par 79 apprenants d'espagnol LE âgés de 20 ans en moyenne (Granena & Yilmaz, 2019). Une des explications avancées par les auteurs est que le niveau de LE au moment de l'instruction peut interagir avec l'aspect implicite/explicite de l'enseignement.

Li (2009) a comparé les effets des *feed-back* explicites et implicites sur l'acquisition des classificateurs du chinois par des apprenants américains de deux niveaux différents de compétence en LE. Les apprenants les plus faibles utilisent plus de classificateurs corrects et

corrigent mieux leur production après un feed-back explicite qu'implicite, et aucune différence n'a été observée entre les deux types de feed-back pour les apprenants avancés. La conséquence pour les résultats de cette étude peut être qu'un effet supérieur de l'enseignement explicite ne soit observable que pour les élèves les plus faibles.

L'influence des caractéristiques de l'apprenant peut interagir avec les demandes spécifiques de la tâche.

1.5.2 Incidence de la tâche sur l'effet du type d'enseignement

Les effets du type d'enseignement, implicite versus explicite, sont influencés d'une part par la nature de la tâche mesurant les performances de l'élève et la compétence observée, et d'autre part par la structure de la progression didactique.

1.5.2.1 Tâche, compétence ciblée et type d'enseignement

Dans mon cadre théorique, les représentations phonétiques et les représentations phonologiques sont des représentations dont l'acquisition devrait permettre de faciliter l'apprentissage du vocabulaire.

L'encodage lexical global ou le phonème contrastif sont deux des représentations phonologiques qui peuvent permettre d'identifier une représentation orthographique (Lahiri & Marslen-Wilson, 1991). Llompert et Reinisch (2017) ont observé les effets d'un entraînement articulatoire sur l'encodage lexical d'un contraste phonémique. L'entraînement visait à développer soit les représentations articulatoires, peu accessibles consciemment, soit l'encodage phonologique lexical global. Il semble que la condition d'écoute ait favorisé un apprentissage plus global de la représentation phonologique lexicale (i.e., par un apprentissage implicite), alors que le développement ciblé des représentations articulatoires a induit un apprentissage ciblé sur le contraste phonémique (i.e., par un apprentissage explicite). En termes de taux d'associations correctes entre les formes orthographiques et phonologiques, aucune différence n'est observée entre les groupes. Les deux formes, globale ou phonémique, devraient pouvoir faciliter l'acquisition du vocabulaire.

La séquence didactique explicite propose de favoriser initialement les représentations phonétiques. L'acquisition des représentations phonétiques est toujours implicite (voir 1.3) mais leur enseignement indépendamment des représentations sémantiques est forcément explicite. Dans la séquence didactique explicite (=E), l'enseignement commence par les représentations phonétiques, ce qui a quatre conséquences didactiques : 1) Bien que les représentations phonétiques ne soient pas accessibles explicitement, l'enseignement de la tâche et le feed-back sont explicites ; 2) Chaque séance débute par le travail de la discrimination et la discrimination précède systématiquement le travail de la prononciation ; 3) Le progrès est construit au fil des séances par l'augmentation de la taille de la représentation (de phonétique à phonologique au niveau du groupe prosodique) ; 4) Les deux premières séances n'utilisent pas

de matériel écrit pour limiter l'influence des autres représentations lexicales. L'influence des autres représentations lexicales est indissociable de l'enseignement de la forme phonologique globale.

La séquence didactique communicative actionnelle (=C) propose une acquisition préliminaire de la représentation phonologique globale du mot. La représentation phonologique globale peut, par un affinage progressif des représentations de taille inférieure, permettre la construction secondaire des représentations phonétiques (Ziegler & Goswami, 2005). L'usage du sens en situation de communication favorise une acquisition des représentations phonologiques et de leur usage (Lyster et al., 2013; Saito & Lyster, 2012). La séquence didactique C utilise directement les représentations phonologiques globales ; ce qui a trois conséquences didactiques : 1) les représentations phonologiques sub-lexicales ne sont jamais travaillées isolément ; 2) l'écrit et toutes les représentations du lexème, y compris les représentations phonologiques, sont travaillées simultanément ; 3) le travail implicite de la discrimination doit être construit didactiquement pour favoriser l'émergence des représentations phonétiques.

Enseigner implicitement les représentations phonétiques repose sur l'application du principe de continuité. L'enseignement implicite de la discrimination devient inefficace quand la tâche permettant l'acquisition propose les éléments à comparer séparés l'un de l'autre par plusieurs autres éléments (DeKeyser, 2003 ; Ellis, N. & Wulff 2012). Par conséquent, dans la séquence didactique C, les représentations phonétiques sont implicitement placées en évidence par le rapprochement au sein de la même phrase des représentations phonémiques de la même paire minimale (e.g., « tu me **donnes** la **maison** s'il te plaît ? »= /tym dɔn la mezɔ sil tɔ ple/). Elle promeut ainsi la saillance de la caractéristique LE qui permet de séparer le sens des lexèmes utilisés. Dans la séquence didactique E, les représentations phonologiques sont présentées sous forme de paires minimales qui juxtaposent les deux phonèmes (e.g., bon-beau = /bɔ̃-bo/). Dans les deux séquences, l'enseignement de la discrimination vise à faire prendre conscience à l'apprenant des particularités des représentations phonétiques de la LE.

Bien qu'un enseignement de la discrimination soit efficace, elle est rarement enseignée (Kissling, 2018). À ma connaissance, aucune autre étude n'a observé l'effet d'une didactique dédiée sur l'évolution des compétences en discrimination en LE en milieu scolaire. Les études qui mesurent cette compétence ont généralement pour but d'établir si son acquisition est une condition préalable ou une conséquence de l'acquisition de la prononciation de la LE.

Dans une étude réalisée avec des étudiants anglophones, Kissling (2014) a comparé l'influence de deux didactiques, explicite ou implicite, sur la prononciation et a évalué la force prédictive de la compétence de discrimination pour l'évolution de la prononciation. Les participants étaient des apprenants de 3 niveaux différents d'espagnol LE (n=74). Ils ont suivi un enseignement ciblé de la prononciation de 8 consonnes espagnoles réputées difficiles pour des apprenants anglophones selon une des deux didactiques proposées. Les participants ont été aléatoirement répartis dans deux groupes ; chaque groupe a suivi l'enseignement correspondant

à une des deux didactiques. L'enseignement ciblé a été dispensé sur support numérique en quatre modules répartis sur 3 semaines. La didactique explicite a proposé un travail explicite des représentations phonologiques, des représentations articulatoires (avec un tractus vocal animé) et des relations graphème-phonème, ainsi que des répétitions de stimuli prononcés par des locuteurs natifs (sans *feed-back*). La didactique implicite a proposé des vidéos de locuteurs natifs de l'espagnol, des traductions en anglais de ces vidéos, des tâches de dictées et une réflexion métalinguistique sur l'accent anglais en espagnol, ainsi que les mêmes répétitions de stimuli que l'autre groupe (sans *feed-back*). Les deux didactiques ont conduit à une amélioration significative moyenne d'environ 7% de la prononciation sans qu'une différence significative entre les deux groupes ne soit observée. Dans les deux groupes, plus les performances en discrimination au pré-test sont hautes, plus le progrès de la prononciation est important, ce qui pourrait indiquer que la discrimination est un préalable à la prononciation.

Dans cette étude de Kissling (2014), la discrimination mesurée au pré-test est un prédicteur significatif mais son influence sur le niveau en prononciation au post-test reste limitée (environ 1,2% de la variance). Ce résultat obtenu en instruction guidée confirme que la discrimination est importante pour l'acquisition de la prononciation sans toutefois être le facteur principal de son développement. Ce résultat rejoint les conclusions de la méta-analyse de Sakai et Moorman (2018). Cette méta-analyse est basée sur trente études réalisées en laboratoire qui visaient à observer les effets d'un entraînement en discrimination sur la prononciation en LE/L2. Les résultats de cette méta-analyse indiquent que l'entraînement de la discrimination améliore de façon modérée les performances en discrimination et de façon limitée la prononciation. L'entraînement de la compétence de discrimination permet toujours une amélioration des performances en discrimination, et souvent une amélioration de la prononciation (revue de Isaacs, 2014 ; méta-analyse de Sakai & Moorman, 2018).

L'efficacité de l'enseignement de la prononciation dépend du niveau de compétence en LE de l'apprenant et du type d'enseignement. Pour un apprenant débutant, un enseignement explicite semble plus efficace pour le développement de la prononciation qu'un enseignement implicite (e.g., revue de Saito, 2021) et les apprenants débutants progressent plus lors d'un travail de la prononciation que les apprenants avancés (méta-analyses de : Lee et al., 2015, Sakai & Moorman 2018). Par conséquent, la séquence didactique E devrait favoriser une augmentation des performances en prononciation supérieure à celle induite par la séquence C, mais principalement pour les élèves les plus faibles. La prononciation est une compétence recrutée dans la tâche de répétition de pseudo-mots qui permet de mesurer les compétences de maintien de l'information verbale en mémoire de travail.

La mémoire de travail est entraînée dans les deux séquences didactiques. D'après la méta-analyse de Gathercole et al. (2019), l'entraînement de la mémoire verbale induit une facilitation des processus de traitement dans l'exécution des nouvelles procédures cognitives qui est transférable à des tâches qui les mobilisent. Dans une première méta-analyse, les auteurs évaluent les résultats de 23 études incluant des adultes et des enfants. Les études incluses

visaient à mesurer les effets d'un entraînement de la mémoire de travail sur ses performances. L'effet de l'entraînement est modéré et significatif pour une tâche de répétition de pseudo-mots communs à l'entraînement et au test. Le transfert de la facilité développée par l'entraînement est régulièrement observé quand ce sont les mêmes procédures qui sont sollicitées pour réaliser la nouvelle tâche, mais ce transfert n'est pas observé entre le domaine visuel et le domaine verbal. Dans une deuxième méta-analyse, les auteurs évaluent les résultats de trois études regroupant des enfants de 8 à 11 ans présentant des capacités réduites de mémoire de travail et des troubles du langage. L'effet de l'entraînement est, de la même façon, significatif avec des effets importants pour certains tests en particulier ceux qui sont liés aux nombres. Les enfants présentant des difficultés avec certaines routines cognitives peu fonctionnelles peuvent être entraînés pour en améliorer le fonctionnement. Par conséquent, les séries contenant uniquement des phonèmes communs L1-LE devraient montrer des performances de répétition des pseudo-mots supérieures au post-test en comparaison du pré-test. Cependant il est possible que l'effet positif de l'entraînement ne soit pas observable dans les séries contenant des phonèmes spécifiques de la LE.

Dans mon cadre théorique, la difficulté engendrée par des nouvelles représentations phonétiques, articulatoires et phonologiques peut entraîner un coût cognitif. Le temps de traitement cognitif est augmenté au détriment des processus de rénovation de l'information verbale ce qui entraîne une perte d'informations. Des représentations phonétiques peu efficaces, que ce soit parce qu'elles sont trop détaillées ou non pertinentes, vont augmenter la charge cognitive. Lorsque les participants ne peuvent pas réaliser de *chunks* sémantiques comme dans la tâche de répétition de pseudo-mots de la présente étude, le traitement procédural est centré sur un encodage phonétique détaillé.

Le focus sur le contraste acoustique influence le traitement automatique de la parole. Dans une étude mesurant des différences de potentiels évoqués, Hisagi et al. (2015), observe que l'effet de l'attention dirigée vers un contraste acoustique modifie le traitement automatique de la parole. Cette étude propose une tâche de discrimination de consonnes japonaises à des participants japonais (L1) et américains (LE), et les auteurs mesurent le MMN² dont la présence témoigne d'un événement sonore surprenant. La direction de l'attention des participants était dirigée soit vers une tâche visuelle (et la tâche de discrimination était inconsciente pour les participants), soit vers la tâche de discrimination en indiquant aux participants de compter les stimuli déviants. Les résultats montrent que quand l'attention est centrée sur le contraste acoustique, le MMN est plus important pour tous les participants sans différence significative entre les deux groupes (L1 et LE), bien que les groupes diffèrent sur le pourcentage de bonnes réponses. Par conséquent dans la présente étude, les apprenants du groupe E ayant

² MMN ou *mismatch negativity effect* est un décalage négatif de « l'onde électrique » cérébrale induit par un événement acoustique occasionnel et bizarre qui diffère en fréquence ou en durée. Le MMN est observable dans les 200ms post-stimulus.

consciemment appris les nouvelles représentations LE sont plus susceptibles de focaliser sur l'encodage phonétique que ceux du groupe C. Ils sont donc plus susceptibles de moins bien maintenir les phonèmes communs L1-LE en présence des phonèmes cibles LE que ceux du groupe C.

La représentation orthographique soutient le développement des représentations phonologiques associées dans certaines conditions (Bhide, Ortega-Llebaria, Fraundorf & Perfetti, 2020; Showalter & Hayes-Harb, 2013). Dans une étude avec 69 participants adultes de langue première anglais, Bhide et al. (2020) ont observé l'acquisition de 3 phonèmes spécifiques du Marathi présentant des différences contrastives de niveaux de difficulté variables (i.e., difficile vs. Facile, voir 1.6). Les participants ont appris chacun des 3 phonèmes dans une des trois conditions orthographiques : pas de support orthographique, écriture selon les règles phonème-graphème de l'anglais, et en écriture Marathi. La condition orthographique associée à un phonème a été aléatoirement répartie entre les participants. Après un pré-test, les participants suivent une phase d'entraînement suivie par un post-test immédiat et un post-test différé le lendemain. Les performances des participants sont mesurées en prononciation, en discrimination et en détection d'un lexème mal-prononcé. Les résultats montrent que le support orthographique en anglais limite la perception du contraste phonémique comparativement à la condition phonologique (sans support orthographique) ou avec un support en Marathi (autre système d'écriture) ; mais qu'au niveau lexical, la détection de la mauvaise prononciation d'un lexème est favorisée par les deux supports orthographiques. Ces résultats ambivalents plaident pour un travail ciblé sur l'interface orthographie-phonologie en particulier au niveau des correspondances phonème-graphème.

La maîtrise des correspondances phonème-graphème est contrainte par les interférences cross-linguistiques : 1) au niveau des correspondances phonème-graphème : par exemple pour un même graphème, un phonème différent est attribué en L1 et en LE, et inversement pour un même phonème, des graphèmes différents le représentent en LE et en L1 ; et 2) au niveau des règles phonotactiques et graphotactiques de la LE : par exemple une suite de lettres, comme une fin de mot en <ag>, est acceptable en allemand et pas en français. En français les fricatives voisées finale comme dans *lavage* = /lavaʒ/ ne sont jamais transformées par un phénomène de sandhi en « *lavache* » = /lavaʃ/ comme c'est le cas en suisse allemand (voir 1.6.2). De plus, la complexité de la relation phonème-graphème (e.g., deux graphèmes <j-g> pour le phonème /ʒ/), la fiabilité des règles LE qui les régissent (nombre d'exceptions comme *nom* pour /ɔ̃/-on) et les limites de leur validité (absence de règles pour expliquer la différence de correspondances entre *gens* /ʒɑ̃/ et *jeux* /ʒø/) s'ajoutent aux interférences cross-linguistiques et rendent encore plus difficile l'acquisition des correspondances phonème-graphème en LE.

Les tâches de mesure des correspondances phonème-graphème et de maintien en mémoire de travail sont dépendantes des performances en discrimination et en prononciation. Par conséquent, ces deux tâches devraient montrer les mêmes sensibilités au type de séquence

didactique (i.e. variation en fonction du niveau de l'apprenant et avantage pour la séquence E), mais elles sont aussi sensibles aux interférences cross-linguistiques.

La séquence didactique explicite (E) devrait être plus efficace pour aider les élèves à gérer les interférences avec la L1. Les interférences entre les langues bloquent l'apprentissage implicite (Ellis, N., 2015). Les interférences avec la L1 sont générées par une compétition entre les représentations L1 et LE dont le résultat est en général une sélection des représentations L1 (e.g., Bassetti, Mairano, Masterson & Cerni, 2020; Ellis, N., 2006a). Pour permettre à l'apprenant de maîtriser les interférences, l'enseignement doit 1) rediriger l'attention sur les caractéristiques spécifiques de la LE ou/et 2) favoriser la prise en compte consciente par l'apprenant des caractéristiques LE (e.g., Ellis, N., 2015). Attirer l'attention de l'apprenant peut être réalisé par un enseignement qui rend la forme spécifique de la LE saillante, ce qui est réalisé dans les deux séquences didactiques de cette étude, et par la promotion d'une prise de conscience par l'apprenant d'une caractéristique LE, ce qui est réalisé seulement dans la séquence didactique E.

La prise de conscience par l'apprenant de la difficulté spécifique LE est induite par un enseignement explicite. Dans la séquence didactique E, l'enseignement explicite utilise la gestuelle, l'explication articulatoire, la visualisation articulatoire, et explique l'opposition des paires minimales et les différences de correspondances phonème-graphème (Detey et al., 2016). Les représentations phonétiques et phonologiques de la L1, qui induisent les interférences entre les deux systèmes phonologiques, sont nommées, imitées et comparées aux représentations LE. Au contraire, la séquence didactique C ne prend pas en charge explicitement les interférences cross-linguistiques. Par conséquent, les interférences devraient être mieux contrôlées dans la séquence E que dans la séquence C. Les élèves bénéficiant de la séquence didactique E devraient plus progresser que ceux de la séquence didactique C pour les tâches de prononciation, de répétition de pseudo-mots et de correspondances phonèmes-graphèmes. L'acquisition du vocabulaire est étroitement liée au développement de ces compétences mais dépend également de l'acquisition de la représentation orthographique du lexème.

Dans son étude doctorale, Rafat (2012) a étudié le transfert par le truchement de la représentation orthographique commune L1-LE (anglais-espagnol) de la représentation phonologique L1 (anglais) à la représentation LE (espagnol). 40 participants adultes anglophones (L1) ont été aléatoirement répartis dans 4 conditions d'enseignement de l'espagnol LE : 3 conditions représentaient différents degrés d'exposition à l'orthographe (2 conditions avec une exposition à la représentation orthographique LE, dont une qui demandait la production de la représentation orthographique LE) et 1 condition à un enseignement LE uniquement phonologique. Les participants ont effectué une tâche de vocabulaire productif en espagnol à partir d'images (i.e., *picture-naming task*) immédiatement après une séance d'enseignement. Chaque stimulus était enseigné par trois présentations de la représentation phonologique globale en espagnol associées à une image représentant sa valeur sémantique. Cet enseignement était effectué dans une des trois conditions préalablement décrites concernant

l'exposition à la représentation orthographique. Trois résultats de cette étude sont intéressants pour ma recherche.

Premièrement, le transfert de la forme phonologique L1 en LE est presque toujours observé dans les conditions orthographiques et exceptionnellement observé dans la condition phonologique, ce qui confirme l'existence d'interférences cross-linguistiques au niveau des correspondances phonèmes-graphèmes. Deuxièmement, la probabilité d'un transfert est d'autant plus haute que la différence acoustique entre les phonèmes L1 et LE est peu saillante, comme pour le phonème /ɔ̃/-on de mon étude qui n'appartient pas au répertoire phonologique des apprenants et est fréquemment assimilé à un des 3 phonèmes L1 /o, u, ɔ/. Dans ce cas, l'enseignement explicite est plus efficace pour surmonter les interférences cross-linguistiques phonologiques (e.g., Ellis, N., 2006a, 2015). Troisièmement, la variabilité de la représentation phonologique associé au graphème en L1 induit un transfert de la forme phonologique L1 la plus fréquente au graphème en LE, comme pour le phonème /ʒ/-j. En allemand la forme la plus fréquente associée au graphème <g> est [g] et pour le graphème <j> est [j], ce qui induit une interférence orthographique qui peut être responsable du haut taux d'erreur orthographique observé pour ce phonème. De la même façon que pour le phonème /ɔ̃/-on, l'enseignement explicite devrait s'avérer être plus efficace pour surmonter les interférences cross-linguistiques orthographiques (e.g., Ellis, N., 2006a, 2015).

L'importance des interférences cross-linguistiques à l'interface orthographie-phonologie peut limiter l'acquisition de nouveaux lexèmes. Hayes-Harb et al. (2010) ont mesuré l'apprentissage de pseudo-mots contenant uniquement des phonèmes anglais par des adultes de L1 anglais. Les 33 participants étaient aléatoirement répartis dans 3 groupes représentant 3 conditions d'apprentissage : input phonologique et input orthographique consistant selon les règles graphème-phonème de l'anglais, input phonologique et input orthographique inconsistant, et input phonologique uniquement. Dans la phase d'apprentissage, les participants ont été oralement exposés 4 fois à chacun des 24 pseudo-mots associés avec une image représentant leur valeur sémantique. Immédiatement après la phase d'apprentissage, les participants ont dû déterminer si le pseudo-mot entendu correspondait à l'image proposée. Les participants qui ont appris les pseudo-mots avec une orthographe inconsistante ont appris significativement moins de pseudo-mots que ceux des deux autres conditions. Aucune différence n'est observée entre les deux groupes de ces deux conditions : apprentissage des pseudo-mots par la forme phonologique uniquement et apprentissage forme phonologique et forme orthographique consistante. Les auteurs concluent que les interférences cross-linguistiques graphème-phonème limitent l'apprentissage du vocabulaire en LE (voir également Bassetti, 2017). Les apprenants ont tendance à accorder peu de confiance au lien entre représentation orthographique et phonologiques.

Les apprenants les moins avancés préfèrent apprendre la représentation orthographique que les représentations phonologiques de la LE. Dans une étude avec 63 participants adultes de langue première finlandais, Veivo, Järvikivi, Porretta et Hyönä (2016) ont observé quelle forme écrite

était choisie par des apprenants de différents niveaux de FLE en suivant les mouvements de leurs yeux lors d'une tâche de reconnaissance lexicale (= *eye-Tracking*) et en mesurant le temps nécessaire pour la formulation d'une réponse. La cible a été présentée visuellement parmi 4 propositions : la bonne réponse, son compétiteur et deux distracteurs. Le compétiteur était soit un compétiteur orthographique (i.e., deux premières lettres communes mais bien qu'écrit avec une lettre similaire le nucleus vocalique correspondait à deux phonèmes différents : *cintr-cidre*), soit un compétiteur phonologique (i.e., deux premiers phonèmes communs et le même nucleus vocalique mais deux écritures différentes : *cidre-cycle*). Les quatre propositions écrites ont été présentées avant la présentation orale du lexème, puis le participant a dû cliquer sur ce qu'il pensait être la forme orthographique correcte du mot entendu. Les résultats confirment que les apprenants les moins avancés mettent plus de temps pour retrouver la forme orthographique correcte que les plus avancés, et réussissent moins bien quand le compétiteur est orthographique.

Le compétiteur orthographique active les deux systèmes de correspondances phonème-graphème et favorise les interférences phonologiques. La charge cognitive est ainsi augmentée en présence du compétiteur orthographique ce qui limite les performances s'appuyant sur ces représentations. Ces résultats confirment les résultats d'études ayant montré que la charge cognitive en LE induit une valorisation accrue des représentations orthographiques au détriment des représentations phonétiques et phonologiques pour les apprenants les moins avancés (Gunnarsson-Largy, Dherbey & Largy, 2019; Mattys & Wiget, 2011).

Le compétiteur phonologique ralentit moins le traitement cognitif et entraîne moins d'erreurs que le compétiteur orthographique. Les auteurs concluent que le compétiteur phonologique est peu pris en compte par les apprenants, et que les apprenants les moins avancés utilisent préférentiellement la représentation orthographique pour identifier un lexème LE.

La forme orthographique doit être liée à la forme phonologique pour que la valeur sémantique soit apprise. Les représentations orthographiques sont les premières apprises (Webb, 2007), mais les études mesurant l'acquisition du lien valeur sémantique-forme orthographique en LE/L2 (i.e., pendant la lecture) montrent que très peu de vocabulaire peut être appris de manière incidente (e.g., Elgort, Brysbaert, Stevens & Assche, 2018; Nation, 2006). Dans une étude avec 50 étudiants d'une université américaine, Chen (2021) compare l'acquisition incidente du vocabulaire en anglais L2 entre deux conditions. Une moitié des participants a écouté le texte en même temps qu'ils le lisaient et l'autre moitié l'a lu en disposant du même temps d'exposition. Les mots cibles étaient des pseudo-mots inclus dans quatre textes narratifs. Les mesures d'acquisition du vocabulaire ont été effectuées dans trois tâches : une tâche de reconnaissance des formes orthographiques parmi des distracteurs, une tâche d'association forme orthographique-valeur sémantique, et une tâche de décision lexicale avec l'effet d'une prime orthographique. Les participants ayant entendu les textes pendant qu'ils les lisaient ont

acquis significativement plus de formes orthographiques et ont réalisé plus d'associations forme orthographique et valeur sémantique, ce qui montre l'importance de la médiation réalisée par les formes phonologiques entre représentations orthographiques et sémantique.

L'enseignement des formes phonologiques est indispensable pour l'apprentissage du vocabulaire en LE car toutes les représentations phonétiques et phonologiques liées interagissent lors du traitement cognitif. Lorsque les représentations phonétiques et phonologiques LE sont imprécises, en particulier lorsqu'elles ne permettent pas d'identifier les spécificités de la LE, l'encodage lexical des caractéristiques phonologiques et phonétiques est imprécis (Cook, Pandža, Lancaster & Gor, 2016). L'accès lexical en LE est alors plus lent et présente un risque d'erreur supérieur (Cook & Gor, 2015). Les représentations phonologiques doivent permettre d'identifier la représentation lexicale.

Plus le lexème est fréquemment travaillé et plus il est travaillé par des exercices visant l'élaboration de la relation forme-valeur sémantique, plus il a de chances d'être appris (e.g. Hulstijn, 2013 ; Nation, 2001). L'apprentissage du vocabulaire en LE semble dépendre du nombre de fois où le lexème est utilisé, mais aussi des modalités de traitement du lexème. Dans une étude sur l'acquisition incidentale du vocabulaire LE lors de la lecture de textes, Webb (2007) a montré que plus le nombre de répétitions augmente, mieux le lexème est écrit et utilisé. Les 121 participants japonais de son étude sont des apprenants adultes d'anglais d'un niveau intermédiaire. Dix pseudo-mots anglais sont présentés dans des textes à lire. Les participants sont répartis dans cinq groupes qui rencontrent chaque pseudo-mot 0, 1, 3, 7 ou 10 fois. Leurs performances, mesurées au pré-test et immédiatement après la lecture, visent à évaluer l'acquisition des représentations orthographiques, syntaxiques et sémantiques au niveau productif et réceptif. Les résultats montrent que la représentation orthographique est la première apprise et que plus la forme est répétée mieux elle est apprise.

Pour certains auteurs, l'acquisition du vocabulaire, c'est-à-dire de toutes les représentations composant la représentation du lexème (représentations formelles : phonétiques, phonologiques, orthographiques, conceptuelles ; représentations pragmatiques, stylistiques...), est considérée comme majoritairement implicite en L1 et explicite en LE (e.g., Hulstijn, 2013). La valeur sémantique est l'élément central de la représentation lexicale lors d'un enseignement communicatif. Dans la présente étude, la valeur sémantique est explicitement et uniquement travaillée dans la séquence C. L'usage du sens devrait conduire à une augmentation plus importante du vocabulaire réceptif pour la séquence C comparativement à la séquence E, sauf si les représentations phonologiques sont primordiales pour l'acquisition du vocabulaire.

Pour certains auteurs, l'apprentissage du vocabulaire en LE serait un processus implicite induit par la compréhension lors d'échanges verbaux oraux ou lors de la lecture (Krashen, 1982, 2003 ; pour une revue Ponniah, 2011). Ces études menées avec des participants étudiants semblent montrer un apprentissage supérieur et plus durable du vocabulaire par la lecture que par un enseignement ciblé sur l'association forme orthographique-valeur sémantique. Dans ces

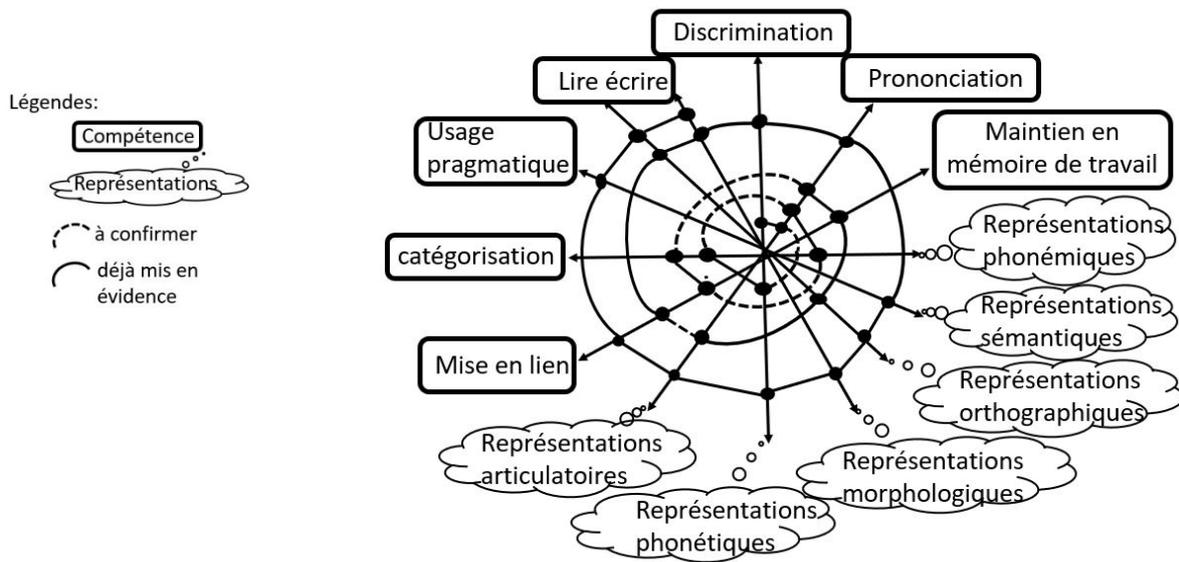
études, les compétences requises (i.e., échanges communicatifs ou lecture) développent le lien entre les représentations phonologiques et phonétiques et les représentations sémantiques. Dans la présente étude, seule la séquence didactique C propose des activités de lecture et d'échanges verbaux basés sur le vocabulaire mesuré par la tâche de vocabulaire réceptif. Par conséquent, dans cet autre cadre théorique, la séquence C devrait conduire à une augmentation supérieure du vocabulaire réceptif en comparaison de la séquence E. Les deux séquences C et E proposent une progression didactique structurée de façon similaire.

1.5.2.2 Développement des compétences dans la progression didactique

De nombreux auteurs considèrent les représentations phonétiques et phonologiques de la LE, et en particulier la compétence de perception, comme l'élément indispensable au développement des autres compétences de la LE (e.g, Flege, 2016; Saito, 2021). Pour les représentations phonologiques, l'effet du *feed-back* est dépendant de la capacité de l'apprenant à percevoir consciemment la différence entre sa production et le son ciblé (Saito, 2021). Pour être conscient de la différence entre l'objectif et sa production, l'apprenant doit pouvoir comparer sa production à des représentations existantes en mémoire à long terme. Le *feed-back* correctif pour la prononciation est plus efficace si l'apprenant a déjà acquis un minimum de représentations phonétiques de la LE (Saito, 2021). Par conséquent dans la présente étude, la discrimination, tâche permettant de mesurer et de travailler les représentations phonétiques, a été enseignée préalablement à la prononciation (Detey, 2016).

Les différents niveaux de représentations phonétiques et phonologiques (définis par leur taille en ordre croissant de la représentation phonétique à la représentation phonologique globale du groupe prosodique) sont étroitement liés et leur apprentissage est favorisé par un parcours progressif et redondant au sein de la séquence didactique (Detey et al., 2016). En effet, l'efficacité de l'utilisation ou de l'apprentissage d'un niveau de représentation dépend en partie de la maîtrise d'un autre niveau. Cependant l'acquisition des différents niveaux ne suit pas un ordre linéaire selon lequel une maîtrise totale des représentation phonétiques par exemple serait un préalable obligatoire à l'acquisition phonémique. L'acquisition des différents niveaux de représentations semble plutôt suivre un parcours spiralaire comme celle de l'acquisition conceptuelle (Barth, 2013). Les représentations phonétiques constituent un support pour l'acquisition de la représentation phonémique, et l'acquisition progressive de la catégorie phonémique affine en retour celles des représentations phonétiques. Le parcours d'apprentissage doit ainsi accompagner des apprentissages interdépendants en favorisant une capitalisation des acquis au sein d'une démarche holistique de mise en réseau des connaissances au fil de leur acquisition (Bosse & Pacton, 2006; Bunting & Engle, 2015). Une même représentation est travaillée plusieurs fois au fil de l'augmentation de la complexité de la tâche qui la met en pratique. La complexification progressive de la tâche exige la mise en œuvre d'un nombre croissant de compétences qui demandent de mettre en lien un nombre croissant de représentations phonétiques, phonologiques et lexicales associées. La mise en œuvre didactique a suivi une complexification progressive des tâches (figure 2).

Figure 2 : Schéma de la progression didactique en fonction des compétences et des représentations



La progression didactique propose le travail des éléments simples avant d'aborder les tâches complexes. Cette progression du simple vers le complexe a été respectée pour toutes les tâches récurrentes d'une séance à l'autre dans la séquence didactique, et également au sein de chaque séance dans la succession des activités et des exercices. Les deux séquences didactiques ont ainsi proposé un parcours d'apprentissage graduel (Vygotsky, 1978). Premièrement dans la séquence explicite (=E), l'oral a d'abord été travaillé seul avant d'introduire la représentation écrite. La séquence communicative (=C) ne pouvait pas proposer à des débutants des activités communicatives sans support écrit. Deuxièmement, les éléments concrets ont été travaillés avant les éléments abstraits. Pour la séquence E, le phonème a d'abord été prononcé, entendu et manipulé avant de devoir être identifié dans la parole. Pour la séquence C, l'objet concret correspondant au lexème qui contenait le phonème cible a été manipulé physiquement et le lexème a été à cette occasion utilisé verbalement dans des tâches communicatives avant qu'il soit demandé à l'apprenant de l'écrire ou de le lire. Troisièmement, les tâches étaient récurrentes d'une séance à l'autre et communes aux deux phonèmes cibles afin de limiter la charge cognitive induite par une nouvelle tâche. Elles ont été complexifiées au fil des séances. Quatrièmement, les activités d'appropriation et de réinvestissement ainsi que les activités métalinguistiques ont été introduites en fin de séquence. En effet ces activités impliquent des élaborations successives de représentations multimodales et une complexité croissante des représentations à utiliser. Elles doivent donc être introduites progressivement en fonction du stade d'apprentissage (Piaget, 1972). Les activités pédagogiques répondent aux différents objectifs didactiques visant l'acquisition et la manipulation des différentes représentations et des compétences qui les utilisent.

Le développement des compétences de discrimination, prononciation, maintien en mémoire de travail, correspondances phonème-graphème et acquisition du vocabulaire dépend aussi en partie de l'objet d'enseignement d'apprentissage.

1.5.3 Incidence des caractéristiques de l'objet d'enseignement/apprentissage sur l'effet du type d'enseignement

Deux qualités de l'objet d'enseignement/apprentissage semblent influencer l'effet de la didactique : sa complexité et sa difficulté. Ces deux qualités ont été définies dans le cadre théorique *Complexity-Accuracy-Fluency* pour l'enseignement de la langue seconde (= CAF, Housen & Kuiken, 2009). La complexité est une qualité inhérente à l'objet d'enseignement/apprentissage (Housen, De Clercq, Kuiken & Vedder, 2019) qui est définie par le nombre d'éléments composant l'objet enseigné, par la variété de ses composants, et par les liens qui les unissent (Housen, Kuiken & Vedder, 2012; Housen & Simoens, 2016; Norris & Ortega, 2009). La difficulté est induite par l'objet d'enseignement/apprentissage lors de son traitement cognitif par un apprenant dans un contexte d'apprentissage (i.e., aussi parfois appelée « relative complexité », Housen et al., 2019). La difficulté relative à l'apprenant (voir 1.5.1) et celle relative à la tâche (voir 1.5.2) ont été traitées précédemment et la notion de difficulté est limitée ici à la partie induite par l'objet d'enseignement/apprentissage. La définition précise des deux valeurs, difficulté et complexité, n'a pas encore été clairement établie en phonologie.

Selon Ellis (N., 2006a), la complexité d'un phonème peut être définie par trois critères : 1) par le nombre de traits articulatoires nécessaires pour le produire, 2) par l'absence de voyelle associée à sa forme de surface (= *syllabicity*), et 3) par une faible saillance (= *relative sonority*). La complexité du phonème influence la plus ou moins grande difficulté rencontrée par le système procédural, mais elle ne peut pas expliquer la difficulté de traitement induite par les interférences cross-linguistiques.

Selon le modèle SLM de Flege (e.g., Flege et al., 2003; McKay, Flege, Piske, & Schirru, 2001), la difficulté est induite par les relations entre les deux grammaires phonologiques et par l'espace disponible dans l'espace général perceptuel de l'apprenant. D'après le SLM, la difficulté d'acquérir une représentation phonémique LE dépend de la probabilité que les exemplaires associés à cette représentation LE puissent être attribués à une représentation L1. La difficulté augmente si deux représentations LE ont une probabilité importante d'être toutes les deux assimilées à la même représentation L1. D'autres auteurs définissent la difficulté dans un cadre théorique communicatif.

Munro et Derwing (2006) évaluent la pertinence de l'enseignement de la prononciation d'un phonème pour déterminer s'il est difficile. Les phonèmes difficiles sont ceux qui ont un important *functional load*. Ce critère détermine quels phonèmes doivent être enseignés. Par exemple en FLE, les voyelles nasales ou les fricatives ont un important *functional load*. Un phonème avec un important *functional load* permet de distinguer un grand nombre de mots

incluant la même paire minimale phonémique (e.g., le phonème *je-/ʒ/* permet de séparer les mots *cage* (=/*kaz/*) et *cache* (=/*kaf/*) et sa prononciation si elle est inadéquate peut entraver la communication (e.g., *mâche* = /*maf/*, *lamb's lettuce salad* ≠ *mage* = /*maz/*, *wizard*) (Derwing & Munro, 2005, 2009; Gatbonton, Trofimovich & Segalowitz, 2011; Isaacs, 2018). L'apprenant qui ne maîtrise pas un phonème avec un fort *functional load* éprouve des difficultés d'apprentissage de la LE et des difficultés de communication en LE. Cependant, le *functional load* n'est pas un critère qui permet à lui-seul de définir la difficulté d'un phonème.

Les objets d'enseignement/apprentissage phonologiques les moins facilement acquis ou les plus difficiles à maîtriser peuvent être identifiés à priori comme étant ceux qui sont le plus fréquemment fossilisés chez les apprenants d'une LE (Abel, 2018). En FLE, les voyelles nasales et les fricatives sont des phonèmes souvent mal perçus et mal prononcés même par des apprenants avancés et sont considérés comme fortement à risque d'être fossilisés. La voyelle nasale /*ɔ̃/-on* et la fricative /*ʒ/-j* sont deux phonèmes complexes, difficiles, facilement fossilisables et avec un important *fonctional load*. Ils constituent les deux objets d'enseignement/apprentissage de cette étude.

L'enseignement explicite semble plus efficace qu'un enseignement implicite pour améliorer la prononciation des phonèmes « difficilement compréhensibles » et fossilisés (Derwing & Munro, 2009; Flege, 2003; Isaacs, 2014; Saito, 2012, 2021). La séquence didactique explicite (=E) devrait donc favoriser un progrès plus grand que la didactique implicite (=C) pour les phonèmes cibles de cette étude.

1.6 Complexité et difficulté des phonèmes cibles

1.6.1. La voyelle nasale /*ɔ̃/* et ses graphèmes <on ; om>

La voyelle nasale /*ɔ̃/-on* est absente du répertoire phonologique d'un apprenant germanophone. Quelques mots d'emprunts au français ont gardé une nasalisation de la voyelle mais cette nasalisation est inconstante et ne témoigne pas d'un contraste phonémique (Laeufer, 2010). La voyelle nasale /*ɔ̃/-on* est un phonème qui appartient uniquement au répertoire phonologique de la LE pour un apprenant germanophone.

Chaque locuteur dispose d'un répertoire phonologique construit au cours de son expérience linguistique (Paradis, 2008). Ainsi les apprenants débutants en FLE ne disposent pas des représentations phonologiques et phonétiques correspondant aux phonèmes qui n'existent pas dans leur L1. La manifestation de cette absence de représentation est que l'apprenant ne distingue pas un phonème LE non inscrit dans son répertoire phonologique d'un phonème de sa L1. Ce phénomène, appelé surdité phonologique, a pour conséquence le non-encodage d'un trait acoustique (ici la nasalité ou la place articulaire). Le trait acoustique peut aussi être encodé en boucle phonologique, mais le recodage l'exclut des représentations multimodales

construites, ou les effets top-down des représentations L1 présentes en mémoire à long terme induisent sa catégorisation comme une variation atypique d'une représentation L1. Le phonème LE est alors assimilé à la catégorie phonémique L1 la plus proche (Best & Tyler, 2007; Dupoux, 2010; Escudero, 2009; Flege, 1995; Peperkamp et al., 2002; van Leussen & Escudero, 2015). Dans le cadre de cette étude, la voyelle nasale /ɔ̃/ est susceptible d'être assimilée à la voyelle orale /O/ par les apprenants. /O/ est considérée comme un archiphonème en français qui englobe les différentes variations de production des deux phonèmes /o ; ɔ/ observées dans les différentes régions francophones (Catach, 1995; Pérez, 2014). Le travail pédagogique vise à travailler la discrimination entre les phonèmes de la paire minimale /ɔ̃ ; O/ et aussi dans une moindre mesure entre les phonèmes de la paire minimale /ɔ̃ ; u/ qui représentent une difficulté pour quelques apprenants.

La discrimination entre voyelle orale et voyelle nasale est sensorielle et non décisionnelle pour tous les auditeurs (Delvaux, 2009), qu'ils soient locuteurs natifs ou apprenants. Dans mon cadre théorique, une différenciation sensorielle repose sur une analyse procédurale des représentations phonétiques par le système de comparaison-contraste du modèle PRIMIR. Une différenciation décisionnelle reposerait sur les traces en mémoire à long terme d'une représentation multimodale (i.e., effet *top-down*) et serait consciente. Delvaux (2009) propose des continuums de voyelle orale à voyelle nasale synthétiques qui varient selon les différents axes de formants. Des auditeurs francophones et anglophones sont soumis à une tâche de jugement basée sur un choix binaire entre deux mots d'une paire minimale orale / nasale. Les deux groupes basent leurs jugements sur une combinaison des variations de formants bien avant toute activation lexicale, ce qui amène l'auteur à conclure que la perception de la différence oral versus nasal est d'ordre sensoriel. Le travail didactique visant à améliorer la compétence de discrimination devra donc s'appuyer sur les habiletés sensorielles. La compétence de discrimination entre voyelle orale et nasale est vraisemblablement aussi liée à la capacité de les prononcer comme deux phonèmes différents (Kartushina & Frauenfelder, 2014).

Discrimination et prononciation sont liées (e.g., Bion, Escudero, Rauber, & Baptista, 2006; Sakai & Moorman, 2018) bien que cette relation ne signifie pas que ces deux habiletés utilisent de façon similaire toutes les représentations phonétiques et phonologiques (Flege, 1999; Flege, Bohn, & Jang, 1997; Huensch & Tremblay, 2015). Pour un apprenant de FLE, l'acquisition de la capacité à discriminer les voyelles nasales précède la capacité à les produire (Detey & Racine, 2015). Dans la progression didactique, la discrimination doit donc être travaillée préalablement à prononciation. Une des difficultés d'apprentissage des voyelles nasales réside en leur grande variabilité.

Les apprenants ont des difficultés à prononcer les voyelles nasales en particulier à cause de l'importante variabilité des prononciations des locuteurs francophones. Les apprenants doivent identifier et reproduire un phonème qui est lié à une grande variété de représentations phonétiques. Cette caractéristique d'instabilité de leurs formes phonétiques est un des deux éléments pouvant expliquer le risque important de fossilisation de ce phonème dans

l'interlangue des apprenants (Han, 2013). Cette difficulté générée par la variabilité des représentations phonétiques dans l'input est accentuée par la nécessaire acquisition d'une nouvelle représentation articulatoire susceptible de permettre de les prononcer.

Les voyelles nasales ne correspondent pas à des voyelles orales qu'il suffirait de modifier légèrement (Amelot, 2004). Des modifications articulatoires majeures sont nécessaires pour transformer la voyelle orale /ɔ/ en /ɔ̃/ (Maeda, 1993; Shosted, Carignan & Rong, 2012). Ces modifications sont essentiellement constituées par le retrait de langue en position postérieure et par la modification de la position des lèvres (arrondies pour /ɔ̃/). Leur production requière donc l'acquisition de mouvements articulatoires spécifiques auxquels s'ajoute l'abaissement du voile du palais produisant la nasalisation.

L'abaissement du voile du palais ne permet pas à lui seul de produire une voyelle nasale (Havy, Bouchon, & Nazzi, 2016) mais il participe à sa production. Cette coarticulation vélaire est spécifique à chaque langue (Clumeck, 1976, cité par Amelot, 2004, p16). Par exemple, lors de la nasalisation de /a/ en anglais (voyelle nasalisée par assimilation régressive, comme par exemple /bānk/-*bank*), le seul changement articulatoire est l'abaissement du voile du palais (Shosted, Carignan & Rong, 2012). D'ailleurs, lors d'emprunts de mots français en anglais canadien, les voyelles nasales sont fréquemment émises sous la séquence de deux phonèmes consécutifs 'voyelle orale + consonne nasale' : par exemple coupon - /kupɔ̃n/ (Paradis & Prunet, 2000). Ce phénomène est lié aux représentations abstraites phonémiques. En effet les voyelles nasalisées n'ont pas de représentation différente de celle de la voyelle orale correspondante dans les langues où les voyelles nasales ne sont pas distinctives (Billières, 2005; Dehaene-Lambertz, Dupoux & Gout, 2000; Pallier et al., 2001; Petersen, 2015). Au moment de leur perception, les voyelles nasales du français sont recodées comme la suite ordonnée des deux phonèmes L1 dont les représentations phonétiques sont les plus ressemblantes (i.e., voyelle+<n>).

L'apprenant de FLE doit construire des représentations phonétiques et phonologiques spécifiques de la LE. Il doit modifier sa mémoire articulatoire construite sur les représentations phonologiques de sa L1, et la compléter au fil de l'apprentissage notamment par des représentations articulatoires spécifiques de la LE (Beddor, Krakow, & Goldstein, 1986; Jusczyk, 1999). La prononciation des voyelles nasales du français nécessite un apprentissage afin de construire la représentation des mouvements articulatoires qui leur correspondent (Beddor et al., 1986). La nouvelle représentation, associant représentations phonétiques et articulatoires, impose l'introduction d'un nouveau trait phonologique contrastif (la nasalité) dans la grammaire de l'apprenant germanophone (Clements, 2003) en lien avec un nouvel assemblage d'ordres moteurs articulatoires. Ce trait contrastif LE est utilisé par le système procédural en réception mais également en production.

La nouvelle représentation articulatoire permet de produire les caractéristiques contrastives (i.e., la nasalité et les modifications articulatoires) mais son acquisition est progressive. Dans un premier temps, le trait phonétique nasal n'est en général pas pris en considération

(impossibilité de discriminer une voyelle orale et une nasale), puis il est encodé grâce à un processus attentionnel dont le coût cognitif diminuera au fil de l'apprentissage (Darcy, Dekydtspotter, Sprouse, Glover..., & Scott, 2012; Masoura & Gathercole, 1999, 2005; Pallier et al., 2001; Strange, 2011). De même pour la prononciation, l'utilisation de la nouvelle représentation articulatoire demande tout d'abord de l'attention, et ce coût cognitif diminuera avec une automatisation progressive des mouvements articulatoires. La manipulation d'une nouvelle représentation phonologique peut aussi dépendre des représentations orthographiques qui lui sont associées (Bassetti et al., 2020).

Le phonème /ɔ̃/ peut être écrit par deux graphèmes : <on> et <om>, mais cette difficulté est limitée car des règles explicites peuvent être établies. Dans le sens de l'oral vers l'écrit, la consistance est de l'ordre de 100% (Sprenger-Charolles, Siegel, & Bonnet, 1998). Lorsque le phonème /ɔ̃/ est entendu, il doit être écrit <on> et l'écriture par <om> suit une règle qui ne souffre aucune exception (i.e., devant les lettres <p, b, m> le graphème complexe [on] s'écrit [om]³). Dans le sens de la lecture, cette relation est consistante à 93% (Peereman, Lété, & Sprenger-Charolles, 2007). Lorsque le graphème <on> doit être lu, s'il est placé avant une consonne qui n'est pas la lettre *n*, ou placé à la fin d'un mot, il doit être prononcé /ɔ̃/. Il est donc possible d'établir des règles contextuelles simples et robustes, et de les enseigner de manière explicite. Cependant en français, en plus des règles contextuelles, il est nécessaire de connaître et d'identifier les lettres muettes lors de la lecture.

Le digramme <on> est fréquemment associé à une lettre muette (e.g., *pont, font, marchons, gond*). Les lettres muettes peuvent gêner l'acquisition de la représentation orthographique car elles compliquent l'apprentissage lorsqu'elles sont fréquemment associées à une relation graphème-phonème (Rentzpopoulos & Kokkinakis, 1992, 1996). Dans la méthode d'enseignement des élèves de la présente étude, le /ɔ̃/ est majoritairement en position finale (75%), position dans laquelle il est souvent suivi par une lettre muette (93% des cas). En effet, la voyelle /ɔ̃/ est fréquemment incluse dans la marque flexionnelle des verbes. Au présent de l'indicatif, il peut s'agir de la première personne du pluriel (e.g., *nous marchons*), ou de la troisième personne du pluriel (ex : *ils font*). Dans cet usage, la voyelle /ɔ̃/ est toujours suivie par une lettre muette soit le <t> soit le <s>, et plus rarement le <d> (e.g., *la poule pond*). La voyelle /ɔ̃/ est aussi présente dans les substantifs et les adjectifs (ex : *bon*) et peut dans ce cas porter la marque silencieuse du pluriel <s> quand elle est en position finale (e.g., *maisons*). Au contraire, la relation graphème phonème est peu compliquée par les règles phonotactiques du français.

Peu de restrictions de distribution du phonème /ɔ̃/ sont observées. Il occupe moins fréquemment l'attaque (e.g., *oncle*) que le noyau de la syllabe (e.g., *longue*) et il n'est jamais suivi d'une consonne sonore appartenant à la même syllabe (i.e., de type consonne-voyelle-consonne).

³ Seuls quelques mots échappent à la régularité de l'association du phonème /ɔ̃/ et du graphème [on], Ces mots (= *nom, prénom*) sont très fréquents et ne posent aucun problème d'apprentissage.

Dans l'enseignement et les tests, il sera donc placé dans la structure syllabique la plus fréquente en français (i.e., =CV) qui est également très fréquente en allemand. Toutes les positions syllabiques sont occupées par le phonème /ʒ/ (e.g., *Longueur* ; *rallonger* ; *bouchon*).

Les difficultés relatives à l'acquisition du phonème /ʒ/ ont été intégrées dans le choix des lexèmes supportant l'enseignement et le choix des stimuli inclus dans les tests.

1.6.2 Le phonème consonantique /ʒ/ et ses graphèmes <j, g>

Le phonème /ʒ/-j est plus ou moins présent dans le répertoire phonologique d'un apprenant germanophone. Le caractère voisé, qui le sépare de /ʃ/-che en français (Dell, 2004), est inconstant en allemand et quasiment absent en suisse allemand. En effet les lexèmes qui comptent [ʒ] parmi ses composants phonétiques sont des mots empruntés au français dont la prononciation est plus ou moins éloignée de la prononciation française en fonction des régions germanophones (Laeufer, 2010). Le contraste phonémique non voisé-voisé n'existe ni en allemand ni en Suisse allemand (Schmid, 2009; Schmid, 2012) et l'opposition des deux fricatives est une opposition lenis-fortis (Fuchs, Brunner & Busler, 2007; Laeufer, 2010).

En allemand standard et suisse, le contraste lenis-fortis repose sur une durée supérieure et sur une intensité supérieure de la friction pour les fortis. Les fortis (/ʃ/) sont toujours non voisées mais les lenis (/ʒ/) sont voisées de façon inconstante (Fleischer & Schmid, 2006). Les apprenants germanophones doivent transformer le trait articulatoire voisé, réalisé superficiellement, en un trait contrastif pour pouvoir discriminer les deux phonèmes français /ʒ-ʃ/. Ils doivent modifier un contraste à priori codé comme une variation allophonique (pour l'allemand standard, Staffeldt, 2010 : 82) en un contraste robuste phonémique.

Les apprenants germanophones de FLE doivent modifier leur grammaire phonologique pour changer l'importance relative des différentes caractéristiques phonétiques des fricatives (i.e., les différents contrastes établis sur les caractéristiques de voisement, de durée, et de force) afin de pouvoir discriminer et prononcer les deux phonèmes de la paire minimale [ʒ ; ʃ] en français. Le trait articulatoire principal séparant /ʒ/ de /ʃ/ en français est le caractère voisé/non voisé (Dell, 2004). En français cette seule différence permet de discriminer et prononcer différemment ces deux phonèmes mais les apprenants germanophones la considèrent comme une variation allophonique d'un même phonème.

Les deux phonèmes français /ʒ-ʃ/ semblent être prononcés comme le phonème suisse allemand /ʒ°. Le phonème français /ʒ/ prononcé par un apprenant suisse allemand est souvent entendu comme /ʃ/. Par exemple, <je vais> - /ʒøve/ est prononcé / ʒ°øfe/ ou / ʃøfe/ même par des apprenants avancés. La prononciation des trois phonèmes /ʒ-ʒ°-ʃ/ est souvent réduite à la forme [ʒ°], ce qui peut être interprété comme le témoignage d'une seule représentation articulatoire pour les trois phonèmes. Le caractère voisé semble être considéré comme non représentatif d'une différence contrastive phonémique. La place d'articulation semble également ne pas pouvoir aider à séparer les catégories phonémiques.

En termes de place d'articulation, le /z°/ suisse allemand semble correspondre au /z/ français (Fleischer & Schmid, 2006; Schmid, 2009) mais il est systématiquement non voisé. Lors de l'apprentissage du FLE, les apprenants doivent apprendre une opposition phonémique entre /z/ et /ʃ/ alors que les deux phonèmes L1 les plus proches (i.e., [z° ; ʃ]) sont susceptibles d'être allophones en suisse allemand, comme ils le sont en Allemand standard (Staffeldt, 2010 : 82). Leur prononciation alterne indépendamment de la représentation lexicale globale. Cette interférence cross-linguistique, l'identification de deux phonèmes LE à des variations allophoniques relatives à une même catégorie L1, est une des plus difficile à dépasser (Best & Tyler, 2007; Escudero, 2005, 2009; Flege, 1995; McKay et al., 2001; Piske et al., 2002; Tyler, Best, Faber, & Levitt, 2014). Cette acquisition est d'autant plus complexe que la variation allophonique en L1 n'est pas robuste (i.e., fiable et régulière) mais au contraire varie en fonction de nombreux critères. L'opposition fortis-lenis /z°-ʃ/ n'indique pas une différence phonémique et les règles phonotactiques du suisse allemand abolissent la robustesse de cette différence (Haas & Hove, 2009; Schmid, 2009).

Les phénomènes de sandhi qui se produisent à la frontière de mot transforment les lenis en fortis en suisse allemand en fonction des contextes consonantiques (e.g., /hæz° zi kse:/ → [hæʃi kse:], 'l'as-tu vu', Fleischer & Schmid, 2006, p.249). Les lenis subissent une neutralisation lorsqu'elles sont précédées dans la phrase par une fortis ou une lenis (Fleischer & Schmid, 2006). La lenis /z°/ suisse allemand est ainsi prononcée /ʃ/ par assimilation progressive ou régressive lorsqu'elle est précédée ou suivie par la lenis /z°/ (Fleischer & Schmid, 2006). La transformation des lenis en fortis est également fréquente lorsqu'elles sont placées en syllabe finale et précèdent une pause.

La tendance linguistique universelle est de dévoiser partiellement ou totalement les lenis placées avant une pause ; règle qui n'est pas respectée en français pour permettre le maintien du contraste phonémique (Bonneau, 2015; Schmid, 2012:631). Cependant en français le caractère voisé de /z/ peut ne pas être réalisé par assimilation régressive lors de la chute du schwa et de sa mise en contact consécutive avec une occlusive ou une fricative (e.g., *je suis parti* - /ʃsq parti/). Mais cette non réalisation du caractère voisé est allophonique, inconstante chez un même locuteur, dépendante du contexte d'élocution et non catégorique (Bürki, Ernestus & Frauenfelder, 2010; Darcy & Kügler, 2007; Niebuhr & Meunier, 2011; Snoeren, Seguí & Hallé, 2008). Elle est une réalisation de surface en français vraisemblablement liée à une facilitation articulaire.

La tendance universelle au dévoisement pré-pausal et les phénomènes de Sandhi sont inscrits dans la grammaire phonologique d'un apprenant suisse allemand et leur application automatique au FLE doit être déconstruite. Les phénomènes de Sandhi et de neutralisation universelle renforcent l'usage d'une catégorie phonémique unique /z°-ʃ/ en suisse allemand et par conséquent l'assimilation à cette catégorie des deux phonèmes français /z-ʃ/. L'inconsistance en suisse allemand lenis/fortis rend encore plus difficile l'acquisition du

contraste phonémique /ʒ-f/ car aucune opposition contrastive « parallèle » n'existe en suisse allemand.

Les restrictions de distribution phonotactique des fortis en suisse allemand compliquent encore les modifications de la grammaire phonologique. Les fortis en suisse allemand ne peuvent pas occuper l'initiale du mot. Les phonèmes /ʃ, s, f/ ne peuvent pas être prononcés comme des fortis à l'initiale du mot et sont dans cette position prononcés comme les lenis /ʒ°, z°, v°/ (Fleischer et Schmid, 2006, p.245). Or cette restriction phonotactique ne s'applique pas au français car le /ʃ/ occupe fréquemment l'initiale du mot (e.g., *chat* /ʃa/, *chalet* /ʃale/...). Il est possible que cette restriction favorise la tendance des apprenants suisses allemands à percevoir le /ʃ/ français comme étant le /ʒ°/ suisse allemand.

Les deux phonèmes de la paire minimale /ʒ-f/ en français sont souvent prononcés comme le phonème /ʒ°/ suisse allemand, l'un pour sa place articulaire similaire et l'autre pour la restriction phonotactique des fortis en début de mot. Les difficultés à entendre et à prononcer le phonème français /ʒ/ peuvent difficilement être compensées par un appui sur la représentation orthographique car le phonème /ʒ/ est complexe à écrire en français.

En français, le phonème /ʒ/ est écrit grâce à deux graphèmes utilisés quasiment aussi fréquemment : /ʒ/- <g> est utilisé dans 51% et /ʒ/-<j> dans 49% des lexèmes qui contiennent le phonème /ʒ/ (Catach et al., 1995). L'usage de l'une ou de l'autre forme graphémique est à la fois régi par un certain nombre de règles et à la fois arbitraire (e.g., /ʒə/ s'écrit *je* mais /ʒəl/ s'écrit *gel*). Cette relation phonème-graphème est donc complexe à enseigner et à apprendre. La relation inverse (graphème vers phonème) est elle aussi complexe.

En lecture le graphème <j> ne pose pas de problème majeur car il est consistant à 100% en français et représente toujours le phonème /ʒ/. Un apprenant suisse allemand doit inhiber l'activation du phonème allemand /j/ lié à ce graphème <j> lorsqu'il lit en français (e.g., en allemand *Jagt* = /jakt/). Par contre, la lecture du graphème <g> est beaucoup plus complexe puisque ce graphème est consistant à seulement 15% en français (Catach et al., 1995) : un graphème [g] a 15 % de chances d'être prononcé [ʒ]. Sa lecture est régie par des règles contextuelles en français qui servent de base pour l'enseignement explicite de cette relation graphème phonème. Le graphème <g> est un des graphèmes les plus fréquents en allemand (Hofmann, Stenneken, Conrad & Jacobs, 2007). De plus, il est totalement consistant et toujours prononcé /g/ lorsqu'il est en position interne au mot (Haas & Hove, 2009). La correspondance <g>-/g/ est donc fortement établie dans la langue de scolarisation des apprenants de cette étude, et est encore renforcée par la fréquente correspondance <g> - /g/ en français (85% ; Catach et al., 1995). L'apprentissage de la relation de la relation <g> - /ʒ/ est donc difficile.

Les difficultés relatives à l'acquisition du phonème /ʒ/ ont été intégrées au choix des lexèmes supportant l'enseignement et au choix des stimuli inclus dans les tests. Les deux objets d'enseignement/apprentissage de cette étude, les phonèmes /ʒ/ et /ʒ/ présentent une complexité différente et induisent des difficultés différentes pour des apprenants suisses allemands.

1.6.3 Complexité et difficulté comparée des deux phonèmes cibles

Ces deux notions, complexité et difficulté, non encore précisément définies pour la représentation phonémique dans la communauté scientifique, permettent une comparaison partielle des deux objets d'enseignement/apprentissage de cette étude.

Le phonème /ð/-on et le phonème /ʒ/-j ont des complexités différentes mais équivalentes en nombre de critères concernés (tableau 2).

Tableau 2 : Complexité : Bilan des modifications de la grammaire phonologique nécessaires pour l'apprentissage des phonèmes /ð/-on et /ʒ/-j du FLE

Critère	Phonème /ð/-on	Complexité	
		Phonème /ʒ/-j	Complexité
Nouveau schéma articulatoire	Oui	+	Non -
Caractéristique contrastive	Nasalisation	+	Voisé +
Assimilation	2 phonèmes LE à 1 phonème L1	+	2 phonèmes LE à 1 phonème L1 +
		-	Variation allophonique L1 +

Le phonème /ʒ/ semble plus difficile pour un apprenant suisse allemand car les difficultés rencontrées en FLE par un apprenant suisse allemand sont plus nombreuses que celles rencontrées pour le phonème /ð/-on (tableau 3).

Tableau 3 : Difficulté : Bilan des difficultés relatives à l'apprentissage des phonèmes /ʃ/-on et /ʒ/-j du FLE

Critère	Phonème /ʃ/-on		Phonème /ʒ/-j	
		Difficulté		Difficulté
Variabilité dans l'input français	Oui	+	Faible	-
Variabilité dans l'input allemand		-	Allophonique avec l'élément de la paire minimale L2+L1	+
Conversion Phonème Graphème	Consistance 100%	-	Consistance 51% [g] 49% [j]	+
	Règles contextuelles	+	Règles contextuelles	+
	Arbitraire	-	e.g., <i>Gel/je</i>	+
Conversion Graphème Phonème	Consistance 93%	-	Consistance [g] 15% [j] 100%	+
	Règles contextuelles	+	Règles contextuelles	+
	Phonotactique			
	Lettres muettes	+	Schwa	+
	Sandhi	-	Sandhi	+
	Variabilité suisse allemand	-	Pas de fortis à l'initiale du mot	+

Le phonème /ʒ/-j présente un plus grand nombre de critères de difficulté que le phonème /ʃ/-on pour les apprenants de cette étude.

Premièrement, selon Ellis (N., 2006a), une différence phonétique entre deux sons LE qui représente aussi une frontière phonémique, comme pour le phonème /ʃ/-on, est plus facile à

percevoir qu'une différence phonétique qui est incluse dans une variation allophonique normale d'un phonème L1 comme c'est le cas pour le phonème /ʒ/-j. La variation allophonique en L1 limite la capacité de noter les particularités des phonèmes LE.

Deuxièmement, au niveau articulatoire, la prononciation du phonème /ʒ/-j est difficile à acquérir car, comme toutes les fricatives voisées, ce phonème est parmi les derniers phonèmes acquis en L1 et ce dans toutes les langues qui contiennent des fricatives voisées dans leur répertoire phonologique (Jarosz, Calamaro & Zentz, 2017; McLeod & Crowe, 2018 ; pour l'anglais : Stoel-Gammon, 2011 ; pour le français québécois : McLeod, Sutton, Trudeau & Thordardottir, 2011). D'ailleurs, les enfants francophones monolingues l'apprennent plus tardivement que les autres phonèmes du français et le phonème /ʒ/-j n'est toujours pas maîtrisé par 90% des enfants francophones de plus de 4 ans (McLeod & Crowe, 2018).

Troisièmement, les relations graphème-phonème sont plus complexes. Au final, je considère le phonème /ʒ/-j plus difficile que le phonème /ʃ/-on.

Plus un phonème est difficile, plus le traitement cognitif prend du temps car l'attention est mobilisée plus longtemps. L'attention peut être accaparée soit par le système procédural, soit par les processus de rénovation.

Le système procédural peut mobiliser l'attention par exemple lorsqu'un conflit doit être résolu entre deux représentations en compétition pour être associées à la représentation multimodale. La séquence didactique explicite en réduisant les interférences cross-linguistiques devrait être plus efficace pour l'apprentissage des représentations liées au phonème /ʒ/-j, au cœur de nombreuses interférences cross-linguistiques, en comparaison du phonème /ʃ/-on. Cependant, la captation de l'attention par le système procédural limite la rénovation de l'information ce qui entraîne une perte d'information.

La rénovation de l'information par le rafraîchissement attentionnel peut accaparer le focus de l'attention et dans ce cas c'est au détriment de l'élaboration des représentations multimodales par le système procédural (Barrouillet & Camos, 2015, p.110). Lorsque la boucle exécutive s'est appuyée sur des états mentaux en mémoire à long terme peu congruents, ou peu activés, ou imprécis, le rafraîchissement attentionnel a conduit à la mémorisation de représentations multimodales peu liées aux réseaux de représentations existantes (Cowan, 2015). Le rafraîchissement attentionnel prend alors plus de temps pour la rénovation des représentations multimodales dans le buffer épisodique.

En boucle phonologique, la rénovation subvocale peut aussi prendre plus de temps en LE soit à cause de procédures peu efficaces, soit à cause d'un encodage très détaillé en boucle phonologique.

Premièrement, tout comme certaines étapes de la production de la parole, la rénovation subvocale peut demander plus d'attention car les procédures sont moins efficaces en LE. La prononciation et la rénovation subvocale en boucle phonologique recrutent les mêmes

structures cognitives au niveau de la planification des ordres moteurs. La planification des ordres moteurs produisant l'articulation de la parole n'est pas automatisée en LE (Skehan, 2015). Il est possible que lors du en/recodage en boucle phonologique, les nouvelles représentations articulatoires LE demandent de l'attention pour être planifiées. Le phonème /z/-j, plus difficile au niveau articulatoire, devrait moins progresser que le phonème /ʒ/-on. La séquence didactique explicite devrait être plus efficace pour l'apprentissage des représentations liées au phonème /z/-j car l'approche pédagogique permet de séparer chaque point de difficulté et de limiter la charge cognitive cumulée.

Deuxièmement, l'en/recodage en boucle phonologique de l'information LE peut être très détaillé au niveau des représentations phonétiques (Hisagi et al., 2015; Hisagi & Strange, 2011; Strange, 2011). L'en/recodage détaillé prend plus de temps qu'un en/recodage uniquement constitué de représentations phonétiques efficaces (i.e., permettant une analyse contrastive par le système procédural tout en ne contenant pas d'éléments superflus) (Majerus, Salmon & Attout, 2013). Un phonème serait ainsi encodé avec plus de détails lorsque l'apprenant débutant a construit consciemment des traces en mémoire à long terme car il inclurait alors un maximum de détails dans les représentations phonétiques auxquelles ce phonème est lié. L'enseignement par la méthode explicite devrait conduire à des représentations plus précises mais aussi plus détaillées des deux phonèmes cibles. Cet encodage détaillé est susceptible à la fois de faciliter l'apprentissage car les représentations multimodales sont plus précises, mais aussi de le ralentir car la rénovation subvocale prend plus de temps. La séquence didactique explicite devrait induire la construction de représentations plus détaillées.

Bilan

Plus un phonème est complexe et/ou difficile, plus les effets de la séquence didactique E devrait surpasser les effets de la séquence C. En prenant en compte les deux dimensions de complexité et difficulté, la séquence didactique E devrait avoir des effets positifs supérieurs à la séquence didactique C, et des effets plus importants pour le phonème /z/ que pour le phonème /ʒ/-on.

2 Définition des questions de recherche

2.1 Détermination des questions de recherche

Dans mon cadre théorique, l'apprentissage du vocabulaire est intimement lié à celui des représentations phonétiques et phonologiques. Toutes les représentations phonétiques et phonologiques peuvent initier l'apprentissage d'un mot de vocabulaire. Dans cette étude, l'enseignement de ces représentations et du vocabulaire repose sur le travail de quelques compétences clés : la discrimination, la prononciation, les correspondances phonème-graphème, le maintien en mémoire de travail et l'usage communicatif du vocabulaire. Des difficultés spécifiques de l'acquisition du FLE ont été observées pour chacune de ses compétences.

Deux phonèmes sont ciblés dans cette étude le phonème /ɔ̃/-on et le phonème /ʒ/-j. Ces deux phonèmes et les représentations qui leur sont associées sont complexes et difficiles. Le phonème /ʒ/-j semble plus difficile que le phonème /ɔ̃/-on. La différence de difficulté entre les deux phonèmes, tout comme le type de compétences observées, peuvent révéler des effets différents en fonction de la séquence didactique enseignée.

De nombreuses études semblent montrer une supériorité de l'enseignement explicite sur l'enseignement implicite pour dépasser les difficultés spécifiques induites par les interférences cross-linguistiques. Cependant de nombreuses questions restent ouvertes.

La comparaison de l'efficacité d'un enseignement implicite versus explicite a souvent été réalisée avec des apprenants adultes, en laboratoire, et sans intégrer l'enseignement implicite ciblé de la difficulté à la didactique communicative. La présente étude propose d'apporter quelques éléments nouveaux en intégrant l'enseignement implicite ciblé dans la didactique communicative, et en prenant place dans le contexte scolaire.

Le contexte scolaire est le cadre expérimental de cette étude. Les participants sont des jeunes adolescents germanophones apprenant le français LE à l'école. Or les enfants apprennent mieux implicitement, alors que les adultes apprennent mieux explicitement (Ellis, R., 2009). Les résultats des quelques autres études réalisées en contexte scolaire semblent montrer un léger avantage de l'enseignement explicite mais le manque d'études et des résultats contradictoires laissent la question du caractère implicite ou explicite de l'enseignement d'une LE en milieu scolaire ouverte. Une étude en milieu scolaire avec des participants jeunes peut amener un nouvel éclairage.

L'objectif de cette étude est de comparer les effets de deux didactiques ciblées, l'une explicite et l'autre implicite, sur l'acquisition des représentations phonétiques, phonologiques et du

vocabulaire qui contiennent deux phonèmes réputés difficiles à acquérir en FLE : *on-/ɔ̃/* et *je-/ʒ/*.

Ma question de recherche est :

Dans quelle mesure une didactique explicite sans support sémantique et une didactique implicite communicative actionnelle peuvent favoriser l'acquisition des représentations phonétiques et phonologiques et celles du vocabulaire réceptif associé ?

D'une façon générale, l'enseignement explicite semble avoir un effet positif supérieur sur l'apprentissage des difficultés spécifiques de la LE. Cependant, les effets d'un enseignement implicite ou explicite peuvent être influencés par la compétence que cet enseignement vise à développer et par la nature de l'objet d'enseignement.

Mes questions secondaires de recherche sont :

Est-il possible d'observer un effet supérieur de la séquence didactique explicite ou de la séquence implicite en fonction de la compétence développée ? Ou en fonction de l'objet d'enseignement ?

2.2 Hypothèses et Prédictions

2.2.1 Question de recherche principale

Dans quelle mesure une didactique explicite sans support sémantique et une didactique implicite communicative actionnelle peuvent favoriser l'acquisition des représentations phonétiques et phonologiques et celles du vocabulaire réceptif associé ?

Mon cadre théorique établit que les représentations phonétiques et phonologiques sont importantes pour l'acquisition du vocabulaire. Par conséquent, mon hypothèse pour l'évolution des performances de chacune des compétences est :

H0 : aucun effet de l'intervention n'est observable pour chacune des compétences observées. Les deux séquences didactiques n'ont pas d'effet sur l'acquisition des représentations phonologiques des deux phonèmes cibles ou du vocabulaire réceptif associé.

H1 : Les deux séquences didactiques ont un effet différent sur l'acquisition des représentations phonologiques des deux phonèmes cibles ou du vocabulaire réceptif associé.

2.2.2 Questions de recherche secondaire

Est-il possible d'observer un effet supérieur de la séquence didactique explicite en fonction de la compétence développée ? Ou en fonction de l'objet d'enseignement ?

En fonction de la compétence développée

La discrimination est une compétence qui utilise les représentations phonétiques. Le fonctionnement de la discrimination ne demande pas d'attention sauf en cas de conflit lors de la vérification de l'appariement des représentations phonétiques et articulatoires. Une tâche de discrimination mesure l'existence et la disponibilité des représentations phonétiques ainsi que leurs relations aux représentations articulatoires. De nouvelles représentations phonétiques en cours d'acquisition implicite peuvent être moins accessibles que lorsque leur acquisition s'appuie sur un entraînement explicite. Un avantage de la séquence didactique explicite pourrait être observé.

La prononciation est une compétence qui utilise les représentations phonétiques, phonémiques et articulatoires. Son fonctionnement est attentionnel au début de l'acquisition de la LE car 1) les biais perceptuels doivent être modifiés par un processus attentionnel, 2) les processus permettant la prononciation ne sont pas automatisés pour les représentations LE (e.g., planification des ordres moteurs, coordination des ordres moteurs), et 3) le système de contrôle de la production est attentionnel et basé sur l'accomplissement d'objectifs phonétiques. Lorsque les besoins attentionnels sont importants, l'enseignement explicite a souvent des effets supérieurs à l'enseignement implicite.

Pour la prononciation, l'apprentissage explicite a montré, dans de nombreuses études, une efficacité supérieure à l'apprentissage implicite pour dépasser les difficultés spécifiques de la LE. Cependant, les *feed-back* implicites et explicites semblent produire des effets presque similaires : les *feed-back* explicites produiraient des effets supérieurs à court terme, mais les *feed-back* implicites semblent avoir un effet plus durable. Un avantage de la séquence didactique explicite pourrait être observé au post-test immédiat (pas de test à 3 mois dans le protocole pour la prononciation).

Les correspondances phonème-graphème utilisent les représentations phonémiques, phonologiques et orthographiques. Les interférences cross-linguistiques sont difficiles à surmonter sans enseignement explicite. Pour l'acquisition des correspondances phonème-graphème, la séquence explicite devrait montrer une efficacité supérieure à la séquence implicite en particulier pour le phonème le plus difficile (i.e., /z/-j). La maîtrise des correspondances phonème-graphème est étroitement liée aux autres compétences travaillées.

Les correspondances phonème-graphème peuvent entraver ou favoriser l'apprentissage des représentations phonétiques et phonologiques. Un progrès des correspondances phonème-graphème peut induire deux types d'effets en fonction de la méthode didactique. Si leur usage est implicite, une meilleure connaissance des correspondances phonème-graphème devrait favoriser les compétences en discrimination, en prononciation, en maintien en mémoire de travail et en vocabulaire réceptif. Au contraire, si l'usage des règles de correspondance est explicite, la charge cognitive induite pourrait limiter le progrès en discrimination, en prononciation, et en maintien en mémoire de travail, tout en facilitant l'apprentissage du

vocabulaire réceptif. Ce double effet pourrait expliquer qu'un progrès des correspondances phonème-graphème induisent peu de différence d'efficacité entre les deux séquences didactiques.

Une tâche de répétition de pseudo-mots permet de mesurer le fonctionnement de la boucle phonologique. Premièrement, la comparaison entre deux séries PC et PIT permet d'évaluer l'amélioration ou le ralentissement du fonctionnement de la boucle phonologique en présence des phonèmes LE cibles qui ont été travaillés. Cette comparaison est réalisée entre la série PC, qui contient les phonèmes communs L1-LE, et la série PIT qui contient les deux phonèmes LE cibles et les phonèmes communs L1-LE. Deuxièmement, la comparaison entre la série PIT et la série SA permet de comparer l'impact de la présence des phonèmes LE cibles travaillés (=PIT) à une condition dans laquelle seul le rafraîchissement attentionnel est possible (=SA). La série SA est réalisée en condition de suppression articulaire et contient uniquement des phonèmes communs L1-LE. La comparaison de l'effet des deux séquences didactiques sera effectuée en comparant les résultats entre les quatre séries de cette tâche.

L'enseignement explicite induit peut-être des représentations multimodales plus complexes dont l'utilisation peut générer une charge cognitive plus lourde en comparaison de représentations multimodales incomplètes induites par l'enseignement implicite. Dans ce cas, la séquence explicite pourrait conduire à des performances inférieures de la séquence PIT en comparaison de la séquence communicative.

L'acquisition du vocabulaire est considérée comme majoritairement explicite en LE, ce qui correspond à la séquence communicative actionnelle (implicite pour les autres compétences mais explicite pour le vocabulaire). La séquence communicative devrait montrer une amélioration du vocabulaire réceptif supérieure à la séquence explicite.

En fonction de l'objet d'enseignement

Cette question est explorée grâce à l'étude de deux phonèmes qui diffèrent par leur complexité et leur difficulté. Leur niveau de complexité semble comparable au vu des connaissances actuelles bien que la source de la complexité diffère. Au contraire, la difficulté induite par le phonème /ʒ/-j semble supérieure à celle induite par le phonème /ʃ/-on. Cette différence de difficulté entre les deux phonèmes peut conduire à un effet positif supérieur de la séquence didactique explicite pour le phonème le plus difficile (i.e., je-/ʒ/) et pas pour le plus facile des deux. La didactique explicite a montré des effets supérieurs à l'apprentissage implicite sur l'apprentissage des structures complexes, en particulier pour les élèves les plus faibles.

3 Méthode expérimentale

L'expérimentation a porté sur l'étude des représentations phonologiques de deux phonèmes réputés difficiles à acquérir en français langue étrangère (Ruellot, 2014), en particulier pour les locuteurs du suisse allemand. Les deux phonèmes choisis sont une voyelle : la nasale /ɔ̃/-on et une consonne : la fricative /ʒ/-j. Ces phonèmes et leurs différentes représentations phonologiques et orthographiques, ainsi que certains lexèmes les contenant, ont été enseignés selon deux séquences didactiques expérimentales en classe de LE à l'école obligatoire en Suisse alémanique.

L'objectif de cette recherche est de déterminer si une didactique ciblée sur les représentations phonétiques et phonologiques permettrait aux apprenants d'acquérir les représentations phonétiques et phonologiques des deux phonèmes cibles et faciliterait ainsi l'acquisition du vocabulaire réceptif qui les contient. Pour déterminer quelle approche pédagogique est la plus appropriée, une didactique ciblée a été développée selon deux types d'enseignements, explicite ou implicite.

3.1 Séquences didactiques enseignées

Deux méthodes didactiques différentes ont été développées ; chaque méthode a été appliquée à une moitié des élèves de chaque classe. Ces deux méthodes diffèrent car l'une, la séquence communicative actionnelle (= C) a eu recours au sens et à la communication, et l'autre, la séquence explicite (= E) s'est appuyée sur un enseignement explicite des représentations phonétiques et phonologiques. Elles ont partagé cependant le même matériel linguistique (lexèmes et représentations orthographiques et phonologiques) et elles ont ciblé les mêmes compétences cognitives impliquées dans l'acquisition du vocabulaire réceptif.

Les objectifs didactiques ont visé à améliorer les représentations phonétiques, favoriser un développement des catégories phonémiques FLE, augmenter la disponibilité de l'information verbale en mémoire de travail, préciser le lien phonème-graphème en FLE et établir des liens lexicaux. Le déroulement des deux séquences didactiques et de leurs séances a suivi la même progression et les mêmes objectifs didactiques. Les activités pédagogiques ont été adaptées en fonction de la compétence travaillée, des difficultés spécifiques du phonème cible, et de la séquence (explicite vs. implicite).

L'enseignement explicite a visé à développer un apprentissage conscient par l'apprenant des nouvelles connaissances phonétiques et phonologiques ciblées. Il a reposé sur des techniques d'enseignement et des *feed-back* explicatifs en prononciation, des jeux à partir de paires minimales et des *feed-back* explicites en discrimination, des répétitions de virelangues et de chansons pour le travail de la mémoire de travail verbale, un travail des règles de

correspondances phonème-graphème et d'orthographe, et un développement de la conscience méta-phonologique autour des phonèmes cibles (tableau 4).

Tableau 4 : Comparaison des actions pédagogiques engagées dans chacune des deux séquences didactiques

Objectifs didactiques		Séquence explicite (=E)	Séquence communicative (=C)
Discrimination	Enseignement	Jeux basés sur les paires minimales	Principe de continuité (voir 5.2)
	<i>Feed-back</i>	Explicite	/
Prononciation	Enseignement	- Méthode articulatoire - Méthode verbotonale	- Lecture à voix haute - Jeux de rôle
	<i>Feed-back</i>	- <i>Recast</i> - Explicatifs	- <i>Recast</i>
Maintien en mémoire de travail	Enseignement	- Virelangues - Chansons	- Mémorisation d'extraits de dialogue - Écriture de parties de dialogue - Jeu du pendu
	<i>Feed-back</i>	Explicatifs (phonologiques)	- <i>Recast</i>
Correspondances graphème-phonème	Enseignement	- Travail des règles de correspondances et des règles phonotactiques - Dictées à trous	- Ciblage sur les formes fréquentes d'associations graphémiques, syllabiques - Lecture à voix haute
	<i>Feed-back</i>	- <i>Recast</i> - Explicatifs	- <i>Recast</i> - Forme correcte écrite au tableau
Lexico-sémantique		/	- Traduction - Support contextuel - Périphrase
Métalinguistique		Conscience phonologique ; élision et blending	Questions de lecture compréhension
	<i>Feed-back</i>	Explicatifs	Sémantique

L'enseignement implicite a été basé sur des activités classiquement développées par les méthodes communicatives (voir tableau 4). Les activités communicatives actionnelles ont placé les apprenants dans des jeux de rôle et des situations d'interaction, utilisé des jeux de réappropriation orale et de compréhension orale. Le lien avec l'écrit a été travaillé par des

lectures à voix haute avec répétition systématique par l'enseignante de la phrase lue par l'apprenant, des jeux de lettres et son (e.g., jeu du pendu), et des jeux de reconstitution de phrase. Toutes ces activités ont eu un but communicatif et le lien sémantique a été en permanence activé. Les deux séquences didactiques, implicites et explicites, ont utilisé les mêmes supports et objectifs didactiques.

3.1.1 Prononciation

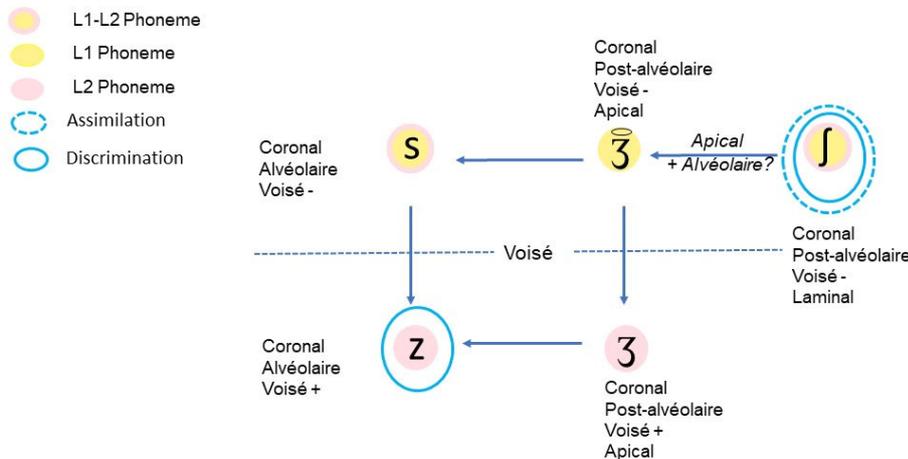
Le travail de la prononciation a été conduit de la même manière pour les deux phonèmes cibles dans chacune des séquences didactiques. Premièrement, dans la séquence explicite, l'attention a été portée sur les traits contrastifs en français (i.e. voisé pour /ʒ/-j et nasal pour /ʒ̃/-on) et sur la place articulaire. Les méthodes employées sont la méthode articulaire qui consiste à donner à l'apprenant un feedback explicatif sur la place des articulateurs, et la méthode verbo-tonale qui permet par l'utilisation d'un contexte spécifique de modifier la place articulaire de la cible. Les *feed-back* ont été des *recast* accompagnés d'explications articulaires. Deuxièmement, dans la séquence implicite, les activités communicatives ont demandé la répétition soutenue de structures linguistiques contenant les deux phonèmes cibles en continuité (i.e., placés au plus proche dans la même phrase, voir 1.5.2). Les *feed-back* donnés par l'enseignante ont consisté à répéter aussi souvent que possible les productions des apprenants contenant les cibles. Pour le travail de la prononciation et de la discrimination du phonème /ʒ/-j, une approche complémentaire, spécifique aux difficultés des élèves suisses allemands, a dû être développée.

Pour le phonème /ʒ/-j, la littérature semble indiquer que seule la différence de prononciation du trait voisé /non voisé permettrait de différencier le phonème /ʒ/-j de /ʃ/-ch en français. Cette seule différence est difficile à faire percevoir puisque 1) ce critère existe dans la grammaire phonologique des apprenants mais n'est pas considéré comme contrastif, et 2) il ne permet pas de séparer les deux phonèmes du FLE entre eux (/ʒ/-j et /ʃ/-ch), mais aussi du phonème suisse allemand auquel ils sont assimilés (i.e., /ʒ̃/). S'il existait une différence de position articulaire entre les deux phonèmes du français /ʒ/ et /ʃ/, un enseignement articulaire pourrait permettre aux apprenants de les considérer indépendamment en fonction de la place articulaire et par conséquent leur permettre d'identifier la différence que chacun des deux phonèmes FLE présente avec le phonème suisse allemand /ʒ̃/.

Pour tenter d'identifier la place articulaire de chacun des 3 phonèmes chez des multilingues, une enquête préalable a été réalisée auprès de 25 trilingues français/ allemand/ suisse allemand. L'enquête visait à leur faire décrire où et comment ils plaçaient leur langue pour les 3 phonèmes. L'analyse des réponses des participants indique que le phonème /ʃ/ serait prononcé avec un placement de la langue légèrement plus postérieur que les phonèmes /ʒ/ (français) et /ʒ̃/ (suisse allemand). Ces deux phonèmes semblent occuper la même position articulaire (comme suggéré par Fleischer et Schmid (2006) et Schmid (2010)). Ils seraient prononcés en posant le bout de la langue (= apical) au lieu des deux bords extérieurs de la langue (=laminal) comme

/ʃ/, et ils seraient prononcés un peu plus antérieurement que /ʃ/ sur le palais dur tout en restant post-alvéolaire (voir figure 3). Les phonèmes /s-z/ sont eux prononcés au niveau alvéolaire, c'est à dire plus antérieurement que les trois phonèmes /ʃ ; ʒ ; ʒ°/.

Figure 3 : Illustration des traits phonologiques qui séparent /ʒ/ des phonèmes les plus proches dans les trois répertoires phonologiques (français, allemand et suisse allemand).



Il serait donc possible lors du travail de prononciation de centrer l'attention sur le caractère voisé et le placement apical de la langue pour permettre une prononciation différenciée de chacun des trois phonèmes. L'absence du trait contrastif voisé dans la grammaire phonologique des apprenants induit des difficultés de production mais aussi de perception (Flege et al., 1997).

3.1.2 Discrimination

Pour faire progresser les compétences de discrimination, les activités pédagogiques se sont appuyées sur la paire minimale formée avec le phonème cible en français. Pour le phonème /ʒ/-j/, l'action pédagogique a travaillé la compétence de discrimination entre /ʒ/ et l'autre phonème constituant la paire minimale (i.e., /ʃ/); mais aussi la discrimination entre /ʒ/ et /s-z/ pour préciser les représentations phonétiques. En effet l'imprécision des représentations phonétiques de /ʒ/ peut favoriser son assimilation au phonème suisse allemand /ʒ°/ par les apprenants débutants, et ce, quelle que soit sa réalisation voisée ou la place de son articulation (voir figure 3). Le phonème /ʒ/ peut également être assimilé à /ʃ/, et un débutant suisse allemand assimile les deux phonèmes du FLE, /ʒ/ et /ʃ/, à /ʒ°/ (catégorisation erronée des phonèmes français comme des variants atypiques d'un unique phonème suisse allemand). L'objectif didactique est donc double : améliorer la discrimination entre /ʒ/ et /ʃ/, et permettre le développement des deux catégories phonémiques FLE de la paire minimale /ʒ-ʃ/ indépendamment de la catégorie L1 /ʒ°/.

Le travail n'a pas été engagé entre /ʒ/ et /ʒ°/ pour limiter l'activation des représentations L1 et les interférences associées. La différenciation entre le phonème suisse allemand /ʒ°/ et le phonème français /ʒ/ est visuellement facilement observable car seuls les phonèmes français

/ʒ-ʃ/ requièrent un arrondissement des lèvres. Il a été possible dans les deux séquences didactiques de constater l'arrondissement des lèvres lorsque l'enseignante a prononcé les deux phonèmes du FLE /ʒ-ʃ/. La discrimination /ʒ-ʃ/ a essentiellement reposé sur la perception du caractère voisé du phonème /ʒ/.

Pour le phonème /ɔ̃/, la discrimination a concerné la paire minimale /ɔ̃-O⁴/ pour favoriser une représentation indépendante de la voyelle nasale dans le répertoire phonologique de l'apprenant. La discrimination a essentiellement visé à faire remarquer le caractère nasal du phonème /ɔ̃/.

Dans la séquence explicite, le travail sensoriel de la discrimination a été soutenu par une gestuelle récurrente et commune aux activités de prononciation, et un *feed-back* immédiat sur l'exactitude de la réponse a été donné (voir tableau 4). Dans la séquence communicative, la solution de continuité des deux éléments de la paire minimale et les *recast* ont permis un enseignement implicite de la discrimination. La compétence de discrimination est étroitement liée à la compétence de prononciation. Dans chaque séance de chacune des séquences, les activités pédagogiques concernant la discrimination et la prononciation se sont suivies et ont prolongé le travail pédagogique de l'une ou de l'autre.

Le travail de la prononciation a été associé à la promotion de la catégorisation phonémique dans les deux séquences didactiques. La catégorisation phonémique a été travaillée à l'aide de jeux d'identification phonémique dans la séquence explicite, et à l'aide de jeux basés sur des paires minimales sémantiques dans la séquence communicative (e.g., choisir le bon mot entre *collection* et *collectionne* dans la phrase : *tu as une ... de drapeaux ?*). La catégorisation phonémique a été réinvestie dans les activités pédagogiques d'apprentissage des relations phonème-graphème.

3.1.3 Correspondances phonème-graphème

L'enseignement des relations phonème-graphème a été réalisé dans la séquence explicite à partir d'un travail en commun réflexif pour dégager les règles puis les appliquer dans des exercices (e.g., dictée à trous, lecture de mots...). Les *feed-back* ont été composés de *recast* et d'explications. Dans la séquence communicative, les lectures à voix haute et les exercices de compréhension associés, les écritures de dialogues et le jeu du pendu ont permis à l'enseignante de concentrer l'attention des élèves sur l'association forme orthographique-forme phonologique. Les *feed-back* ont été composés de *recast* et destinés à permettre la compréhension (voir tableau 4).

L'enseignement des règles phonotactiques a visé à définir les positions légalles du phonème en français et les associations phono-graphémiques les plus fréquentes. Un certain nombre de

⁴ /O/ est l'archiphonème qui regroupe les phonèmes /o ; ɔ/ du français. Ils ont été prononcés par l'enseignante selon les injonctions de la prononciation canonique : /o/ en syllabe ouverte et /ɔ/ en syllabe fermée.

régularités phonotactiques entre la distribution du phonème /ʒ/ et la forme orthographique correcte peuvent être utilisées lors de l'enseignement. Le phonème /ʒ/ s'écrit <j> surtout lorsqu'il est à l'attaque du mot (87% de la distribution de <j> dans Mille feuilles3 ; 89% dans Manulex CP (Peereman et al., 2007)). Il s'écrit <g> surtout au milieu ou en syllabe finale (92% de la distribution de <g>-/ʒ/ dans Mille feuilles3 ; 82 % dans Manulex CP). Enfin dans Mille Feuilles3 (méthode d'apprentissage des apprenants de cette étude) la terminaison en <age> représente 56% des lexèmes qui contiennent /ʒ/-<g> dans la syllabe finale (40 % dans Manulex CP) et 100% des lexèmes finissant par [aʒ] s'écrivent <age>. Ces régularités en français sont intéressantes à didactiser car l'acquisition du phonème /ʒ/ dépend de sa position dans le lexème.

Dans la séquence explicite, le travail des relations entre /ʒ/ et ses graphèmes a inclus une approche explicite de la fréquence phonotactique (e.g., /aʒ/ s'écrit toujours <age> en fin de mot). Dans la séquence communicative, les lexèmes représentatifs ont été mis en évidence sémantiquement ; il s'agit des lexèmes dont l'écriture est arbitraire (e.g., *gens* / *j'en* ; *geler* / *jeter*) ou dont la forme représente la distribution phonotactique la plus fréquente (e.g., *cage*). Le lien prononciation et discrimination avec l'acquisition des relations phonème-graphème a été travaillé par l'intermédiaire de dictée à trous (séquence explicite) ou de questions de compréhension (séquence communicative). Entendre et pouvoir écrire la part manquante du lexème lors d'une tâche de complétion des trous, ou entendre et interagir lors d'une interaction est possible grâce au maintien en mémoire de travail de la représentation phonologique.

3.1.4 Mémoire de travail

L'objectif didactique est de favoriser l'établissement des liens en mémoire de travail entre les représentations phonologiques et les autres représentations liées au lexème. Des chansons et des virelangues dans le groupe explicite et des bouts de dialogues dans le groupe communicatif ont été appris par les apprenants pour entraîner le maintien en mémoire des lexèmes qui contenaient les phonèmes cibles ou le membre de leur paire minimale.

La compétence de relier une représentation phonologique à d'autres représentations a été travaillée de la même manière pour les deux phonèmes. Dans la séquence explicite, les liens avec les représentations orthographiques ont été activés au niveau méta-phonologique à l'aide d'exercices d'élimination, d'assemblage ou de permutations des clusters, syllabes ou phonèmes cibles. Dans la séquence communicative, la compétence pour établir des liens avec la valeur sémantique des lexèmes a été en permanence sollicitée (e.g., L'enseignante a décrit quelle partie d'un drapeau était coloriée en rouge comme par exemple : le rond, en haut... ; les élèves ont dû colorier la partie concernée le plus vite possible). Les compétences métalinguistiques ont été travaillées à l'aide de jeux de reconstruction de phrases, de lexèmes ou de courts textes narratifs. La progression des performances de toutes les compétences travaillées a été mesurée à l'aide de tâches spécifiques.

Tout le matériel verbal utilisé pour l'enseignement et les tests a été extrait d'une liste établie à partir de l'analyse de la méthode d'enseignement des apprenants (i.e., *Mille Feuilles*). Une analyse de la fréquence des lexèmes dans la méthode d'enseignement des apprenants a permis d'établir une liste des lexèmes les plus fréquents qui contenaient les deux phonèmes cibles. Les lexèmes choisis pour atteindre un objectif didactique précis étaient identiques dans les deux séquences didactiques. Bien qu'ils aient été travaillés par des actions pédagogiques distinctes, les objectifs didactiques des deux séquences sont majoritairement similaires.

Pour évaluer et comparer l'effet de chacune des méthodes sur l'apprentissage du vocabulaire réceptif, les performances des compétences cognitives impliquées dans l'usage de représentations phonétiques et phonologiques et le vocabulaire réceptif ont été mesurés.

3.2 Tâches utilisées pour mesurer l'évolution des performances des compétences enseignées

Les compétences travaillées pour l'apprentissage des phonèmes cibles sont la discrimination, la prononciation, les correspondances phonème-graphème et le maintien en mémoire de travail. Ces compétences requérant l'usage des représentations phonétiques et phonologiques ont été identifiées comme prenant part au processus d'acquisition du vocabulaire. Afin de répondre à ma question de recherche, ces quatre compétences et le vocabulaire réceptif ont été mesurés.

Tous les stimuli sonores ont été enregistrés dans une cabine insonorisée avec un microphone de studio électrostatique à grande capsule et filtre anti-pop par une francophone native. Ils ont été prononcés et enregistrés enchâssés dans une phrase support : « Pour sortir je mets la ...X... sur mon dos ». Puis le stimulus le plus compréhensible pour les deux francophones présents a été extrait pour chaque pseudo-mot. L'ensemble des stimuli choisis a alors été lissé à une amplitude de 64Hz (Ludusan & Dupoux, 2015) pour qu'aucune différence d'intensité entre les stimuli ne puisse expliquer les différences observées dans les résultats grâce au logiciel Adobe AuditionCC. Dans une dernière étape, les stimuli ont été inclus dans une bande son à l'aide du logiciel Audacity.

3.2.1 Tâche de discrimination

L'objectif de cette tâche est d'évaluer si les apprenants perçoivent la différence phonétique entre les phonèmes cibles et les phonèmes auxquels ils sont susceptibles d'être assimilés. La tâche consiste à déterminer si les deux pseudo-mots présentés oralement dans la même paire sont identiques (e.g., /bo/ et /bo/) ou différents (e.g., *achi-/aʃi/* et *agi-/azi/*). Les participants ont dû cocher une des deux cases proposées en face du numéro identifiant le stimulus entendu (voir Annexe 2). Une moitié des élèves avait en premier la colonne « Gleich » *pareil* (grille A) et

l'autre moitié en premier la colonne « Verschieden » *différent* (grille B) sur la feuille de réponse. Les deux grilles A et B ont été aléatoirement distribuées parmi les participants.

3.2.1.1 Procédure de la tâche de discrimination

Le passage du test a été collectif. Cette tâche dure 4 minutes 20 secondes. Une bande son a été diffusée via un ordinateur et une enceinte de qualité. Chaque paire minimale a été annoncée par un numéro énoncé en allemand suivi par la paire testée après 500ms de silence. Un beep de 300ms a annoncé le début et la fin de la tâche.

La tâche a débuté par une phase d'entraînement. Après l'écoute des consignes, les participants ont évalué deux paires minimales dans les mêmes conditions que durant la tâche : il s'agit de /ku/-/ku/ (bonne réponse = *Gleich*) et de /mu/-/mo/ (bonne réponse = *Verschieden*). Les participants ont reçu un *feed-back* immédiat.

Les consignes ont été traduites et enregistrées par une germanophone à l'aide d'un enregistreur Tascam D40. Les consignes ont été : « Nach einem Beep hörst du eine Serie von Wörtern. Du hörst immer zwei Wörter auf einmal. Nach jeder Serie musst du feststellen, ob die zwei Wörter gleich oder verschieden sind und du musst die richtige Antwort auf dem Blatt ankreuzen. Zuerst trainieren wir nur.» (*Après un beep, tu vas entendre une série de mots. Tu vas entendre les mots deux par deux. Après chaque série de deux mots, tu dois déterminer s'ils sont pareils ou différents et tu dois cocher sur la feuille la bonne réponse.*)

La méthode de comptage est simple : pour chaque stimulus, une bonne réponse est égale à 1 point, une réponse fautive à 0.

3.2.1.2 Les stimuli

Les pseudo-mots monosyllabiques utilisés sont de structure consonne-voyelle. Ils sont communs avec la tâche de prononciation. Ils sont associés par paire pour composer un stimulus de la tâche de discrimination (e.g., /ʒa/-/ʃa/). Ils constituent un total de 30 paires proposées dans la tâche : 10 paires identiques (e.g., /mo/-/mo/) et 20 paires différentes (e.g., /ʒe/-/ʃe/). Les paires identiques sont construites avec les phonèmes /ʃ/ (2), /o/ (3), /ʒ/ (3), /z/ (1) et /f/ (1) (voir Annexe 2).

La durée de l'intervalle entre deux pseudo-mots de la même paire semble influencer les résultats des tests de discrimination (Strange, 2011). Le temps séparant les deux pseudo-mots d'un stimulus (=ISI) a donc été fixé à 250ms pour 6 stimuli (discrimination rapide et ISI de la tâche de mémoire de travail) et à 500ms pour 24 stimuli (ISI couramment utilisé dans les tâches de discrimination et ISI entre les deux répétitions du stimulus dans la tâche phonème graphème). Entre chaque stimulus le silence dure 6s pour permettre à la boucle phonologique de se libérer et ainsi limiter le risque d'interférences entre les stimuli (Cowan, 2015). Ces trente stimuli composés d'une paire de pseudo-mots ont été construits autour des deux phonèmes cibles.

Les paires de pseudo-mots sont composées par les phonèmes cibles présentés dans 4 contextes différents : 4 contextes consonantiques pour /ð/ et 4 vocaliques pour /ʒ/. Les phonèmes qui sont fréquemment confondus avec les phonèmes cibles constituent la variante différente dans les paires différentes (voir tableau 5).

Tableau 5 : Pseudo-mots créés pour explorer le phonème /ʒ/

Contexte	Phonème cible /ʒ/	/ʃ/	/s/, /z/
/a/	/ʒa/	/ʃa/	/za/
/a # i/	/aʒi/	/aʃi/	/assi/
/e/	/ʒe/	/ʃe/	/ze/
/a # u/	/aʒu/	/aʃu/	/azu/

Le phonème /ʒ/ est fréquemment confondu avec /ʃ/, et beaucoup moins souvent /s/ ou /z/. Ces trois phonèmes du FLE ont donc constitué la variante des paires différentes. Les pseudo-mots d'une paire identique sont deux fois le même extrait sonore du même pseudo-mot. Le contexte vocalique est commun aux deux pseudo-mots appartenant à une même paire qu'elle soit différente ou identique. Le schwa est source de difficulté en discrimination et a donc été remplacé dans cette tâche par des voyelles sonores.

Le schwa est difficile à percevoir et le voisement est mieux discriminé en position intervocalique (Schmid, 2012). Pour faciliter la discrimination /ʒ-ʃ/, les stimuli communs avec la tâche de prononciation qui contenaient un schwa (e.g., contexte /a # ə/) ont été modifiés pour un contexte /a # u/.

Aucune modification des stimuli n'a été effectuée pour le phonème /ð/ et tous les stimuli sont communs entre prononciation et discrimination. Les deux voyelles fréquemment confondues avec /ð/ sont /u/ et /o/ et constituent donc la variante des stimuli différents (voir tableau 6). Le contexte consonantique est commun aux deux pseudo-mots, que la paire soit identique ou différente.

Tableau 6 : Pseudo-mots créés pour explorer le phonème /ð/

Contexte	Phonème cible /ð/	/ɔ/	/u/
/k/	/kð/	/kɔ/	/ku/
/b/	/kð /	/bɔ/	/bu/
/m/	/mð /	/mɔ/	/mu/
/s/	/sð /	/sɔ/	/su/

3.2.2 Tâche de prononciation

La tâche de prononciation vise à évaluer la possibilité d'identifier les deux phonèmes cibles de cette étude : /ʃ/ et /z/ lorsqu'ils sont prononcés par les apprenants. Une tâche d'imitation de mots et pseudo-mots isolés a permis cette évaluation sans requérir un savoir lexical ou phonotactique du participant tout en induisant une charge cognitive limitée. Les stimuli sont des monosyllabiques pour limiter la charge cognitive induite par la mémorisation ponctuelle demandée par la tâche de répétition.

3.2.2.1 Procédure de la tâche de prononciation

Le passage a été individuel. La tâche a duré 2 minutes 21 s. Une bande son a été diffusée via un ordinateur et une enceinte de qualité. L'apprenant a dû répéter le stimulus immédiatement après l'avoir entendu.

La phase préalable d'entraînement a consisté à écouter les consignes puis à répéter deux stimuli dans les mêmes conditions que durant la tâche : il s'agit de /ku/ et de /mu/. Un beep de 300ms a annoncé le début de la tâche. Un temps de 6 secondes a séparé la fin d'un stimulus du début du suivant pour permettre de limiter les interférences en boucle phonologique.

Deux bandes son (listes A et B) qui diffèrent uniquement par l'ordre de présentation des stimuli ont été attribuées au hasard parmi les participants.

Liste A : [aze mo ko sɔ̃ za afe bɔ̃ ze ʃa su aze ʃe azi mɔ̃ so za bo asi ze bu kɔ̃ aʃi]

Liste B : [asi za bo mɔ̃ ʃe aze kɔ̃ sɔ̃ afe bu azi ko ze mo aʃi ze bɔ̃ za su aze so ʃa]

Les consignes ont été traduites et enregistrées par une germanophone. Les consignes ont été : « Nach einem Beep, hörst du Wörter. Du musst jedes Wort wiederholen, nachdem du es gehört hast. Zuerst trainieren wir nur. » (*Après un beep, tu vas entendre des mots. Tu dois répéter chaque mot juste après l'avoir l'entendu*). Les prononciations élicitées ont été enregistrées dans une salle de classe calme au format « .wav » (44.1Hz ; 16 bit) par un Tascam DR40 comportant deux micros multidirectionnels.

3.2.2.2 Les stimuli

Les 24 stimuli sont majoritairement constitués des mêmes pseudo-mots que ceux construits pour la tâche de discrimination autour des deux phonèmes cibles (voir 3.2.1). La principale différence concerne les pseudo-mots contenant /z/ et construits autour du contexte /a # u/.

Le phonème /z/ est fréquemment suivi d'un schwa quand il appartient à la dernière syllabe du mot en français. Le contexte /a # u/ utilisé pour la tâche de discrimination est donc remplacé dans cette tâche par le contexte /a # ə/ ce qui permet d'observer la prononciation de /z/ en présence du schwa. La réalisation voisée de /z/ est plus aisée en position intervocalique qu'en présence d'un schwa et est observable dans le contexte /a # i/. Les deux autres pseudo-mots sont monosyllabiques et place le /z/ à l'attaque de la syllabe.

La structure monosyllabique est toujours maintenue pour éliciter la prononciation du /ɔ̃/. La nasalisation régressive peut induire le participant à percevoir la voyelle nasale /ɔ̃/ comme les voyelles orales /o/ ou /ɔ/. Lorsque la langue ne contient pas de voyelle nasale phonémique (e.g., l'anglais, l'allemand, le mandarin), la nasalisation des voyelles est la conséquence de la coarticulation avec une consonne nasale (i.e., /n ; m ; ŋ/ en français). Pour éviter une interprétation de la nasalité vocalique comme le résultat d'une nasalisation régressive, la voyelle nasale /ɔ̃/ a été placée à la fin des stimuli. L'observation d'une prononciation identifiable de la voyelle nasale peut indiquer une intégration du trait nasal dans la grammaire des apprenants. Cependant une partie des stimuli sont des lexèmes du français et la prononciation correcte peut aussi être attribuée à l'acquisition d'une représentation phonologique globale.

La majorité des stimuli sont des lexèmes du français (19/24). Cette situation est très courante lors de l'étude de la prononciation (e.g., Delvaux, 2003) car les monosyllabiques de structure consonne-voyelle sont très fréquents en français (Meynadier, 2013). Cependant, aucune référence à la valeur sémantique n'est donnée et les stimuli sont présentés à l'aide de listes décontextualisées. La représentation lexicale a peu de chance de constituer une aide articulatoire car les participants sont débutants (lexique limité). Dans le cas où la représentation phonologique globale du lexème ait soutenu les performances, les résultats sont néanmoins toujours valides. Les représentations articulatoires utilisées pour la prononciation du lexème sont communes à celles utilisées en l'absence d'une représentation phonologique globale liée au lexème.

3.2.2.3 Méthode d'évaluation

Deux méthodes de mesure sont théoriquement possibles : une méthode en lien avec l'intelligibilité qui demande à des personnes francophones d'identifier ce qui a été prononcé, et une méthode acoustique réalisée grâce au logiciel PRAAT (Boersma & Weenink, 2018). Cependant il manque encore une définition claire des valeurs que doivent avoir les phonèmes prononcés en LE pour pouvoir considérer que ces phonèmes sont identifiables en situation de communication. La méthode acoustique n'est donc pas applicable. La méthode d'évaluation de la prononciation retenue a donc été l'évaluation par des codeurs francophones.

Trois codeuses ont évalué toutes les bandes son. Les codeuses ne sont pas musiciennes car le codage effectué par des musiciens diffère de celui des non musiciens (Isaacs & Thomson, 2013; Isaacs & Trofimovich, 2010; Saito, Trofimovich, Isaacs & Webb, 2017). Les trois codeuses ont chacune un profil différent pour essayer d'évaluer la prononciation des participants par des interlocuteurs variés : deux francophones natives dont une est une linguiste expérimentée et l'autre non-expérimentée, et une locutrice du français niveau C1/C2 (L1 Brésilien : langue dans laquelle les deux phonèmes cibles existent). En effet en situation de communication, le participant peut parler à un locuteur natif ou non natif, habitué à comprendre un fort accent étranger ou non. Le locuteur doit être compréhensible en LE quelle que soit l'expérience

linguistique de son interlocuteur. Pour favoriser la cohérence entre les codeuses une échelle d'évaluation a été établie.

L'évaluation a été basée sur une échelle de Likert, comme habituellement dans les mesures de la compréhensibilité lors de l'évaluation de la prononciation en situation de communication (Derwing & Munro, 1997). Chacune des codeuses a travaillé indépendamment après une séance commune d'entraînement au codage. Les différences entre les codages respectifs ont été discutées pendant une heure (Isaacs, 2014). L'échelle d'évaluation préétablie a alors été affinée qualitativement.

L'échelle d'évaluation préétablie a, tout comme l'approche didactique, été construite autour de l'analyse des difficultés cognitives liées à la perception et la manipulation d'un phonème LE qui n'appartient pas au répertoire phonologique de la L1. Les erreurs segmentales qui sont liées à des contrastes phonémiques fréquents, comme les deux phonèmes cibles de cette étude, sont particulièrement délétères pour l'intelligibilité et la compréhensibilité (Munro & Derwing, 2006; Zielinski, 2008). L'approche didactique utilisée dans cette étude tout comme l'échelle d'évaluation sont focalisées sur le niveau segmental.

L'objectif de l'évaluation doit être celle de l'enseignement reçu. Un enseignement qui viserait une prononciation proche de celle des natifs n'est ni réaliste ni souhaitable en LE. En effet, accent et identité sont deux caractéristiques de l'apprenant étroitement imbriquées (Gatbonton et al., 2011). Un focus sur l'accent peut être délétère pour la motivation de l'apprenant. De plus il n'est pas nécessaire de parler la langue du pays d'accueil sans accent pour s'intégrer socialement et réussir professionnellement (Derwing & Munro, 2009). L'évaluation des participants a donc été réalisée sur la compréhensibilité au niveau segmental.

Une échelle de Lickert a été établie pour mesurer l'évaluation de la compréhensibilité au niveau segmental. D'après les définitions de Derwing et Munro (2009) la compréhensibilité représente la perception par l'expérimentateur de combien il est facile de comprendre le participant. Cette variable est généralement mesurée sur une échelle de Likert basée sur 9 niveaux (de très difficile à très facile à comprendre). L'échelle que j'utilise est plus précise pour laisser moins de variation inter-juge et correspondre aux objectifs des séquences didactiques expérimentales.

L'échelle de Lickert construite pour évaluer les performances en prononciation comporte 5 niveaux qui visent à mesurer la compréhensibilité du phonème :

1 : Le phonème prononcé n'est pas compréhensible ou est erroné.

2 : L'interlocuteur ne peut pas décider entre deux phonèmes, dont le phonème cible, celui qui correspond à la production de l'apprenant.

3 : Le phonème cible semble pouvoir être identifié mais il présente une forte déformation /accent qui laisse un doute important à l'auditeur.

4 : Le phonème cible est identifiable, mais il n'est pas parfaitement prononcé.

5 : Le phonème cible est identifiable, totalement compréhensible et prononcé sans accent LE marqué.

La valeur retenue pour l'analyse statistique a été l'évaluation établie sur cette échelle par chaque codeuse pour chaque prononciation du phonème cible.

3.2.3 Tâche de maintien en mémoire de travail

L'objectif de cette tâche est de déterminer si le travail de la prononciation et du maintien de l'information verbale en FLE a permis une amélioration de la capacité d'encodage et du maintien en boucle phonologique de l'information verbale. La question de recherche cherche à évaluer l'efficacité didactique pour l'acquisition de la LE, et non les habiletés exécutives nécessaires à l'apprenant pour l'acquisition de la LE (voir Annexe 3 pour plus de détails). Cette tâche de mémoire de travail est donc simple et non complexe. Une tâche simple peut présenter des stimuli verbaux, visuels, numériques... Les résultats des tâches de mémoire de travail utilisant un support verbal sont plus fortement corrélés au développement du vocabulaire en LE (Engle, Kane & Tuholski, 1999). Une tâche simple de répétition de l'information verbale a par conséquent été utilisée.

3.2.3.1 Procédure de la tâche de mémoire de travail

Le passage du test s'est déroulé individuellement dans une pièce calme. La tâche complète est composée de 4 séries de tests et d'une série d'entraînement (durée d'un test =1.30mn). La tâche complète a duré environ 10 minutes. L'ordre des 4 séries de test étaient différents en fonction de l'attribution aléatoire d'un groupe A ou B à chaque apprenant.

Chaque série est composée de 8 listes de pseudo-mots présentées par ordre croissant du nombre de pseudo-mots : 2 listes de 2 pseudo-mots, puis 2 listes de 3 pseudo-mots, puis 2 listes de 4, puis 2 listes de 5. Les apprenants ont dû répéter chaque liste de pseudo-mots immédiatement après l'avoir entendue. La bande son délivrant les stimuli a été diffusée via un ordinateur et une enceinte de qualité. Un beep (=300ms) a annoncé le début de chaque liste et le début de la tâche de rappel. Un temps de silence (=10s) a séparé la fin d'une liste du début de la suivante pour permettre de limiter les interférences (Cowan, 2015). Une courte pause a été faite entre chaque série.

Les séries contiennent le même nombre de phonèmes (=68) mais la nature des phonèmes est différente. Les phonèmes communs L1-LE composent les séries de « phonèmes communs » (=PC) et d'entraînement, ainsi que la série de suppression articulatoire (=SA). Pour la série SA, la tâche est un peu différente car l'apprenant doit répéter [ja] sans interruption pendant l'écoute des listes et ne s'interrompre qu'au moment de parler. La série « Phonèmes inconnus travaillés » (=PIT) est composée des phonèmes communs L1-LE et des deux phonèmes cibles (2 x 8 exemplaires); tandis que la série « Phonèmes inconnus non travaillés » (=PInT ou série « parallèle » à PIT) est composée des phonèmes communs L1-LE et des deux phonèmes

parallèles (i.e. /ɛ/-in et /ɲ/-gn). Les phonèmes parallèles ont été choisis pour leur proximité articulatoire et des fréquences proches dans le discours oral (/ɛ/= 1.9% et /ɔ/= 2.2%), ou pour leur ressemblance graphique (/z/-<ge> et /ɲ/-<gn>). Les phonèmes parallèles occupent les mêmes places syllabiques et les mêmes places dans les listes que les phonèmes cibles. Le détail des séries est disponible dans l'annexe 3.

La tâche a commencé par une série d'entraînement. Après l'écoute des consignes, les apprenants ont restitué les 8 listes de la série d'entraînement.

Les consignes ont été entendues une fois au début de la tâche d'entraînement et l'expérimentatrice les a répétées en français avant de débiter le test. Les consignes ont été : «Du wirst ein «beep» hören, dann zwei Wörter und ein Schluss-«beep». Nachdem du beide Wörter gehört hast, wiederholst du sie in der selben Reihenfolge. Dies wird sich 8 mal wiederholen aber du hörst dann 3, 4 und zuletzt 5 Worte, die du jedes Mal wiederholen musst. Auch wenn du nicht sicher bist, sage einfach spontan was kommt, dein Hirn registriert mehr als du glaubst. Zuerst üben wir nur! Jetzt geht's richtig los... OK? » ; « *Tu vas entendre des mots que tu dois répéter dans l'ordre où tu les entends. Tu entends d'abord un beep puis deux mots puis un beep final. Tu répètes les deux mots dans l'ordre après avoir entendu le beep final. Tu vas recommencer huit fois : la liste contiendra d'abord 2 mots, puis 3, puis 4, et pour finir 5 mots. Même si tu n'es pas sûr de répéter correctement tous les mots essaie de dire tout ce qui te vient à l'esprit ! Ton cerveau a enregistré plus de mots que tu ne crois ! En premier tu fais un entraînement. Tu es prêt.e ?* ».

La série de suppression articulatoire a toujours été proposée en dernier. Cette série a commencé par une série d'entraînement dédiée (la bande son est celle de la série d'entraînement du début de test) car les consignes sont légèrement différentes : « Während du die Worte hörst, musst du fortlaufend « ja ja ja » sagen. Nur sobald du dann die Worte wiederholen musst, hörst du mit dem «ja ja ja» sagen auf. Nach der Wiederholung, fängst du sofort wieder mit «ja ja ja» sagen an. Zuerst üben wir nur! Jetzt geht's richtig los... OK? » ; « *Pendant que tu vas entendre le mot, tu dois dire tout haut et tout le temps «ya ya ya». Tu ne dois pas arrêter de dire «ya ya ya», jusqu'à ce que tu doives répéter les mots entendus. Après avoir répété les mots entendus, tu dois tout de suite recommencer à dire «ya ya ya» à voix haute. En premier tu fais un entraînement. Tu es prêt.e ?* »

3.2.3.2 Structure de la tâche

Pour mesurer le maintien et la répétition de l'information verbale dans une tâche de mémoire de travail, les stimuli peuvent être des mots ou des pseudo-mots. La restitution des mots est généralement supérieure à celle des pseudo-mots (effet de lexicalité). En effet, le maintien et le rappel sont favorisés par toutes les représentations liées au lexème : phonologiques, orthographiques et sémantiques (Hulme, Maughan & Brown, 1991). D'ailleurs les empan de mots sont sensibles à la fréquence lexicale (Hulme et al., 1991, 1997; Roodenrys et al., 1994) et à l'effet sémantique : les mots concrets sont mieux rappelés que les mots abstraits (Taylor,

Francis, Borunda-Vazquez & Carbajal, 2019; Walker & Hulme, 1999) et les listes de mots sémantiquement liées sont mieux restituées (Kowialiewski, Lemaire & Portrat, 2021; Majerus, 2010; Majerus & van der Linden, 2003). Cependant en LE, les représentations lexicales de la L1 peuvent générer de nombreuses interférences qui limitent l'évaluation du maintien et de la répétition de l'information verbale LE (Bartolotti, Bradley, Hernandez & Marian, 2017).

Afin de limiter les interférences cross-linguistiques lexicales et l'accès lexical, la tâche de mémoire de travail utilise des pseudo-mots respectant les règles phonotactiques du français. Les pseudo-mots permettent de mesurer l'accessibilité et la connaissance des représentations phonologiques de la LE (Carroll & Snowling, 2004; Coady & Evans, 2008; Foy & Mann, 2001). En effet l'usage de pseudo-mots induit un traitement en mémoire de travail relativement indépendant des représentations lexicales et qui donne plus d'importance aux représentations phonétiques (voir 1.1, 1.2, 1.3). La frontière entre deux catégories phonémiques (L1 et LE) est alors plus fidèle à leurs différences phonétiques car la limite entre ces catégories est établie sur un même continuum phonétique (mesure de *goodness of fit*) et pas sur des critères phonétiques indépendants (mesure d'une identification catégorielle) (Piske et al., 2002). Néanmoins le maintien des pseudo-mots peut aussi être renforcé par les représentations lexicales car certains monosyllabiques utilisés dans cette tâche sont des mots de français (51% en moyenne). Il est aussi très probable que certains soient aussi des lexèmes dans une des langues de l'apprenant. Or l'objectif de cette étude est d'évaluer la capacité de maintien et de codage au niveau phonétique et phonémique, indépendamment du niveau de vocabulaire de l'apprenant.

Pour limiter les possibilités d'accès lexical, la durée du silence entre les pseudo-mots a été limitée. Un intervalle inter-stimuli de 250ms permet la mise en place des processus de rénovation et l'activation de toutes les représentations phonologiques. Avant 300ms post stimulus, l'information verbale est traitée cognitivement essentiellement au niveau phonologique. Il est couramment admis en neurologie que le traitement de l'information verbale est réalisé parallèlement selon deux axes d'activation corticaux : un flux dorsal qui est phonologique et qui précède le flux ventral, lui-même sémantique (e.g., Hickok, 2012). Le flux dorsal est parcouru en 250ms post stimulus au maximum. D'un autre côté, il faut seulement 50 à 80ms pour que la rénovation de l'information contenue dans la boucle phonologique soit fonctionnelle. Pour ces raisons, dans cette tâche, deux pseudo-mots sont toujours séparés par un intervalle de 250ms.

La période de silence est fondamentale pour séparer les monosyllabiques et ainsi faciliter leur stockage en boucle phonologique (Billières, 2005, p.3). Cette période de silence segmente clairement les éléments acoustiques afin qu'ils ne puissent pas être interprétés comme le résultat d'une interaction entre les phonèmes qui se suivent (i.e., le résultat de phénomènes d'assimilation ou de Sandhi). L'individualisation des monosyllabiques à la place d'un long polysyllabique (Gathercole, Willis, Baddeley & Emslie, 1994) permet de supprimer les ambiguïtés phonémiques. En effet, la nasalisation d'une voyelle peut être interprétée par l'apprenant comme le résultat de l'influence régressive d'une consonne nasale sur la voyelle

qui la précède comme en anglais par exemple /ba~nk/-*bank* et non comme un phonème du français comme par exemple /bāk/- *banque*. De même des effets prosodiques transforme un /ʒ^o/ en /ʃ/ en suisse allemand alors que ce n'est pas le cas en français où les deux phonèmes /ʒ ; ʃ/ ne sont pas allophoniques. La tâche a donc été construite pour supprimer les ambiguïtés phonémiques et ainsi répondre au mieux à la question de recherche.

La question de recherche vise à évaluer la participation des représentations phonologiques dans l'acquisition du vocabulaire réceptif, et pour cette tâche à déterminer si le travail didactique permet un meilleur maintien en mémoire de travail du matériel verbal LE. La tâche utilisée pour répondre aux questions de recherche a donc été une tâche simple, utilisant des pseudo-mots monosyllabiques séparés par un silence de 250ms dans une liste à répéter.

3.2.3.3 Les stimuli

La tâche de répétition de pseudo-mots a proposé 8 listes contenant un nombre croissant de monosyllabiques : 2 puis 3, 4 ou 5 pseudo-mots (Demont & Botzung, 2003). Les listes consécutives deux par deux ont le même nombre de monosyllabiques et de phonèmes (ex. listes 1 et 2 : 2 pseudo-mots ; ... listes 7 et 8 : 5 pseudo-mots). Les phonèmes composant les pseudo-mots sont soit des phonèmes LE qui n'existent pas en L1, soit des phonèmes communs L1-LE.

Les phonèmes communs au répertoire phonologique du français et de l'allemand sont les voyelles : /a, e, i, o, y, u/ et les consonnes : /b, d, l, m, n, p, r, t, k, f, s, ʃ/. Ils constituent la totalité des séries des phonèmes connus (=PC), de suppression articulatoire (=SA) et d'entraînement. Ils occupent également les places vacantes des séries des phonèmes inconnus travaillés (=PIT) et des phonèmes inconnus non travaillés (=PInT).

Les phonèmes cibles de cette étude /ɔ̃/-*on* et /ʒ/-*j* et leur phonèmes « parallèles » /ɛ̃/-*in* et /ɲ/-*gn* appartiennent exclusivement au répertoire du français, Ces phonèmes sont présents une fois dans chaque liste de la série PIT (i.e., 8 listes par série = 8 exemplaires de chaque phonème cible par série). Les phonèmes inconnus /ɛ̃/ et /ɲ/ sont placés exactement aux mêmes places que les phonèmes cibles (respectivement /ɔ̃/ et /ʒ/) dans les listes de la série PInT. Ils occupent la même place dans le monosyllabique : attaque ou coda ; Ils appartiennent au monosyllabique placé au même rang dans une liste (e.g., le troisième monosyllabique de la liste). Ces phonèmes LE non travaillés et « parallèles » devraient me renseigner sur les transferts possibles entre phonèmes LE pendant leur acquisition, en particulier si j'observe ou non une facilitation de leur apprentissage implicite. La voyelle /ɛ̃/ est nasale comme le phonème cible /ɔ̃/ et partage certaines caractéristiques phonétiques et phonologiques. Le lien est moins évident entre les deux consonnes /ʒ/ et /ɲ/ qui partagent partiellement la représentation graphique (= <g>) et sont deux phonèmes consonantiques absents du répertoire germanophone. La distribution phonémique dans les monosyllabiques doit respecter un certain nombre de règles.

Tous les monosyllabiques respectent les règles phonotactiques du français (Meynadier, 2001). Ils sont composés de deux ou trois phonèmes. Ils sont tous construits sur un squelette consonne-voyelle +/-consonne. Il n'y a pas de clusters consonantiques. Cette structure simple représente

77% des syllabes du français et 66,8 % des syllabes de l'allemand (Léon, 1992, p.96) et par conséquent ne constitue pas une difficulté pour les apprenants.

La place du monosyllabique qui contient le phonème cible dans la liste présentée influence sa restitution. Un effet de récence est observé dans chaque liste : les derniers monosyllabiques sont mieux restitués que les antérieurs (Baddeley & Hitch, 1993; Gupta, 2005; Gupta, McWhinney, Feldman & Sacco, 2003). Les phonèmes cibles ont donc été répartis dans les listes d'une même série pour représenter toutes les positions : par exemple liste 5 : en position 2 et 3 ; liste 6 en position 3 et 4 ; liste 7 en position 1 et 4 ; liste 8 en position 2 et 5. La place du phonème cible dans le monosyllabique influence aussi sa restitution.

Les phonèmes ont des réalisations acoustiques différentes en fonction de leur position syllabique (Voir 1.2 ; Bürki, 2018; Meynadier, 2001; Steinschneider, 2013). Les phonèmes cibles ont été distribués dans toutes les positions syllabiques (i.e. l'attaque, le centre ou la coda de la syllabe). Un exemplaire de chaque phonème cible est présent par liste, ce qui pour un total de 8 listes par série représente 8 phonèmes cibles par série (i.e. 8x/5/ et 8x/3/ pour une série PIT). Le nombre total de phonèmes par liste a été déterminé par les caractéristiques des participants.

L'âge des participants peut être un facteur limitant pour la tâche. Vers 11-12 ans, un individu a acquis, en L1, les mêmes capacités de maintien en boucle phonologique qu'un adulte, soit environ 7 objets (Gaonac'h & Larigauderie, 2000, p.198 ; Raguene, 2011; Tam, Jarrold, Baddeley & Sabatos-DeVito, 2010), ce qui représente 3 à 4 syllabes (Gineste & Le Ny, 2002). Cependant en LE, l'empan est généralement réduit en fonction du degré de maîtrise de la LE (Gaonac'h & Larigauderie, 2000, p.514, Gathercole, 2006 ; Majerus, 2010; Masoura & Gathercole, 1999; Service & Kohonen, 1995). Ces critères ont déterminé la longueur maximale possible d'une liste (=5).

La liste la plus longue contient cinq monosyllabiques (douze phonèmes) et ne devrait donc pas induire d'effet plafond des mesures. Cette longueur de liste a déjà été testée dans mon étude de master et dans le pilote de cette étude sans effet plafond (Dherbey-Chapuis, 2016).

3.2.3.4 Méthode d'évaluation

L'intégralité des séries de répétition de pseudo-mots a été transcrite par une codeuse francophone dans un fichier « word » sans qu'elle puisse se référer aux transcriptions des stimuli. Elle n'a jamais transcrit les productions des participants pendant plus d'une heure. La lecture automatique de la transcription a déterminé le nombre de phonèmes correctement restitués par série (i.e., correctement = identifiés comme étant ceux des stimuli). Cette mesure a été utilisée pour l'analyse statistique.

3.2.4 Tâche de mesure des relations phonème-graphème

Les tests habituellement utilisés pour mesurer le décodage sont des tests de lecture de non-mots. Or le français est plus inconstant dans le sens inverse, du phonème vers le graphème, et cette identification du son à la forme orthographique est essentielle pour apprendre le vocabulaire en LE (Ziegler, Jacobs, & Stone, 1996). Un nouveau test a donc été créé pour mesurer la capacité des apprenants à utiliser des graphèmes pour écrire une suite de phonèmes en FLE. La tâche consiste à identifier la forme orthographique d'un pseudo-mot bisyllabique présenté sous forme orale. Le participant doit cocher la forme correcte parmi 4 propositions (voir Annexe 4).

Ce test a fait l'objet d'une étude préliminaire qui visait à vérifier sa validité (Dherbey-Chapuis & Berthele, 2020). Il dure 11 minutes. Un beep a annoncé la présentation d'un numéro en allemand puis après 300ms d'un stimulus qui est répété 700ms plus tard. Un temps de 10 secondes est laissé au participant pour qu'il choisisse sa réponse parmi les quatre possibilités.

3.2.4.1 Procédure

Le test a été passé collectivement. Après l'écoute des consignes deux exemples ont été proposés pour s'entraîner et les participants ont reçu un *feed-back* immédiat. Les consignes ont été : « Du wirst ein «beep» hören, dann ein Wort, das wiederholt wird. Nachdem du es zweimal gehört hast, musst du die richtige Schreibweise (Orthographie) auf Französisch wählen und diese auf dem Blatt ankreuzen.» *Tu vas entendre un beep puis un mot qui est répété juste après. Après avoir entendu deux fois le mot, tu dois choisir comment s'écrit ce mot en français. Lis bien les quatre orthographes qui te sont proposées et choisis celle que tu penses être l'orthographe en français du mot entendu.*

3.2.4.2 Les stimuli

Les stimuli utilisés sont des pseudo-mots. L'utilisation de lexèmes aurait présenté de nombreux désavantages. Premièrement, la lecture de mots peut se réaliser par la voie directe, de la forme globale du lexème à sa forme phonologique, et dans ce cas ce ne sont pas les correspondances phonème-graphème (=CPG) qui sont mesurées mais l'acquisition des formes orthographique et phonologique globales (Nakayama et al., 2012). Deuxièmement, la lecture de lexèmes est influencée par leur fréquence, et il est toujours difficile de contrôler la fréquence exacte d'un lexème dans l'environnement d'apprentissage d'un apprenant. Troisièmement, utiliser des lexèmes aurait introduit un biais pour la comparaison entre les deux séquences didactiques car seule la séquence communicative utilise la valeur sémantique. Quatrièmement, les lexèmes ne permettaient pas de contrôler les paramètres phonotactiques souhaités.

Les stimuli sont bisyllabiques car cette structure est fréquente en français et en allemand (Sprenger-Charolles, 2004). Ils ne contiennent aucun cluster consonantique qui complexifierait la lecture des formes orthographiques et ont majoritairement un squelette CV-CV (CV=consonne-voyelle). Le nombre d'interférences cross-linguistiques a été limité à une par stimulus.

Les interférences cross-linguistiques entre le français et l'allemand se produisent quand un graphème correspond à un phonème différent dans chacune des langues, ou quand un phonème du français est assimilé à un autre phonème au moment de sa perception. Trois types de relations unissent les deux systèmes de CPG et déterminent la nature des interférences.

Premièrement, les graphèmes correspondent à des phonèmes communs. Pour les voyelles, c'est le cas de <a, e, i, o> et pour les consonnes de <m, n, l, r>. Ces graphèmes sont utilisés pour les locus neutres des pseudo-mots. Les consonnes occlusives médiales et antérieures auraient pu être explorées également pour la réalisation ou non du voisement mais elles ont été utilisées comme support de pseudo-mot seulement (<b, d, p, t>).

Deuxièmement, les phonèmes français n'appartiennent pas au répertoire d'un germanophone, comme les deux phonèmes cibles de cette étude /ʃ/ et /ʒ/, mais aussi pour les voyelles, les trois autres voyelles nasales /œ - ɛ̃ - ã/, les semi-voyelles <oi> - /w/ et <ui> - /ɥ/ et pour les consonnes <gn> - /ɲ/. Des erreurs phonologiques d'assimilation sont attendues.

Troisièmement les phonèmes sont communs aux deux systèmes mais leur correspondant graphémique diffère (voir tableau 7). Des erreurs orthographiques et phonotactiques sont attendues.

Tableau 7 : Comparaison des représentations graphémiques dans les deux langues d'un phonème commun

Graphème allemand	Phonème commun	Graphèmes français correspondant au phonème					
		% d'utilisation décroissant -> (Catach 1995)					
sch	/ʃ/	ch					
j	/j/	i+V	V- y- V	V/C +ill			
k	/k/	qu	c+ a/o/u	cqu			
Ss ou s (int)	/s/	s-	sc-	cs	ce ; ci	ç	-tion -tier
w	/v/	v					
s	/z/	V-s-V	Z intérieur mot	x-ième			
g /ie	/g/	gu					
e	/E/	é, è	ai	es*, est*, et*	ez, er, ei	ex	Ê, ë
o	/o/	o	au (x)	eau (x)			

Légendes : * Formes les plus fréquentes dans la méthode des apprenants ; V = voyelle ; C = consonne

Les erreurs produites par les différents types d'interférences cross-linguistiques sont proposées dans les trois réponses fausses. Premièrement, les interférences phonologiques à l'origine des erreurs d'assimilation constituent les erreurs phonologiques (leurre 1). La différence entre deux

voyelles nasales n'a pas été mesurée car même des apprenants avancés peinent à l'entendre. De même la correspondance graphémique <ch> - /k/ (e.g., *chorale*) n'a pas été incluse car l'objet de cette étude est l'interférence représentée par la correspondance erronée <ch>-/ʒ/. Deuxièmement, les interférences induites par des CPG différentes entre les deux langues constituent les erreurs de type phonotactique (leurre 2). Troisièmement, les erreurs orthographiques (leurre 3) ne sont pas liées au système germanophone mais sont induites par les règles orthographiques, graphotactiques et phonotactiques du français (e.g., /ʒi/ ne peut pas s'écrire <ji>, en fin de mot /aʒ/ s'écrit toujours <age> et jamais <agge>, <aje>, <ag> ou <aj>). Le quatrième type de pseudo-mot proposé est la bonne réponse.

Les réponses justes respectent les règles graphotactiques et phonotactiques du français. Elles contiennent les phonèmes cibles dans toutes les positions autorisées sur le squelette (toutes les positions dans une syllabe et toutes les syllabes). Le phonème /ʒ/ est étudié grâce à cinq stimuli dont un qui contient un <t> final muet. Le phonème /ʒ/ est étudié grâce à six stimuli dont deux se terminent par <ge>. Réponses justes et leures sont présentés en ordre aléatoire sur la ligne des réponses. Deux ordres aléatoires des quatre choix de réponses par stimulus constituent les deux types de test (A et B) qui ont été répartis aléatoirement entre les participants. La non-existence de tous les pseudo-mots proposés a été vérifiée sur larousse.org et leo.org.

3.2.4.3 Méthode d'évaluation

Le participant a dû cocher une seule réponse. Chaque pseudo-mot coché est marqué 1 point. L'analyse statistique de cette tâche a été conduite à l'échelle de chaque pseudo-mot ayant servi de leurre phonologique, orthographique ou phonotactique (=erreurs).

3.2.5 Tâche de mesure du vocabulaire

De nombreux tests existent pour mesurer le vocabulaire (voir Annexe 5). Il est possible de mesurer la taille ou la profondeur du vocabulaire. La mesure de la taille du vocabulaire est fortement corrélée au niveau de compétences en LE (Laufer & Goldstein, 2004). Elle est par conséquent adaptée pour mesurer l'impact d'une intervention didactique visant à développer plusieurs compétences liées à l'acquisition du vocabulaire. De plus, les participants de cette étude sont tous des débutants, qui disposent d'une taille de vocabulaire inférieure à deux mille lemmes. Par conséquent, la valeur de la taille et celle de la profondeur du vocabulaire sont deux indicateurs équivalents pour évaluer le niveau global de vocabulaire (voir Schmitt, 2014, pour une revue). La taille du vocabulaire peut être mesurée en production ou en réception.

Une tâche mesurant la taille du vocabulaire réceptif est la plus facile à mettre en œuvre avec des participants ayant un niveau réduit de vocabulaire. Deux types de test sont disponibles qui font appel à deux types de tâches : Il s'agit soit de relier la valeur conceptuelle représentée par une image au lexème (e.g., Peabody), soit de décider si le mot est connu ou non (les « yes/no » tests comme par exemple le X-Lex). Les deux types de tests ont été comparés selon des critères de validité, fiabilité et leurs aspects pratiques (voir annexe 5 pour plus de détails). Le format

« yes/no » semble le plus adapté à mon étude car le choix ciblé des lexèmes permet d'évaluer plus précisément la contribution des représentations phonologiques acquises dans l'acquisition du vocabulaire réceptif. D'ailleurs, O'Dell, Read et McCarthy (2000) recommande de tester le vocabulaire sous forme de lexème isolé avec un format Yes/No pour réaliser un suivi de l'acquisition du vocabulaire en classe.

Les caractéristiques du test Yes/No utilisé dans cette étude sont identiques à celles du X-LEX (Meara & Milton, 2003). Il est composé d'une liste de 80 lexèmes et de 20 pseudo-mots présentés sous forme écrite. Le sujet doit déterminer s'il connaît le stimulus et en cas de réponse positive cocher le oui. Le test créé a été adapté pour répondre aux questions de recherche et à mes apprenants.

3.2.5.1 Procédure

Le test a été passé collectivement. Après avoir écouté les consignes en allemand, deux exemples ont été réalisés avec les participants à titre d'entraînement. Ils ont été suivis d'un *feed-back* immédiat. Les consignes, écrites sur le test, ont été écoutées : « Alle untenstehenden Wörter sind in französischer Sprache, mit Ausnahme von einigen, die sehr ähnlich sind jedoch nichts bedeuten. Lese jedes Wort, spreche es in deinem Kopf aus, und wenn du denkst, dass es etwas bedeutet, kreuze «Ja» an. Hast du es nie gehört, kreuze «Nein» an. Wir antworten zusammen: zum Beispiel lesen wir das erste Wort, *pour* und das zweite Wort *fatine*.» *Tous les mots ci-dessous sont des mots du français... sauf quelques-uns qui leur ressemblent fortement! Lis chaque mot, écoute-le dans ta tête et si tu penses que tu le connais, coche oui. Si tu ne l'as jamais entendu, coche non.*

3.2.5.2 Les Stimuli

Les stimuli sont des lexèmes et des pseudo-mots présentés isolément oralement et écrits sans majuscule sur la feuille de réponse. Les lexèmes retenus (total= 80) ont été choisis parmi les lexèmes les plus fréquents de la méthode d'enseignement des participants de cette étude. La moitié des lexèmes (=40) contiennent le phonème cible /ʒ/ et l'autre moitié le phonème cible /z/. La moitié des lexèmes de chaque phonème cible (=20) est uniquement utilisée pour le test (lexèmes réservés) et l'autre moitié sert de support aux interventions pédagogiques. La mesure de l'apprentissage des lexèmes non enseignés peut indiquer l'effet facilitateur des nouvelles connaissances phonologiques dans l'acquisition du vocabulaire réceptif.

Le transfert de connaissances lexicales de la L1 vers la LE n'a pas été favorisé. Les cognats ont été dans la mesure du possible écartés car il est difficile d'estimer le niveau de vocabulaire lorsque leur proportion dans le test est supérieure à celle qu'ils occupent dans la langue (Meara, Lightbown & Halter, 1994). Le mot peut alors être reconnu pour sa valeur L1 ou LE. Cependant le faible nombre de lexèmes présents dans la méthode d'enseignement et contenant l'association /z/- <g> m'a obligée à intégrer trois cognats. Ils ont été inclus dans les lexèmes travaillés et testés ce qui a permis de travailler la prononciation de <gi>-/zi/ spécifique du français. Les

participants de cette étude n'ont en général pas semblé les reconnaître pendant les séquences d'enseignement. La fréquence des lexèmes a aussi été un critère de sélection.

La fréquence des lexèmes dans la méthode des apprenants a été prise en compte pour qu'elle soit similaire entre lexèmes travaillés et non travaillés pour chacun des phonèmes cibles et des phonèmes de leur paire minimale (voir Annexe 5). En effet, un lexème plus fréquent a plus de chances d'être appris et un déséquilibre de fréquence aurait constitué un biais pour l'interprétation des résultats. La catégorie grammaticale est aussi un facteur influençant l'apprentissage du lexique.

Les lexèmes ont été choisis pour que les catégories grammaticales soient à peu près similairement représentées dans les lexèmes travaillés des séquences didactiques et ceux réservés pour le test (voir Annexe 5). De la même façon, cinq pseudo-mots présentent des terminaisons qui laissent penser qu'ils pourraient être des verbes (i.e., <lépont, lidons, galter, faronné, sapoît>). Les pseudo-mots répondent également à certains des autres critères de sélection des lexèmes.

Les pseudo-mots sont 20 bisyllabiques de squelette CV-CV. La moitié (=10) sont porteurs des deux phonèmes cibles. Les phonèmes cibles dans les pseudo-mots occupent toutes les positions autorisées. Tous les pseudo-mots respectent les règles phonotactiques du français. Ils ont été construits à partir de syllabes existantes. Leur inexistence en français et en allemand a été vérifiée dans les dictionnaires en ligne (larousse.fr et leo.org).

L'ordre de présentation des pseudo-mots et des mots dans le test a été établi selon deux ordres aléatoires distincts A et B. Les deux listes A et B contenant les 80 lexèmes et 20 pseudo-mots ont été aléatoirement réparties entre les participants.

3.2.5.3 Méthode d'évaluation

La façon de calculer le score du vocabulaire réceptif a été largement discutée dans la littérature et reste encore un sujet de débats (voir annexe 5 pour un récapitulatif de la littérature).

Les lexèmes reconnus et les pseudo-mots erronément reconnus marquent 1 point. Le stimulus non reconnu marque zéro. Pour l'analyse de la tâche, l'analyse statistique est faite à l'échelle du stimulus c'est-à-dire des leurres erronés qui ont été choisis (=erreur). La correction par les pseudo-mots erronément reconnus constitue un facteur inclus dans l'analyse statistique.

3.3 Protocole expérimental

L'objectif est de répondre à la question de recherche définie préalablement : « Dans quelle mesure une didactique explicite sans support sémantique et une didactique implicite communicative actionnelle peuvent favoriser l'acquisition des représentations phonologiques et celles du vocabulaire réceptif associé ? ». Pour répondre à cette question, 127 participants ont été recrutés et l'expérimentation s'est déroulée en milieu scolaire.

3.3.1 Participants

Dans un contexte scolaire, l'échantillon de cette étude longitudinale a inclus un nombre de participants important en comparaison des études présentes dans la littérature. Pour un pouvoir statistique suffisant, l'objectif était d'inclure une centaine de participants. La taille de l'échantillon devait cependant être compatible avec un enseignement par une seule enseignante pour limiter les biais induits par la variation de l'enseignement. Le nombre de participants recrutés est de 127 élèves. La moyenne d'âge est de 12;6 (sd= 6). Les participants sont 78 filles et 49 garçons. Ils appartenaient à 3 classes de 8H (n=64) et de 3 classes de 9H (n=63). Les 63 élèves de 9H représentaient 3 niveaux différents : pré-gymnasial (n=23) ; général (n=22) et exigences de base (n=18). Les élèves de niveau général et de niveau exigences de base étaient mélangés dans des classes mixtes. Quels que soient la classe et le niveau, tous les élèves ont suivi les mêmes séquences didactiques.

Le nombre de langues des participants et leur diversité (3 participants dyslexiques ont été inclus), ainsi que le temps disponible ne permettait pas de prendre en compte toutes les interférences cross-linguistiques dans les séquences didactiques. Tous les participants étaient bilingues Schwytzerdeutsch et allemand standard suisse. 39% ont déclaré parler une troisième langue plus d'une fois par semaine en dehors de l'école. Seules les interférences cross-linguistiques français-Schwytzerdeutsch et français-allemand standard suisse ont été prises en charge dans les séquences d'apprentissage.

Le système phonologique et grapho-phonémique de l'allemand est en permanence fortement activé par la pédagogie proposée par la méthode scolaire d'enseignement du FLE. L'enseignement dédié au vocabulaire est centré sur les lexèmes complexes du texte. La méthode encourage deux stratégies des apprenants : la recherche de cognats et la traduction en allemand. Ces deux stratégies, centrales dans cette méthode scolaire, accompagnent avec un support écrit tous les apprentissages et en particulier ceux du vocabulaire.

Le vocabulaire utilisé dans la méthode utilise toute la complexité du système phono-graphémique dès la première leçon. Le vocabulaire n'appartient pas uniquement aux groupes de lexèmes les plus fréquents du français (tranches des mille et deux mille lexèmes les plus fréquents). En effet la méthode scolaire est basée sur des textes authentiques d'un niveau C1 (i.e., demandant pour être compris de connaître les 9000 lexèmes les plus fréquents, Nation, 2006) et qui, par conséquent, utilise du vocabulaire contenant des structures orthographiques complexes (e.g., *côtoyons* en première année de FLE). L'opacité du système phono-graphémique du français n'est pas prise en compte dans cette méthode alors que le support écrit est présent dès le premier jour d'enseignement. La demande cognitive est très importante pour gérer les interférences entre les systèmes phonologiques et phono-graphémiques, d'autant plus que les apprenants sont débutants en FLE.

Les deux phonèmes cibles de cette étude représentent deux types d'interférences difficiles à surmonter. Les interférences phonologiques affectent les deux phonèmes cibles. Les apprenants

font des erreurs de prononciation et ont des difficultés de discrimination entre /z/ et /ʒ/ et entre /O/ et /ɔ/. Les interférences phono-graphémiques sont particulièrement complexes pour /z/ et sont renforcées par la méthode d'enseignement car les apprenants sont encouragés à utiliser l'interface de leur langue première et à étayer leur apprentissage préférentiellement avec le support écrit. Aucune étude, à ma connaissance, a reporté combien d'heures d'enseignement sont nécessaires pour dépasser ces difficultés.

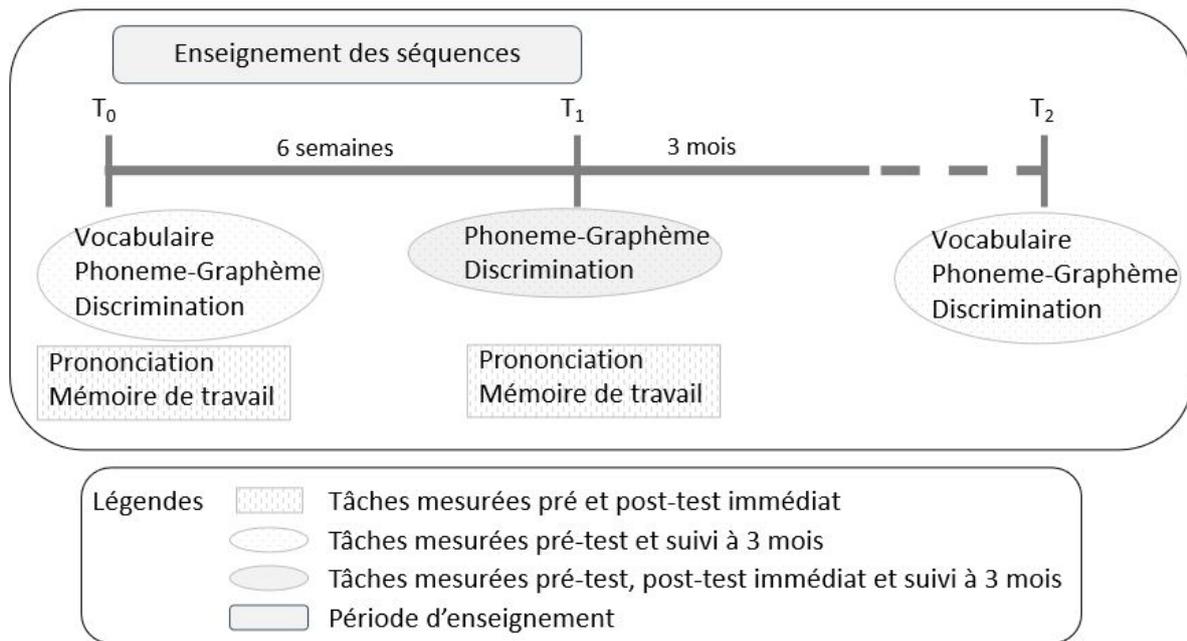
Les participants ont déjà reçu entre 312 heures (8H) et 390 heures (9H) d'enseignement du FLE au début de l'expérimentation. Or les interférences étudiées persistent après 330 heures d'enseignement du FLE (Dherbey Chapuis & Berthele, 2020). Pour tenter de dépasser ces interférences deux séquences didactiques ont été construites et les participants ont été aléatoirement répartis dans chaque séquence.

3.3.2 Procédure expérimentale

Les participants de chaque classe ont été aléatoirement répartis dans deux groupes. Chaque groupe a suivi l'intégralité des douze séances d'une séquence didactique (i.e., soit E (= explicite), soit C (= implicite ciblée communicative actionnelle). Une séance dure 20 minutes. L'ensemble des douze séances a représenté 4 heures d'enseignement ciblé et leur enseignement s'est étalé sur 6 semaines de classe. L'enseignant.e en charge du cours a travaillé en demi-classe pendant que l'expérimentatrice enseignait une séance d'une des deux séquences. Les groupes ont été échangés à la moitié de l'heure de cours. L'ordre des groupes pris en charge par l'expérimentatrice a varié aléatoirement, mais chaque participant a suivi une seule des deux séquences didactiques. Les participants du groupe E ont toujours suivi une séance de la séquence explicite, et ceux du groupe C de la séquence communicative. L'effet de l'enseignement des deux séquences didactiques a été mesuré par différentes tâches.

Les mesures de performance des compétences cognitives utilisant des représentations phonétiques et phonologiques et du vocabulaire réceptif ont été effectuées avant l'enseignement, juste après et trois mois après la fin de l'intervention (figure 4). Ces trois phases de tests ont donc permis de mesurer cinq variables : la discrimination, la prononciation, le maintien et la répétition de pseudo-mots, la maîtrise des relations phonème-graphème et le vocabulaire réceptif. Les tâches permettant la mesure d'une même variable sont identiques entre les différents moments de mesure.

Figure 4 : Planification des tests



Les tâches collectives ont été effectuées au pré-test (T₀), dans les trois jours après la fin de l'intervention (T₁) et trois mois après la fin de l'intervention (T₂). Les tâches collectives sont le vocabulaire réceptif, les relations phonème-graphème et la discrimination. Seul le vocabulaire n'a pas été mesuré au post-test immédiat car il fallait écarter le risque que les apprenants se souviennent des stimuli 6 semaines après avoir effectué le test de vocabulaire. L'ordre de passage des tâches collectives étaient : 1) questionnaire de données personnelles (au pré-test uniquement), 2) discrimination, 3) relations phonème-graphème et 4) vocabulaire.

Les tâches individuelles ont été effectuées au pré-test (T₀), dans les trois jours après la fin de l'intervention (post-test immédiat = T₁). Elles n'ont pas été effectuées trois mois après la fin de l'intervention car leur organisation pendant les heures de classe n'a pas été possible trois fois. Les tâches individuelles mesurent la prononciation et le maintien en mémoire de travail. L'ordre de passage des tâches individuelles étaient : 1) échanges verbaux en français, 2) mémoire de travail, 3) prononciation.

Chaque tâche a été effectuée sous deux formes (A et B) qui diffèrent par un changement de l'ordre aléatoire des stimuli. Un participant a toujours passé le test sous une même forme (i.e., A ou B). Les deux formes de la même tâche ont été aléatoirement réparties parmi les participants lors de l'expérimentation.

L'expérimentation s'est déroulée sur deux années scolaires. La première année une classe de 8H et une classe de 9H ont participé (n= 45). Le protocole a pu être partiellement réalisé la première année et les tâches de prononciation et de discrimination n'ont pas été effectuées. La deuxième année, deux classes de 8H et deux classes de 9H ont participé (n= 82) et le protocole complet a pu être réalisé.

3.3.3 Pilote de l'expérimentation

Un pilote de l'expérimentation a été conduit dans une classe double de 7H et 8H. Les 18 élèves ont suivi le protocole expérimental. L'analyse du pilote a été faite pendant sa réalisation grâce à un carnet de bord, et après sa réalisation par l'analyse des résultats des élèves et par des échanges avec l'enseignante. Deux modifications principales ont été conduites suite aux conclusions de cette analyse. Premièrement, les séquences didactiques ont été améliorées en intégrant les observations des progrès ou non des élèves, leur motivation et leur participation. Les critiques et suggestions de l'enseignante ont également été prises en compte. Deuxièmement, les tests ont été légèrement modifiés. Le temps entre les stimuli a été réduit pour maintenir l'attention des participants concentrée sur le test mais ne pas introduire de pression cognitive induite par le manque de temps. Les pseudo-mots qui n'avaient pas été cochés par un seul élève ont été modifiés dans la tâche de vocabulaire et dans la tâche de mesure des correspondances phonème-graphème.

L'analyse descriptive du pilote esquissait déjà les résultats de l'analyse statistique des tests relatifs à l'expérimentation.

4 Résultats

Pour chaque tâche l'analyse des résultats a été descriptive et observatoire avant d'être modélisée dans R (version 3-6-3) (R Foundation for Statistical Computing, 2020). L'ensemble de ces résultats est disponible pour chaque tâche (voir OSF). Cependant seuls les résultats pertinents pour la question de recherche sont présentés. La question de recherche est : « Dans quelle mesure une didactique explicite sans support sémantique et une didactique implicite communicative actionnelle peuvent favoriser l'acquisition des représentations phonologiques et celles du vocabulaire réceptif associé ? ». Les résultats sont présentés par tâche. Les analyses statistiques inférentielles ont pour but de pouvoir généraliser les observations à la population présentant les mêmes caractéristiques que les apprenants de cette étude.

Les apprenants sont inclus dans des classes et des écoles différentes mais ces effets sont contrôlés par le protocole car chaque classe a été divisée en deux groupes C et E. Par convention les apprenants appartenant au groupe ayant suivi la didactique communicative actionnelle sont nommés groupe C et au groupe ayant suivi la didactique explicite groupe E. Les différences individuelles sont prises en compte par les effets aléatoires (=1|participant). Les effets des différences entre stimuli sont inclus dans les effets aléatoires pris en compte au niveau individuel. Tous les stimuli ont été présentés à tous les participants (*nested*) et dans ce cas la recommandation est de ne pas les inclure dans les effets aléatoires. "If items are truly nested, then the simpler model with only *Subject* as random effect is appropriate" (Baayen, 2008, p.274).

Les effets des variables indépendantes pertinentes ont été inclus dans les effets fixes des modèles.

La construction du modèle adéquat et sa vérification ont suivi certaines règles. Seuls les facteurs pertinents pour la question de recherche ont été inclus dans les modèles statistiques. Les interactions observées lors de l'analyse exploratoire ont été incluses. Lorsque plusieurs modèles étaient possibles la sélection a été effectuée par la méthode de sélection du critère Aikai A.I.C. (Burnham & Anderson, 2002; Myung, Forster & Browne, 2000). La normalité des résidus, les effets des outliers et les résidus du modèle par variable indépendante ont été contrôlés.

La modélisation compare deux hypothèses :

H0 : Aucun effet de l'intervention n'est observable pour chacune des compétences observées. Les deux séquences didactiques n'ont pas d'effet sur l'acquisition des représentations phonologiques des deux phonèmes cibles ou du vocabulaire réceptif associé.

H1 : Les deux séquences didactiques ont un effet différent sur l'acquisition des représentations phonologiques des deux phonèmes cibles ou du vocabulaire réceptif associé.

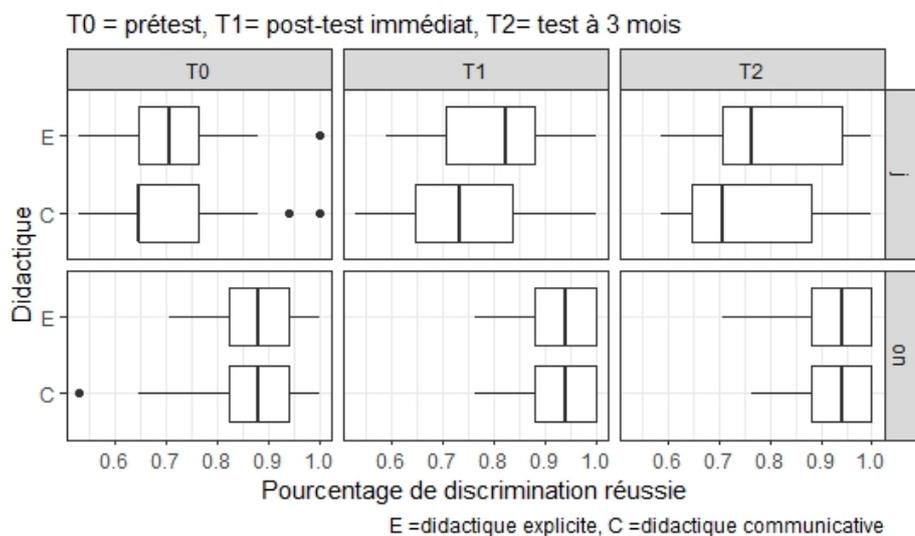
4.1 Résultats pour la tâche de discrimination

Lors de la présentation orale d'une paire de pseudo-mots, le participant a dû déterminer si les deux éléments du stimulus étaient similaires ou différents (voir 3.2.1). Une réponse juste a été valorisée par un point ($eva = 1$ point) et fautive zéro point ($eva = 0$). Quinze paires de pseudo-mots ont été présentées par phonème cible dont un tiers de paires identiques. Le même test a été conduit 3 fois au pré-test (T0), post-test immédiat (T1) et trois mois après la fin de l'enseignement (T2).

4.1.1 Analyse descriptive et observatoire

Les résultats recueillis ont été représentés sous forme de box-plot visualisant la médiane et le premier et troisième quartile de la distribution du taux de réussite (figure 5).

Figure 5 : Discrimination moyenne par séquence didactique et par phonème



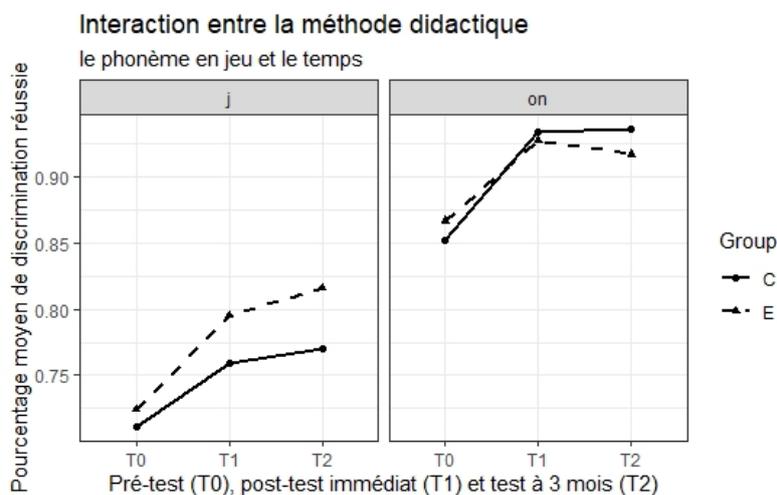
Au pré-test (=T0), le groupe E réussit légèrement mieux en moyenne la tâche de discrimination pour les deux phonèmes (figure 5 et tableau 8).

Tableau 8 : Pourcentage observé moyen de discrimination réussie par groupe, par test et par phonème

Séquence didactique	Phonème	Pré-test (=T0)	Post-test Immédiat (=T1)	Différence T1 - T0	Post-test à 3 mois (=T2)	Différence T2 - T0
C	/ʃ/-on	85,2	93,4	+ 8,2	93,5	+ 8,3
E	/ʃ/-on	86,7	92,7	+ 6	91,7	+ 5
C	/ʒ/-j	71,2	75,9	+ 4,7	77,0	+ 5,8
E	/ʒ/-j	72,5	79,5	+ 7	81,6	+ 8,1

Le progrès (différence avec T0) est comparé entre les deux groupes. Au post-test immédiat (=T1), le progrès est supérieur dans le groupe C pour le phonème /ʃ/-on (C = 8 % > E = 6 %) et inversement le progrès est supérieur dans le groupe E pour le phonème /ʒ/-j (C = 5% < E = 7 %). Au post-test à 3 mois (=T2), les résultats sont à peu près ceux de T1. Le progrès est supérieur dans le groupe C pour le phonème /ʃ/-on (C = 8 % > E = 5 %), et inversement le progrès est supérieur dans le groupe E pour le phonème /ʒ/-j (C = 6 % < E = 9 %). On observe une interaction entre le groupe, le phonème et le temps (voir figure 6).

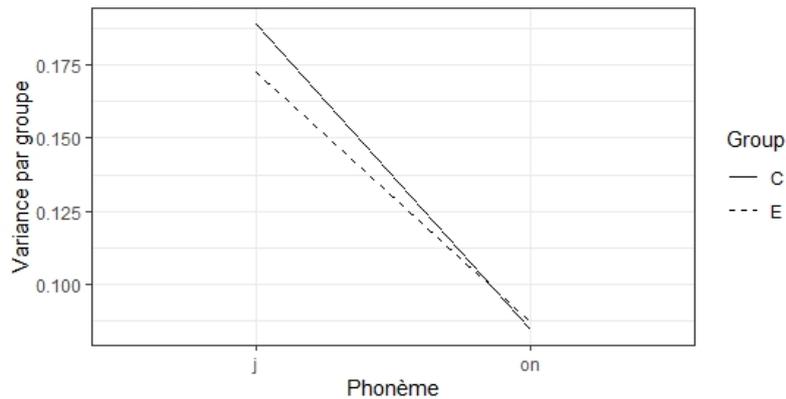
Figure 6 : Interaction entre la méthode didactique, le phonème en jeu et le temps



Il est important de vérifier que la variance au sein de chaque groupe est équivalente pour pouvoir comprendre et modéliser cette interaction.

La variance est différente entre les phonèmes mais est similaire dans chaque groupe (figure 7). La comparaison des groupes pour chaque phonème n'est donc pas problématique.

Figure 7 : Variances de la probabilité d'une discrimination réussie



4.1.2 Analyse statistique⁵

Deux types de variables sont possibles pour l'analyse statistique d'une tâche de discrimination : soit à l'échelle de chaque réponse, soit en suivant les recommandations de la *Signal Detection Theory* et en calculant un score par participant (e.g., Green & Swets, 1966; Macmillan & Creelman, 2005; McNicol, 1972). Les modèles d'analyse utilisant la *Signal Detection Theory* sont une sous-classe de modèle logistique (DeCarlo, 1998) mais présentent certains désavantages.

L'analyse statistique à partir des valeurs définies par la *Signal Detection Theory* n'a pas été choisie pour plusieurs raisons. Un score global ne permet pas de tenir compte des différences de variance (DeCarlo, 2010, 2012), et ne permet pas d'inclure l'influence d'autres prédicteurs (Wright et al., 2009) comme par exemple les différences entre les deux phonèmes cibles. Les données sont en général mieux décrites par un modèle mixte généralisé (DeCarlo, 2012).

L'analyse statistique a été conduite par un modèle linéaire mixte généralisé (=glmer, Bolker, 2015; Bolker, Brooks, Clark, Geange, ... & White, 2009). Le modèle évalue la probabilité d'une réponse juste par stimulus. La variable dépendante (=eva) est la valeur attribuée à la réponse (juste $eva = 1$ ou fausse $eva = 0$) lors de la présentation d'une paire de pseudo-mots. Les facteurs fixes, variables indépendantes, sont l'appartenance du participant au groupe (groupe= {E ; C}), le moment du test (Time = {T0 ; T1 ; T2}) et l'interaction entre le phonème et le groupe observée dans l'analyse descriptive (groupe*phone). Les effets aléatoires représentent la part de variance inhérente à chaque individu (1|participant).

Les contrastes sont codés :

- la séquence didactique comme une *dummy* variable : $n.group = \{ E = -0.5 ; C = 0.5 \}$
- le phonème comme une *dummy* variable : $n.phone = \{ j = -0.5 ; on = +0.5 \}$
- le test *simple coding* : $Time = \{ T0 = 0 \text{ (référence)} ; T1 = (0 ; 1 ; 0) ; T2 = (0 ; 0 ; 1) \}$

L'interaction séquence didactique et phonème est significative quand elle est exprimée en log-odds (0.3 ; 95% confidence intervalle : [0.04; 0.56], $Z = 2.295$, $p = .022$).

⁵ Voir Annexe 6 pour l'analyse statistique et descriptive complète

Le tableau donne les valeurs estimées des effets fixes (tableau 9). L'intercept représente la chance relative (*Odds ratio*) ou la probabilité (*Probabilités*) de discriminer les phonèmes (en moyenne pour les deux phonèmes), au pré-test et pour l'ensemble des apprenants.

Tableau 9 : Effets fixes transformés en odds ratios et en probabilités d'une discrimination réussie

<i>Prédicteurs</i>	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>	<i>Probabilités</i>	<i>CI</i>	<i>p</i>
(Intercept)	4.00	3.55 – 4.50	< .001	0.80	0.78 – 0.82	< .001
n.group	0.94	0.78 – 1.13	0.487	0.48	0.44 – 0.53	0.487
n.phone	3.19	2.80 – 3.64	< .001	0.76	0.74 – 0.78	< .001
Time [T1]	1.57	1.36 – 1.82	< .001	0.61	0.58 – 0.65	< .001
Time [T2]	1.68	1.45 – 1.95	< .001	0.63	0.59 – 0.66	< .001
n.group * n.phone	1.36	1.05 – 1.76	.022	0.58	0.51 – 0.64	.022
Effets aléatoires						
σ^2	3.29			3.29		
τ_{00}	0.09			0.09		
ICC	0.03			0.03		
N	76			76		
Observations	7752			7752		
Marginal R ² / Conditional R ²	0.105 / 0.128			0.105/ 0.128		

Modèle : `glmer(formula = eva ~ n.group*n.phone + Time + (1|participant), family = binomial(link="logit"), data = d3O)`

Légendes : CI = intervalle de confiance ; n.group = { E= -0.5 ; C= 0.5 } ;
n.phone = { j= -0.5 ; on= +0.5 } ;
T1 = post-test immédiat ; T2 = post-test à 3 mois

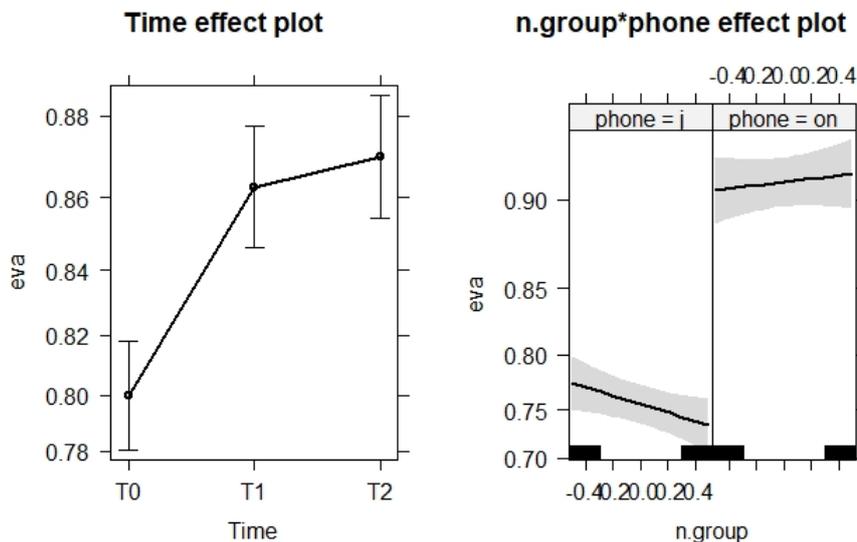
L'*odd* (i.e., la chance de donner une bonne réponse par groupe) est calculé en faisant la somme des réponses justes divisée par le nombre total de réponses pour chaque groupe. L'*odd ratio* est le ratio des chances de donner une réponse juste, et est obtenu en divisant l'*odd* d'un groupe par celle de l'autre. La probabilité de donner une bonne réponse est pour chaque groupe égale à $odd/(1+odd)$. Les deux résultats du modèle *odds ratio* et *probabilités* représentent des estimations similaires mais les probabilités sont plus faciles à comprendre.

Les estimations du modèle exprimées en $\log(odds)$ ont permis de calculer la probabilité de discriminer correctement une paire minimale au post-test immédiat (=T1) et au post-test à trois

mois (voir Annexe 6 pour l'étude statistique complète « discrimination » et le fichier excel pour le calcul « calcul_estimates_discrimination »).

Entre le pré-test et les post-tests, les participants ont significativement progressé ($p < .001$) (voir tableau 9 et figure 8, « Time effect plot »).

Figure 8 : Représentation graphique des effets fixes du modèle⁶



Légendes : eva = probabilité d'une discrimination réussie
n.group = -0.5= groupe E ; +0.5= groupe C
T0 = Pré-test, T1 = Post-test immédiat, T2 = Post-test à 3 mois

Le progrès est stable et durable car il se maintient à T2.

Aux post-tests, les participants ont, en moyenne pour les deux groupes, progressé de 9% (immédiat= T1) et 10% (à 3 mois= T2) de chances supplémentaires de réussir à discriminer le phonème /ʒ/-j et de 4% pour le phonème /ʃ/-on. Les intervalles de confiance entre T0 et T1 ne se chevauchent pas pour le phonème /ʒ/-j, et entre T0 et T2 pour les deux phonèmes ce qui confirme la progression (tableau 10).

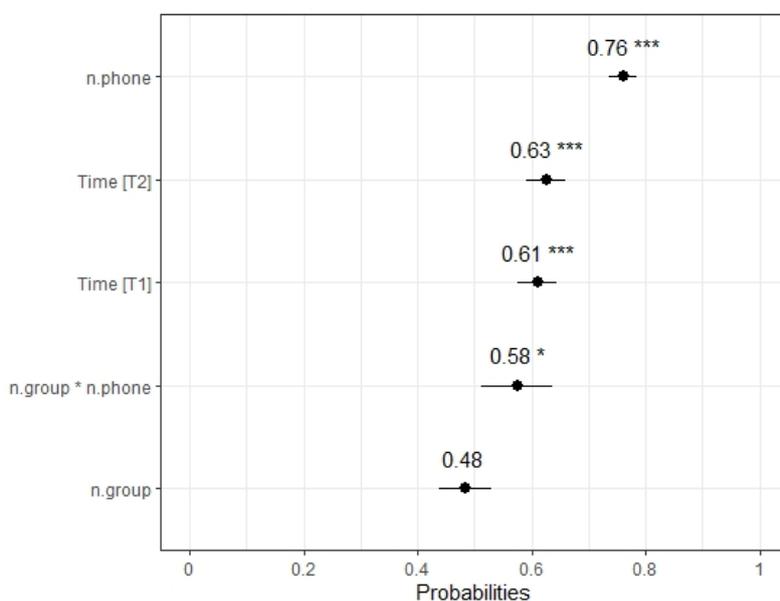
⁶ La partie n.group*phone correspond au même modèle mais avec un codage phone= { j= 0 ; on =1} car le codage utilisé pour la modélisation prend des valeurs continues pour n.phone entre [-0.5 ; 0.5] et les représentations graphiques sont illisibles.

Tableau 10 : Probabilités estimées d'une discrimination réussie et leur intervalle de confiance

Séquence didactique	Phonème	Pré-test (=T0)	Post-test immédiat (=T1)	Différence T1-T0	Post-test à 3 mois (=T2)	Différence T2-T0
E	/z/-j	70 [68 ;71]	80 [77 ;82]	+ 10%	81 [78 ;83]	+ 11%
E	/ʒ/-on	88 [87 ;89]	92 [90 ;93]	+ 4%	92 [91 ;93]	+ 4%
C	/z/-j	69 [65 ;71]	76 [72 ;80]	+ 7%	77 [73 ;81]	+ 8%
C	/ʒ/-on	87 [84 ;90]	92 [89 ;94]	+ 5%	93 [88 ;95]	+ 6%
moy	/z/-j	69 [68 ;69]	78 [74 ;81]	+9%	79 [75 ;82]	+10%
moy	/ʒ/-on	88 [86 ;90]	92 [89 ;94]	+4%	92 [90 ;94]	+ 4%

Le phonème /ʒ/-on a 19% de chances en plus en moyenne d'être discriminé correctement que le phonème /z/-j au pré-test et 13 % au post-test à trois mois. Cette différence entre les phonèmes est significative ($p < .001$) (voir figure 8 « n.group*phone effect plot » et figure 9).

Figure 9 : Estimations des effets fixes du modèle relatif aux performances en discrimination

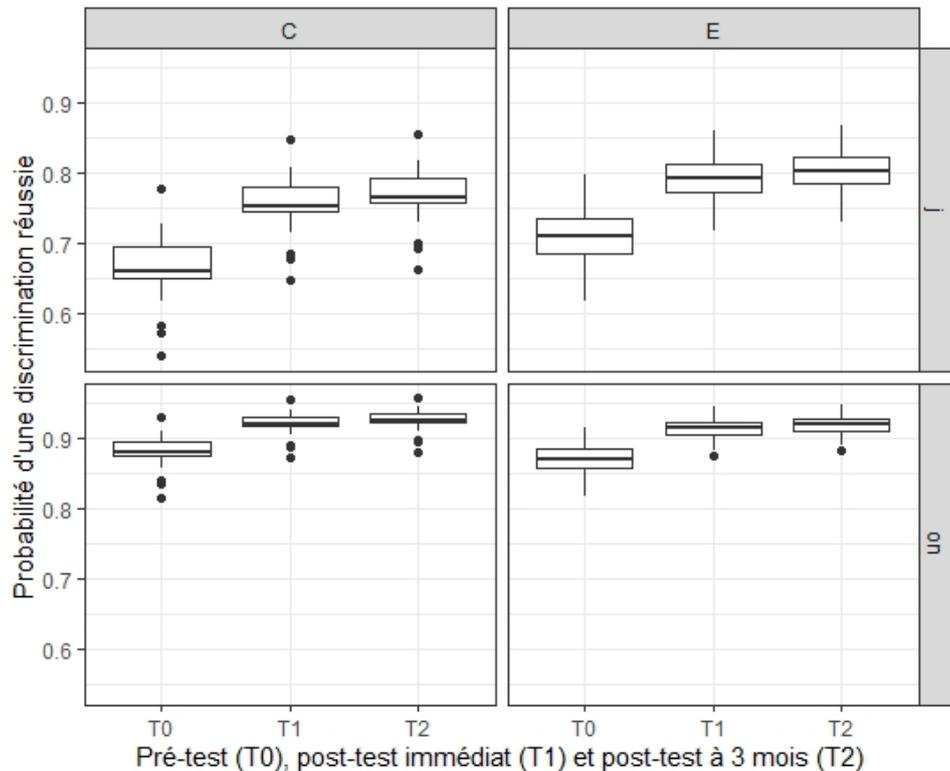


Légendes : $n.phone = -0.5 = /z/-j ; +0.5 = /ʒ/-on$
 $Time [T1] = \text{Post-test immédiat}, [T2] = \text{Post-test à 3 mois}$
 $n.group = -0.5 = \text{groupe E} ; +0.5 = \text{groupe C}$

La différence entre les groupes est non significative ($p > .05$), mais les groupes diffèrent significativement dans leur interaction avec le phonème enseigné ($p > .05$). Pour le phonème /ʒ/-on, le progrès est supérieur de 2% dans le groupe C en comparaison du groupe E. Pour le

phonème /ʒ/-j, le progrès est supérieur de 3% dans le groupe E en comparaison du groupe C (tableau 10, figure 8 « n.group*phone effect plot », figure 9 et figure 10).

Figure 10 : Prédictions par le modèle de la probabilité d'une discrimination réussie en fonction de l'interaction du phonème et de la séquence didactique sur les deux post-tests



Légendes : C = Groupe C, séquence communicative
E = Groupe E, séquence explicite

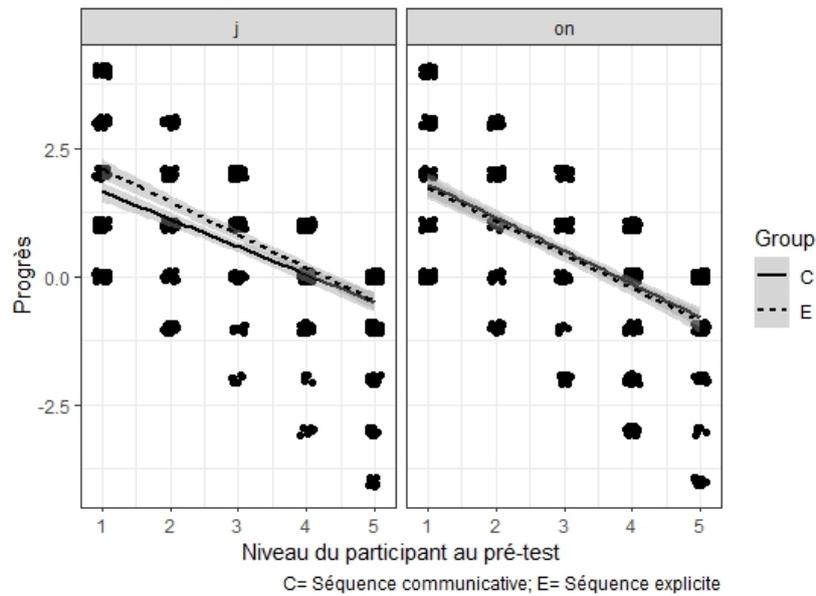
4.2 Résultats pour la tâche de prononciation

Le participant a répété un stimulus immédiatement après l'avoir entendu : au total 4 stimuli par phonème (voir 3.2.2). Son imitation du stimulus a été évaluée sur une échelle de Likert en 5 points par trois codeuses. La variable observée est la valeur de cette évaluation (=eva). Le même test a été conduit 2 fois au pré-test (=T0) et au post-test immédiat (=T1).

4.2.1 Analyse descriptive et observatoire

La valeur de l'évaluation au pré-test (eva à T0) a été soustraite à la valeur de l'évaluation au post test immédiat (eva à T1) pour calculer une nouvelle variable représentant le progrès (evaT1 - evaT0 = progrès). Le progrès observé a été représenté sur un graphique de points en fonction du niveau du participant au pré-test, de la séquence didactique et du phonème (figure 11).

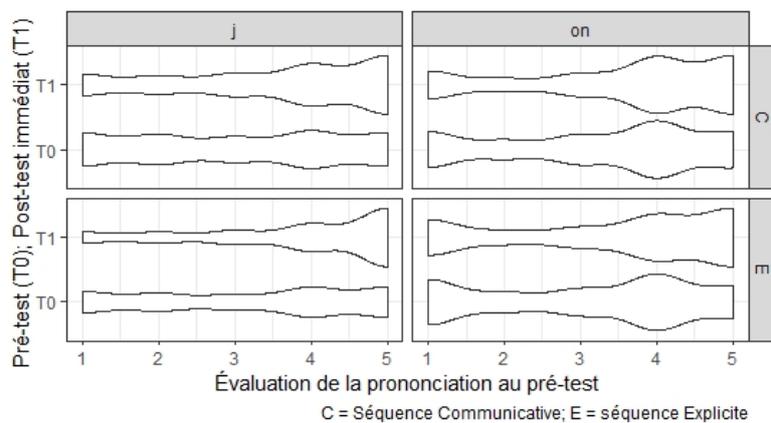
Figure 11 : Progrès de la prononciation en fonction de la séquence didactique et du phonème



Les progrès observés semblent identiques pour les deux séquences didactiques pour le phonème /ɔ̃/-on. Pour le phonème /ʒ/-j, la séquence explicite semble permettre un progrès plus important que la séquence communicative, sauf pour les apprenants obtenant la note maximale au pré-test (évaluation au pré-test >4), qui eux ont tendance à régresser dans les deux groupes. L'évaluation de l'apprenant au pré-test est donc un facteur indépendant à inclure dans le modèle.

La distribution des évaluations pour les deux tests a été représentée sous forme de violon en fonction des séquences didactiques et des phonèmes. Le progrès semble plus important pour le phonème /ʒ/-j (figure 12).

Figure 12 : Distribution de l'évaluation de la prononciation en fonction de la séquence didactique, du phonème et du test (pré-test ou post-test immédiat)



Pour les deux phonèmes, le groupe E a une moyenne de progrès supérieure au groupe C (tableau 11).

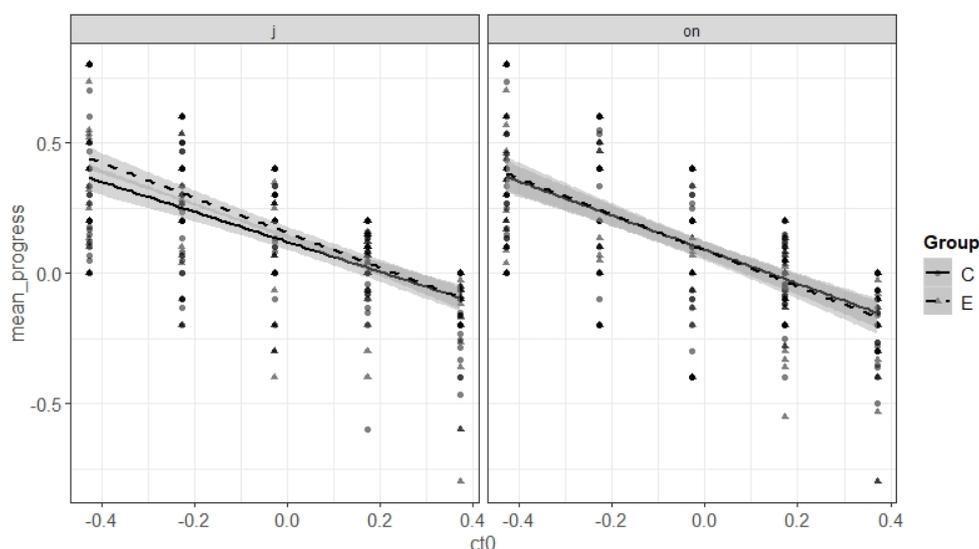
Tableau 11 : Moyenne du progrès entre pré- et post-test par séquence didactique et par phonème

Phonème	Séquence didactique	Évaluation moyenne T0	Évaluation moyenne T1	Progrès moyen
/ʃ/-on	C	3.22 (1.41)	3.58 (1.41)	+ 7 % (32)
/ʃ/-on	E	3.04 (1.46)	3.44 (1.52)	+ 8 % (34)
/ʒ/-j	C	3.05 (1.45)	3.6 (1.4)	+11 % (29)
/ʒ/-j	E	3.26 (1.46)	3.9 (1.33)	+13 % (31)

Légendes : Déviation standard = ()

En fonction du niveau au pré-test et du phonème, on observe de meilleures performances d'une ou de l'autre des séquences didactiques. Une interaction existe entre le phonème et la séquence didactique et le niveau à T0 (figure 13).

Figure 13 : Interaction entre la séquence didactique, le phonème et le niveau au pré-test en fonction de la moyenne centrée (=ct0)



Légendes : ct0 = moyenne centrée au pré-test

C = Séquence Communicative ; E = Séquence Explicite

4.2.2 Analyse statistique⁷

L'analyse statistique a été conduite par un modèle linéaire mixte (=lmer). La variable dépendante dans le modèle est le progrès. Le progrès est la différence entre le pré-test et le post-test de la valeur attribuée par chacune des trois codeuses à la prononciation d'un stimulus par

⁷ Voir Annexe 7 pour l'analyse statistique complète

un participant. La variable dépendante (=progrès) a été transformée en pourcentage pour permettre une modélisation par un modèle linéaire.

Les facteurs fixes, ou variables indépendantes, sont l'appartenance du participant au groupe (Group= {E ; C}), le phonème inclus dans le stimulus (Phoneme= {on ; j}), le niveau au pré-test du stimulus prononcé par un participant (=ct0) et les interactions entre le niveau au pré-test et la séquence didactique observée dans l'analyse descriptive (cT0*n.group) et entre le phonème et la séquence didactique (n.phone*n.group).

Le niveau du stimulus prononcé au pré-test par un participant (cT0) est la moyenne des évaluations des trois codeuses, qui a été standardisée par la moyenne de la population. À cause des contraintes liées aux modalités d'évaluation (échelle de Likert à 5 niveaux), ct0 peut prendre seulement cinq valeurs { -0.43 ; -0.23 ; 0 ; +0.17 ; + 0.37}. Comme la moyenne de l'échantillon est environ égale à 3, les valeurs de ct0 correspondent à l'évaluation au pré-test {1 ; 2 ; 3 ; 4 ; 5}. La variable ct0 est considérée comme continue pour la modélisation et les valeurs prises sont utilisées pour interpréter et calculer les estimations du modèle.

L'interaction entre la séquence didactique (=n.group) et le phonème (n.phone) est significative (.05 ; 95% confidence intervalle : [0 ; 0.10], t= 1.976, p= .01).

L'interaction entre la séquence didactique (=n.group) et le niveau au pré-test (ct0) est non significative (.06 ; 95% confidence intervalle : [- 0.03; 0.15], t= 1.347, p > .05).

Les effets aléatoires représentent la part de variance inhérente à chaque individu (1|participant).

Après retraits des outliers, le modèle a toujours des résidus non normaux. Un modèle robuste a confirmé les estimations du modèle. Le modèle robuste permet une estimation correcte des effets fixes même en présence d'hétéroscédasticité ou de résidus non normaux (Jacqmin-Gadda, Sibillot, Proust, Molina & Thiébaud,2007). Le modèle robuste a été conduit à l'aide des package lavaan (Rosseel, 2012), robustlmm (Koller, 2016) et WRS2 (Mair & Wilcox, 2020).

Les contrastes sont codés :

- la séquence didactique comme une *dummy* variable : $n.group = \{E = -0.5 ; C = +0.5\}$.
- le phonème comme une *dummy* variable : $n.phone = \{j = -0.5 ; on = +0.5\}$

Les estimations du modèle ont permis d'évaluer le pourcentage d'acquisition de la prononciation des phonèmes au post-test immédiat (voir Annexe 7 pour l'étude statistique complète « Prononciation »). Le tableau donne les valeurs estimées des effets fixes pour le modèle de base et le modèle robuste (voir tableau 12).

Les analyses sont décrites à partir des résultats du modèle robuste.

L'intercept représente le progrès + 15% [12 ;19] en moyenne pour les deux phonèmes et les deux groupes.

Tableau 12 : Estimations des effets fixes du modèle sur le progrès de la prononciation

Prédicteurs	Modèle linéaire mixe			Modèle robuste		
	Estimations	CI	p	Estimations	CI	p
(Intercept)	0.14	0.11 – 0.17	< .001	0.15	0.12 – 0.19	< .001
ct0	-0.66	-0.72 – -0.60	< .001	-0.63	-0.69 – -0.56	< .001
n.group	-0.05	-0.09 – -0.00	.043	-0.05	-0.10 – -0.00	.043
n.phone	-0.07	-0.10 – -0.03	< .001	-0.06	-0.10 – -0.03	< .001
ct0 * n.group	0.06	-0.02 – 0.15	0.142	0.06	-0.03 – 0.15	0.178
n.group * n.phone	0.05	0.00 – 0.09	.034	0.05	0.00 – 0.10	.048
Random Effects						
σ^2	0.06			0.07		
τ_{00}	0.01	Participant		0.01	Participant	
ICC	0.08			0.09		
N	77	Participant		77	Participant	
Observations	1848			11848		
Marginal R ² / Conditional R ²	0.335 / 0.385			0.297 / 0.361		

Modèle : formula (progress~ ct0 + n.group *n.phone + n.group*ct0 + (1|Participant), data=dm)

Légendes : CI = intervalle de confiance ; ct0 = niveau au pré-test
n.phone = -0.5 = j ; +0.5 = on
n.group = -0.5 = groupe E ; +0.5 = groupe C

Le progrès est calculé à partir des estimations du modèle robuste en fonction de chacune des 5 valeurs de ct0, en fonction de la séquence didactique et du phonème (voir tableau 13 et annexe 7, fichier excel « synthese_estimations_prononciation »).

Tableau 13 : Estimations calculées à partir du modèle robuste

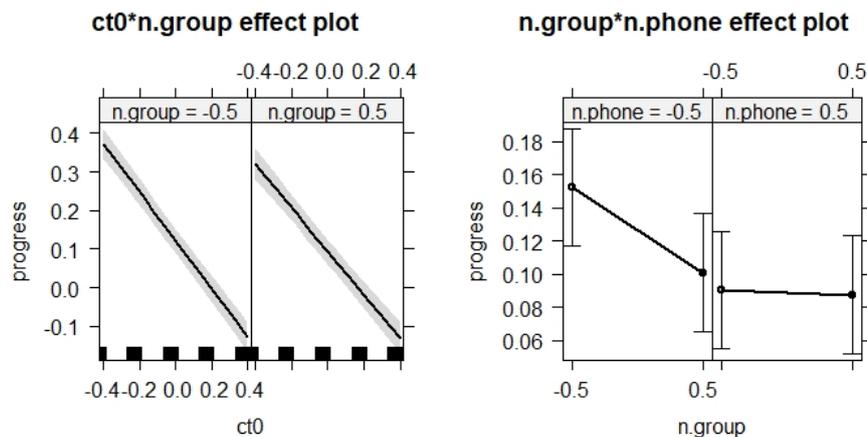
Niveau au pré-test	Séquence didactique	Phonème	Progrès	Intervalle de confiance (0.95)
1 ct0= -0.43	E	<i>/ʒ/-j</i>	+ 50	[50 ; 51]
		<i>/ʒ̃/-on</i>	+ 42	[41 ; 42]
	C	<i>/ʒ/-j</i>	+ 40	[39 ; 42]
		<i>/ʒ̃/-on</i>	+ 37	[32 ; 41]
2 ct0= -0.23	E	<i>/ʒ/-j</i>	+ 37	[37 ; 38]
		<i>/ʒ̃/-on</i>	+ 28	[28 ; 30]
	C	<i>/ʒ/-j</i>	+ 28	[28 ; 29]
		<i>/ʒ̃/-on</i>	+ 25	[18 ; 31]
3 ct0= 0 Moyenne de la classe	E	<i>/ʒ/-j</i>	+ 22	[22 ; 23]
		<i>/ʒ̃/-on</i>	+ 13	[12 ; 15]
	C	<i>/ʒ/-j</i>	+ 14	[12 ; 18]
		<i>/ʒ̃/-on</i>	+ 11	[2 ; 20]
4 ct0= +0.17	E	<i>/ʒ/-j</i>	+ 11	[11 ; 12]
		<i>/ʒ̃/-on</i>	+ 2	[1 ; 4]
	C	<i>/ʒ/-j</i>	+ 4	[0 ; 10]
		<i>/ʒ̃/-on</i>	+ 1	[-10 ; 12]
5 ct0= +0.37	E	<i>/ʒ/-j</i>	- 3	[-3 ; 0]
		<i>/ʒ̃/-on</i>	- 11	[-13 ; -8]
	C	<i>/ʒ/-j</i>	- 8	[-14 ; 0]
		<i>/ʒ̃/-on</i>	- 11	[-24 ; 2]

Le progrès maximal est de +50% [50 ; 51] pour les participants ayant reçu la note de « 1 » (minimum) au pré-test et le groupe E. Un progrès de + 50 % représente plus de deux niveaux de l'échelle de Lickert utilisée pour l'évaluation (1 niveau=20%) et signifie que les participants ont réussi à prononcer un phonème identifiable en fin d'expérimentation.

Lorsque le niveau à T0 est égal à 4, les deux groupes progressent pour le phonème */ʒ/-j* / (l'intervalle de confiance n'inclut pas zéro), mais pour le phonème */ʒ̃/-on*, seul le groupe E progresse +2 [1 ;4].

Les participants progressent pour les deux phonèmes tant que leur niveau au pré-test est inférieur à 4. Plus le niveau de l'évaluation est faible au pré-test plus le progrès est important ($p < .001$) (voir tableau 13 et figure 14, « ct0*n.group effect plot »).

Figure 14 : Ensemble des effets fixes de la modélisation du progrès de la prononciation



Légendes : Phonème : n.phone = {j= -0.5 ; on= +0.5} ; ct0= moyenne centrée au pré-test
Séquence didactique : n.group = {E= -0.5 ; C= 0.5}.

Lorsque le niveau à T0 est au maximum de l'échelle d'évaluation (i.e., évaluation =5), une régression est observée pour les deux groupes pour les deux phonèmes, mais elle est plus importante pour le phonème /ʃ/-on en comparaison du phonème /z/-j pour les deux groupes, et pour le phonème /z/-j, la régression est plus importante pour le groupe C. Cette régression signifie cependant que le participant prononce toujours (avant et après l'expérimentation) un phonème identifiable (évaluation ≥ 4).

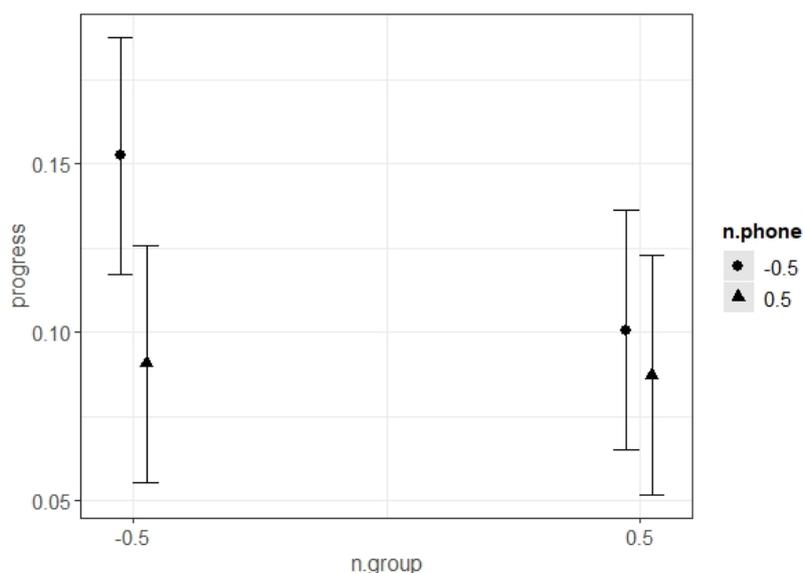
En moyenne pour tous les niveaux, les participants progressent plus dans la séquence didactique explicite (=groupe E) en comparaison de la séquence communicative (= groupe C) de +2.5 % [0 ; +5%] (voir tableau 13 et figure 14). La différence entre les deux groupes est significative (voir tableau 12, *n.group*, $p < .05$). Les intervalles de confiance des deux groupes ne se chevauchent pas (tableau 14 ; figure 14, « n.group*n.phone effect plot »). La séquence didactique communicative (= groupe C) n'a jamais un effet significativement supérieur à celui de la séquence explicite (= groupe E).

Tableau 14: Bilan de la comparaison des deux séquences didactiques en fonction du niveau au pré-test

Séquences didactiques	Progrès	Phonèmes	Niveau au pré-test
E > C	50 [50 ; 51] > 40 [39 ; 42]	/ʒ/-j	1
	37 [37 ; 38] > 28 [28 ; 29]		2
	22 [22 ; 23] > 14 [12 ; 18]		3
	11 [11 ; 12] > 4 [0 ; 10]	4	
	42 [41 ; 42] > 37 [32 ; 41]	/ʒ/-on	1

Pour le phonème /ʒ/-j, la séquence didactique E a toujours un effet supérieur (figure 15) sauf pour les participants qui ont le niveau maximal au pré-test pour lesquels la différence entre les deux séquences didactiques n'est pas significative.

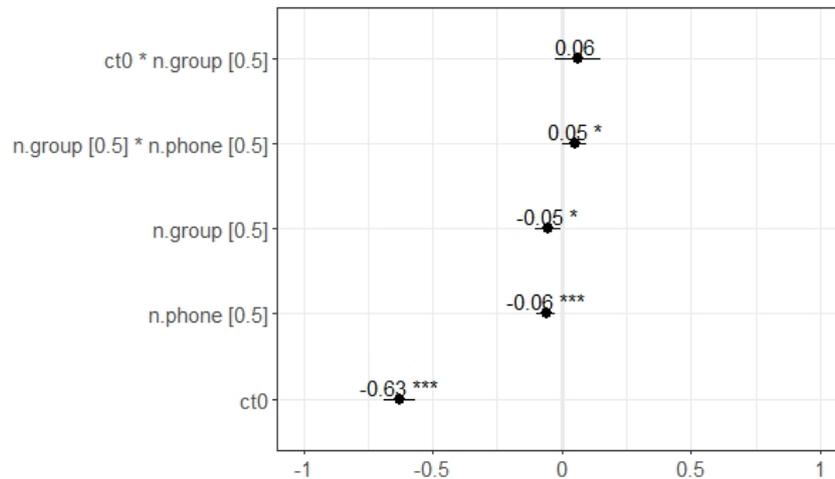
Figure 15 : Prédications du modèle de l'effet des deux séquences didactiques en fonction du phonème (autres prédicteurs maintenus constants).



Légendes : Phonème : n.phone = {j= -0.5 ; on= +0.5}
 Séquence didactique : n.group = {E= -0.5 ; C= +0.5}.

Les participants progressent plus dans la prononciation du phonème /ʒ/-j, dans les deux groupes, que dans celle du phonème /ʒ/-on. En moyenne pour les deux groupes, les participants progressent plus de 3% [+1.5 ; +5] pour le phonème /ʒ/-j en comparaison de /ʒ/-on (figure 16). Cette différence est significative ($p < .001$).

Figure 16 : Estimations par le modèle des effets fixes



Légendes : Phonème : n.phone = {j= -0.5 ; on= +0.5}
 Séquence didactique : n.group = {E= -0.5 ; C= +0.5}
 ct0 = moyenne centrée au pré-test

4.3 Résultats pour la tâche de maintien en mémoire de travail

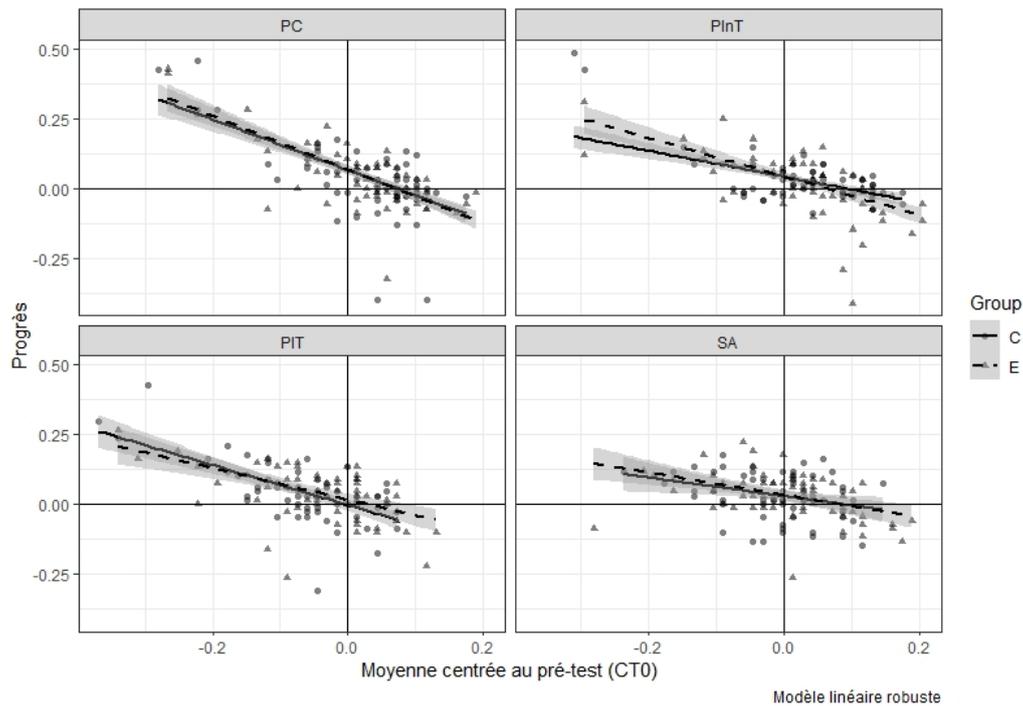
La tâche complète correspond à quatre séries d'une tâche de répétition de pseudo-mots séparées par une pause de quelques minutes (voir 3.2.3). Chaque série proposait 8 listes de pseudo-mots de longueur croissante, à répéter immédiatement après avoir été entendues. La tâche complète a été conduite 2 fois : au pré-test (T0) et au post-test immédiat (T1).

La mesure de chaque série est le pourcentage de phonèmes correctement répétés par tâche (=recp) (i.e., correctement = identiques au stimulus).

4.3.1 Analyse descriptive et observatoire

Le nombre de phonèmes correctement répétés au pré-test (=T0) a été soustrait au nombre de phonème correctement répétés au post test immédiat (=T1) pour calculer une nouvelle variable représentant le progrès (= *progrès*). Le progrès observé a été représenté sur un graphique de points en fonction de la séquence didactique (=Group) et du résultat du participant au pré-test pour chaque série (=CT0). Le nombre de phonèmes correctement répétés au pré-test (=CT0) a été centré par rapport à la moyenne de l'échantillon (figure 17).

Figure 17 : Progrès du nombre de phonèmes répétés en fonction de la série et de la séquence didactique (= Group)



- Légendes :*
- Group C = séquence didactique communicative
 - Group E = séquence didactique explicite
 - PC = série composée des phonèmes communs L1-LE
 - PInT = série composée des phonèmes communs L1-LE et des phonèmes non travaillés / $\tilde{\epsilon}$; η /
 - PIT = série composée des phonèmes communs L1-LE et des phonèmes travaillés / δ ; ζ /
 - SA = série composée des phonèmes communs L1-LE et condition de suppression articulaire

Une partie des participants semble progresser dans les quatre séries (progrès >0). Les participants qui semblent progresser le plus dans les quatre séries semblent être ceux qui sont en dessous de la moyenne de l'échantillon au pré-test (CT0 < 0). Cette variable doit donc être prise en compte dans la modélisation. Les participants se situent au pré-test dans un intervalle de moyenne centrée égal à [-0.3 ; +0.2]. Les deux séquences didactiques semblent produire des effets différents en fonction de la série.

Les moyennes de progrès par séquence didactique et par série sont calculées en incluant ou non le nombre de phonèmes cibles répétés dans le calcul du progrès (tableau 15).

Tableau 15 : Moyennes du progrès en pourcentage de phonèmes répétés correctement par séquence didactique et par série

Séries	Séquence Explicite Group=E		Séquence Communicative Group=C	
	Phonème cibles inclus	Phonème cibles Non inclus	Phonèmes cibles inclus	Phonèmes cibles Non inclus
PC	+4% (12)		+ 3.3% (14)	
PINT ; /ɛ̃ ; ɲ/	+0.8% (11)	+2.4% (12)	+ 3.7% (9)	+ 4.6% (11)
PIT ; /ʃ ; ʒ/	+3.3% (10)	+2.4% (11)	+ 4% (11)	+ 4% (12)
SA	+2.9% (9)		+ 2.7% (8)	

Légendes : PC = série composée des phonèmes communs L1-LE
PInT = série composée des phonèmes communs L1-LE et des phonèmes non travaillés /ɛ̃ ; ɲ/
PIT = série composée des phonèmes communs L1-LE et des phonèmes travaillés /ʃ ; ʒ/
SA = série composée des phonèmes communs L1-LE en condition de suppression articuloire
() = Déviation standard

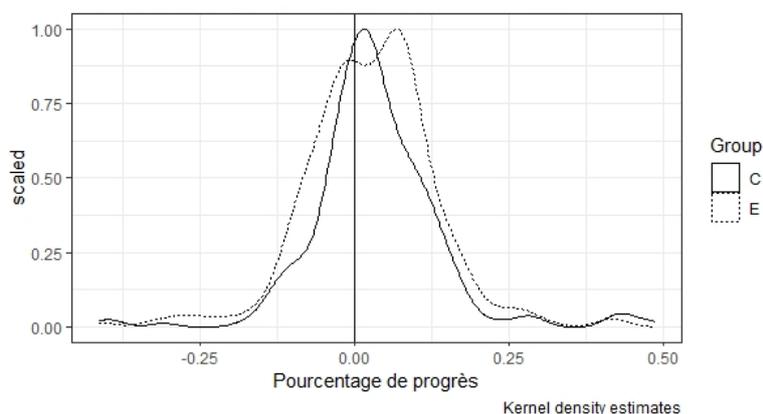
Les deux groupes diffèrent pour les séries incluant les deux phonèmes cibles (=PIT) ou les deux phonèmes « parallèles » des phonèmes cibles (=PInT).

La séquence communicative a un effet supérieur pour la série incluant les deux phonèmes « parallèles » (PInT 3.7>0.8).

Pour la série incluant les phonèmes cibles (=PIT), les deux séquences induisent presque le même progrès lorsque les phonèmes cibles sont inclus dans le compte du nombre de phonèmes correctement répétés (E=3.3% et C=4%). Cependant les phonèmes cibles représentent un progrès de 0.9% pour le groupe E et de 0% pour le groupe C.

La représentation de la probabilité d'un progrès par séquence didactique peut sembler indiquer une supériorité de la séquence explicite (figure 18).

Figure 18 : Probabilité d'un progrès par séquence didactique pour l'ensemble des séries



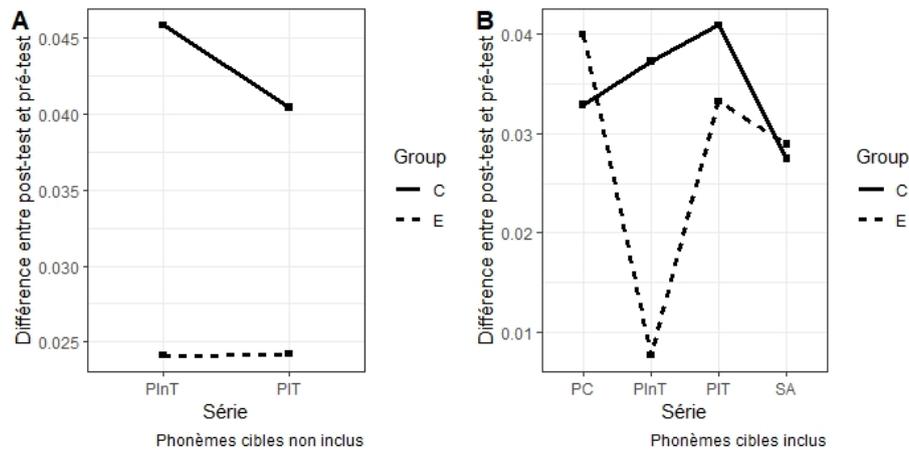
Légendes : Group C = séquence didactique communicative
Group E = séquence didactique explicite

La probabilité d'un progrès est supérieure pour la séquence explicite (ligne pointillée) pour une partie des participants seulement : elle est clairement biphase. Deux hypothèses sont envisageables :

- Hypothèse 1 : une des deux séquences didactiques permet de mieux progresser dans une série mais pas dans les/l'autre ;
- Hypothèse 2 : les résultats sont influencés par l'effet de la séquence didactique sur les phonèmes cibles en particulier. L'étude des progrès moyens (voir tableau 15) indique que le progrès des phonèmes cibles constitue une part importante du progrès de la séquence PIT pour le groupe E.

Analyse hypothèse 1 : Que le nombre de phonèmes cibles correctement répétés soit inclus (B) ou non (A) dans le calcul du progrès, on observe une interaction entre l'effet de la séquence didactique et le progrès moyen par série (figure 19).

Figure 19 : Interactions entre les séquences didactiques et la série sur le nombre total de phonèmes répétés en fonction de l'inclusion ou non des phonèmes cibles



- Légendes :**
- Group C = séquence didactique communicative
 - Group E = séquence didactique explicite
 - PC = série composée des phonèmes communs L1-LE
 - PInT = série composée des phonèmes communs L1-LE et des phonèmes non travaillés /ɛ̃ ; ɲ/
 - PIT = série composée des phonèmes communs L1-LE et des phonèmes travaillés /ʃ ; ʒ/
 - SA = série composée des phonèmes communs L1-LE en condition de suppression articulaire

La séquence didactique communicative (=C) induit un progrès supérieur du pourcentage de phonèmes communs L1-LE répétés (phonèmes cibles et parallèles non comptabilisés ; PIT, partie A de la figure 19). La différence entre les séquences didactiques diminue lorsque les phonèmes cibles sont inclus dans les stimuli (PIT, partie B de la figure 19).

La séquence didactique explicite (= E) induit un progrès plus important du pourcentage de phonèmes cibles répétés qui est observable en comparant les performances entre les parties A et B pour la série PIT (phonèmes cibles inclus /ʃ ; ʒ/), mais pas des phonèmes parallèles (PInT, phonèmes parallèles /ɛ̃ ; ɲ/). Cet effet n'est pas observé pour la séquence communicative.

La séquence didactique explicite limite le nombre de phonèmes communs L1-LE répétés lorsque les phonèmes cibles sont inclus dans les stimuli (série PIT) mais ne limite pas les phonèmes communs L1-LE lorsqu'ils sont présentés sans les phonèmes cibles (série PC). Cet effet n'est pas observé pour la séquence communicative.

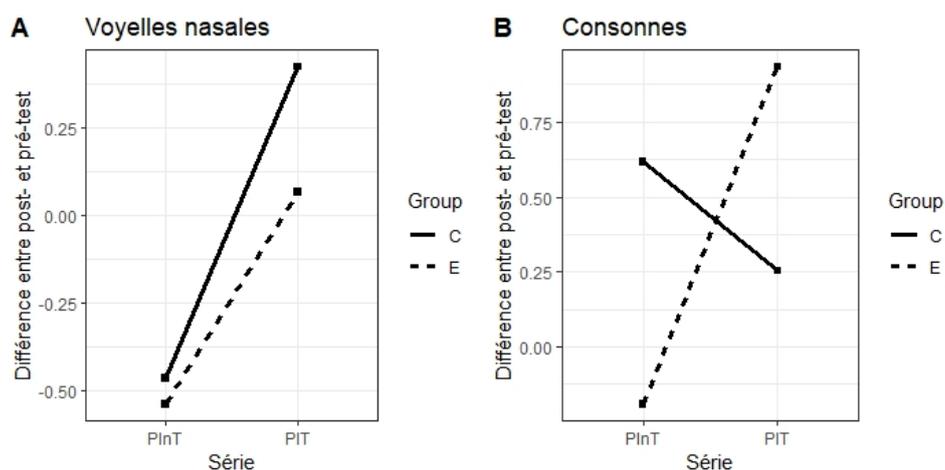
Analyse hypothèse 2 : Les deux séquences didactiques influencent différemment le maintien de chacun des deux phonèmes cibles dans la même série (tableau 16).

Tableau 16 : Moyenne du progrès du nombre de phonèmes cibles correctement restitués par séquence didactique (max=8)

Séquence didactique	Voyelle cible <i>/ɔ̃/-on</i>	Voyelle proche <i>/ɛ̃/-in</i>	Consonne cible <i>/ʒ/-j</i>	Consonne « parallèle » <i>/ɲ/-gn</i>
Communicative	+0.42	-0.47	+0.25	+0.62
Explicite	+0.07	-0.54	+0.94	-0.19

La didactique communicative favorise le maintien du phonème */ɔ̃/-on* (figure 20, partie A, série PIT) et la didactique explicite du phonème */ʒ/-j* (figure 20, partie B, série PIT).

Figure 20 : Évolution du nombre moyen de phonèmes cibles et « parallèles » répétés en fonction des séries et des séquences didactiques



- Légendes :
- Group C = séquence didactique communicative
 - Group E = séquence didactique explicite
 - PInT = série composée des phonèmes communs L1-LE et des phonèmes non travaillés */ɛ̃ ; ɲ/*
 - PIT = série composée des phonèmes communs L1-LE et des phonèmes travaillés */ɔ̃ ; ʒ/*

Les participants du groupe E ont répété en moyenne presque un phonème */ʒ/-j* de plus (sur un total de 8 possibles) au post-test (PIT, partie B, figure 20). La différence entre les deux groupes est faible pour les voyelles nasales mais favorable au groupe C (PIT, partie A, figure 20). Néanmoins, l'influence du nombre de phonèmes répétés sur le progrès global par série est négligeable (8 phonèmes cibles /68 phonèmes au total $\approx 12\%$). Cette interaction n'est pas prise en compte dans la modélisation.

4.3.2 Analyse statistique⁸

Analyse statistique principale

L'analyse statistique a été conduite par un modèle linéaire mixte (=lmer). Le modèle évalue le progrès du nombre total de phonèmes répétés (les phonèmes cibles sont inclus). La variable dépendante (=progrès) a été transformée en pourcentage pour permettre une modélisation par un modèle linéaire. Le progrès est la différence entre le post-test et le pré-test du nombre total de phonèmes répétés par série par un participant. Il est exprimé en pourcentage.

Les facteurs fixes, ou variables indépendantes, sont l'appartenance du participant à la séquence didactique (Group= {E ; C}), la série (Task= {PC; SA ; PInT ; PIT}), et le nombre de phonèmes répétés au pré-test par un participant (= CT0). Le nombre de phonèmes répétés au pré-test par un participant (CT0) a été transformé en pourcentage et standardisé par la moyenne de la population.

L'interaction entre la série (=Task) et la séquence didactique (=Group) observée dans l'analyse descriptive (Task*Group) n'est pas significative et sa présence ne modifie pas les estimations du modèle (voir Annexe 8). De plus la différence entre le modèle incluant l'interaction et celui qui ne l'inclut pas n'est pas significative. Le modèle retenu est donc le modèle le plus simple (i.e. sans interaction). Les effets aléatoires représentent la part de variance inhérente à chaque individu (1|participant).

Après retraits des outliers, le modèle a toujours des résidus non normaux. Un modèle robuste a confirmé les estimations du modèle. Le modèle robuste permet une estimation correcte des effets fixes même en présence d'hétéroscédasticité ou de résidus non normaux (Jacqmin-Gadda et al., 2006). Le modèle robuste a été conduit à l'aide des package lavaan (Rosseel, 2012), robustlmm (koller 2016) et WRS2 (Mair & Wilcox, 2020).

Les contrastes sont codés en dummy variable pour la séquence didactique et en *simple coding* pour la série :

- la séquence didactique, : $n.group = \{E = -0.5 ; C = 0.5\}$.
- la série: $Task = \{ PC = 0 \text{ (référence)} ;$
PInT = (1 ; 0 ; 0) ;
PIT = (0 ; 1 ; 0) ;
SA = (0 ; 0 ; 1) }

Le tableau donne les valeurs estimées des effets fixes (tableau 17).

L'intercept représente le progrès entre le pré-test et le post-test immédiat du pourcentage de phonèmes correctement répétés pour la série PC (incluant que des phonèmes communs L1-LE).

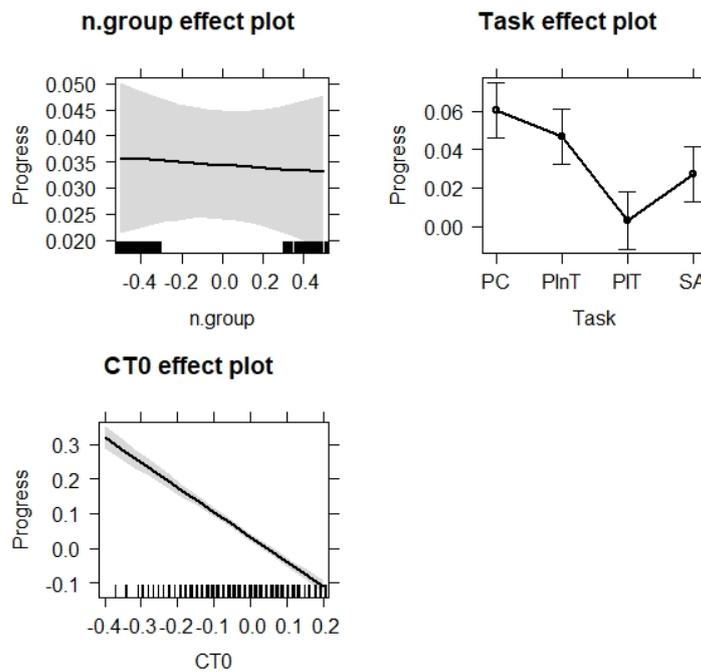
⁸ Voir Annexe 8 pour l'analyse statistique complète

Tableau 17 : Estimations des effets fixes sur le progrès lors de la répétition de pseudo-mots

<i>Prédicteurs</i>	Modèle linéaire mixte			Modèle robuste		
	<i>Estimations</i>	<i>CI</i>	<i>p</i>	<i>Estimations</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.06	0.04 – 0.07	< .001	0.06	0.05 – 0.07	< .001
n.group	-0.00	-0.02 – 0.02	0.793	-0.01	-0.02 – 0.01	0.557
Série [PInT]	-0.01	-0.03 – 0.00	0.084	-0.01	-0.03 – 0.00	0.113
Série [PIT]	-0.06	-0.07 – -0.04	< .001	-0.06	-0.07 – -0.04	< .001
Série [SA]	-0.03	-0.05 – -0.02	< .001	-0.03	-0.04 – -0.02	< .001
CT0	-0.72	-0.80 – -0.64	< .001	-0.70	-0.77 – -0.63	< .001
Effets aléatoires						
σ^2	0.00			0.00		
τ_{00}	0.00	Participant		0.00	Participant	
ICC	0.37			0.34		
N	121	Participant		121	Participant	
Observations	480			480		
Marginal R ² / Conditional R ²	0.419 / 0.635			0.476 / 0.656		
<i>Modèle :</i>	rmlmer/lmer (progress ~ n.group + Task+ CT0 + (1 Participant), data= dm_outL)					
<i>Légendes :</i>	Séquence didactique = n.group E = -0.5; C = +0.5					
	PC = série composée des phonèmes communs L1-LE					
	PInT = série composée des phonèmes communs L1-LE et des phonèmes non travaillés /ɛ̃ ; ɲ/					
	PIT = série composée des phonèmes communs L1-LE et des phonèmes travaillés /ʃ ; ʒ/					
	SA = série composée des phonèmes communs L1-LE en condition de suppression articulaire					
	CT0 = Moyenne centrée au pré-test					
	CI = Intervalle de confiance					

La séquence didactique n'influence pas les résultats ($p > .05$; voir « n.group effect plot », figure 21). Les résultats sont donc présentés pour les deux séquences didactiques (l'estimation de l'effet groupe est égale à zéro, tableau 17).

Figure 21 : Représentation graphique des effets fixes du modèle



Légendes : Séquence didactique = n.group E = -0.5; C = +0.5
 PC = Série composée des phonèmes communs L1-LE
 PInT = Série composée des phonèmes communs L1-LE
 et des phonèmes non travaillés /ɛ̃ ; ɲ/
 PIT = Série composée des phonèmes communs L1-LE
 et des phonèmes travaillés /ɔ̃ ; ʒ/
 SA = Série composée des phonèmes communs L1-LE
 en condition de suppression articulaire
 CT0 = Moyenne centrée au pré-test

Les estimations du modèle ont permis d'évaluer le pourcentage de phonèmes correctement répétés au post-test immédiat (voir Annexe 8 pour l'étude statistique complète « WM_lmer_synthese »).

Plus le niveau de l'évaluation est faible au pré-test plus le progrès est important (figure 21, « CT0 effect plot », $p < .001$) dans les quatre séries. Le progrès théorique maximal est de +72 % ($p < .001$) pour les participants étant à un niveau inférieur d'un point à la moyenne centrée de l'échantillon au pré-test (tableau 17). Cependant, les participants se situent au pré-test dans l'intervalle de la moyenne centrée [-0.3 ; +0.2].

L'estimation du progrès du pourcentage de phonèmes correctement répétés a été calculée en fonction des valeurs prises réellement par la moyenne centrée au pré-test (= CT0) (tableau 18). Les intervalles de confiance sont calculés à partir des résultats du modèle robuste⁹.

⁹ Voir fichier excel « Estimations_WM » dans Annexe 8

Tableau 18 : Évolution du nombre de phonèmes répétés par série en %

Valeur de CT0	PC	PInT	PIT	SA
- .3	+ 28 [26 ; 28]	+ 27 [26 ; 28]	+ 22 [21 ; 22]	+ 24 [23 ; 24]
- .2	+ 20 [20 ; 20]	+ 19 [17 ; 20]	+ 14 [13 ; 16]	+ 17 [15 ; 18]
- .1	+ 13 [12 ; 13]	+ 12 [9 ; 13]	+ 7 [5 ; 9]	+ 10 [7 ; 11]
0	+ 6 [4 ; 7]	+ 5 [1 ; 7]	0 [-3 ; 3]	+ 3 [-1 ; 5]
+ .1	- 1 [-4 ; 1]	- 2 [-7 ; 1]	- 7 [-11 ; -3]	- 4 [-9 ; -1]
+ .2	- 8 [-12 ; -6]	- 9 [-15 ; -6]	- 14 [-19 ; -10]	- 11 [-17 ; -8]

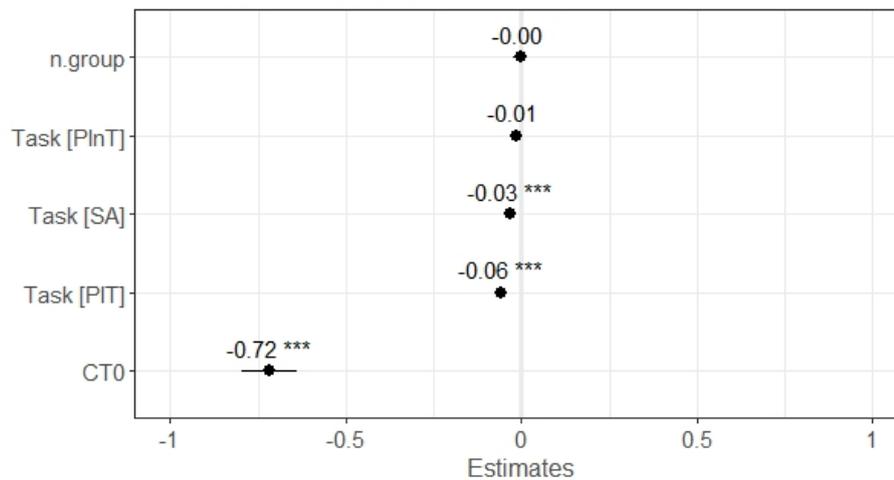
Légendes :

- PC = Série composée des phonèmes communs L1-LE
- PInT = Série composée des phonèmes communs L1-LE et des phonèmes non travaillés /ɛ̃ ; ɲ/
- PIT = Série composée des phonèmes communs L1-LE et des phonèmes travaillés /ɔ̃ ; ʒ/
- SA = Série composée des phonèmes communs L1-LE en condition de suppression articulaire
- CT0 = Moyenne centrée au pré-test

En **gras** la zone sans progrès ni régression

Les séries PInT et PC ne montrent pas de différence significative ($p > .05$). La série PIT est significativement inférieure à celle de PC ($p < .001$). La série PIT, incluant les phonèmes travaillés, est celle qui enregistre le progrès le plus faible en comparaison des autres séries (« Task effect plot », figure 21 et tableau 18). Elle montre des résultats inférieurs de 3% à la série de suppression articulaire (=SA) (figure 22).

Figure 22 : Estimations des effets fixes du modèle



Légendes : Séquence didactique= n.group : E = -0.5; C = +0.5
 PC = Série composée des phonèmes communs L1-LE
 PInT = Série composée des phonèmes communs L1-LE et des phonèmes non travaillés / $\tilde{\epsilon}$; η /
 PIT = Série composée des phonèmes communs L1-LE et des phonèmes travaillés / $\tilde{\sigma}$; ζ /
 SA = Série composée des phonèmes communs L1-LE en condition de suppression articuloire
 CT0 = Moyenne centrée au pré-test

Pour les participants les plus faibles au pré-test, le progrès maximal est de +28% pour la série PC (tableau 18). Les participants dont le niveau au pré-test est inférieur ou égal à la moyenne de l'échantillon progressent significativement dans chacune des 4 séries (les intervalles de confiance n'incluent pas zéro). Pour les deux séries PIT et SA, et les participants à la moyenne de l'échantillon au pré-test, le progrès ou la régression ne sont pas significatifs.

Pour les participants dont le niveau au pré-test est supérieur à la moyenne du groupe de 10%, le progrès ou la régression ne sont pas significatifs pour les séries PC et PInT. Pour les autres séries, une régression significative est observée.

Pour les participants dont le niveau au pré-test est supérieur à la moyenne du groupe de 20%, une régression significative est observée pour les 4 séries (tableau 18).

Analyses statistiques des questions secondaires

A) L'observation des données montre que le phonème parallèle / $\tilde{\epsilon}$ -in (série PInT) ne progresse pas contrairement au phonème / $\tilde{\sigma}$ -on (série PIT, voir figure 20, partie A et tableau 16). Pour pouvoir généraliser cette observation, elle a été modélisée.

L'analyse statistique a été conduite par un modèle linéaire mixte (=lmer). Le modèle évalue le progrès du nombre de voyelles nasales répétées. La variable dépendante (=Prog_nv) est la

différence entre le post-test et le pré-test du nombre de voyelles nasales répétées par série par participant.

Les facteurs fixes, ou variables indépendantes, sont l'appartenance du participant à la séquence didactique (Group= {E ; C}), et la série (Task= {PInT ; PIT}). La série PInT représente le nombre de voyelles / $\tilde{\epsilon}$ /-in répétées et la série PIT le nombre de voyelles / $\tilde{\delta}$ /-on.

Les contrastes sont codés en dummy variable pour la séquence didactique et en *simple coding* pour la série :

- la séquence didactique, : $n.group = \{E = -0.5 ; C = 0.5\}$
- la série: $Task = \{PIT = 0 \text{ (référence)} ; PInT = 1 \}$.

Le tableau donne les valeurs estimées des effets fixes (tableau 19).

L'intercept représente la différence du nombre de voyelles nasales répétées entre le pré-test et le post-test immédiat pour la série PIT, c'est-à-dire le phonème / $\tilde{\delta}$ /-on.

Tableau 19 : Estimations des effets fixes sur l'évolution de la répétition des voyelles nasales

<i>Predictors</i>	Modèle linéaire mixte			Modèle robuste		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
Intercept= / $\tilde{\delta}$ /-on	0.24	-0.10 – 0.59	0.170	0.26	-0.10 – 0.61	0.155
Task [PInT] = / $\tilde{\epsilon}$ /-in	-0.74	-1.23 – -0.26	.003	-0.73	-1.23 – -0.23	.004
n.group	0.21	-0.27 – 0.70	0.389	0.25	-0.25 – 0.74	0.332
Random Effects						
σ^2	3.76			3.73		
τ_{00}	0.02	Participant		0.00	Participant	
ICC	0.00			0.00		
N	123	Participant		123	Participant	
Observations	244			244		
Marginal R ² / Conditional R ²	0.038 / 0.043			0.038 / 0.038		

Modèle : `rlmer/lmer (Prog_vn~ Task + n.group + (1|Participant), data=dio)`

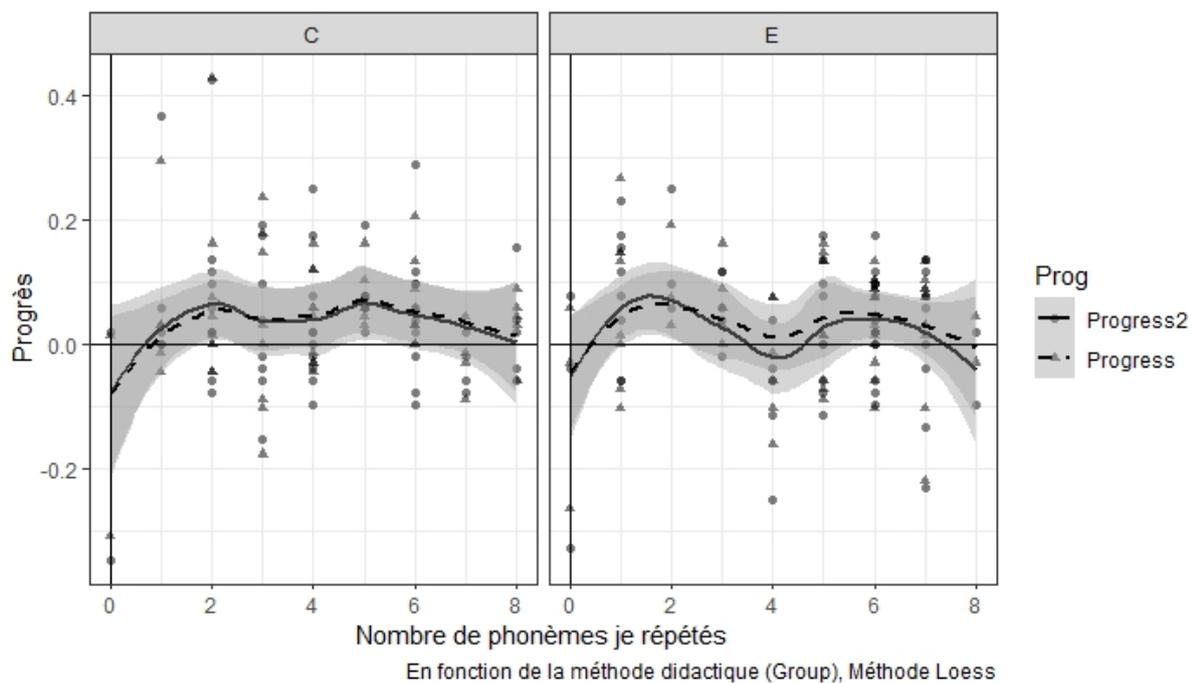
L'effet groupe n'est pas significatif ($p > .05$).

Le phonème / $\tilde{\epsilon}$ /-in ne progresse pas et régresse entre le pré-test et le post-test, contrairement au phonème / $\tilde{\delta}$ /-on qui progresse. Cette différence entre les deux voyelles nasales est significative ($p < .01$).

B) L'observation des données montre que dans la série PIT, la part des phonèmes cibles répétés dans le total des phonèmes répétés est nettement plus importante pour le groupe E que pour le groupe C, et la part des phonèmes communs L1-LE moins importante (voir tableau 15 et figure 19).

La représentation du progrès (=différence entre pré-test et post-test) du nombre de phonèmes répétés en fonction du nombre de fois où le phonème /z/-j est répété au post-test immédiat est similaire dans les deux groupes jusqu'à 2 phonèmes répétés (figure 23).

Figure 23 : Progrès du nombre de phonèmes répétés entre le pré-test et le post-test immédiat



Légendes : Séquence didactique : C : séquence communicative ; E : Séquence explicite
 Progress2 : pourcentage supplémentaire de phonèmes communs L1-LE
 Progress : pourcentage supplémentaire de phonèmes communs L1-LE et du phonème cible /z/-j.

Entre 2 et 6 phonèmes /z/-j répétés, les deux groupes diffèrent.

Pour le groupe C, le nombre de phonèmes cibles (= /z/-j et /ʃ/-on) répétés ne semble pas influencer le nombre total de phonèmes répétés (=phonèmes cibles + phonèmes communs L1-LE) ; les courbes des Progress2 et Progress sont confondues et les deux courbes se maintiennent au-dessus de zéro.

Pour le groupe E, la courbe qui inclut seulement les phonèmes communs L1-LE (=Progress2) est inférieure à la courbe qui inclut tous les phonèmes répétés (=Progress). La courbe qui inclut tous les phonèmes répétés est même négative (valeurs inférieures à 0) lorsque le participant répète autour de 4 phonèmes /z/-j.

4.4 Résultats pour la tâche de correspondance phonème-graphème

La tâche consiste à identifier la forme écrite correcte d'un pseudo-mot entendu parmi 4 propositions (voir 3.2.4). Trois propositions sont des erreurs représentant chacune un type d'erreur : orthographique (=orth), phonologique (=phon) ou phonotactique (=pta). Une seule des quatre formes écrites proposées peut correspondre au pseudo-mot entendu.

Chaque pseudo-mot coché par le participant a été valorisé par 1 point. L'analyse statistique de cette tâche a été conduite par l'analyse du risque d'erreur : une bonne réponse a un score égal à 0 ; une erreur a un score égal à 1.

Le même test a été conduit 3 fois au pré-test (=T0), post-test immédiat (=T1) et trois mois après la fin de l'enseignement (=T2).

L'analyse ne porte que sur les items incluant les phonèmes cibles : 5 stimuli pour le phonème /ʒ/-on et 6 stimuli phonème /z/-j par participant (nombre total de stimuli pour cette tâche =47).

4.4.1 Analyse descriptive et observatoire

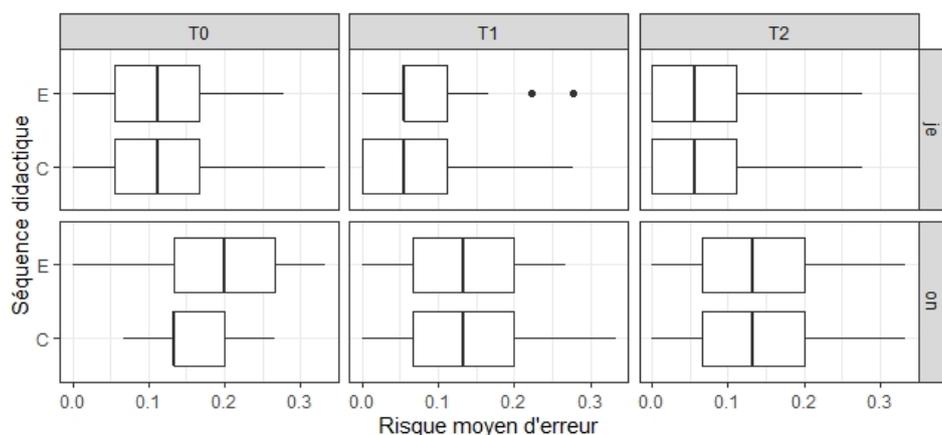
Pour l'ensemble des deux phonèmes et des deux séquences didactiques, le risque d'erreurs a été réduit de 3% environ au post-test à 3 mois en comparaison du pré-test (Tableau 20).

Tableau 20 : Observation de la réduction de la probabilité de faire une erreur

Test	En moyenne des deux séquences didactiques	Pour la séquence didactique C	Pour la séquence didactique E
Pré-test	33.1 %	32.5 %	33.7 %
Post-test immédiat	30.8 %	31.1 %	30.4 %
Post-test à 3 mois	30.3 %	30.2 %	30.3 %

La séquence didactique E semble induire une réduction plus importante du risque d'erreur en moyenne pour les deux phonèmes. Les résultats recueillis sont représentés sous forme de box-plot visualisant la médiane et le premier et troisième quartile de la distribution du risque d'erreur pour les deux phonèmes cibles (voir figure 24).

Figure 24 : Représentation de l'évolution du risque d'erreur par phonème cible et en fonction du test



Légendes : C = Séquence didactique communicative ; E = Séquence didactique explicite
 T0 = Pré-test ; T1 = Post-test immédiat ; T2 = Post-test à 3 mois

Pour le phonème /j/, les deux groupes sont au même niveau au pré-test (=T0). Au post-test immédiat (=T1), le groupe C a plus progressé que le groupe E (i.e., réduction du risque d'erreur supérieure ; voir figure 24 et tableau 21).

Tableau 21 : Probabilité observée du risque d'erreur par groupe, par test et par phonème

Séquence didactique	Phonème	T0	T1	Différence T1 - T0	T2	Différence T2 - T0
C	/j/	11 [0 ; 42]	7 [0 ; 32]	- 4	7 [0 ; 33]	- 4
E	/j/	11 [0 ; 43]	9 [0 ; 38]	- 2	8 [0 ; 35]	- 3
C	/ɔ/	17 [0 ; 54]	15 [0 ; 51]	- 2	14 [0 ; 49]	- 3
E	/ɔ/	18 [0 ; 57]	13 [0 ; 47]	- 5	14 [0 ; 49]	- 4

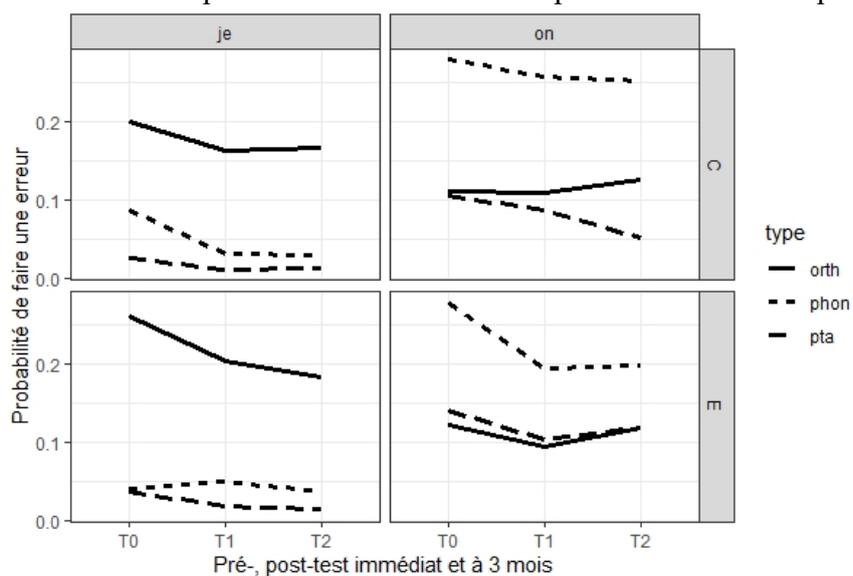
Au test à 3 mois (=T2) pour le phonème /j/, le groupe E a presque autant réduit le risque d'erreur que le groupe C (E= - 3 vs. C= - 4).

Pour le phonème /ɔ/, le groupe E a un risque d'erreur légèrement supérieur à T0. Au post-test immédiat (=T1), le groupe E a plus progressé (risque d'erreur diminué de 5%) que le groupe C (= -2). Une évolution est observée entre le post-test immédiat et celui à 3 mois et les performances deux groupes semblent se rapprocher. Le taux d'erreur peut être lié au type d'erreur.

L'évolution du risque d'erreur est représentée en fonction du type d'erreurs¹⁰ (figure 25).

¹⁰ Exemple de type d'erreurs : orthographique /zi/ ne peut pas s'écrire < ji > ; phonologique /zi/ écrit < chi > ; phonotactique /zi/ écrit < schi >

Figure 25 : Réduction du risque d'erreur en fonction du phonème et de la séquence didactique



Légendes : C = séquence didactique communicative ; E = séquence didactique explicite
 T0 = pré-test ; T1 = post-test immédiat ; T2 = post-test à 3 mois
 orth = erreur orthographique ; phon = erreur phonologique ;
 pta = erreur phonotactique

Pour le phonème /*ʒ*-, les erreurs orthographiques sont les plus fréquentes. Elles sont réduites au post-test à trois mois pour les deux groupes, mais la réduction est plus importante pour le groupe E (figure 25 et tableau 22).

Tableau 22 : Probabilité observée du risque d'erreur par groupe, par type d'erreur et par phonème

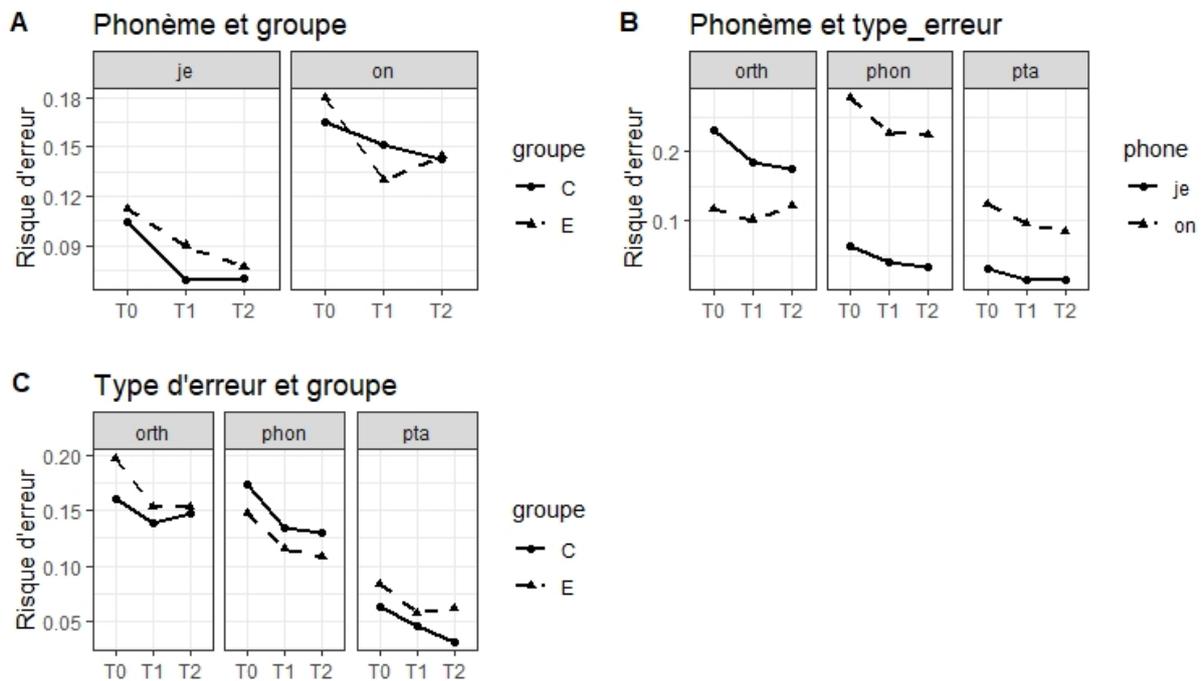
Séquence didactique	Phonème	Type d'erreur	T0		T2		Différence T2 - T0
C	/ʒ/-j		20	(40)	17	(37)	- 3
E	/ʒ/-j	Orthographique	26	(44)	18	(39)	- 8
C	/ʃ/-on		11	(31)	13	(33)	+ 2
E	/ʃ/-on		12	(33)	12	(32)	0
C	/ʒ/-j		9	(28)	3	(17)	- 6
E	/ʒ/-j	Phonologique	4	(19)	4	(19)	0
C	/ʃ/-on		28	(45)	25	(43)	- 3
E	/ʃ/-on		28	(45)	20	(40)	- 8
C	/ʒ/-j		3	(16)	1	(12)	- 2
E	/ʒ/-j	Phonotactique	4	(19)	1	(12)	- 3
C	/ʃ/-on		10	(31)	5	(22)	- 5
E	/ʃ/-on		14	(35)	12	(32)	- 2

Légendes : () = déviation standard

Pour le phonème /ʃ/-on, les erreurs phonologiques sont les plus fréquentes. Elles sont réduites au post-test à trois mois pour les deux groupes, mais la réduction est plus importante pour le groupe E (figure 25 et tableau 22).

Une interaction est observable entre la séquence didactique, le phonème et le type d'erreur (figure 26).

Figure 26 : Interactions observées sur la réduction du risque d'erreur : entre la séquence didactique et le phonème, entre le phonème et le type d'erreur, et entre la séquence didactique et le type d'erreurs.



Légendes : C = séquence didactique communicative ; E = séquence didactique explicite
T0 = pré-test ; T1 = post-test immédiat ; T2= post-test à 3 mois
orth= erreur orthographique ; phon = erreur phonologique ;
pta = erreur phonotactique

Les erreurs phonologiques sont plus fréquentes que les erreurs orthographiques pour le phonème /ɔ̃/-on, et inversement pour le phonème /ʒ/-j. Les erreurs phonologiques sont réduites après l'intervention didactique pour les deux phonèmes et cet effet se maintient à trois mois (=T2).

Les erreurs orthographiques sont réduites au post-test immédiat (=T1), mais pour le phonème /ɔ̃/-on, elles ont retrouvé leur niveau du pré-test au post-test à 3 mois (figure 26, partie B « Phonème et type d'erreur »).

Les erreurs phonotactiques sont réduites pour les deux phonèmes et cet effet se maintient à trois mois (=T2).

Les deux groupes réduisent les erreurs orthographiques (= orth) entre le pré-test et le post-test immédiat, et l'effet se maintient en comparaison du pré-test au post-test à 3 mois (figure 26, partie C « Type d'erreur et groupe »). Le groupe E semble réduire plus fortement les erreurs orthographiques que le groupe C. Les deux groupes réduisent les erreurs phonologiques (= phon) et cet effet se maintient à trois mois (=T2). Les deux groupes réduisent les erreurs phonotactiques (= pta) au post-test immédiat (=T1) mais seul le groupe C continue de réduire le nombre d'erreurs au post-test à 3 mois (=T2).

L'interaction est triple (figure 26) : 1) entre le phonème en jeu et la séquence didactique, observable sur la partie A, 2) entre le phonème et le type d'erreur, observable sur la partie B, et 3) entre la séquence didactique et le type d'erreurs, observable sur la partie C.

Ces interactions sont prises en compte dans l'analyse statistique.

4.4.2 Analyse statistique¹¹

L'analyse statistique a été conduite par un modèle linéaire mixte généralisé (=glmer, Bolker, 2015). Cette modélisation est complexe car elle doit inclure la triple interaction observée. La comparaison de la probabilité de faire une erreur a donc été réduite à une comparaison entre pré-test et post-test à trois mois.

Le modèle évalue la probabilité d'une erreur par stimulus. La variable dépendante (=score) est la valeur attribuée à la réponse (erreur choisie *score* =1 ou pas d'erreur *score*=0) lors de la présentation de 4 pseudo-mots dans la tâche de correspondance phonème-graphème (1 réponse juste pour 3 réponses fausses).

L'interaction observée dans l'analyse descriptive entre le type d'erreur, la séquence didactique et le phonème s'est avérée être non significative, mais deux interactions liées à cette triple interaction sont significatives : 1) l'interaction phonème et type d'erreur (*type*phone*) et 2) l'interaction séquence didactique et type d'erreur (*type*groupe*). Ces deux interactions sont incluses dans le modèle.

L'interaction phonème et type d'erreur est significative pour évaluer le pourcentage d'erreurs moyen des deux tests, $F(2, 8241) = 142.431$, $p = 2e-61$, $\eta^2[g] = 0.033$. L'analyse des effets principaux pour le type d'erreur au post-test a été effectuée avec une signification statistique recevant un ajustement de Bonferroni. L'effet du type d'erreurs dans les scores moyens d'erreurs est statistiquement significatif pour chacun des phonèmes au post-test (= T2) : pour /ʒ/-j ($F(2, 8241) = 57,1$, $p = 2e-25$) et pour /ʒ/-on ($F(2, 8241) = 32.271$, $p = 1e-14$).

L'interaction séquence didactique (=groupe) et type d'erreur est significative pour évaluer le pourcentage d'erreurs moyen des deux tests, $F(2, 8241) = 5.148$, $p = .006$, $\eta^2[g] = 0.001$. L'analyse des effets principaux au post-test pour le type d'erreur a été effectuée avec une signification statistique recevant un ajustement de Bonferroni. L'effet du type d'erreurs dans les scores moyens d'erreurs est statistiquement significatif pour chacun des deux groupes à T0 et au post-test : pour le groupe C ($F(2, 8241) = 25.665$, $p = 8e-12$) et pour le groupe E ($F(2, 8241) = 13.725$, $p = 1e-6$).

Les facteurs fixes, ou variables indépendantes, sont l'appartenance du participant à une séquence didactique (groupe= {E ; C}), l'appartenance du stimulus au type de phonème

¹¹ Voir Annexe 9 pour l'analyse statistique complète

(*phone* = {on ; j}), le moment du test (*Time* = {T0 ; T2}), et les deux interactions (*type***phone* et *type***groupe*). Les effets aléatoires représentent la part de variance inhérente à chaque individu (1|participant).

Les contrastes sont codés :

- la séquence didactique (*groupe*) comme une *dummy* variable : $n.group = \{ E = - 0.5 ; C = + 0.5 \}$
- le test en *simple coding* : $test = \{ T0 = 0 \text{ (référence)} ; T2 = 1 \}$
- le type d'erreur en *simple coding* : $type = \{ orth0 = 0 \text{ (référence)} ; phon = 1 \text{ et } pta = 0 ; phon = 0 \text{ et } pta = 1 \}$
- le phonème en *simple coding* : $phone = \{ on = 0 \text{ (référence)} ; j = 1 \}$ ¹²

Le tableau donne les valeurs estimées transformées en odds ratio et en probabilités¹³ des effets fixes (voir tableau 23). L'intercept représente le phonème /z/-j, au pré-test, pour les erreurs orthographiques et pour une moyenne des deux groupes.

Tableau 23 : Effets fixes estimés sur le risque d'erreur au post-test à 3 mois

<i>Prédicteurs</i>	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>	<i>Probabilités</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.29	0.25 – 0.34	< .001	0.22	0.20 – 0.25	< .001
phone [on]	0.52	0.42 – 0.64	< .001	0.34	0.30 – 0.39	< .001
type [phon]	0.19	0.15 – 0.25	< .001	0.16	0.13 – 0.20	< .001
type [pta]	0.09	0.06 – 0.12	< .001	0.08	0.06 – 0.11	< .001
n.group	0.84	0.67 – 1.07	0.162	0.46	0.40 – 0.52	0.162
test [T2]	0.72	0.62 – 0.82	< .001	0.42	0.38 – 0.45	< .001
phone [on] * type [phon]	13.26	9.39 – 18.72	< .001	0.93	0.90 – 0.95	< .001
phone [on] * type [pta]	9.78	6.29 – 15.20	< .001	0.91	0.86 – 0.94	< .001
type [phon] * n.group	1.48	1.09 – 2.01	.012	0.60	0.52 – 0.67	.012
type [pta] * n.group	0.72	0.49 – 1.06	0.100	0.42	0.33 – 0.52	0.100

¹² Créer une deuxième *dummy* variable en interaction dans un *glmer* est impossible.

¹³ Voir Annexe 9 pour les estimations du modèle en log(odds)

Effets aléatoires

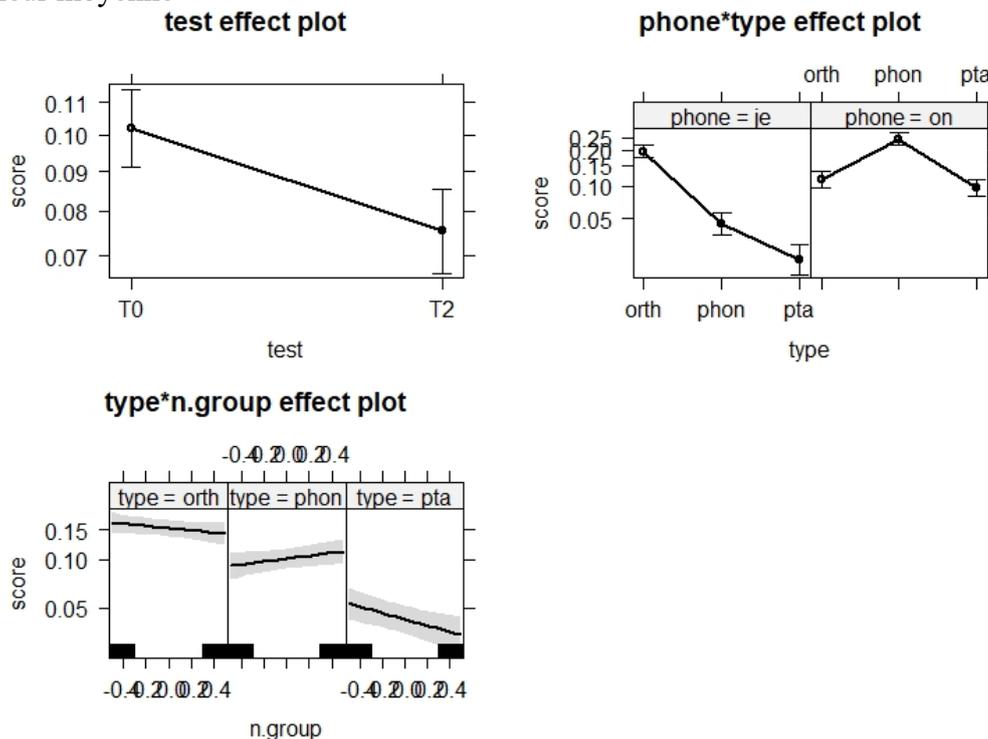
σ^2	3.29	3.29
τ_{00}	0.12	0.12
ICC	0.03	0.03
N	127	127
Observations	8247	8247
Marginal R ² / Conditional R ²	0.220 / 0.247	0.220 / 0.247

Modèle : `glmer(score~ phone*type + phone*n_group + test + (1|Participant), family = binomial, data= d2to)`

Légendes : phon = erreur phonologique ; pta = erreur phonotactique
 Séquence didactique : n.group = { E = -0.5 ; C = +0.5 }
 CI = intervalle de confiance ; p = probabilité

La diminution du risque de faire une erreur est significative au post-test à 3 mois (= T2) pour les deux groupes (*test*[T2], $p < .001$) (tableau 23 et figure 27 « test effect plot »).

Figure 27 : Estimations des effets des prédicteurs lorsque les autres prédicteurs sont fixés à leur valeur moyenne



Légendes : C = séquence didactique communicative ; E = séquence didactique explicite
 T0 = pré-test ; T1 = post-test immédiat ; T2= post-test à 3 mois
 Séquence didactique : n.group = -0.5 = E ; n.group = +0.5 = C
 Type d'erreurs : orth= erreur orthographique ; phon = erreur phonologique ;
 pta = erreur phonotactique

Les deux groupes ne diffèrent pas significativement sur la réduction totale du nombre d'erreurs (*n.group*, $p > .05$) ; mais ils diffèrent significativement sur la réduction du nombre d'erreurs par type d'erreur : en particulier pour les erreurs phonologiques et orthographiques (*type[phonologique]*n.group*, $p < .05$) .

Les estimations du modèle exprimées en log(odds) ont permis de calculer la probabilité de faire une erreur au post-test à trois mois (tableau 24)¹⁴.

Tableau 24 : Estimations du modèle en termes de probabilités de faire une erreur au post-test à 3 mois et comparaison avec les valeurs observées à T0

Phonème	Type d'erreur	Séquence didactique	Probabilité estimée à T2	Probabilité observée à T0	Différence T2 - T0
/ʒ/-j	Orthographique	E	18 [16 ; 21]	26 (44)	- 8
		C	16 [11 ; 22]	20 (40)	- 4
/ʃ/-on	Orthographique	E	11 [7 ; 15]	12 (33)	- 1
		C	9 [5 ; 16]	11 (31)	- 2
/ʒ/-j	Phonologique	E	3 [3 ; 5]	4 (19)	- 1
		C	4 [2 ; 9]	9 (28)	- 5
/ʃ/-on	Phonologique	E	20 [9 ; 36]	28 (45)	- 8
		C	23 [7 ; 55]	28 (45)	- 3
/ʒ/-j	Phonotactique	E	2 [2 ; 3]	4 (19)	- 2
		C	1 [1 ; 4]	3 (16)	- 2
/ʃ/-on	Phonotactique	E	10 [4 ; 24]	14 (35)	- 4
		C	7 [1 ; 26]	10 (31)	- 3

Légendes : () = déviation standard

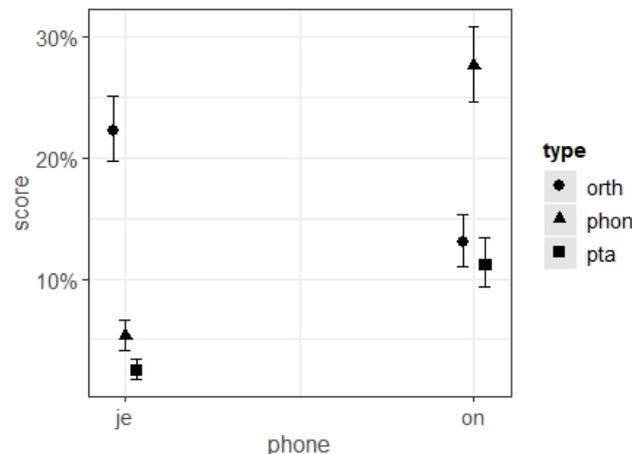
L'effet estimé de chacune des séquences didactiques sur la réduction du nombre d'erreur est calculé par phonème et par type d'erreur. Pour le phonème /ʃ/-on, avec des niveaux de départ assez similaires dans les deux groupes, l'effet estimé de la séquence didactique explicite

¹⁴ Voir Annexe 9 pour l'étude statistique complète « grapheme_phoneme » et le fichier excel pour le calcul « calcul_estimates_GP »

(E = - 8) est une réduction des erreurs phonologiques supérieure en comparaison de la séquence communicative (C = - 3). De même, pour le phonème /ʒ/-j, l'effet estimé de la séquence didactique explicite (E = -8) est une réduction des erreurs orthographiques supérieure en comparaison de la séquence didactique communicative (C = - 4). Les deux phonèmes diffèrent par le nombre et le type d'erreurs générés.

Pour l'ensemble des deux séquences didactiques au post-test à trois mois, le phonème /ʒ/-on a un risque d'erreur phonologique significativement plus élevé que le phonème /ʒ/-j ($p < .001$) : on = 21 [8 ; 46] > je = 4 [2 ; 6] (figure 28).

Figure 28 : Comparaison du risque et du type d'erreurs entre les deux phonèmes étudiés au post-test à 3 mois.



Légendes : Type d'erreurs : orth= erreur orthographique ; phon = erreur phonologique ; pta = erreur phonotactique

Pour l'ensemble des deux séquences didactiques au post-test à trois mois, le phonème /ʒ/-on a un risque d'erreur phonotactique significativement plus élevé que le phonème /ʒ/-j ($p < .001$) : on= 8 [2 ; 25] > je= 2 [1 ; 3].

Pour l'ensemble des deux séquences didactiques au post-test à trois mois, le phonème /ʒ/-j a un risque d'erreur orthographique significativement plus élevé que le phonème /ʒ/-on ($p < .001$) : je= 17 [13 ; 22] > on= 10 [6 ; 15].

4.5 Résultats pour la tâche de vocabulaire réceptif

La tâche consiste à déterminer si le mot présenté est un mot connu (voir 3.2.5). Lorsque le mot est déclaré : connu le score est de 1, et inconnu le score est de 0.

La même tâche a été conduite 2 fois : au pré-test (T0) et au post-test à 3 mois (T2).

L'analyse ne porte que sur les stimuli incluant les phonèmes cibles : 31 stimuli pour le phonème /ʒ/-on et 34 stimuli phonème /ʒ/-j par participant dont la moitié a été enseignée et l'autre moitié n'a pas été enseignée dans l'expérimentation. Les stimuli écartés de l'analyse sont ceux qui

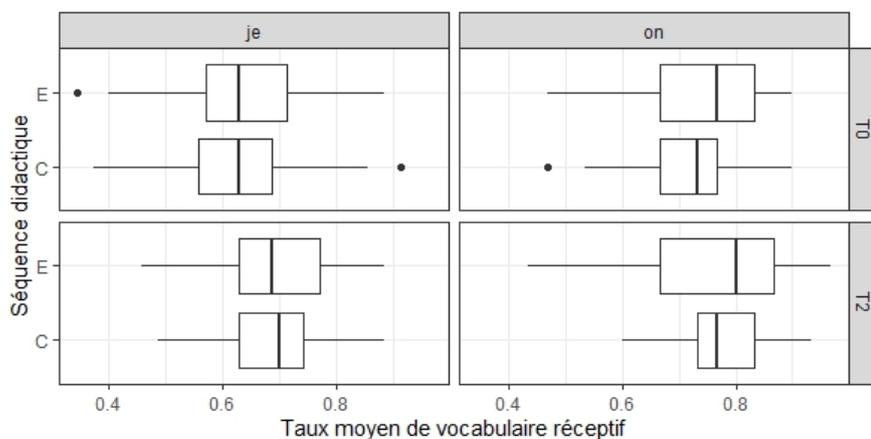
n'incluaient pas les phonèmes cibles : soit ils incluaient le phonème de la même « paire minimale écrite » (i.e., < onn, ou > pour /ɔ̃/ et < g > pour /ʒ/), soit ils équilibraient la fréquence des classes grammaticales et ne contenaient pas le phonème cible (e.g., « nous », voir annexe 5).

Les pseudo-mots (=NM, n=20) ont été inclus dans le modèle statistique pour contrôler le fait que certains participants essaient de deviner la réponse.

4.5.1 Analyse descriptive et observatoire

Les résultats recueillis sont représentés sous forme de box-plot visualisant la médiane et le premier et troisième quartile de la distribution du taux de réussite (figure 29).

Figure 29 : Représentation du taux de vocabulaire réceptif en fonction du test, du phonème et de la séquence didactique



Légendes : C = séquence didactique communicative ; E = séquence didactique explicite
T0 = pré-test ; T2 = post-test à 3 mois

Au pré-test (=T0), le groupe E réussit très légèrement mieux la tâche de vocabulaire en moyenne pour le phonème /ɔ̃/-on : taux de vocabulaire réceptif pour E = 73 % > C = 71% ; pour le phonème /ʒ/-j, les deux groupes sont au même niveau E = C = 63%. (tableau 25 et figure 29).

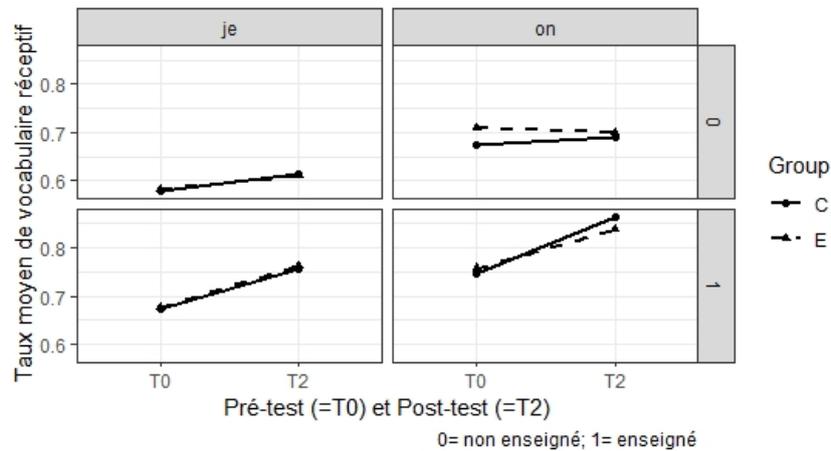
Tableau 25 : Taux de vocabulaire réceptif observés

Test	Séquence didactique	Phonème	Enseigné	Taux moyen	Taux moyen () erreur standard	Différence T2 - T0	
Pré-test (=T0)	C	<i>/ʒ/-j</i>	0	63	58 (49)		
			1		67 (47)		
	E		0	63	58 (49)		
			1		68 (47)		
	Post-test à 3 mois (=T2)	C	<i>/ʒ/-j</i>	0	71	68 (47)	
				1		75 (44)	
		E	<i>/ʒ/-on</i>	0	73	71 (45)	
				1		76 (43)	
C		<i>/ʒ/-j</i>	0	69	61 (49)	+3	
			1		76 (43)	+9	
E		0	69	61 (49)	+3		
		1		76 (43)	+8		
E	<i>/ʒ/-on</i>	0	77	70 (46)	-1		
		1		84 (37)	+8		

Au post-test à 3 mois (=T2), le progrès est supérieur dans le groupe C pour le phonème */ʒ/-on* (C = +12 % > E = 8 %) pour les lexèmes enseignés; pour le phonème */ʒ/-j* le progrès du taux de vocabulaire réceptif est presque identique dans les deux groupes pour les lexèmes enseignés et non-enseignés.

Le progrès dépend surtout du fait que les lexèmes aient été enseignés ou non (tableau 25 et figure 30).

Figure 30 : Évolution du taux moyen de vocabulaire réceptif en fonction du temps, du phonème et de la séquence didactique



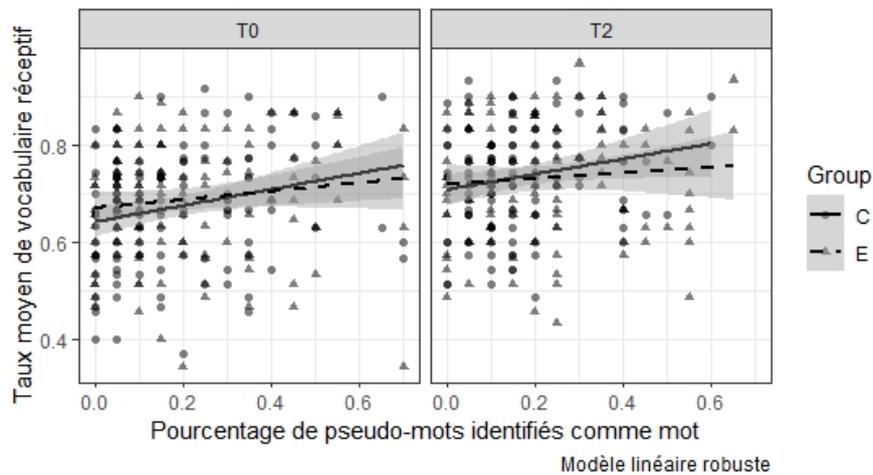
Légendes : C = séquence didactique communicative ; E = séquence didactique explicite

Une interaction entre le phonème et le fait d'avoir été enseigné est observable. Lorsque les lexèmes n'ont pas été enseignés, le taux de vocabulaire réceptif ne progresse pas pour le phonème /ɔ̃/-on (+ 1% pour C et - 1% pour E). Un progrès est observable pour les lexèmes non enseignés pour le phonème /ʒ/-j dans les deux groupes (+ 3%).

Lorsque les lexèmes ont été enseignés, le taux de vocabulaire réceptif progresse dans les deux groupes pour les deux phonèmes. Pour le phonème /ɔ̃/-on, le groupe C progresse de 12% et le groupe E de 8%. Pour le phonème /ʒ/-j, les deux groupes progressent de façon presque similaire : le groupe C de 9% et le groupe E de 8%.

Le taux de pseudo-mots cochés par erreur ($FA = \%NM$) est positivement corrélé au taux de vocabulaire réceptif pour la séquence didactique C (figure 31). Il est donc introduit comme facteur dans la modélisation linéaire pour contrôler le fait que certains élèves essaient de deviner la réponse. Il est légèrement supérieur au post-test à 3 mois (=T2) dans le groupe C.

Figure 31 : Influence du taux de pseudo-mots cochés par erreur (=FA) sur le taux de vocabulaire réceptif en fonction de la séquence didactique (=Group)



Légendes : C = séquence didactique communicative ; E = séquence didactique explicite
T0 = pré-test ; T2 = post-test à 3 mois

Aucune interaction n'est significative.

4.5.2 Analyse statistique¹⁵

L'analyse statistique a été conduite par un modèle linéaire mixte généralisé (=glmer, Bolker, 2015).

Le modèle évalue la probabilité de la reconnaissance d'un lexème comme « connu » par le participant. La variable dépendante (=score) est la valeur attribuée à la réponse (lexème déclaré « connu » score =1 ou « inconnu » score =0).

Les facteurs fixes, ou variables indépendantes, sont la séquence didactique ($groupe = \{E ; C\}$), le phonème inclus dans le stimulus ($phone = \{on ; j\}$), le moment du test ($Time = \{T0 ; T2\}$), le fait d'avoir été enseigné ou non ($taught = \{0 ; 1\}$) et le pourcentage de pseudo-mots reconnus comme des mots par erreur ($\% NM = FA$).

L'interaction entre le phonème et le fait d'avoir été enseigné est non significative.

Les effets aléatoires représentent la part de variance inhérente à chaque individu (1|participant).

Les contrastes des variables catégorielles sont codés :

- la séquence didactique, $n.group = \{E = -0.5 ; C = +0.5\}$
- le phonème, $n.phone = \{\text{« je »} = -0.5 ; \text{« on »} = +0.5\}$; l'absence d'interaction rend possible ce codage et la différence entre les phonèmes est plus précisément explorée
- le fait d'avoir été enseigné, $n.taught = \{\text{non enseigné} = -0.5 ; \text{enseigné} = +0.5\}$

¹⁵ Voir pour l'analyse statistique complète Annexe 10

- le test, $test = \{T0 = 0 ; T2 = 1\}$

Le tableau donne les valeurs estimées transformées en odds ratio et en probabilités¹⁶ des effets fixes (tableau 25).

L'intercept représente le taux de vocabulaire réceptif au pré-test, pour une moyenne des deux phonèmes et pour une moyenne des deux groupes.

Tableau 26 : Effets fixes estimés sur l'évolution de l'identification d'un lexème comme connu

<i>Prédicteurs</i>	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>	<i>Probabilités</i>	<i>CI</i>	<i>p</i>
(Intercept)	1.80	1.63 – 2.00	< .001	0.64	0.62 – 0.67	< .001
n.phone	1.58	1.48 – 1.70	< .001	0.61	0.60 – 0.63	< .001
n.group	1.01	0.87 – 1.17	0.880	0.50	0.47 – 0.54	0.880
test [T2]	1.32	1.23 – 1.41	< .001	0.57	0.55 – 0.59	< .001
n.taught	1.78	1.66 – 1.91	< .001	0.64	0.62 – 0.66	< .001
FA	2.60	1.85 – 3.65	< .001	0.72	0.65 – 0.78	< .001
Effets aléatoires						
σ^2	3.29					
τ_{00} Participant	0.14					
ICC	0.04					
N Participant	126					
Observations	16184					
Marginal R ² / Conditional R ²	0.049 / 0.088					

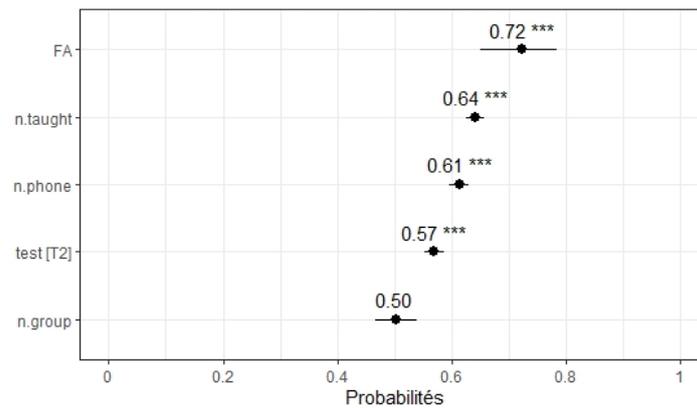
Modèle : `glmer(score ~ n.phone + n.group + test + n.taught + FA + (1|Participant), family= binomial, dph)`

Légendes : CI = intervalle de confiance ;
n.phone = { j = -0.5 ; on = 0.5 } ; n.group = { E = -0.5 ; C = +0.5 } ;
n.taught = { non enseigné = -0.5 ; enseigné = +0.5 }

Il est important d'intégrer le taux de pseudo-mots reconnus comme mots (=FA) car il est responsable d'une part de variance importante et son effet est significatif ($p < .001$) (tableau 26 et figure 32).

¹⁶ Voir Annexe 10 pour les estimations du modèle en log(odds) : « calcul_estimates_vocabulaire »

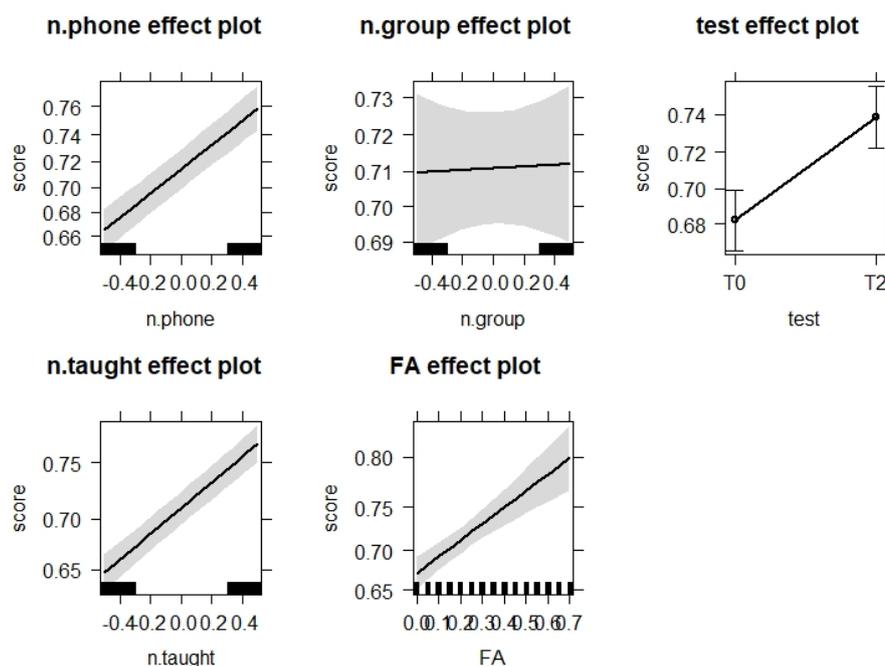
Figure 32 : Estimation des effets fixes sur le vocabulaire réceptif



Le test à 3 mois (=T2) montre une augmentation significativement différente par rapport au pré-test ($p < .001$) de 7% de chances supplémentaires de reconnaître un lexème pour la moyenne des deux séquences didactiques et les autres facteurs étant maintenus constants (tableau 26, figure 32, figure 33 « test »).

La séquence didactique n'induit pas un effet différent sur l'acquisition du vocabulaire réceptif (estimation 0.5 ; $p > .05$) (tableau 26 et figure 33 « n.group effect plot »).

Figure 33 : Représentation graphique des effets fixes du modèle



Légendes : T0 = pré-test, T2 = post-test à 3 mois ; FA = taux de pseudo-mots cochés
 n.phone = { $j = -0.5$; $on = 0.5$ }
 n.group = { E = -0.5 ; C = 0.5 }
 n.taught = { non enseigné = -0.5 ; enseigné = $+0.5$ }

Les estimations du modèle exprimées en log(odds) ont permis de calculer la probabilité de déclarer un lexème comme connu au post-test à trois mois (tableau 27)¹⁷.

¹⁷ Voir Annexe 10 pour l'étude statistique complète « vocabulaire » et le fichier excel pour le calcul « calcul_estimates_vocabulaire »

Tableau 27 : Estimations du taux de vocabulaire réceptif

Test	Phonème	Enseigné ou non	Estimation	Progrès
Pré-test (=T0)	<i>/ʒ/-j</i>	Non	52 [51 ; 52]	
	<i>/ʒ/-j</i>	Oui	66 [63 ; 68]	
	<i>/ʃ/-on</i>		63 [61 ; 65]	
	<i>/ʃ/-on</i>	En moyenne	75 [72 ; 78]	
	<i>/ʒ/-j</i>		59 [57 ; 60]	
		<i>/ʃ/-on</i>		69 [66 ; 72]
Post-test à 3 mois (=T2)	<i>/ʒ/-j</i>	Non	59 [56 ; 61]	+ 7
	<i>/ʒ/-j</i>	Oui	72 [68 ; 75]	+ 6
	<i>/ʃ/-on</i>		69 [65 ; 73]	+ 6
	<i>/ʃ/-on</i>	En moyenne	80 [76 ; 84]	+ 5
	<i>/ʒ/-j</i>		65 [62 ; 68]	+ 6
		<i>/ʃ/-on</i>		75 [71 ; 79]

Les mots enseignés et non enseignés sont significativement différents ($p < .001$). Les intervalles de confiance ne se chevauchent pas entre pré-test et post-test à 3 mois ce qui confirme la progression. La progression entre pré- et post-test à 3 mois est en moyenne de 6%. Cependant le modèle ne reflète pas l'interaction observée qui se révèle être non significative. Cette interaction observée indique que : 1) pour les lexèmes non enseignés, aucun progrès n'est observé pour le phonème */ʃ/-on* pour les deux groupes, et un progrès de 3% est observé pour le phonème */ʒ/-j* pour les deux groupes ; 2) pour les lexèmes enseignés, le progrès observé est supérieur dans le groupe C pour le phonème */ʃ/-on*, et identique pour le phonème */ʒ/-j* pour les deux groupes. Une modélisation par phonème permettrait peut-être d'estimer cette légère différence entre les groupes.

Les lexèmes contenant le phonème */ʒ/-j* sont moins connus au pré-test que ceux contenant le phonème */ʃ/-on* ($59 < 69$, $p < .001$, voir tableau 27 et « n.phone effect plot » dans figure 33). L'écart entre les lexèmes contenant les deux phonèmes est maintenu au post-test à 3 mois ($65 < 75$, $p < .001$). Lorsque le phonème */ʃ/-on* appartient à un lexème, ce lexème a 10% de chances supplémentaires d'être déclaré connu en comparaison d'un lexème contenant le phonème */ʒ/-j* ($p < .001$).

5 Discussion

Cette thèse vise à évaluer les effets comparés d'une didactique explicite et d'une didactique communicative actionnelle (implicite ciblée) sur l'acquisition des représentations phonologiques et du vocabulaire réceptif qui leur est associé. Les résultats exposés au chapitre précédent montrent que ces effets ne peuvent pas être analysés sans prendre en compte les interactions entre didactique et compétence observée (e.g., prononciation ou discrimination), mais aussi entre didactique et objet d'enseignement (i.e., phonème en jeu). Le présent chapitre propose donc de répondre à la question de recherche principale : « Dans quelle mesure une didactique explicite sans support sémantique et une didactique implicite communicative actionnelle peuvent favoriser l'acquisition des représentations phonologiques et celles du vocabulaire réceptif associé ? », sous l'angle de chacune des interactions observées dans les résultats. Ces interactions sont aussi l'objet des sous-questions de recherche : « Est-il possible d'observer un effet supérieur de la séquence didactique explicite ou de la séquence implicite en fonction de la compétence développée ? Ou en fonction de la difficulté de l'objet d'enseignement ? ». Ainsi dans une première partie la comparaison des résultats des deux didactiques est discutée en fonction de chaque compétence, puis en fonction de chaque phonème avant de proposer une synthèse des résultats.

La deuxième partie aborde les recommandations susceptibles d'être pertinentes pour la salle de classe, la troisième partie les implications théoriques de cette étude. La quatrième et dernière partie de cette discussion analyse la méthode pour essayer d'en identifier les biais et les limites afin d'évaluer l'intérêt de cette étude.

5.1 Comparaison des effets des deux didactiques

Cinq variables ont été mesurées pour suivre l'évolution de cinq compétences utilisant différents niveaux de représentations. Les représentations phonétiques ont été évaluées par une tâche de discrimination (i.e, AX tâche), les représentations phonétiques, articulatoires et phonologiques par une tâche de prononciation (i.e., tâche d'imitation de mots ou de pseudo-mots), les représentations phonémiques et graphémiques par une tâche de correspondances phonème-graphème, et les représentations lexicales par une tâche de vocabulaire réceptif (i.e., yes/no tâche). Le fonctionnement et le traitement des représentations phonétiques, phonologiques et lexicales en mémoire de travail a été mesuré par une tâche de répétition de pseudo-mots. Les effets des deux didactiques sont d'abord comparés pour chaque compétence.

5.1.1 Effets comparés des deux didactiques par compétence

5.1.1.1 Compétence de discrimination

En moyenne des deux phonèmes, les deux didactiques ne montrent pas de différence significative. Elles améliorent toutes les deux la discrimination des phonèmes cibles entre le pré-test et le post-test immédiat, et également entre le pré-test et le post-test à 3 mois. Pour l'ensemble des participants et pour les deux phonèmes, le progrès est significatif ($p < .001$) et en moyenne égal à 7%. Les résultats de cette étude montrent qu'améliorer la discrimination est possible en classe de LE, que ce soit par une didactique explicite ou par une didactique implicite. La didactique explicite a attiré l'attention de l'apprenant et a expliqué ce qui différencie deux membres d'une paire minimale. La didactique implicite a juxtaposé les éléments lexicaux porteurs de la paire minimale phonémique (e.g., « **donne-moi le ballon** ») pour augmenter la capacité du système procédural des apprenants à remarquer les régularités phonétiques spécifiques de la LE (i.e., =*noticing*). Contrairement à mes prédictions, les deux didactiques ont permis des progrès équivalents de la discrimination.

La méta-analyse de Sakai et Moorman (2018) indique que l'entraînement de la discrimination améliore de façon modérée les performances en discrimination et de façon limitée la prononciation. Le type d'entraînement, explicite ou implicite, ne semble pas influencer l'importance du progrès en discrimination, comme observé dans la présente étude où les deux didactiques induisent un progrès similaire. Au contraire, les résultats de cette méta-analyse indiquent que le type de *feed-back* correctif influence le développement de la compétence de discrimination.

Contrairement à mes prédictions et aux résultats de la méta-analyse de Sakai et Moorman (2018), les *feed-back* de la séquence didactique explicite n'ont pas été plus efficaces que ceux de la séquence didactique implicite. La séquence didactique explicite a fourni des *feed-back* composés de *recast* comme la séquence didactique implicite, mais a procuré en plus des explications. Cependant, les deux séquences didactiques n'ont pas induit d'effets différents.

L'absence de différences entre les deux séquences didactiques peut être expliquée par deux hypothèses :

- 1) Le *recast* peut avoir été la seule part efficace du *feed-back*. L'âge de mes participants peut expliquer ce résultat car les enfants et adolescents profitent mieux des *feed-back* implicites en particulier des *recast* systématiques (Lyster et al., 2013).
- 2) Le progrès observé peut peut-être aussi être induit par les autres particularités des séquences didactiques et non pas par les *recast*, bien que les *recast* soient décrits comme très efficaces dans la littérature (Lyster et al., 2013 ; Saito, 2021).

L'effet supérieur de la partie explicite du *feed-back* dans la séquence explicite peut avoir été contrebalancé par la structure de l'input dans la séquence communicative. Dans la séquence implicite ciblée communicative actionnelle, la progression de la discrimination pourrait avoir

été induite par la structure saillante des cibles dans l'input. La présentation rapprochée des deux éléments de la paire minimale dans la même phrase a pu aider les apprenants à affiner leurs représentations phonétiques. La présentation rapprochée dans la séquence implicite pourrait avoir égalé l'effet positif de la partie explicite du *feed-back* dans la séquence explicite ce qui pourrait expliquer que les deux didactiques aient induit le même progrès en discrimination.

Prononciation et discrimination sont liées et le progrès de la discrimination a pu favoriser une meilleure prononciation. Kissling (2014) a montré que l'amélioration des performances en discrimination a conduit à une amélioration significative moyenne d'environ 7% de la prononciation sans qu'une différence significative entre enseignement implicite et explicite n'ait été observée. Dans l'étude de Kissling (2014), plus les performances des participants en discrimination au pré-test étaient hautes, plus le progrès de la prononciation a été important.

Dans la présente étude, bien que les deux groupes aient développé des compétences similaires en discrimination, un progrès supérieur en prononciation devrait être observé dans le groupe explicite. En effet, le progrès en prononciation induits par les *feed-back* correctifs est supérieur lorsque l'apprenant dispose de représentations phonétiques et lorsque qu'il a conscience des différences perceptuelles caractéristiques des cibles travaillées (Saito, 2021). Dans la présente étude, les deux groupes ont disposé des mêmes représentations phonétiques (mesurées par la compétence de discrimination) mais seul le groupe explicite a travaillé consciemment les différences perceptuelles.

5.1.1.2 Compétence de prononciation

Entre le pré-test et le post-test immédiat, les deux didactiques ont conduit à un progrès significatif ($p < .05$) de la prononciation. Les élèves dont le niveau au pré-test était celui de la moyenne de la classe ont progressé en moyenne pour les deux phonèmes de 15%. Ce progrès, quoique modérément important, est néanmoins intéressant car il varie en fonction du niveau du pré-test jusqu'à +50% pour les élèves les plus faibles. Ces élèves faibles dont le niveau au pré-test témoigne de la prononciation d'un phonème non identifiable ont progressé suffisamment pour que ce phonème soit identifiable au post-test. La difficulté potentielle engendrée par la non-acquisition de la prononciation phonémique sur le processus communicatif a donc été fortement limitée grâce à l'intervention, même pour les élèves les plus faibles après quatre heures d'enseignement ciblé. Ces résultats sont conformes à ceux de la méta-analyse de Lee et al. (2015) portant sur 86 études dont l'objectif était d'évaluer l'effet de l'enseignement de la prononciation à des apprenants adultes. Enseigner la prononciation entraîne généralement une amélioration modérée à importante de cette compétence (Saito, 2012; Thomson & Derwing, 2015). Bien que conduisant à des résultats dont les conséquences pratiques sont similaires, les effets des deux didactiques sont légèrement différents.

La didactique explicite a été significativement plus efficace que la didactique implicite ($p < .05$) mais cette différence est très limitée ($\approx 2.5\%$). Cette supériorité observée de la séquence explicite est en accord avec les prédictions émises après ma revue de littérature consacrée aux

effets comparés des *feed-back* explicites et implicites sur l'acquisition de la prononciation (e.g., Saito, 2021). Cependant, la petite taille de la différence n'a pas de signification concrète pour l'enseignement et limite l'importance de ce résultat. Cette différence peut avoir différentes origines.

La différence entre les deux didactiques peut avoir été induite par : premièrement, les caractéristiques de chaque séquence d'enseignement, ou deuxièmement être liée à l'interaction de la didactique et de la difficulté induite par le phonème (cette interaction est significative, $p < .001$, voir 4.2.2).

Premièrement, l'effet de l'enseignement pourrait expliquer la différence entre les deux séquences didactiques soit par les différences entre les *feed-back*, soit par le type d'enseignement explicite ou implicite. Concernant la nature du *feed-back*, les études en contexte scolaire n'ont montré aucune différence entre les effets comparés des *feed-back* explicites et implicites (Granena & Yilmaz, 2019; S. Li, 2009; Papachristou, 2011). Concernant la nature de l'enseignement, il faut noter qu'un enseignement implicite ciblé de la prononciation n'a, à ma connaissance, pas encore été éprouvé pour étudier l'évolution de la prononciation au niveau segmental. L'enseignement explicite semble plutôt favoriser le développement de la compétence de prononciation au niveau segmental bien que l'effet d'un enseignement explicite soit inconstant (Derwing & Munro, 2005; Zhang & Yuan, 2020).

Deuxièmement, la difficulté induite par le traitement cognitif a pu être différente pour chaque phonème. Un phonème plus difficile devrait être mieux appris de façon explicite (voir 5.1.2). La prononciation est une compétence difficile à améliorer en classe de LE et comparer les effets de deux didactiques avec une différence si faible pour la moyenne des deux phonèmes cibles interroge sur la pertinence d'une telle comparaison au niveau pédagogique. Le recours à l'une ou l'autre des didactiques serait plus judicieusement guidé par l'objectif didactique (i.e., la/les représentations phonétiques et phonologiques visées), les apprenants et leurs caractéristiques et le contexte d'apprentissage.

Les compétences de prononciation et de discrimination sont fréquemment étudiées conjointement comme le montre la méta-analyse de Sakai et Moorman (2018). Au contraire, le lien entre prononciation et mémoire de travail est rarement observé alors même que la tâche la plus commune pour mesurer la mémoire de travail verbale, la répétition de non-mots, mobilise la compétence de prononciation.

5.1.1.3 Compétence de maintien de l'information verbale en mémoire de travail

Entre le pré-test et le post-test immédiat, les deux didactiques ont conduit à un progrès significatif ($p < .001$) dans les 4 séries pour les apprenants dont le niveau au pré-test est inférieur à la moyenne de la classe. Les deux didactiques n'ont pas montré de différence significative contrairement à mes prédictions. L'entraînement du traitement des sons de la LE semble donc possible soit par un travail des virelangues (groupe E) soit par un travail de mémorisation de *chunks* lexicaux ou de dialogues courts (groupe C). L'entraînement au maintien de nouveaux

phonèmes LE a permis d'augmenter significativement le nombre de phonèmes maintenus pour les deux séquences didactiques et pour les élèves les plus faibles.

Les résultats de la présente étude sont similaires à ceux de la méta-analyse de Gathercole et al. (2019). Cette méta-analyse a montré que l'entraînement de la mémoire verbale permettait de faciliter l'exécution des nouvelles procédures cognitives dans des nouvelles tâches qui les mobilisent. De même, dans la présente étude, l'entraînement des procédures cognitives à manipuler les représentations phonétiques et phonologiques a augmenté les performances des tâches qui les mobilisent. L'entraînement a facilité le traitement cognitif des phonèmes communs L1-LE mais les performances globales ont été limitées par la présence des phonèmes LE en cours d'acquisition.

La présence des phonèmes LE travaillés en classe a induit une surcharge cognitive comme annoncé dans mes prédictions. Les quatre séries de répétition de pseudo-mots diffèrent par la présence ou l'absence de phonèmes LE dans leurs listes. Deux séries ont montré des performances significativement inférieures à deux autres séries. La série qui ne contient que des phonèmes communs L1-LE (= PC) et celle qui contient des phonèmes communs L1-LE et des phonèmes LE non travaillés (PInT) n'ont pas montré de différence significative ($p > .05$). Par contre la série PC (=phonèmes communs L1-LE) a montré une différence significative en comparaison de deux autres séries ($p < .001$):

- 1) de la série qui contient les deux phonèmes travaillés (= PIT), et
- 2) de la série où la rénovation subvocale est supprimée par une articulation concurrente (= SA) et qui contient uniquement des phonèmes communs.

L'absence de différence entre PC et PInT et la différence significative entre PC et PIT indique que ce n'est pas la présence de phonèmes LE qui induit la surcharge cognitive, mais c'est celle de phonèmes LE travaillés.

Les apprenants ont vraisemblablement mobilisé leur attention dans le maintien en détail des nouvelles représentations phonétiques LE qu'ils venaient d'apprendre. Le maintien en détail de ces nouvelles représentations LE s'est fait au détriment du maintien des autres phonèmes. La présence des nouvelles représentations LE a limité les deux processus de rénovation de l'information. Premièrement, la présence de représentations phonétiques et phonologiques en cours d'acquisition (i.e., qui sont liées aux phonèmes LE travaillés) a fortement limité la répétition subvocale, comme cela a été le cas dans la condition de suppression articulatoire (=série SA). Le progrès pour la série PIT a toujours été inférieur à celui de la série SA, et cette différence est même significative pour les élèves les plus faibles (les intervalles de confiance ne se chevauchent pas). Deuxièmement, la présence des phonèmes LE travaillés semble même avoir limité le rafraîchissement attentionnel des représentations car la série PIT montre des performances inférieures d'environ 3% à la série SA. Les nouvelles représentations LE semblent avoir induit une charge cognitive au niveau du re/encodage en boucle phonologique.

Le traitement cognitif des nouvelles représentations LE et des phonèmes communs L1-LE qui les accompagnaient a été ralenti ce qui a entraîné une perte d'information verbale. Le focus acoustique, induit par la tâche de répétition de pseudo-mots, a favorisé l'usage de nouvelles représentations LE encore mal définies et difficiles à lier entre elles ce qui a vraisemblablement diminué la rénovation des représentations communes L1-LE. L'observation de la diminution du nombre de phonèmes communs L1-LE et de l'augmentation du nombre de phonèmes cibles répétés indique que l'attention a été focalisée par le maintien des phonèmes cibles LE au détriment du maintien des phonèmes communs L1-LE. Cependant les élèves dont le niveau était inférieur à la moyenne de classe au pré-test ont réussi à augmenter le nombre total de phonèmes maintenus malgré ce focus sur les phonèmes cibles LE.

Les bénéfices d'un entraînement des routines et procédures cognitives ont été supérieurs aux difficultés de traitement induites par les nouveaux phonèmes LE pour les élèves les plus faibles. Le travail pédagogique de répétition des lexèmes LE dans la séquence didactique communicative ou de répétition des représentations phonologiques dans la séquence explicite a permis la facilitation du fonctionnement des procédures cognitives. Ces résultats sont en accord avec ceux de Webb (2007) qui a montré que plus le nombre de répétitions augmente, plus les représentations orthographiques, syntaxiques et sémantiques du lexème progressent.

Dans mon cadre théorique, le support apporté par les autres représentations développées, en particulier sub-lexicales, peut aussi expliquer l'effet positif observé de l'entraînement du maintien en mémoire de travail. L'enseignement des correspondances phonème-graphème (groupe E) et les activités de lecture avec *recast* (groupe C) dans mon étude ont vraisemblablement permis une amélioration des routines et procédures cognitives pour toutes les paires phonème-graphème.

5.1.1.4 Compétence phono-graphémique

Entre le pré-test et le post-test à 3 mois, les deux didactiques ont conduit à une réduction significative du nombre total d'erreurs d'environ 3% ($p < .001$) pour les deux phonèmes et pour les 3 types d'erreurs. Les deux didactiques n'ont pas montré de différence significative sur le nombre total d'erreurs ($p > .05$), contrairement à mes prédictions.

L'absence de différences entre les deux séquences didactiques pour la moyenne des deux phonèmes et des trois types d'erreurs indique que les correspondances phonème-graphème et les règles phonotactiques ont pu être enseignées en salle de classe avec succès. Les progrès observés pour la séquence implicite suggèrent que lire à voix haute et entendre immédiatement après la lecture la forme phonologique correcte a favorisé l'acquisition des correspondances phonème-graphème et des règles phonotactiques, tout comme l'enseignement explicite. En revanche, les effets des deux didactiques ont été significativement différents sur le type d'erreurs qu'elles ont permis de réduire ($p < .05$).

Les deux didactiques ont eu des effets différents sur la réduction ciblée de la difficulté. Chacun des deux phonèmes a induit des types différents d'erreurs ($p < .001$) et l'effet comparé des

didactiques est à analyser en fonction de chaque phonème cible et de chaque type d'erreur. Pour le phonème /z/-j, le type d'erreur le plus fréquent a été orthographique. La séquence didactique explicite (groupe E = - 8%) a permis une réduction plus importante de ces erreurs que la séquence communicative (groupe C = - 4%). Pour le phonème /ʃ/-on, le type d'erreur le plus fréquent a été phonologique. La séquence didactique explicite (groupe E = - 8%) a permis une réduction des erreurs plus importante que la séquence communicative (groupe C = - 3%). La séquence didactique explicite a donc été plus efficace que la didactique implicite communicative pour une réduction ciblée des erreurs induites par la difficulté spécifique du phonème, comme préalablement annoncé dans mes prédictions. Cette difficulté spécifique est vraisemblablement étroitement liée aux interférences cross-linguistiques.

L'importance des interférences cross-linguistiques à l'interface orthographe-phonologie limite l'acquisition de nouveaux lexèmes (Hayes-Harb, Nicol, & Barker, 2010). Le travail explicite des correspondances phonème-graphème dans la séquence didactique explicite a vraisemblablement permis de mieux réduire les interférences cross-linguistiques induites par l'alphabet commun en français et en allemand que l'enseignement implicite (e.g., Ellis, N., 2006a, 2015). La didactique explicite pourrait ainsi favoriser l'acquisition du vocabulaire LE. Cependant, la revue de littérature indique une importance supérieure de la représentation sémantique sur les représentations phonologiques ; ce qui devrait avantager la didactique implicite communicative actionnelle.

5.1.1.5 Compétence en vocabulaire réceptif

Pour l'enseignement du vocabulaire, la comparaison n'est plus explicite versus implicite. Dans la séquence communicative actionnelle (i.e., appelée implicite pour les autres compétences), le vocabulaire a été enseigné dans son contexte sémantique, avec des traductions lorsque l'élève l'a demandé et il a toujours été travaillé dans un but communicatif. Par conséquent, l'enseignement du vocabulaire dans le groupe C a été explicite alors qu'il a été implicite pour toutes les représentations phonétiques, phonologiques et les correspondances phonème-graphème. Au contraire dans la séquence didactique appelée « explicite », l'enseignement a reposé uniquement sur les représentations orthographiques, phonétiques et phonologiques liées au lexème sans jamais utiliser la représentation sémantique. L'absence d'enseignement de la valeur sémantique fait qu'il ne peut pas être qualifié d'explicite pour le vocabulaire. La comparaison entre les deux séquences didactiques est plutôt explicite et communicative pour le groupe C, versus phonétique et phonologique pour le groupe E.

Entre le pré-test et le post-test à 3 mois, les deux didactiques ont permis une croissance significative du vocabulaire réceptif d'environ 6% ($p < .001$) pour les deux phonèmes et pour les 2 groupes. Les deux didactiques ne montrent pas de différence significative sur l'acquisition du vocabulaire ($p > .05$).

Les prédictions d'une supériorité de la séquence didactique communicative ne sont pas vérifiées. Les résultats de l'analyse statistique des performances de la tâche de vocabulaire

réceptif sont conformes aux effets des deux séquences didactiques précédemment décrits qui ont montré un développement assez similaire des compétences étudiées. Cependant, à ma connaissance, cette étude est la première à observer l'effet d'un enseignement du vocabulaire uniquement basé sur les formes phonétiques et phonologiques qui lui sont associées. L'enseignement explicite des représentations phonétiques et phonologiques a permis une augmentation du vocabulaire réceptif aussi importante qu'après un enseignement communicatif. Ces résultats montrent l'importance des représentations phonétiques et phonologiques pour l'acquisition du vocabulaire.

5.1.1.6 Synthèse de la comparaison de l'effet des deux séquences didactiques sur les compétences observées

Toutes les compétences observées ont significativement progressé, y compris les compétences rarement travaillées en salle de classe et peu observées telles que la discrimination, le maintien en mémoire de travail et les correspondances phonème-graphème. Les deux séquences didactiques ont donc permis aux apprenants de progresser et les élèves les plus faibles sont ceux qui ont le plus progressé.

Pour la moyenne des deux phonèmes, aucune différence significative entre les deux séquences didactiques n'a été observée dans l'évolution des compétences de discrimination, de maintien en mémoire de travail et de vocabulaire. La didactique explicite a été significativement plus favorable au développement de la compétence de prononciation mais la taille de l'effet est petite ce qui limite ses implications didactiques. Par contre, la didactique explicite a été également significativement plus favorable au développement de la maîtrise des correspondances phonème-graphème par une réduction des interférences cross-linguistiques ce qui a des implications didactiques importantes.

Peu (ou pas ?) d'études ont observé les effets d'une didactique implicite ciblée en prononciation. La présente étude indique qu'une didactique implicite ciblée a permis un progrès presque équivalent à celui de la didactique explicite et a permis à la majorité des apprenants de produire un phonème identifiable après la phase d'enseignement.

Les effets supérieurs de la didactique explicite dans certaines compétences peuvent probablement expliquer que la didactique communicative, par ailleurs implicite ciblée pour les représentations phonétiques et phonologiques, n'ait pas montré de résultats supérieurs dans l'augmentation du vocabulaire réceptif. Ce résultat souligne, pour la moyenne des deux phonèmes cibles de cette étude et leurs représentations associées, l'importance des représentations phonétiques et phonologiques dans l'acquisition du vocabulaire en FLE. L'enseignement de la prononciation et de la discrimination, habituellement négligées en classe de LE, a probablement contribué à favoriser l'acquisition du vocabulaire réceptif.

L'apprentissage du vocabulaire a peut-être également été favorisé par l'amélioration du fonctionnement des procédures de traitement cognitif des phonèmes cibles LE. Cependant, la

manipulation des phonèmes LE nouvellement acquis a induit une diminution du nombre de phonèmes communs L1-LE répétés dans la tâche de maintien en mémoire de travail.

Comme observé pour les correspondances phonème-graphème, les deux didactiques ont des effets différents en fonction du phonème.

5.1.2 Effets comparés des deux didactiques par phonème

Les résultats d'analyse décrits par compétence doivent être nuancés par la prise en compte de l'interaction fréquemment observée entre le phonème et la tâche, entre le phonème et la séquence didactique ou entre les 3 variables.

La différence entre les deux phonèmes est significative ($p < .001$) dans toutes les compétences étudiées.

5.1.2.1 Comparaison de l'effet des deux didactiques sur les représentations associées au phonème /ʒ/-on

Pour la compétence de discrimination, l'interaction entre la séquence didactique et le phonème est significative ($p < .05$). Au pré-test les deux groupes ont un niveau équivalent, alors qu'au post-test à 3 mois la didactique communicative a permis un progrès supérieur de 2% dans le groupe C en comparaison du groupe E (cette différence est significative car les intervalles de confiance ne se chevauchent pas). La différence, quoique faible, signifie que la probabilité d'une discrimination réussie a été légèrement plus favorisée par une didactique implicite ciblée et communicative que par une didactique explicite pour ce phonème.

Pour la compétence de prononciation, l'interaction entre la séquence didactique et le phonème est significative ($p < .05$). Pour le phonème /ʒ/-on, le progrès réalisé au post-test immédiat dépend essentiellement du niveau des apprenants au pré-test. Seulement pour les élèves les plus faibles, la séquence didactique explicite a permis un progrès significativement supérieur de la prononciation du phonème /ʒ/-on (cette différence est significative car les intervalles de confiance ne se chevauchent pas). Cette différence observée pour les élèves les plus faibles est vraisemblablement uniquement attribuable au fait que la séquence didactique explicite a mieux convenu aux élèves les plus faibles. Cependant à l'exception des élèves les plus faibles, aucune différence significative n'a été observée entre les deux groupes.

Un très léger progrès a été observé seulement pour le groupe E uniquement lorsque les apprenants ont prononcé au pré-test un phonème parfaitement identifiable (i.e., $ct0 = 4$). Ce léger avantage de la séquence didactique explicite est peut-être négligeable (= + 2%) car il n'a pas de signification concrète dans l'échelle de Lickert en 5 points qui a servi à l'évaluation des productions des apprenants (chaque étape d'acquisition est représentée par 20%).

Pour la compétence de maintien en mémoire de travail, les résultats de l'analyse descriptive ont indiqué que la séquence didactique communicative a permis un progrès de +42% dans le maintien du phonème /ʒ/-on, alors que la séquence explicite a permis un progrès de seulement

+7%. Ces résultats observés indiquent que ce phonème est mieux maintenu après un enseignement ciblé communicatif actionnel qu'après un enseignement explicite dans mon échantillon.

Le meilleur maintien du phonème / $\tilde{\text{ɔ}}$ -on est vraisemblablement lié aux effets de la séquence didactique implicite ciblée communicative. Les activités communicatives ont été construites pour favoriser le *noticing* en rendant plus saillantes les régularités phonétiques du phonème / $\tilde{\text{ɔ}}$ -on (et du phonème / ʒ -j). Le phonème parallèle / $\tilde{\text{ɛ}}$ -in était également présent dans l'input de la séquence didactique communicative mais pas dans un input structuré pour mettre en évidence ses traits saillants. Le maintien du phonème / $\tilde{\text{ɛ}}$ -in n'a pas été favorisé par la séquence implicite ciblée communicative bien qu'il partage de nombreuses particularités LE avec le phonème / $\tilde{\text{ɔ}}$ -on.

Le phonème parallèle / $\tilde{\text{ɛ}}$ -in n'a pas été mieux maintenu après l'intervention pour les deux groupes alors qu'il partage avec / $\tilde{\text{ɔ}}$ -on un grand nombre de caractéristiques. Le phonème parallèle / $\tilde{\text{ɛ}}$ -in partage les règles phonotactiques des voyelles nasales du français avec le phonème / $\tilde{\text{ɔ}}$ -on, ainsi que la structure de leur forme orthographique la plus fréquente <V+n>. Ces deux phonèmes ont également des fréquences dans la parole assez proches (i.e., phonème / $\tilde{\text{ɛ}}$ -in : 1.9% vs. phonème / $\tilde{\text{ɔ}}$ -on : 2.2% (Bédard, Audet, Drouin, Roy..., & Tremblay, 2017; Tremblay & Bédard, 2013; Wioland, 1991).

Le phonème / $\tilde{\text{ɛ}}$ -in a été significativement moins bien maintenu que le phonème / $\tilde{\text{ɔ}}$ -on ($p < .01$, voir 4.3.2, Analyse complémentaire A), et ce pour les deux séquences didactiques (pas de différence significative entre les groupes, $p > .05$). Le phonème parallèle / $\tilde{\text{ɛ}}$ -in a même régressé entre le pré-test et le post-test, alors que le phonème / $\tilde{\text{ɔ}}$ -on progresse dans les deux groupes. L'apprentissage implicite du / $\tilde{\text{ɔ}}$ -on a vraisemblablement été favorisé par la didactique implicite ciblée communicative actionnelle.

Pour l'acquisition des correspondances phonème-graphème, l'interaction phonème et séquence didactique n'est pas significative ($p > .05$). Cependant l'interaction phonème et type d'erreurs est significative ($p < .001$), tout comme l'interaction séquence didactique et type d'erreurs ($p < .01$). Les erreurs phonologiques ont représenté le type d'erreur le plus fréquent pour le phonème / $\tilde{\text{ɔ}}$ -on ($p < .001$). La didactique explicite a permis une diminution des erreurs phonologiques significativement supérieure à la didactique implicite ciblée ($p < .01$). Pour les deux autres types d'erreurs (i.e., orthographiques et phonotactiques), aucune différence n'a été observée entre les deux séquences didactiques pour ce phonème.

Pour l'acquisition du vocabulaire réceptif, la modélisation n'a pas pu rendre compte de l'interaction observée entre séquence didactique, phonème, temps et le fait d'avoir été enseignés ou non. En considérant juste l'analyse descriptive, les lexèmes enseignés ont été plus appris dans la séquence communicative (= + 12%) que dans la séquence explicite (= + 8%). Les lexèmes non-enseignés n'ont pas progressé dans les deux groupes. Bien que non généralisables, ces observations ont indiqué, pour les données récoltées, un effet supérieur de la didactique

communicative pour les lexèmes incluant le phonème /ʃ/-on comme prédit par le cadre théorique.

Bilan

En considérant l'ensemble des compétences étudiées pour le phonème /ʃ/-on, la séquence didactique communicative a permis un progrès supérieur à celui induit par la séquence didactique explicite en discrimination et en vocabulaire réceptif pour les lexèmes enseignés. Peu de différences sont observables entre les séquences didactiques pour la prononciation. En mémoire de travail, la différence significative du maintien entre le phonème /ʃ/-on et le phonème /ɛ/-in a montré que la progression du phonème /ʃ/-on est vraisemblablement le résultat du *noticing* induit par la séquence didactique implicite ciblée communicative. Au contraire pour les relations phonème-graphème, la séquence didactique explicite a permis une réduction supérieure du taux d'erreurs phonologiques en comparaison de la séquence communicative.

5.1.2.2 Comparaison de l'effet des deux didactiques sur les représentations associées au phonème /ʒ/-j

Pour la compétence de discrimination, au pré-test les deux groupes ont un niveau équivalent alors qu'au post-test à 3 mois, la didactique explicite a permis un progrès supérieur de 3% en comparaison de la séquence didactique implicite ciblée pour le phonème /ʒ/-j. Cette différence est significative car les intervalles de confiance ne se chevauchent pas et indique un effet supérieur de la séquence explicite.

Pour la compétence de prononciation du phonème /ʒ/-j, le progrès réalisé au post-test immédiat dépend du niveau des apprenants au pré-test. La séquence didactique explicite a permis un progrès significativement supérieur pour tous les apprenants (cette différence est significative car les intervalles de confiance ne se chevauchent pas), y compris les apprenants avancés (i.e., $ct_0 = 4$). Par contre, aucune différence significative dans l'évolution de leurs performances n'a été observée entre les deux groupes pour les élèves qui prononcent un phonème parfaitement identifiable et sans accent au pré-test (i.e., $eva = 5$, niveau maximal de l'échelle d'évaluation). La prononciation du phonème /ʒ/-j a donc été plus développée par une didactique explicite (e.g., Escudero et al., 2008) que par une didactique implicite ciblée communicative actionnelle.

Pour la compétence de maintien en mémoire de travail, les résultats de l'analyse descriptive ont indiqué que la séquence didactique explicite induit un progrès de +94% dans le maintien du phonème /ʒ/-j, alors que la séquence implicite communicative induit un progrès de seulement +25% (voir 4.3.1, tableau 16). Bien que non généralisables, ces résultats observés indiquent que ce phonème a été mieux maintenu après un enseignement explicite qu'après un enseignement communicatif ciblé dans mon échantillon. Cependant l'enseignement communicatif ciblé a permis un meilleur maintien des phonèmes communs L1-LE.

Pour la séquence PIT (qui contient le phonème /ʒ/-j), l'analyse descriptive a indiqué que la didactique communicative a amélioré un peu plus le maintien de l'ensemble des phonèmes

communs L1-LE en mémoire de travail (groupe C= + 4%) que la didactique explicite (groupe E= +3.3%) (voir 4.3.1, tableau 15, figure 19).

Ces deux effets opposés ont été induits par l'interaction entre la séquence didactique et le phonème /ʒ/-j. Le maintien du phonème /ʒ/-j semble avoir limité le maintien des autres phonèmes L1-LE.

La présence du phonème /ʒ/-j dans les stimuli semble avoir induit une charge cognitive quand ils ont été explicitement enseignés. La prise en compte consciente de la difficulté relative au phonème /ʒ/-j par le groupe E (séquence didactique explicite) a probablement permis un meilleur maintien de ce phonème (voir 4.3.1, tableau 16, figure 20). Le nombre de phonèmes /ʒ/-j répétés a ainsi augmenté le score global de phonèmes répétés (voir 4.3.1, figure 19). Cependant, la présence des phonèmes /ʒ/-j dans les stimuli semble avoir induit une charge cognitive qui a limité le taux de maintien des phonèmes communs L1-LE (voir 4.3.1, Analyse complémentaire B, figure 23). Lorsque l'apprenant a pu répéter entre 2 et 6 phonèmes /ʒ/-j, le nombre de phonèmes communs L1-LE répétés est devenu inférieur au post-test par rapport au pré-test. Au contraire, cette régression n'a pas été observée dans le groupe C.

La séquence didactique implicite ciblée communicative n'a pas produit une diminution du nombre de phonèmes communs répétés. Cependant, la séquence implicite ciblée n'a pas favorisé un maintien prioritaire du phonème /ʒ/-j. Le phonème /ʒ/-j a été moins bien restitué dans la série PIT pour le groupe C en comparaison du groupe E. La non priorisation des détails des représentations phonétiques lors de l'en/recodage de ce phonème en boucle phonologique peut peut-être expliquer qu'il ait été moins bien restitué. Par contre, la charge cognitive induite par le traitement cognitif du phonème /ʒ/-j semble avoir été inférieure dans le groupe C (ayant suivi la séquence implicite ciblée communicative), ce qui a pu libérer de l'attention pour le traitement par le système procédural des phonèmes communs L1-LE. Les phonèmes communs L1-LE ont été mieux restitués dans le groupe C (voir 4.3.1, tableau 15, figure 19). L'enseignement implicite semble avoir permis l'acquisition de représentations phonémiques dont la manipulation n'a pas induit de surcharge cognitive observable au contraire de l'enseignement explicite.

Pour les deux séquences didactiques, l'hypothèse d'une charge cognitive induite par la seule présence de phonème LE est infirmée par les résultats de la série PInT (série qui inclut des phonèmes communs L1-LE et des phonèmes LE « parallèle »). La série PInT n'a montré aucune différence significative avec la série PC (contenant uniquement des phonèmes communs L1-LE), au contraire de la série PIT incluant les phonèmes cibles. La présence des phonèmes parallèles (i.e., /ɛ/-in et /ɲ/-gn) dans la série PInT n'a pas limité le nombre de phonèmes communs L1-LE répétés.

Pour l'acquisition des correspondances phonème-graphème, il est nécessaire de considérer la réduction de chaque type d'erreur. Les erreurs orthographiques ont représenté le type d'erreur le plus fréquent pour le phonème /ʒ/-j ($\approx 17\%$, $p < .001$). Pour le phonème /ʒ/-j, la séquence

didactique explicite a permis une diminution significative et importante des erreurs orthographiques ($\approx - 8\%$, $p < .01$), supérieure aux élèves ayant suivi la séquence communicative ($\approx - 4\%$). Pour le phonème /z/-j, les erreurs phonologiques ont été beaucoup moins fréquentes ($\approx 4\%$), mais pour ce type d'erreur, la séquence didactique communicative semble avoir été plus efficace. Cependant le groupe C a fait plus d'erreurs de ce type au pré-test ($\approx 9\%$) que le groupe E ($\approx 4\%$) et les valeurs estimées au post-test à 3 mois ont été similaires dans les deux groupes. La réduction supérieure du nombre d'erreurs dans le groupe C peut être liée à la marge de progression plus importante qui existait au pré-test. Pour les erreurs phonotactiques, aucune différence n'a été observée entre les deux séquences didactiques.

Pour l'acquisition du vocabulaire réceptif, la modélisation n'a pas pu rendre compte de l'interaction observée entre la séquence didactique, le phonème et le fait d'avoir été enseignés ou non. Les deux groupes n'ont montré aucune différence significative ou observée pour le progrès du vocabulaire lié au phonème /z/-j, que le vocabulaire ait été enseigné ou non.

Dans les deux groupes, les progrès des différentes représentations travaillées a permis une augmentation de 3% du vocabulaire non enseigné lié au phonème /z/-j. Ce résultat observé n'est pas généralisable mais laisse penser que les progrès sur les représentations travaillées ont peut-être pu faciliter le traitement cognitif des représentations liées à ce phonème. La facilitation du traitement cognitif a peut-être permis aux apprenants de considérer comme du vocabulaire connu les lexèmes non-enseignés par l'expérimentatrice et qui incluaient ce phonème.

Bilan

En considérant l'ensemble des compétences étudiées pour le phonème /z/-j, la séquence didactique explicite a permis un progrès supérieur à la séquence didactique communicative pour la discrimination, la prononciation et les relations phonème-graphème. Seul l'apprentissage du vocabulaire réceptif a semblé ne pas montrer de différences entre les séquences didactiques pour ce phonème.

Les performances du maintien en mémoire de travail de stimuli incluant le phonème /z/-j peuvent être interprétées comme indiquant que la présence du phonème /z/-j a induit une charge cognitive supérieure lorsqu'il a été enseigné explicitement. Cependant l'enseignement explicite a permis un meilleur maintien de ce phonème.

5.1.3 Synthèse de la comparaison des deux didactiques

Les interactions observées dans l'analyse statistique entre l'objet d'enseignement/apprentissage, la nature explicite ou implicite de l'enseignement et la compétence observée trouvent en partie leur explication dans les difficultés cognitives. La difficulté est l'expérience dynamique d'une charge cognitive induite par la manipulation d'un objet d'enseignement/apprentissage par l'apprenant. Les deux didactiques ont limité la difficulté et ainsi favorisé l'apprentissage, mais en fonction de la tâche et/ou de l'objet d'enseignement/apprentissage, une des deux didactiques a été plus efficace.

D'un côté, les différences d'effets des deux didactiques observées peuvent être expliquées par le fonctionnement cognitif en comparant les résultats obtenus pour chacun des deux phonèmes observés. Ces différences d'effets didactiques ont des implications pour la salle de classe et pour la théorie de l'apprentissage d'une LE. D'un autre côté, en généralisant les résultats au niveau de l'ensemble des deux objets d'enseignement/ apprentissage observés, ces deux didactiques ont toutes deux eu une action ciblée qui a permis à la majorité des apprenants de progresser dans la plupart des tâches.

5.1.3.1 Analyse cognitive

L'analyse théorique de la difficulté des deux phonèmes étudiés est une des clés pour mieux comprendre les effets différents des deux didactiques. Le phonème /z/-j est considéré comme plus difficile que le phonème /ʒ/-on dans la présente étude.

Le traitement cognitif des phonèmes difficiles prend du temps. Les deux processus de rénovation en mémoire de travail peuvent prendre plus de temps en LE. Dans la présente étude, lors de la tâche de mémoire de travail, le rappel des phonèmes communs L1-LE dans la série PIT a vraisemblablement été limité non seulement par une absence de rénovation subvocale comme dans la série SA, mais aussi par une altération supplémentaire du rafraîchissement attentionnel. Les résultats de la série PIT, qui contient seulement 12% de phonèmes LE travaillés, semble indiquer que leur présence a limité les deux processus de rénovation. Les processus de rénovation ont pu être ralentis par des représentations très détaillées des phonèmes LE travaillés.

L'enseignement explicite a attiré l'attention des apprenants sur toutes les caractéristiques spécifiques de la LE et aurait ainsi favorisé un en/recodage détaillé des représentations phonétiques en boucle phonologique (Hisagi et al., 2015; Hisagi & Strange, 2011; Majerus et al., 2013; Strange, 2011). Pour la discrimination, compétence utilisant les représentations phonétiques et articulatoires, l'enseignement explicite a induit un progrès supérieur à un enseignement implicite pour l'objet d'enseignement/apprentissage le plus difficile (i.e., le phonème /z/-j). L'effet *top-down* des traces des représentations multimodales incluant certaines de ses caractéristiques a pu renforcer leur en/recodage détaillé. Pour la tâche de maintien en mémoire de travail, le phonème /z/-j a été mieux maintenu et pour la tâche de prononciation, il a été légèrement mieux prononcé après un enseignement explicite qu'après un enseignement implicite. La rénovation du phonème /z/-j a été vraisemblablement basée sur un en/recodage en boucle phonologique plus détaillé. Plus la représentation phonétique est détaillée plus sa rénovation prend du temps.

La rénovation des représentations phonétiques détaillées du phonème /z/-j lors de la tâche de répétition de pseudo-mots a pris plus de temps après un enseignement explicite qu'implicite. La rénovation du phonème /z/-j semble s'être faite au détriment du traitement procédural des phonèmes communs L1-LE qui ont été moins bien maintenus dans le groupe E en comparaison

du groupe C (voir 4.3.1 et 4.3.2). Le traitement cognitif du phonème /ʒ/-j semble avoir entraîné une charge cognitive.

La charge cognitive, observée lors de la tâche de répétition de pseudo-mots, est uniquement observable après l'enseignement explicite. Cette charge cognitive indique vraisemblablement la construction de représentations en mémoire déclarative dont la manipulation ralentit le traitement cognitif. L'enseignement explicite du phonème /ʒ/-j qui est lié à d'importantes interférences cross-linguistiques n'a pas permis de faciliter son traitement cognitif. L'enseignement explicite semble plutôt avoir induit de nouvelles représentations en mémoire déclarative sans au préalable faciliter le fonctionnement du système procédural. Cette observation suggère une possible implication théorique importante : l'enseignement explicite serait plus à même de modifier les représentations multimodales existantes directement ; c'est à dire sans modifier au préalable les routines du système procédural. L'enseignement explicite induirait des nouvelles représentations en mémoire déclarative, par une orientation du système procédural vers un but défini, en focalisant l'attention de l'apprenant sur les éléments linguistiques à mémoriser ou à manipuler dans la tâche en cours.

Au contraire de l'enseignement explicite, l'enseignement implicite ne semble pas avoir induit d'encodage détaillé. La discrimination d'un phonème plus facile, comme le phonème /ʒ/-on a plus progressé sous l'effet d'un enseignement implicite qu'explicite (voir 4.1.2). Les représentations phonétiques ont été en/recodées avec moins de détails et il suffit qu'un critère contrastif ait été encodé pour avoir permis une différenciation efficace des deux membres de la paire minimale lors de la tâche de discrimination. Les apprenants ont a priori accordé moins d'attention à l'encodage des caractéristiques de la LE car ils n'avaient pas conscience de l'objet enseigné. L'enseignement implicite ciblé a vraisemblablement permis au système procédural d'extraire les régularités LE et de les contraster pour induire la construction des représentations phonétiques LE. L'enseignement implicite pourrait modifier les procédures cognitives qui permettraient ensuite de construire de nouvelles représentations LE. Il n'a pas induit de charge cognitive lors de la manipulation de ces représentations, ce qui permet de supposer qu'elles étaient peu détaillées.

L'apprentissage implicite n'a pas induit une charge cognitive importante lors de l'usage des nouvelles représentations au contraire de l'enseignement explicite. Les phonèmes qui ont été acquis implicitement n'ont vraisemblablement pas été liés à des représentations phonétiques détaillées. En effet, leur présence n'a pas limité le maintien des autres phonèmes L1-LE en mémoire de travail (voir 4.3.1 et 4.3.2). Le traitement cognitif des phonèmes semble donc être différent après un enseignement implicite ou un enseignement explicite.

L'observation d'un traitement cognitif différent en fonction de l'aspect implicite ou explicite de l'enseignement, et en fonction du phonème amène à trois implications théoriques importantes. Premièrement, les effets différents de l'enseignement implicite ou explicite sur le traitement cognitif des représentations phonétiques et phonologiques plaident en faveur d'une

interface dynamique entre système procédural et mémoire déclarative. Deuxièmement, la négation de l'existence du phonème en tant que représentation fonctionnelle n'est pas compatible avec les différences observées du traitement cognitif entre les deux phonèmes présentant des niveaux de difficulté phonologique différents. Troisièmement, la difficulté de l'objet d'enseignement/apprentissage doit être considérée pour choisir le caractère explicite ou implicite de l'enseignement.

Le phonème /ð/-on qui induit peu d'interférences cross-linguistiques a plus progressé par un enseignement implicite ciblé pour les compétences dont le fonctionnement est principalement procédural. L'enseignement implicite ciblé du phonème /ð/-on a favorisé des performances supérieures à celles induites par la didactique explicite en discrimination, en prononciation et en maintien en mémoire de travail. Ces résultats amènent à mieux valoriser l'intérêt de l'enseignement implicite ciblé en salle de classe en fonction de l'objectif didactique et appellent à vérifier les résultats de cette étude avec d'autres phonèmes de difficultés théoriques variées.

L'enseignement explicite semble avoir permis aux apprenants de mieux maîtriser les interférences cross-linguistiques liées au phonème /ʒ/-j. Le phonème /ʒ/-j semble avoir été encodé par des représentations plus détaillées après un enseignement explicite. Plus la représentation phonétique est détaillée, plus elle est congruente avec les représentations LE et moins elle est susceptible d'activer des représentations L1. Après un enseignement explicite, le phonème /ʒ/-j a été mieux discriminé, mieux prononcé, mieux maintenu et mieux orthographié qu'après un enseignement implicite. L'enseignement explicite de la discrimination, par exemple, a facilité le contrôle des interférences cross-linguistiques qui masquent les critères contrastifs essentiels à la discrimination des trois phonèmes L1 et LE (i.e., /ʒ-ʒ°-ʃ/). Lorsque les interférences cross-linguistiques masquent les caractéristiques de la LE, le système procédural persiste dans des routines cognitives erronées et la compétence observée progresse peu ou pas. D'une façon générale, les interférences cross-linguistiques semblent avoir été mieux contrôlées après un enseignement explicite.

La séquence didactique explicite a montré qu'elle était plus à même de lutter contre les interférences cross-linguistiques dans la maîtrise des relations phonème-graphème. La supériorité de la séquence didactique explicite sur la séquence implicite ciblée communicative a été observée pour les deux phonèmes au niveau de leurs difficultés spécifiques. L'enseignement explicite a permis de plus réduire les erreurs principales de chacun des deux phonèmes observés : les erreurs orthographiques pour le phonème /ʒ/-j, et les erreurs phonologiques pour /ð/-on. L'enseignement implicite, bien que ciblé, n'a pas permis de réduire autant ces erreurs pour les deux phonèmes observés dans cette tâche. L'enseignement explicite a travaillé indépendamment chaque type de difficulté (i.e., phonologique, orthographique et phonotactiques). L'attention est susceptible d'avoir été moins monopolisée par les interférences pendant l'enseignement, et d'être restée suffisamment disponible pour permettre la construction de représentations multimodales contenant les spécificités de la LE.

L'effet de la didactique, explicite ou implicite dépend de la tâche. Pour le phonème /ʒ/-on, l'enseignement implicite induit un progrès supérieur à l'enseignement explicite pour les tâches de discrimination et de maintien en mémoire de travail, qui sollicitent essentiellement le système procédural et un progrès inférieur pour la tâche mesurant la maîtrise des correspondances phonème-graphème qui implique obligatoirement les représentations déclaratives stockées en mémoire à long terme. Ces résultats plaident eux aussi pour une interface dynamique entre système procédural et mémoire déclarative.

La réduction supérieure des effets des interférences par l'enseignement explicite a trois implications théoriques et une conséquence pour la salle de classe. Premièrement, la difficulté de l'objet d'enseignement/ apprentissage /ʒ/-j est mieux maîtrisée par un enseignement explicite. La didactique explicite est connue pour favoriser l'acquisition de la prononciation des représentations phonologiques LE les plus difficiles à percevoir (e.g., Ellis, N., 2006a) et à prononcer (e.g., Derwing & Munro, 2009 ; Isaacs, 2014 ; Zhang & Yuan, 2020). Deuxièmement, l'enseignement explicite a été plus à même de favoriser l'apprentissage en présence de multiples sources d'interférences cross-linguistiques. La séquence didactique explicite a induit une réduction supérieure des erreurs de correspondances phonème-graphème qui ont été générées par la difficulté et complexité de chacun des deux phonèmes observés. Troisièmement, les interférences à l'interface oral-écrit ont été limitées par un enseignement explicite pour les deux phonèmes cibles en contexte scolaire. Quatrièmement, ces résultats montrent l'intérêt d'établir une échelle de difficulté des objets d'enseignement/apprentissage pour mieux choisir la didactique la plus pertinente.

Face à une difficulté induite par l'objet d'enseignement/apprentissage comme celle générée par le phonème /ʒ/-j, ou par la tâche (e.g., les correspondances phonème-graphème) ou par une caractéristique de l'apprenant (comme les élèves dont la moyenne centrée au pré-test est inférieure à celle de la classe), l'enseignement explicite a été plus efficace pour induire un progrès des performances de l'apprenant.

La difficulté individuelle de l'apprenant a semblé être mieux prise en charge par une didactique explicite pour les deux objets d'enseignement apprentissage difficiles explorés dans cette étude. Pour les compétences de prononciation et de maintien en mémoire de travail, les analyses statistiques ont montré une influence significative de l'évaluation de l'élève au pré-test ($p < .001$). Pour les élèves les plus faibles, l'effet de la séquence explicite a été systématiquement supérieur à celui de la séquence didactique implicite que ce soit pour la compétence de prononciation ou de maintien en mémoire de travail. Bien que le progrès ait été supérieur avec une didactique explicite, la progression des élèves les plus faibles a toujours été plus importante que celle des élèves avancés pour les deux didactiques, comme observé par Derwing et Munro dans leur revue de littérature (2005 ; méta-analyse de Lee et al., 2015, Sakai & Moorman, 2018). Au final, les différentes sources de difficulté observées dans cette étude (i.e., phonème, tâche, difficulté de l'élève) ont été mieux prises en charge par une didactique explicite.

La variété des tâches est vraisemblablement importante pour expliquer les progrès observés après les deux types d'enseignement. Les différentes tâches ont mobilisé de nombreuses représentations liées à un même phonème. La variété des tâches et la diversité des représentations liées au même phonème ont vraisemblablement favorisé l'établissement rapide de liens entre les représentations, et facilité leur prise en charge par les routines du système procédural. Ces résultats semblent indiquer qu'un enseignement est plus efficace quand il propose des tâches variées et qu'il favorise l'utilisation de différentes représentations liées aux phonèmes cibles.

5.1.3.2 Importance d'un enseignement ciblé

Les effets équivalents des deux didactiques ciblées peuvent être expliqués par un équilibre de leurs avantages et inconvénients dans les interactions séquences didactiques et difficultés du phonème. Les deux phonèmes étudiés diffèrent sur trois points principaux de difficulté : les interférences en FLE, la difficulté articulatoire et le traitement comme une variation allophonique en L1. Les deux didactiques ciblées ont permis un progrès similaire de la prononciation de /*ʃ*-*on*. Ce phonème n'est pas difficile à acquérir en français L1 (i.e., =pas de difficulté articulatoire) mais il est connu pour être difficile en FLE. Dans la présente étude, le traitement procédural du phonème /*ʃ*-*on* a semblé être modifié grâce à l'enseignement implicite ciblé. Au contraire, les interférences et la difficulté articulatoire observées pour le phonème /*ʒ*-*j* ont été mieux prises en charge par un enseignement explicite ciblé qui semble avoir permis une modification des représentations multimodales. Les deux séquences didactiques présentent des avantages complémentaires qui s'équilibrent lors d'une comparaison reposant sur la moyenne des deux phonèmes. Elles ont cependant en commun d'être toutes deux des didactiques ciblées.

L'enseignement ciblé a attiré l'attention de l'apprenant sur les caractéristiques spécifiques de la représentation travaillée qu'il soit explicite ou implicite. D'une part, enseignées explicitement les caractéristiques spécifiques de la LE ont pu permettre de modifier les représentations déclaratives en mémoire à long terme de l'apprenant. L'enseignement explicite de ces caractéristiques a visé à favoriser l'intégration attentionnelle des caractéristiques LE dans les représentations multimodales pour induire un effet top-down en mémoire de travail. Le système procédural pourrait adapter ses routines de fonctionnement secondairement. D'autre part, l'enseignement implicite ciblé a visé à modifier les routines du système procédural en rendant plus saillantes les caractéristiques LE. La modification des routines et procédures aurait permis secondairement la construction de représentations incluant les caractéristiques LE en mémoire déclarative. Le cadre théorique permet d'expliquer comment ces deux didactiques ciblées ont permis l'apprentissage des représentations liées aux deux phonèmes cibles de cette étude.

L'absence de progrès du maintien du phonème /*ɛ*-*in* a confirmé l'importance de l'aspect ciblé de la didactique implicite ciblée communicative. L'input dans la séquence didactique implicite

ciblée communicative contenait le phonème / $\tilde{\epsilon}$ -in. La présentation du phonème / $\tilde{\epsilon}$ -in n'était cependant pas didactisée pour rendre saillantes ces particularités. Si l'apprentissage observé du phonème / $\tilde{\sigma}$ -on était uniquement le fruit de sa présence dans l'input de la séquence communicative, alors les deux phonèmes, / $\tilde{\sigma}$ -on et / $\tilde{\epsilon}$ -in, auraient dû être maintenus et appris de façon similaire ; ce qui n'est pas le cas. Le progrès du maintien du phonème / $\tilde{\epsilon}$ -in a été significativement différent de celui du phonème / $\tilde{\sigma}$ -on et a montré une régression entre pré-test et post-test pour les deux séquences didactiques.

Ces résultats ont une implication théorique. Les deux didactiques sont ciblées et le fait qu'elles aient favorisé toutes deux le progrès de la majorité des élèves plaide fortement pour l'importance de faire remarquer les particularités de la LE lors de son apprentissage en contexte scolaire (=noticing). Cette conclusion est renforcée par le fait que les deux séquences didactiques ont pu avoir des effets significatifs sur l'apprentissage des représentations liées aux deux phonèmes cibles en seulement 4H d'enseignement, alors que les apprenants avaient déjà reçu entre 300H et 400H d'input communicatif.

L'enseignement implicite ciblé a favorisé le progrès des performances des élèves dans toutes les compétences et pour les deux phonèmes cibles. L'enseignement implicite ciblé a même conduit à des progrès supérieurs à un enseignement explicite pour certaines compétences comme observé pour le phonème / $\tilde{\sigma}$ -on. Il a permis un meilleur maintien en mémoire de travail du phonème / $\tilde{\sigma}$ -on que l'enseignement explicite ciblé, sans pour autant limiter le maintien de l'information qui l'accompagnait au contraire des représentations acquises par l'enseignement explicite. L'enseignement implicite ciblé a également induit des progrès supérieurs en discrimination et pour l'acquisition du vocabulaire réceptif pour le phonème / $\tilde{\sigma}$ -on. L'enseignement implicite ciblé communicatif est à valoriser et à explorer car il ouvre des applications pour la salle de classe particulièrement intéressantes.

L'absence de différence entre les deux séquences didactiques sur le vocabulaire réceptif acquis peut s'expliquer par l'importance des représentations phonologiques au sein du lexique LE. L'enseignement explicite des représentations phonétiques et phonologiques a permis au groupe E d'égaliser les performances d'acquisition du vocabulaire réceptif du groupe C qui a lui bénéficié d'un enseignement explicite de la valeur sémantique et d'un enseignement implicite des représentations phonologiques et phonétiques.

Le lien entre la valeur sémantique et la représentation phonologique est crucial pour l'acquisition du vocabulaire. Les études mesurant l'acquisition de l'association sens-forme orthographique en LE/L2 montrent que très peu de vocabulaire peut être appris de manière incidente en établissant un lien entre ces deux représentations (Chen, 2021; Elgort et al., 2018; Nation, 2006). Bien que les représentations orthographiques soient les premières apprises (Webb, 2007), elles semblent devoir être associées aux représentations phonologiques pour que l'apprenant puisse utiliser la valeur sémantique. La première hypothèse pour expliquer les résultats du groupe E (n'ayant pas reçu d'enseignement sémantique) est que les représentations

sémantiques étaient vraisemblablement déjà liées à des représentations phonétiques, phonologiques et orthographiques LE mais que ces représentations étaient parcellaires, pas forcément construites avec uniquement des informations congruentes et probablement peu liées entre elles. Les progrès des compétences utilisant les représentations phonétiques et phonologiques ont permis aux apprenants de les affiner et de mieux les lier entre elles. L'autre hypothèse est que ce progrès ait permis de reconnaître et de faire le lien des représentations phonologiques avec la valeur sémantique lors d'un enseignement dispensé par l'enseignante responsable.

Ces résultats ont une implication théorique et induisent une recommandation pour la salle de classe. Premièrement, l'enseignement des représentations phonétiques et phonologiques semble avoir permis de restructurer les liens entre les représentations liées à un même lexème LE. D'ailleurs dans l'analyse descriptive des données, même les lexèmes non enseignés ont progressé pour le phonème /ʒ-/j dans les deux séquences didactiques. Ces résultats renforcent la théorie basant l'apprentissage sur une restructuration permanente des représentations linguistiques. Deuxièmement, la possibilité de faciliter l'acquisition du vocabulaire réceptif uniquement par le travail des représentations phonétiques et phonologiques soutient la nécessaire réintroduction de l'enseignement de ces représentations en classe de LE.

Les représentations phonétiques et phonologiques ne sont pas ou très peu enseignées aujourd'hui en classe de FLE. La comparaison des deux didactiques, pour les représentations liées aux deux phonèmes cibles, amène à repenser l'enseignement du vocabulaire qui est actuellement centré uniquement sur l'acquisition de la valeur sémantique.

5.2 Recommandations pour l'enseignement du FLE en contexte scolaire

Trois points me semblent apporter un éclairage intéressant pour émettre des recommandations pour l'enseignement du FLE en milieu scolaire. Premièrement, la présente étude a confirmé qu'une didactique ciblée s'appuyant sur un enseignement explicite peut favoriser l'apprentissage des représentations phonétiques et phonologiques, comme observé pour les représentations liées aux deux phonèmes de cette étude dans les cinq compétences observées. L'élément nouveau est l'efficacité d'une didactique implicite ciblée de ces représentations, comme constaté pour les représentations liées aux deux phonèmes de cette étude dans quatre des cinq compétences observées. Deuxièmement, les interférences cross-linguistiques à l'interface oral-écrit (correspondances phonème-graphème) ont été mieux contrôlées après un enseignement explicite pour les deux phonèmes de cette étude. L'enseignement explicite des correspondances phonème-graphème est peu présent (voire absent ?) aujourd'hui dans les méthodes ou les salles de classes. Troisièmement, l'acquisition du vocabulaire réceptif, par les participants n'ayant pas travaillé les représentations sémantiques et ayant seulement travaillé

les représentations phonétiques, phonologiques et orthographiques associées, égale celle de ceux qui ont suivi une didactique communicative actionnelle. Ces résultats illustrent l'importance des représentations phonologiques pour l'acquisition du vocabulaire. La prise en compte de la triple interaction, didactique, objet d'enseignement/apprentissage et compétence, est essentielle pour comprendre les effets d'une séquence didactique (DeKeyser, 2003 ; Ellis, N., 2006b).

La présente étude a montré un progrès des performances des apprenants dans des compétences qui favorisent l'apprentissage du vocabulaire LE (comme observé par les mesures du vocabulaire réceptif). Le fonctionnement de ces compétences a été amélioré par l'acquisition des nouvelles représentations LE phonétiques et phonologiques enseignées. Après un enseignement ciblé, explicite ou implicite, j'observe un progrès significatif des compétences de prononciation et de discrimination, du maintien des représentations phonologiques en mémoire de travail et de la maîtrise des correspondances phonème-graphème. Un travail de ces compétences en salle de classe serait profitable à l'apprentissage du FLE.

Les résultats de la présente étude indiquent que quatre heures d'enseignement ciblé ont permis d'améliorer suffisamment la prononciation pour permettre à la grande majorité des apprenants de prononcer un phonème identifiable. Les rares études réalisées en contexte scolaire ont montré que l'enseignement de la prononciation a des effets beaucoup plus limités en salle de classe que dans les études réalisées en laboratoire (e.g., Li, 2010 ; Lee et al., 2015 ; Papachristou, 2011 ; Zhang & Yuan, 2020). Dans la présente étude, j'observe un effet supérieur à ce qui a été observé auparavant en salle de classe, peut-être parce que l'enseignement a concerné des phonèmes difficiles dont la marge de progression était importante et dont l'apprentissage est fréquemment entravé par des interférences cross-linguistiques, mais probablement aussi parce que cet enseignement est ciblé et s'est appuyé sur d'autres propositions didactiques.

Le progrès observé dans les deux séquences didactiques peut être lié à la structure de la progression didactique commune aux deux séquences. La progression didactique a organisé un développement coordonné des différents types de représentations liées. Les nouvelles représentations ont été sollicitées par différentes compétences tout au long de la progression didactique. Le progrès d'un type de représentation (i.e., orthographique, phonologique ou sémantique) a pu étayer le développement d'un autre type ou d'un autre niveau de représentation (e.g, phonétique ou phonologique) (Perfetti & Hart, 2002) et permettre de faire progresser l'apprentissage de ces phonèmes difficiles. La structure de la progression didactique a peut-être ainsi contribué au développement des représentations en renforçant les effets d'un enseignement ciblé.

Le progrès observé après les deux types d'enseignement peut avoir été soutenu par les effets ciblés des deux didactiques sur les représentations liées aux deux phonèmes FLE de cette étude. L'enseignement ciblé met en évidence plus facilement les particularités de la LE que

l'apprenant doit remarquer pour construire une nouvelle présentation LE. Il aide à faire remarquer consciemment ou inconsciemment à l'apprenant les caractéristiques spécifiques de la LE (liées à l'objet d'enseignement /apprentissage) et favorise ainsi son acquisition (Housen & Pierrard, 2005; Schmidt, 2001, 2012). L'enseignement ciblé est généralement explicite.

Les enseignements ciblés, explicites versus implicites, sont peu (pas ?) comparés dans des études visant à mesurer les effets d'un enseignement des représentations phonologiques en salle de classe, en dehors des études sur le feed-back. Les études mesurant les effets d'un enseignement visant à favoriser un apprentissage implicite (= incidental) ont surtout concerné le vocabulaire ou la grammaire. L'acquisition incidentale de la grammaire semble être moins efficace qu'un apprentissage explicite (e.g., Denhovska et al., 2016) et l'acquisition incidentale du vocabulaire en LE très faible (Chen, 2021; Elgort et al., 2018; Nation, 2006). Contrairement aux résultats obtenus lors de l'acquisition incidentale de la grammaire ou du vocabulaire, l'enseignement implicite ciblé des représentations phonétiques et phonologiques a permis dans la présente étude aux apprenants de progresser en particulier pour le phonème /ð/-on.

La première recommandation pour la salle de classe est de reconsidérer l'enseignement ciblé et de revaloriser ses apports. La présente étude a montré que l'enseignement ciblé des représentations associées aux deux phonèmes choisis a été pertinent et efficace en salle de classe pour de jeunes adolescents, que cet enseignement soit explicite ou implicite.

L'aspect ciblé de l'enseignement, commun aux deux séquences didactiques, a permis de structurer efficacement l'enseignement. Le rejet des méthodes audio-orales et structuro-globale-audio-visuelles (SGAV) a écarté des méthodes didactiques l'enseignement des représentations phonologiques. Puis les méthodes communicatives ont écarté la progression structurée de la grammaire ou du lexique au profit d'une approche par tâches dont la complexification est censée accompagner l'acquisition des compétences de communication (Atmaca, 2017). Cependant dans la présente étude, l'enseignement ciblé a pu être efficace tout en étant inclus dans une didactique communicative actionnelle.

L'enseignement ciblé des représentations phonétiques et phonologiques a été efficace tout en maintenant l'objectif communicatif. L'enseignement ciblé et implicite est rarement (pas ?) observé dans des études visant à mesurer les effets d'un enseignement des représentations phonétiques et phonologiques et est, à ma connaissance, non exploré en salle de classe. Pourtant dans la présente étude, il a été facilement inclus au sein d'une didactique communicative actionnelle pour travailler des compétences spécifiques (Celce-Murcia, Dornyei & Thurrell, 1995; Trofimovich & Gatlinton, 2006).

Pour être fonctionnelle, la représentation lexicale doit lier les représentations phonétiques et phonologiques, orthographiques et sémantiques (e.g., Hulstijn, 2013 ; Perfetti & Hart, 2002). De nombreuses études indiquent que les apprenants les moins avancés préfèrent s'appuyer sur les représentations orthographiques plutôt que phonologiques pour acquérir la valeur sémantique car ils considèrent peu fiables les représentations phonologiques (Gunnarsson-

Largy et al., 2019; Mattys & Wiget, 2011; Veivo et al., 2016). Le lien entre représentation orthographique et représentation phonologique est bidirectionnel. Il peut tout autant gêner l'apprentissage en favorisant les interférences cross-linguistiques, que favoriser l'apprentissage de la représentation liée.

Lorsque les systèmes d'écriture sont communs, le support orthographique limite la perception du contraste phonémique LE (Bhide, Ortega-Llebaria, Fraundorf & Perfetti, 2020). Cependant le contraste phonémique a pu être développé dans la présente étude tout en favorisant le développement des représentations orthographiques. L'enseignement explicite ciblé des correspondances phonème-graphème a permis de mieux maîtriser les difficultés orthographiques spécifiques des phonèmes observés tout en observant un progrès de la perception de leur contraste phonémique en discrimination. L'enseignement implicite avec support orthographique a lui aussi favorisé le développement de la perception du contraste phonémique et vraisemblablement aussi de l'encodage lexical. Encodage lexical et contraste phonémique sont deux représentations phonologiques liées à la même représentation orthographique dont les processus d'acquisition semblent relativement indépendants. Le lien entre perception et représentations orthographiques est particulièrement complexe et soumis à de nombreuses interférences cross-linguistiques.

L'activation des représentations phonologiques par les représentations orthographiques génèrent inévitablement des interférences cross-linguistiques. Les représentations phonologiques sont communes entre les langues et sont liées aux systèmes d'écriture sans que la langue d'appartenance soit spécifiée dans le lien qui relie représentations phonologiques et système d'écriture (e.g., McAllister, Flege & Piske, 2002; Nakayama, Sears, Hino & Lupker, 2012; Piske et al., 2002). L'activation d'une représentation phonologique active toutes les représentations orthographiques, phonologiques et phonétiques, et sémantiques qui lui sont associées indépendamment de la langue en cours d'utilisation ou de la langue à laquelle cette représentation appartient. La représentation orthographique active toutes les représentations phonologiques qui lui sont liées.

L'incongruence entre les deux systèmes de correspondances phonème-graphème L1 et LE limite l'encodage lexical (Showalter, 2020; Showalter & Hayes-Harb, 2015), alors que le support orthographique congruent ne gêne pas l'encodage phonologique global des lexèmes LE (Bhide et al., 2020). Showalter (2020) a montré que la présentation simultanée d'une forme orthographique dont le système d'écriture n'est pas maîtrisé par l'apprenant et d'une nouvelle représentation phonologique rend très difficile l'apprentissage simultané des représentations orthographiques, phonologiques et sémantiques. La représentation orthographique incongruente induit des interférences qui génèrent une charge cognitive limitant les possibilités d'apprentissage lexical.

La deuxième recommandation pour la salle de classe est que les correspondances phonème-graphème soient enseignées explicitement lorsque le phonème est difficile à

apprendre comme les deux phonèmes de la présente étude. L'enseignement explicite a favorisé la maîtrise des interférences cross-linguistiques à l'interface orthographe-phonologie dans la présente étude. Les résultats de la présente étude montrent que l'enseignement explicite des correspondances phonème-graphème du FLE a permis une réduction significative des erreurs spécifiques de chaque phonème. Les apprenants avaient au moment de l'expérimentation déjà reçu plus de 300 heures de FLE mais un travail de quatre heures seulement a permis de réduire significativement le taux d'erreurs. L'enseignement explicite des correspondances phonème-graphème a été très profitable pour ces deux phonèmes difficiles à percevoir ou à écrire. Cependant l'enseignement explicite semble être plus efficace lorsqu'il inclut également un travail de l'ensemble des représentations phonétiques et phonologiques (Showalter, 2020).

La présente étude a montré que le travail de plusieurs compétences recrutant les représentations phonétiques et phonologiques a pu favoriser le lien entre représentations orthographique et phonologiques même par les apprenants débutants. D'après Hulstijn (2013), l'association la plus difficile à acquérir en LE pour apprendre un mot de vocabulaire est celle entre la valeur sémantique et ses représentations phonologiques.

La troisième recommandation pour la salle de classe propose de mieux valoriser l'importance du lien entre d'une part les représentations phonétiques et phonologiques, et d'autre part la représentation sémantique. Les représentations phonétiques et phonologiques devraient être de nouveau enseignées et les compétences qui les mobilisent travaillées en salle de classe pour faciliter l'acquisition du vocabulaire. La présente étude a montré la contribution importante des représentations phonologiques pour faciliter l'acquisition du vocabulaire réceptif associé aux deux phonèmes cibles de cette étude. La séquence didactique explicite, dédiée à l'acquisition des représentations phonétiques, phonologiques et à leurs liens avec les représentations orthographiques, a permis le même accroissement du taux de vocabulaire réceptif que la séquence communicative actionnelle centrée sur le processus communicatif et la valeur sémantique. Les liens bidirectionnels entre représentation sémantique et représentations phonétiques et phonologiques semblent avoir étayé leur apprentissage mutuel.

La valeur sémantique peut également faciliter l'acquisition des autres représentations lexicales, en particulier phonologiques (e.g., Hawkins, Astle & Rastle, 2015), ce qui expliquerait les résultats observés dans la séquence didactique communicative actionnelle. Le travail communicatif a pu servir de support au développement des représentations phonétiques et phonologiques grâce 1) à l'acquisition de la valeur sémantique et 2) à l'instrumentalisation didactique qui a placé les deux éléments des paires minimales au plus près dans la même phrase, et qui immédiatement après leur production, a fait entendre aux apprenants la forme phonologique correcte. Pour les deux phonèmes du FLE observés, la présente étude a montré que l'enseignement du vocabulaire a été renforcé par une intégration dans la didactique des relations entre les différentes représentations sémantiques, phonétiques et phonologiques, et orthographiques.

Les avantages complémentaires des types d'enseignement, implicite et explicite, ont pu permettre de soutenir l'acquisition des représentations phonétiques et phonologiques, sémantiques et orthographiques du vocabulaire LE. Des représentations phonologiques plus précises ont peut-être pu permettre un encodage lexical plus aisé de l'association formes phonologiques-formes orthographiques (Perfetti & Hart, 2002). Des études complémentaires permettraient d'établir quelles représentations phonétiques et phonologiques devraient être enseignées implicitement (les plus faciles ?) et quelles sont celles qui doivent impérativement être enseignées explicitement (lorsque les apprenants montrent des difficultés d'apprentissage ?). Tout en organisant l'enseignement autour du processus communicatif, il serait peut-être profitable de réintroduire un enseignement planifié et structuré des représentations phonétiques et phonologiques dans les méthodes d'enseignement en FLE.

5.3. Implications théoriques

Les implications théoriques de la présente étude apportent un éclairage différent sur la représentation phonémique dans un modèle basé sur la mémorisation d'exemplaires et confirment un certain nombre de théories de l'apprentissage en LE. Dix implications théoriques ont pu être identifiées :

- 1) les résultats sont compatibles avec l'hypothèse d'une interface dynamique entre système procédural et déclaratif,
- 2) Le phonème /ɔ̃/-on, objet d'enseignement/apprentissage le moins difficile de cette étude, peut être enseigné implicitement,
- 3) le phonème /ʒ/-je, objet d'enseignement/apprentissage le plus difficile de cette étude, doit être enseigné explicitement,
- 4) l'enseignement explicite permet de limiter les effets des interférences cross-linguistiques,
- 5) les représentations multimodales apprises explicitement sont plus difficiles à utiliser car elles entraînent une charge cognitive plus importante,
- 6) les résultats confirment l'hypothèse de restructuration des représentations linguistiques,
- 7) les deux séquences didactiques montrent l'importance de remarquer la caractéristique spécifique LE (=noticing),
- 8) une définition de l'intensité de la difficulté phonémique serait très utile pour orienter le choix de la didactique,
- 9) la représentation phonémique est une représentation phonologique importante,
- 10) le facteur décisif pour un enseignement efficace est vraisemblablement la variété des tâches et le nombre de manipulations des nouvelles représentations.

L'effet différent des séquences didactiques en fonction de la difficulté de l'objet d'enseignement/apprentissage est aligné avec la théorie d'une interface dynamique entre le système procédural et la mémoire déclarative (*dynamic interface hypothesis*, Ellis, N., 2005, 2006a, 2006b, 2015). Les résultats observés dans la présente étude indiquent que la séquence explicite a été plus efficace que la séquence implicite lorsque le phonème étudié était le plus difficile, et au contraire, la séquence didactique implicite a été plus efficace pour le phonème le plus facile. Les résultats de la séquence implicite avec l'objet d'enseignement/apprentissage le plus facile ont confirmé la théorie de Ellis (N., 2006a), selon laquelle l'apprentissage est d'abord implicite (=procédural) et basé sur l'analyse statistique de l'input pour établir les exemplaires représentatifs de la LE. Cependant le traitement procédural semble avoir peiné à surmonter les biais des filtres dynamiques de la L1 (voir 1.2). Au contraire, l'enseignement explicite, en attirant l'attention et en faisant prendre conscience à l'apprenant de la spécificité de la LE, a pu faciliter la construction des représentations multimodales qui incluent des caractéristiques LE (voir 5.1.3). Les traces de ces représentations multimodales ont vraisemblablement servi ensuite de support aux effets *top-down* qui ont permis des modifications du traitement procédural afin que le système procédural prenne en compte les caractéristiques LE. Ces effets différents en fonction du caractère explicite ou implicite de l'enseignement plaident en faveur de l'hypothèse d'une interface dynamique entre système procédural et mémoire déclarative tout comme les effets différents induits par chacune des séquences d'enseignement sur les performances de certaines compétences.

Pour une compétence dont le fonctionnement mobilise principalement les représentations construites par le système procédural, les deux séquences didactiques ont conduit en moyenne pour les deux phonèmes à des progrès similaires ; au contraire pour une compétence dont le fonctionnement implique les savoirs déclaratifs en mémoire à long terme, la didactique explicite a conduit à des progrès supérieurs. L'absence de différences entre la séquence explicite et implicite sur le progrès des performances en discrimination, en mémoire de travail et en vocabulaire réceptif pourrait être interprétée comme le résultat d'un apprentissage uniquement implicite des représentations. Cependant, le non apprentissage du phonème parallèle /ɛ/-in dans le groupe C et la supériorité de la séquence explicite pour la tâche de maîtrise des correspondances phonème-graphème, ou pour le phonème le plus difficile, ou encore pour les élèves les plus faibles, semble indiquer que l'attention accordée à des particularités LE a été capable d'induire des modifications du savoir déclaratif. Une absence théorique d'interface entre le système procédural et le savoir déclaratif comme défendu par Krashen (1982) ne peut pas expliquer ces résultats. Ces résultats sont alignés avec de nombreux autres résultats et avec mon cadre théorique (voir 1.1) soutenant l'hypothèse d'une interface dynamique entre système procédural et mémoire déclarative (Ellis, N., 2005. 2006a, 2006b, 2015).

Squire et Dede (2015) dans leur revue de littérature, montrent que la mémoire est composée de plusieurs systèmes cérébraux activant des zones cérébrales identifiables, plus ou moins connectées entre elles. Comme dans mon cadre théorique, la mémoire de travail est organisée

autour d'un système procédural, et la mémoire à long terme inclut un savoir déclaratif qui est explicitement utilisable et un savoir non-déclaratif implicite. Il est intéressant de noter que le savoir non-déclaratif décrit par Squire et Dede inclut les habitudes, les compétences, les systèmes de prime, et les formes de conditionnement. Ce savoir non-déclaratif pourrait être dans mon cadre théorique les traces du fonctionnement procédural observées dans mon expérimentation par la modification des procédures et routines cognitives relatives aux représentations phonologiques. Dans la présente étude, les deux séquences didactiques, explicite et implicite, ont permis aux apprenants de faire progresser leurs représentations phonétiques et phonologiques. Elles ont cependant toutes deux montré des limites.

Un objet d'enseignement/apprentissage moins difficile comme le /ʃ/-on a pu être enseigné implicitement avec succès car la modification des procédures peut avoir été facilitée par la saillance de ses traits distinctifs et/ou la faiblesse des interférences (DeKeyser, 2003 ; Ellis, N., 2015). La séquence didactique implicite ciblée a renforcé la saillance des traits distinctifs de /ʃ/-on et la simplicité de sa transcription orthographique a vraisemblablement favorisé un apprentissage procédural basé sur les régularités. La régularité remarquée des caractéristiques des représentations liées au phonème /ʃ/-on a probablement permis aux apprenants de mémoriser des exemplaires de plusieurs niveaux de représentation (e.g., Ellis, N., 2015). Ce phonème étant absent du répertoire phonologique des apprenants germanophones, les interférences à l'interface orthographe-phonologie ont été suffisamment limitées pour ne pas empêcher cette acquisition. La séquence didactique implicite ciblée a probablement d'abord induit une modification des routines et procédures avant de permettre une acquisition d'un savoir plus déclaratif constitué des nouvelles représentations multimodales incluant les caractéristiques LE. La didactique implicite ciblée a montré son efficacité pour l'objet d'enseignement/apprentissage le moins difficile de cette étude (i.e., le phonème /ʃ/-on). Les résultats de la séquence implicite sont cependant inférieurs à ceux de la séquence explicite pour le phonème /ʒ/-j.

Interférences et faible saillance ont été désignées dans la littérature comme des facteurs limitant l'acquisition implicite (Ellis, N., 2015). Le phonème /ʒ/-j est moins saillant que le phonème /ʃ/-on car son critère contrastif (i.e., le caractère voisé) est un critère de variation allophonique dans la L1 des apprenants. De plus, le phonème /ʒ/-j est plus difficile à prononcer, à lire et à écrire (les règles ne suffisent pas pour produire l'orthographe correcte ; les deux graphèmes (i.e., <j ;g>) représentant le phonème ont la même fréquence d'usage ; il existe nombreuses interactions entre la L1 et la LE au niveau des correspondances phonème-graphème). Les résultats de la présente étude ont confirmé que l'enseignement explicite a été plus à même de faciliter l'acquisition de cet objet d'enseignement/apprentissage lié à de nombreuses interférences cross-linguistiques (e.g., Ellis, N., 2006b ; Ellis, R., 2015 ; Escudero et al., 2008 ; Norris & Ortega, 2000).

L'enseignement explicite permet de limiter les interférences cross-linguistique dans l'apprentissage (Ellis, N., 2015). Dans la compétence de maîtrise des correspondances

phonème-graphème, compétence dans laquelle les interférences cross-linguistiques sont les plus nombreuses, la méthode explicite a été plus efficace que la méthode implicite pour les deux phonèmes. L'enseignement explicite a permis de réduire plus efficacement l'erreur la plus fréquente des apprenants lors de la manipulation des objets d'enseignement/apprentissage observés (i.e., orthographique pour /ʒ/-j et phonologique pour /ʒ/-on). Les difficultés d'apprentissage induites par des interférences cross-linguistiques ont été mieux maîtrisées après un enseignement explicite. Cependant pour les erreurs non spécifiques, les deux séquences didactiques, explicite et implicite, ont eu des effets similaires.

L'enseignement explicite, quoique plus efficace pour dépasser les difficultés spécifiques liées à l'interaction entre l'objet d'enseignement/apprentissage, la tâche et l'apprenant, a été moins performant qu'un enseignement implicite lorsque son bénéfice a été contrebalancé par la surcharge cognitive qu'il a induit. Tagarelli, Mota et Rebuschat (2011) ont observé l'apprentissage par des apprenants adultes anglophones de la grammaire d'une langue semi-artificielle, composée de mots anglais dans une syntaxe allemande, dans deux conditions : soit une acquisition implicite, soit un apprentissage explicite par la recherche de la règle. Les deux conditions ont permis un apprentissage des objets grammaticaux, bien que la condition explicite ait induit des progrès significativement supérieurs. Les capacités en mémoire de travail, mesurée par deux tâches complexes de répétition, sont prédictives du taux de jugements grammaticaux corrects seulement pour les participants ayant reçu un enseignement explicite. L'apprentissage explicite semble avoir induit la formation de représentations dont l'usage a été attentionnel. Dans la présente étude, les représentations multimodales apprises explicitement semblent aussi avoir entraîné une charge cognitive plus importante, et par conséquent avoir été plus difficiles à utiliser.

Dans la présente étude, la charge cognitive induite par les représentations multimodales acquises explicitement a été observée lors de la tâche de répétition de pseudo-mots. L'enseignement explicite des représentations phonétiques et phonologiques a favorisé le maintien du phonème cible /ʒ/-j au détriment des autres phonèmes contenus dans l'information verbale. Ce résultat n'a pas été observé pour le groupe ayant reçu un enseignement implicite. Les représentations multimodales acquises explicitement ont induit une charge cognitive plus importante dans la compétence observée. Les routines procédurales semblent n'avoir pas été encore complètement établies au moment de l'utilisation des représentations multimodales acquises explicitement. Au contraire, l'enseignement implicite semble avoir favorisé d'abord la modification des routines et procédures. Les procédures modifiées ont probablement ensuite permis de produire des représentations multimodales qui incluent des particularités de la LE. L'enseignement explicite du phonème /ʒ/-j a permis la construction d'un savoir déclaratif intégrant des caractéristiques LE qui ont été consciemment mobilisées dans la tâche de maintien en mémoire de travail.

Les résultats de la présente étude semblent confirmer l'hypothèse de restructuration des représentations linguistiques (Garlock, Walley & Metsala, 2001; Housen & Pierrard, 2005;

Housen & Simoens, 2016; McKay et al., 2001; Pajak & Levy, 2014; Ventura, Kolinsky, Fernandes, Querido & Morais, 2007). L'enseignement explicite semble avoir réussi à faire modifier la représentation commune L1-LE pour individualiser celle du phonème /ʒ/-j. La charge cognitive observée dans la tâche de maintien en mémoire de travail indique que le maintien de ce phonème a été renforcé par une représentation en mémoire déclarative. Pour ce phonème, le contraste phonémique LE semble avoir été inscrit dans la mémoire déclarative comme un caractère spécifique de la LE uniquement pour les apprenants ayant suivi l'enseignement explicite. Les représentations phonologiques L1 qui lui étaient liées avant l'enseignement ont donc été modifiées sous l'influence de l'enseignement explicite. L'enseignement explicite des représentations phonétiques et phonologiques a induit le même progrès du taux de vocabulaire réceptif que l'enseignement implicite des représentations phonétiques et phonologiques avec un enseignement explicite de la valeur sémantique du vocabulaire. L'écart entre les deux phonèmes observés du taux de vocabulaire réceptif les incluant n'a pas évolué après les séquences d'enseignement.

La différence d'acquisition du vocabulaire réceptif en fonction du phonème LE inclus dans les lexèmes peut être expliquée par la difficulté de la restructuration du savoir déclaratif L1. Le taux de vocabulaire réceptif contenant le phonème /ʒ/-j est significativement inférieur à celui contenant le phonème /ʃ/-on au pré-test et cet écart est maintenu au post-test à 3 mois. Les séquences didactiques n'ont pas permis aux lexèmes contenant le phonème /ʒ/-j d'être autant reconnus que ceux contenant le phonème /ʃ/-on. Une explication possible est que l'encodage global d'un lexème soit plus difficile quand les représentations phonologiques qui composent sa représentation lexicale demandent une restructuration plus importante des représentations L1. Dans les deux modèles qui fondent le cadre théorique de la présente étude (i.e., le modèle TBRS et le modèle PRIMIR), les représentations multimodales peuvent être modifiées soit sous l'influence d'une caractéristique remarquée par le système procédural ou soit d'une caractéristique LE imposée dans la représentation multimodale par l'association attentionnelle.

Les résultats de la présente étude semblent corroborer que sans être remarquée la caractéristique LE peut difficilement être apprise. Le phonème /ɛ/-in n'est pas mieux maintenu après la période d'enseignement bien qu'il partage les caractéristiques phonologiques, phonotactiques et orthographiques spécifiques du français avec le phonème /ʃ/-on. Ces deux phonèmes ont des fréquences proches dans la parole. Par conséquent si seul l'input est responsable de l'acquisition, les performances des apprenants de la séquence implicite auraient vraisemblablement dû montrer une progression entre le pré-test et le post-test, ce qui n'a pas été observé (voir 4.3.1, tableau 16 et 4.3.2 Analyse complémentaire A). L'acquisition d'une caractéristique LE qui n'est pas suffisamment saillante ou qui induit des interférences importantes est limitée (Lichtman, 2013; Schmidt, 2001, 2012). L'importance de faire remarquer, consciemment ou inconsciemment, à l'apprenant la spécificité de la représentation LE peut expliquer qu'en seulement 4 heures d'enseignement les progrès soient significatifs alors que les apprenants ont déjà reçu plus de 300 H d'enseignement de FLE. Le « faire

remarquer » (i.e., induire le *noticing* par l'apprenant) est un objectif didactique communs aux deux séquences d'enseignement. Il semble avoir été un élément central pour induire une modification du savoir déclaratif ou des routines du système procédural. Les limites du « faire remarquer » implicitement sont observées avec le phonème le plus difficile dans le groupe d'élèves ayant suivi la séquence implicite ciblée communicative actionnelle.

Housen et Simoens (2016) appellent à considérer l'objectif didactique avec le contexte de son apprentissage pour pouvoir limiter les difficultés de l'apprenant. Pouvoir évaluer la difficulté induite par l'apprentissage phonémique en fonction des particularités de l'apprenant serait très utile pour choisir d'enseigner implicitement ou explicitement une représentation phonémique ciblée. Les modèles actuels d'acquisition phonémique de la LE n'intègrent pas les effets des phénomènes de Sandhi ou les interactions cross-linguistiques possibles entre les systèmes de correspondances phonème-graphème. La difficulté phonémique doit être explorée par des études empiriques pour établir les critères fondamentaux susceptibles de fonder le choix didactique. L'enseignement explicite de l'objet d'enseignement/apprentissage le plus difficile à utiliser (i.e., le phonème /ʒ/-j dans cette étude) a conduit à un progrès supérieur des performances dans la majorité des compétences observées. Cependant, certains auteurs, également dans un cadre théorique basé sur l'acquisition d'exemplaires, ont contesté l'existence de la représentation phonémique.

Les premiers modèles d'acquisition basés sur l'usage de la langue et l'acquisition d'exemplaires ont été construits par l'étude des changements phonétiques et phonologiques dans des corpus longitudinaux. Bybee (2002, abstract) affirme « *change that is both phonetically and lexically gradual presents a serious challenge with phonemic underlying forms* » (p261) (=les changements (linguistiques) qui sont graduels à la fois au niveau phonétique et au niveau lexical présentent un sérieux défi pour (l'existence) de formes phonémiques sous-jacentes¹⁸). Bybee utilise comme argument central qu'un changement phonétique d'une représentation phonémique devrait diffuser dans toutes les représentations lexicales contenant le phonème car toutes les représentations orthographiques, sémantiques, phonétiques et phonologiques d'un même lexème sont liées. Cet argument fait les hypothèses : 1) qu'un phonème est indépendant de sa distribution phonotactique, 2) que l'encodage lexical est obligatoirement phonémique, et 3) que l'acquisition d'un lexème doit être précédé par l'acquisition de ses phonèmes.

Bybee (2002) étudie par exemple la déletion du /t/ ou /d/ lors du marquage flexionnel des verbes réguliers au prétérit en anglais. Elle constate que les verbes les plus fréquents sont les plus susceptibles de perdre la marque phonétique du prétérit. La non-diffusion à des verbes moins fréquents serait un argument contre l'existence d'une représentation phonémique. Cependant les phonèmes finaux sont susceptibles d'être modifiés par les phénomènes de Sandhi.

¹⁸ Traduction personnelle

Les phénomènes de Sandhi ne sont pas la conséquence d'un non-encodage lexical du phonème mais de l'influence du contexte d'occurrence comme observés pour le dévoisement final des fricatives en français (Jatteau, Vasilescu, Lamel, & Adda-Decker, 2019). Le même lexème (e.g., *partage*) est traité différemment en fonction du contexte d'occurrence entre un haut taux de dévoisement avant une pause, et très peu de variation avant une voyelle ou une consonne sonore (Jatteau et al., 2019). Deux hypothèses compatibles ensemble et en faveur de l'existence de la représentation phonémique peuvent expliquer ces résultats : 1) la représentation phonémique est différente en fonction des positions phonotactiques légales qu'elle occupe ; ce qui explique qu'un changement de la prononciation d'un phonème initial ne diffuse pas aux lexèmes qui le contiennent en position médiale, 2) une partie des variations phonétiques observées sont induites par le principe d'économie (i.e., le locuteur équilibre l'effort articulatoire et la volonté d'être compris). Ces variations sont le résultat d'une modification de la prononciation au niveau du buffer de sortie. Dans les deux cas, les variations phonétiques les plus fréquentes du lexème sont mémorisées comme des exemplaires phonétiques possibles du lexème sans que ces variations remettent en question l'existence d'une représentation phonémique.

L'étude de l'acquisition de la liaison en français montre que l'encodage lexical n'est pas obligatoirement préalablement phonémique. Chevrot, Dugua et Fayol (2009) étudie l'acquisition de la liaison en français L1 par des enfants de 2 à 6 ans au moyen d'une approche expérimentale et d'une analyse de corpus. La liaison en français est un phénomène de Sandhi qui s'observe entre deux mots dans le flux de parole. Par exemple, <les arbres> se prononce [le.zarbr] avec une re-syllabification et une inclusion de la consonne finale du mot 1 dans le mot 2. Le même lexème utilisé dans <un arbre> se prononce [œ.narbr]. Pour observer une liaison, le mot 2 suivant le mot 1 doit commencer par une voyelle quand il est prononcé isolément. Le mot 1 doit se terminer par une consonne de liaison qui n'est jamais prononcée quand le mot est à la fin d'une phrase ou quand le mot 2 commence par une consonne. Le modèle d'acquisition proposé est un modèle composé de quatre stades successifs d'acquisition. Premièrement, la structure « consonne-voyelle », très fréquente dans toutes les langues, guide la segmentation de la parole. Par conséquent, les exemplaires phonétiques et phonologiques mémorisés incluent la consonne de liaison dans la représentation lexicale (e.g., [zarbr]). La deuxième étape est induite par la variation de la consonne de liaison dans la parole ce qui induit la mémorisation de plusieurs exemplaires phonologiques d'un même mot lié à une unique représentation sémantique. La troisième étape est une étape d'acquisition de la relation correcte entre les mots (i.e., un +[narbr], deux + [zarbr]). La quatrième et dernière étape est la généralisation de la règle d'usage « un +/nX/ » et deux +/zX/ ». Le développement d'une segmentation plus précise indique l'acquisition de schémas syntaxiques, mais aussi de certains des éléments abstraits qui constituent les invariants du lexème, c'est à dire des phonèmes (Bowers, Kazanina & Andermane, 2016; Kazanina, Bowers & Idsardi, 2018). Ce modèle d'acquisition est totalement compatible avec le modèle PRIMIR qui affirme que la représentation phonémique est une abstraction des régularités entre représentations phonétiques

et sémantiques. D'ailleurs Bybee (2002 : 282) utilise le même argument pour justifier de l'existence de la représentation morphémique qui serait une catégorie émergente de l'analyse de la régularité des liens entre phonétique et sémantique. En L1 la représentation phonémique ne précéderait donc pas l'encodage lexical.

En LE, les représentations L1 sont préexistantes aux représentations LE et les modèles actuels d'acquisition phonologique sont basés sur l'hypothèse que les représentations L1, en particulier phonémiques, sont les supports de la construction du système phonologique LE (e.g., Best & Tyler, 2007; Escudero, 2009; McAllister et al., 2002; van Leussen & Escudero, 2015). Les représentations phonétiques et phonologiques LE sont considérées comme étant le fruit d'une restructuration des représentations L1. Les résultats de la présente étude sont compatibles avec l'existence des représentations phonémiques LE et une élaboration des représentations phonémiques spécifiques de la LE à partir des représentations phonémiques L1.

La représentation phonémique est indispensable à la compréhension de l'organisation lexicale. Dans la présente étude, les deux phonèmes observés ont été appris différemment après la même séquence didactique ce qui montre que leurs caractéristiques interagissent avec la compétence et le mode d'enseignement. Les représentations phonémiques ont donc été identifiables. Leurs régularités, identifiées par le système procédural ont conduit à la construction d'une représentation dédiée par représentation phonémique. Si la représentation phonémique n'existait pas, les séquences didactiques pourraient induire uniquement un encodage lexical phonétique et les deux didactiques conduiraient à des résultats similaires quel que soit le phonème. Or, les deux didactiques ont eu des effets différents sur le même phonème. De plus, l'observation de la charge cognitive induite par le phonème /z-/j après un enseignement explicite devrait également être observée après un enseignement implicite ce qui n'est pas le cas. D'ailleurs, certains modèles de reconnaissance lexicale utilisent la représentation phonémique comme le niveau de représentation permettant de lier les représentations phonétiques aux mots (e.g., le modèle TRACE : McClelland & Elman, 1986). L'importance de la représentation phonémique a pu être modélisée mathématiquement.

Les modélisations neuronales de l'identification des mots de la parole, basées sur l'appariement entre une image et des représentations phonétiques, montrent l'importance de la représentation phonémique dans un lexique de petite taille. Havard, Chevrot et Besacier (2019) ont utilisé un modèle neuronal récurrent qui a été entraîné à segmenter implicitement des phrases parlées en unité de type « mot » et à associer ces mots de façon fiable à leur image de référence. Ils montrent que l'activation des mots candidats pour cette association n'est possible que si le premier phonème est disponible. L'absence du premier phonème entraîne une diminution très importante et très rapide de la précision de l'association entre le mot et son référent visuel. Au contraire, une grande partie des phonèmes de fin de mot peut manquer sans que la précision de l'identification ne soit compromise. Ce résultat indique que le phonème est à considérer en fonction de sa distribution phonotactique. L'importance du premier phonème peut décroître si un phonème saillant est présent en fin de mot. Ces résultats montrent que la représentation

phonémique et ses caractéristiques (i.e., distribution et saillance) sont essentielles pour que les liens existants soient fonctionnels entre les représentations phonétiques et les représentations sémantiques dans un lexique de petite taille. Les résultats de la présente étude amènent aux mêmes conclusions que la modélisation de la liaison ou la modélisation neuronale de l'identification lexicale à partir de la parole. Les deux phonèmes observés sont des représentations clé pour la structure du lexique et les enseigner a pu faciliter l'acquisition lexicale en LE. Un schéma d'acquisition de la LE, basé sur la mémorisation d'exemplaires et l'usage de la langue, peut être cohérent en intégrant la représentation phonémique et sa contribution à l'apprentissage du vocabulaire.

La revue de littérature de Hulstijn (2013) conclut que la question de l'efficacité supérieure d'un apprentissage implicite ou explicite n'est pas pertinente mais que le facteur décisif est la variété des tâches et le nombre de manipulations des nouvelles représentations (Elgort et al., 2018; Robinson & Gilabert, 2007). La présente étude a confirmé que les deux types d'enseignement sont complémentaires et que la meilleure méthode didactique est celle qui prend en charge les interactions complexes entre l'objet d'enseignement/apprentissage, les caractéristiques des apprenants, les interférences cross-linguistiques en jeu (DeKeyser, 2016; DeKeyser, 2005), et le niveau et le type de compétence visée (Robinson, 2015).

5.4. Analyse de la méthode utilisée

5.4.1 Difficultés, biais et limites de cette étude

Les différents biais, limites et difficultés identifiés sont induits par cinq éléments constitutifs de la présente étude : le cadre théorique, les différences inéluctables qui existent entre les séquences didactiques, le choix des phonèmes, le choix des compétences observées et l'expérimentatrice/enseignante.

Les modèles théoriques choisis sont cohérents avec l'hypothèse d'une interface entre le système procédural et la mémoire déclarative ; il est donc peu surprenant que les résultats de la présente étude soient en faveur de cette hypothèse. Les deux modèles choisis, le TBRS pour la mémoire travail et le modèle PRIMIR pour la perception, ont défini le cadre expérimental dans lequel les systèmes procédural et déclaratif sont en contact lors de l'apprentissage des représentations phonétiques et phonologiques de la LE. Les résultats ont soutenu l'hypothèse d'une interface dynamique entre le système procédural et la mémoire déclarative. Cependant, la présente étude n'a pas observé l'évolution d'une tâche productive qui aurait peut-être apporté d'autres informations. Les enseignements implicite ciblé et explicite pourraient conduire à des observations différentes en fonction du mode réceptif ou productif (e.g., Hulstijn & de Graaff, 1994), comme observé lors de l'acquisition de certains éléments de morphosyntaxe (Spinner, 2013). Une prochaine étude pourrait viser à comparer les performances d'une même

compétence, telle que la maîtrise des correspondances phonème-graphème, selon les deux modes réceptif et productif avec les deux phonèmes observés dans la présente étude.

Les deux séquences didactiques ont utilisé le même matériel verbal et tenté de mobiliser des compétences approximativement similaires. Les facteurs linguistiques sont identiques dans les deux séquences didactiques : la fréquence des cibles et du vocabulaire lors de leur présentation, les représentations linguistiques travaillées, et la saillance des éléments ciblés. Les deux séquences didactiques sont également équivalentes sur le temps consacré à chaque compétence dans une même classe. Elles suivent aussi la même progression didactique. Certaines différences, qui ne sont pas induites par le matériel verbal support, subsistent néanmoins entre les séquences didactiques.

Une première différence pourrait exister entre les deux groupes recevant chacun un type d'enseignement car certaines demi-classes ont été plus ou moins rapides, plus ou moins intéressées ou motivées. Cependant, comme les deux groupes d'une même classe se suivaient, l'équilibre a pu être maintenu entre les séquences didactiques d'une même classe. Les différences entre les classes ne posent pas de problème puisque chaque moitié de classe a suivi une séquence didactique. Deuxièmement, certaines tâches, comme la lecture, ne sont pas exactement similaires au niveau cognitif dans les deux séquences. La lecture d'un texte (avec *recast*) et l'échange communicatif autour de ce texte n'est pas comparable à des jeux de décodage ou de travail de la conscience phonologique. Les activités communicatives ont par exemple mobilisé plus de vocabulaire. Cependant j'ai construit un enseignement relativement parallèle dans les séquences didactiques en équilibrant le temps passé par type et niveau de représentation. Les compétences visées par les exercices pédagogiques et observées ont été les mêmes, structurées dans la même progression avec les mêmes objets d'enseignement/apprentissage.

La comparaison des deux séquences didactiques aurait été fortement enrichie si une troisième séquence didactique mixte, implicite et explicite, avait pu être expérimentée. Il serait intéressant d'observer les effets d'une séquence communicative actionnelle proposant des activités intégrées, explicites et ciblées sur une forme. Cependant, pour construire une séquence mixte efficace, les résultats de la présente étude sont essentiels car il faut adapter le type d'enseignement à la difficulté de la cible et à la compétence visée. Cette troisième séquence didactique pourrait faire l'objet d'une prochaine étude avec un groupe contrôle suivant une simple didactique communicative actionnelle non ciblée. La troisième séquence didactique enseignerait explicitement les objets difficiles et implicitement les objets faciles à utiliser en fonction des compétences et en fonction des résultats par compétence obtenus dans la présente étude.

Bien que je ne puisse pas exclure que les apprenants aient travaillé consciemment l'objectif didactique dans la séquence implicite ciblée communicative actionnelle, il est peu probable que l'ensemble des participants aient consciemment construit un apprentissage de leurs

représentations phonétiques et phonologiques tout en participant activement aux activités communicatives. Dans l'étude de Denhovska et al. (2016), un tiers seulement des participants adultes avaient identifié la nature grammaticale de l'enseignement implicite. Le plus probable, dans la présente étude avec de jeunes adolescents, est que l'apprentissage ait été principalement implicite et inconscient mais que l'instrumentalisation didactique ait permis aux mécanismes procéduraux d'être réorientés par les incohérences entre les informations *bottom-up* et *top-down* lors de la construction des représentations multimodales en mémoire de travail. Les progrès observés ont pu être induits par une progression spontanée à laquelle s'est ajouté l'effet de la didactique.

La présente étude n'inclut pas de groupe contrôle qui ne recevrait aucune intervention. Il est possible que les effets observés aient été en partie liés à une progression spontanée. Cependant, plusieurs éléments indiquent que les progrès observés ont été induits par les séquences didactiques : 1) les effets observés sont présents dès le post-test immédiat, 2) une progression spontanée ne pourrait pas expliquer les effets différents observés entre les deux didactiques pour un même phonème, 3) il est peu probable qu'après 300H d'enseignement de FLE, les apprenants puissent progresser rapidement en 4H sans que ce progrès soit induit par l'enseignement. De plus, l'enrôlement de participants qui sont supposés ne retirer aucun bénéfice de l'étude me pose un problème éthique et j'ai choisi de ne pas inclure de groupe contrôle « inerte ». Chaque groupe représente un élément de référence pour l'autre groupe. L'attribution au hasard des participants dans les groupes, les analyses statistiques incluant les participants comme *random factors*, et la cohérence globale des résultats entre compétences pour un même phonème ont tout de même permis l'interprétation des résultats.

Les deux séquences didactiques ont été construites autour de deux phonèmes dont la prononciation est connue pour être difficiles à acquérir en FLE pour des apprenants germanophones. Peut-être que les résultats auraient été différents pour d'autres phonèmes enseignés aux mêmes apprenants. Cette étude a révélé cependant un point important : les effets du type d'enseignement dépendent de la difficulté phonémique.

Une des difficultés de la présente étude a été de ne pas avoir pu choisir les phonèmes en fonction des critères de difficulté établis dans la littérature. J'ai proposé un certain nombre de critères qui peuvent peut-être constituer la base d'une telle grille d'évaluation de la difficulté phonémique. Cependant, il faudrait cibler un critère qui constitue l'unique différence entre deux phonèmes et ainsi, pas à pas, comparer les critères de difficulté les uns avec les autres. Le rôle de la complexité du phonème et de son interaction avec la difficulté reste encore à établir. Observer des phonèmes qui n'ont pas la réputation de difficulté des deux phonèmes choisis apporterait peut-être un autre éclairage sur la difficulté, en particulier en fonction de la compétence impliquée dans cette observation. Par exemple, dans quelle mesure une difficulté dans une compétence particulière, comme les règles de correspondances phonème-graphème qui impliquent l'application d'une règle contextuelle (i.e., le contexte vocalique <i ; e> transforme la relation graphème-phonème <s>-/s/ en <s>-/z/ en fonction de la position du

graphème et de son contexte dans le mot), ou de l'influence d'une lettre muette fréquente associée au graphème représentant le phonème, doivent-elles être prise en compte dans une échelle de difficulté de l'acquisition phonémique ? La relation étroite qui semble unir la représentation lexicale et la représentation phonémique est une des limites de la présente étude.

La présente étude aurait été plus complète si l'encodage phonétique lexical avait été mesuré. Une tâche de décision lexicale orale, dans laquelle les participants doivent rapidement déterminer si le mot entendu est un mot ou non, aurait pu présenter les lexèmes déclarés connus par les participants avec des pseudo-mots différents de ces lexèmes uniquement au niveau du phonème cible. La mesure dans ce type de tâche est le temps de réponse et la justesse des réponses. Un temps de réponse court et l'acceptation erronée d'un pseudo-mot indique un encodage lexical phonétique. Une autre solution consiste à utiliser les méthodes de suivi du regard (Weber & Cutler, 2004). Cependant, en raison du caractère écologique de la présente étude (i.e., effectuée en contexte scolaire avec les moyens coutumiers d'enseignement) et des contraintes temporelles qui ont limité le temps de test et le temps d'enseignement par élève, l'encodage phonétique lexical n'a pas pu être directement observé.

Une seule tâche par compétence a permis de suivre le progrès des apprenants. La présente étude a choisi d'explorer les représentations phonétiques et phonologiques au travers de l'observation du fonctionnement de cinq compétences. Une exploration de chaque compétence de façon approfondie aurait été plus satisfaisante. La tâche de prononciation, par exemple, gagnerait à être complétée par une tâche de lecture et une tâche de production spontanée. Il n'a pas été possible de mobiliser les élèves plus longtemps pour les tests et une évaluation plus précise par tâche a donc été écartée au profit d'une vue d'ensemble des compétences mobilisant les représentations travaillées. Pour la même raison, les tâches passées individuellement (i.e., la prononciation et la répétition de pseudo-mots) n'ont pu être réalisées qu'au pré- et post-test immédiat car le passage de ces tâches perturbait le fonctionnement de la classe pendant 3 jours.

Il est possible que les apprenants aient été plus engagés dans leur apprentissage car l'expérimentatrice a été l'enseignante de toutes les séquences didactiques. L'intérêt pour l'évènement nouveau et une nouvelle enseignante a pu favoriser leur apprentissage. Une des six classes participantes a eu une remplaçante de l'enseignante en poste pendant trois des six semaines expérimentales. Cet effet de nouveauté n'a pas entraîné des résultats différents pour cette classe. Néanmoins, bien qu'un effet bénéfique de la nouveauté ne puisse pas être exclu dans l'observation du progrès, cet effet aura bénéficié aux deux séquences didactiques. Un autre effet susceptible de favoriser le progrès est le fait que l'expérimentatrice soit francophone.

Habituellement les enseignants de LE ne sont pas natifs de la langue enseignée et des études ont montré que leur enseignement est tout autant efficace que celui des natifs (e.g., Levis, Sonsaat, Link & Barriuso, 2016). Cependant, Ellis (N., 2006b) a montré que le système de la L1 contraint la perception et la production de la LE. Il serait intéressant de mesurer la variabilité de la prononciation d'un locuteur non-natif car sa volonté de bien prononcer devrait en

permanence lutter contre les influences cross-linguistiques. En admettant que les productions verbales d'un locuteur non-natif varient plus que celles d'un natif, l'absence de variabilité dans la production de l'expérimentatrice a pu favoriser l'acquisition implicite par une plus grande régularité dans le matériel verbal produit.

La variation peut interagir avec l'effet des séquences didactiques. La méta-analyse de Sakai et Moorman (2018) montre que les entraînements phonétiques et phonologiques sont plus profitables (= taille d'effet supérieure) aux apprenants lorsqu'ils sont massifs et peu variables, c'est-à-dire ciblés autour d'un prototype (Giannakopoulou, Brown, Clayards & Wonnacott, 2017). L'exemplaire entendu est encodé en représentations phonétiques en même temps que celle de la marque acoustique du locuteur (Lachs, McMichael & Pisoni, 2003). Chaque locuteur a des traits spécifiques acoustiques qui sont encodés dans les représentations phonétiques. Une variation induite par différents locuteurs est profitable aux apprenants avec de hautes capacités de maintien en mémoire de travail, et est délétère pour les apprenants avec de faibles capacités (Antoniou & Wong, 2015). La variation entre locuteurs, et vraisemblablement aussi celle d'un locuteur LE, devrait interagir avec l'effet des séquences didactiques. Cependant, aucune étude à ma connaissance n'a mesuré la variabilité des productions verbales des enseignants de LE.

5.4.2 Intérêts de cette étude

Cette étude apporte plusieurs éclairages aux études existantes. Les intérêts de cette étude concernent sa conception et sa réalisation (contexte, participants, durée, quantité de données), ses résultats au sujet de l'efficacité didactique, les compétences observées qui sont largement absentes des études en milieu scolaire, ses implications en classe de LE et ses implications théoriques.

Cette étude a été réalisée en milieu scolaire, au sein du cours normal de FLE. Peu d'études ont cette validité écologique, et peu d'études concernent le contexte scolaire. Quatre études seulement sur 86 se sont déroulées en contexte scolaire dans la méta-analyse de Lee et al. (2015). Les résultats de la présente étude sont probablement plus proches de ce que l'enseignant.e peut induire chez ses apprenants, que ceux d'une étude réalisée sur une journée pendant laquelle entraînement et test se succèdent. La présente étude connecte recherche et pratique enseignante (Levis & Wu, 2018).

La présente étude a montré qu'une didactique ciblée permet en seulement quatre heures d'enseignement de faire progresser la majorité des élèves. Toutes les compétences étudiées ont significativement progressé. Certaines des compétences travaillées dans cette étude sont absentes ou très peu travaillées en classe de LE.

La discrimination, le maintien en mémoire de travail et les correspondances phonème-graphème ne constituent pas aujourd'hui des compétences développées en LE à l'école. La présente étude a montré que l'enseignement peut les faire progresser et que ces compétences sont complémentaires pour permettre le développement de l'ensemble des représentations

phonétiques et phonologiques et du vocabulaire réceptif. Ces compétences renforcent également le développement de la prononciation qui est fondamentale pour la réussite du processus de communication.

Un enseignement implicite construit pour augmenter la saillance des éléments distinctifs du phonème /ʒ/-on a pu favoriser l'acquisition du vocabulaire et des représentations phonétiques et phonologiques qui lui sont associées, même lorsque le contraste phonémique LE concerné est réputé difficile à acquérir. Les méthodes d'enseignement peuvent soutenir les enseignant.e.s en préparant le matériel didactique en conséquence pour chaque compétence.

Les enseignant.e.s considèrent que la prononciation est importante et qu'il faut la travailler en classe de LE. Cependant, une grande majorité des enseignant.e.s manque de confiance et pense manquer de formation (Levis et al., 2016). La compétence de prononciation est au final peu ou pas travaillée en contexte scolaire. La présente étude apporte une solution en montrant l'efficacité d'un enseignement implicite de la prononciation. Aucun savoir technique n'est nécessaire pour proposer cet enseignement implicite. Au contraire, l'enseignement explicite demande une maîtrise technique.

Les méthodes utilisées dans la séquence didactique explicite (i.e., la méthode articulatoire et la méthode verbo-tonale) demandent une formation préalable qui est peu (pas ?) dispensée aujourd'hui dans les centres de formation. Former les enseignant.e.s à ces techniques leur permettrait d'exploiter pleinement les résultats de la présente étude.

La présente étude a montré l'importance des interactions entre séquence didactique, difficulté du phonème, difficultés de l'apprenant et compétence en jeu. Plus l'enseignant.e disposera de techniques différentes, plus il pourra s'adapter aux besoins spécifiques de ses apprenants (Levis & Wu, 2018). L'enseignant.e gagnerait également à mieux comprendre la nature des interférences cross-linguistiques vécues par ses apprenants pour pouvoir construire des séquences didactiques efficaces.

Les interactions cross-linguistiques sont spécifiques des langues en contact mais la langue natale ou la langue à apprendre par les participants est principalement l'anglais dans les études en LE (83 études sur 86 dans la méta-analyse de Lee et al., 2015). Cette étude concerne des apprenants Suisses Allemands de FLE. Elle apporte un éclairage nouveau sur comment ces deux langues interagissent. Par exemple, un point de difficulté du phonème français /ʒ/-j est la perception de son caractère voisé comme une variation allophonique en allemand, alors que le phonème /ʃ/-ch est l'élément de sa paire minimale en français. L'enseignement des représentations phonétiques et phonologiques par une didactique explicite a été le plus à même de permettre à l'apprenant de dépasser les interférences phonologiques cross-linguistiques, mais il faut qu'elles soient identifiées et reconnues dans les erreurs des apprenants par l'enseignant.e. D'autres interférences ont été mise en évidence dans cette étude.

Les interférences cross-linguistiques sont nombreuses à l'interface orthographe-phonologie. Cependant, les correspondances graphème-phonème ne sont jamais enseignées explicitement.

La présente étude a montré qu'un enseignement explicite des correspondances phonème-graphème est plus favorable à la réduction des interférences les plus importantes pour les phonèmes /z/-j et /ʒ/-on. L'effet bénéfique ou délétère du support orthographique en LE est sujet à controverse dans la littérature. La présente étude montre que l'effet du support orthographique doit être pensé en fonction des langues en contact, du niveau des apprenants et du type implicite ou explicite d'enseignement. L'effet du support orthographique semble dépendre d'une part, de l'existence d'un système d'écriture commun ou non, et d'autre part de la charge cognitive induite par les interférences entre les systèmes d'écriture. La question ne peut vraisemblablement pas trouver une seule réponse, mais des auteurs appellent à repenser l'introduction systématique du support orthographique dès le premier jour d'enseignement de la LE (Hayes-Harb et al., 2010).

Cette étude a suivi une cohorte importante d'élèves longitudinalement et a inclus un post-test à 3 mois. Le nombre d'élèves inclus (=127), l'âge des élèves (= 12-13 ans), l'enseignement d'une séquence didactique complète (= 6 semaines) et la durée du suivi des performances (= 18 semaines) distinguent cette étude d'une grande majorité d'études dans lesquelles : 1) la phase d'enseignement et la phase de test sont effectuées le même jour, 2) l'âge des participants est autour de 25 ans, 3) une ou deux compétences seulement sont explorées conjointement. Les résultats de cette étude sont vraisemblablement assez représentatifs des processus d'apprentissage et d'acquisition en classe de LE et ils permettent d'émettre quelques recommandations.

La présente étude s'inscrit dans un cadre théorique de l'apprentissage par la mémorisation d'exemplaires. Elle montre que ce cadre théorique peut expliquer des résultats présentés comme contradictoires dans la littérature. Le cadre théorique et les résultats de la présente étude sont cohérents avec un apprentissage implicite des représentations phonétiques et un apprentissage explicite de représentations phonologiques. Les deux types d'apprentissage peuvent advenir simultanément ou indépendamment et induire des effets différents. Ainsi il serait possible d'apprendre les représentations phonétiques et phonologiques LE soit par l'acquisition des représentations phonétiques soit par celle des représentations phonologiques. Il est possible de les apprendre implicitement ou explicitement. Enfin le cadre théorique de la présente étude permet également de définir la charge cognitive et de lier son observation aux effets sur l'apprentissage du caractère implicite ou explicite de l'enseignement, à la difficulté de l'objet d'enseignement/apprentissage et à la compétence observée.

6 Conclusions

L'enseignement des représentations phonétiques et phonologiques a été délaissé avec l'avènement des méthodes communicatives. Cependant, une didactique communicative ne permet pas toujours de surmonter les difficultés cross-linguistiques en particulier phonétiques et phonologiques. Depuis une quinzaine d'années, l'enseignement de la prononciation est de nouveau recommandé (e.g., Derwing & Munro, 2005) et bénéficie d'un document descriptif et prescriptif attiré dans le cadre du CECR (Piccardo, Goodier & North, 2018). De nombreuses études ont montré qu'en laboratoire et avec des apprenants adultes, un entraînement de la perception permettait d'améliorer la prononciation (e.g., Sakai & Moorman, 2018), en particulier quand cet entraînement comprend des feed-back correctifs (e.g., Saito, 2021). La salle de classe est néanmoins peu étudiée (DeKeyser, 2003 ; Norris & Ortega, 2000).

La présente étude vise à évaluer les effets d'un enseignement des représentations phonétiques et phonologiques en contexte scolaire. Les élèves recrutés dans six classes ont suivi un enseignement ciblé pendant six semaines durant les heures de cours dédiées à l'apprentissage du FLE. Les 127 participants (âge moyen = 12;6) sont des élèves ordinaires de dernière année de primaire (=8H) et de première année de cycle d'orientation (=9H) en Suisse germanophone. Aucune sélection des participants n'a été effectuée et tous les élèves ont été inclus quelle que soit leur langue première ou qu'ils aient reçu un diagnostic éventuel de trouble du langage. L'objectif était de mener cette étude dans les conditions de la classe ordinaire. L'enseignement a visé les représentations phonétiques et phonologiques de deux phonèmes cibles.

Certaines erreurs segmentales sont très délétères pour le processus communicatif (Isaacs, 2018; Thomson & Derwing, 2015). Les deux phonèmes enseignés dans cette étude sont connus pour induire des difficultés même chez des apprenants avancés. Ces deux phonèmes, le /ʒ/-j et le /ʃ/-on ont un *functional load* important (Isaacs, 2018 ; Munro & Derwing, 2006) car ils permettent de séparer de nombreuses paires minimales, et le phonème /ʃ/-on participe à la morphologie flexionnelle verbale. Les deux phonèmes ont des degrés de complexité assez similaires, mais le phonème /ʒ/-j est plus difficile.

La difficulté est induite par le traitement cognitif du phonème par un apprenant dans un contexte d'apprentissage (Housen et al., 2019). Le traitement cognitif des phonèmes implique leur perception, leur maintien en mémoire de travail, et leur association dans des représentations multimodales. J'ai fait l'hypothèse que l'enseignement ciblé de la discrimination et de la prononciation en salle de classe pouvait améliorer la perception des phonèmes. De plus, le maintien de ces phonèmes en mémoire de travail a été spécifiquement travaillé. Dans une étude préliminaire (Dherbey-Chapuis & Berthele, 2020), nous avons constaté que les correspondances phonème-graphème n'étaient pas maîtrisées même après 300 heures d'enseignement de FLE. Dans la présente étude, inclure un enseignement de ces correspondances phonème-graphème a vraisemblablement pu faciliter l'intégration de

l'information phonologique au sein des représentations multimodales « orthographique-phonologique-sémantique ». Cette intégration s'est montrée essentielle pour l'apprentissage du vocabulaire lié. L'enseignement ciblé des représentations phonétiques et phonologiques a visé l'amélioration de cinq compétences.

L'enseignement des représentations phonétiques et phonologiques a travaillé les compétences de discrimination, de prononciation, de maintien en mémoire de travail, de maîtrise des correspondances graphème-phonème et de vocabulaire réceptif. Les performances des élèves dans les cinq compétences ont été suivies dans un cadre pré-test, post-test immédiat et test à 3 mois. Un matériel didactique dédié a été élaboré pour la présente étude.

Deux séquences d'enseignement équivalentes ont été construites, l'une explicite et l'autre implicite. À ma connaissance, la littérature n'a pas décrit d'enseignement implicite ciblé des représentations phonétiques et phonologiques, bien que l'acquisition incidentale du vocabulaire ait fait l'objet de nombreuses études. L'acquisition du vocabulaire doit obligatoirement inclure les représentations phonétiques et phonologiques même si l'association représentation phonologique-valeur sémantique semble l'association la plus difficile à acquérir (e.g., Hulstijn, 2013). Or, il est possible de mettre en évidence implicitement les caractéristiques phonologiques LE pour qu'elles soient remarquées par le système procédural de l'apprenant. Le système procédural modifierait alors en conséquence les routines et les procédures permettant le traitement cognitif des caractéristiques LE. Premièrement, j'ai fait l'hypothèse qu'une didactique implicite ciblée, c'est-à-dire orientée par la sémantique, mais construite pour mettre en évidence les caractéristiques de la LE sans que l'apprenant en soit conscient, pourrait permettre l'amélioration des performances de chacune des cinq compétences étudiées. Une telle didactique implicite ciblée devrait faciliter l'acquisition des représentations phonétiques, phonologiques et du vocabulaire associé. Deuxièmement, l'enseignement des représentations phonétiques et phonologiques est traditionnellement explicite et ciblé sur la forme.

De nombreuses études ont montré l'efficacité d'un enseignement ciblé explicite des représentations phonétiques et phonologiques, mais aussi la supériorité des effets d'un enseignement explicite en comparaison d'un enseignement implicite (e.g., Ellis, N., 2015 ; Norris & Ortega, 2000). Toutefois, les rares études menées en contexte scolaire ne montrent pas de résultat clair (e.g., Papachristou, 2011). J'ai donc construit deux séquences didactiques similaires au niveau du matériel utilisé, des compétences travaillées et du temps passé. Dans la séquence didactique explicite, les apprenants ont été conscients des objectifs et des représentations visées mais la valeur sémantique du vocabulaire n'a jamais été enseignée ni utilisée; dans la séquence didactique implicite ciblée, les apprenants ont travaillé selon une didactique communicative actionnelle et la mise en évidence des spécificités de la LE s'est faite sans qu'ils en aient conscience. Les deux séquences didactiques ont été enseignées par l'expérimentatrice.

Chaque classe a été aléatoirement divisée en deux groupes recevant chacun un type d'enseignement. L'enseignante régulière et l'expérimentatrice ont échangé les groupes au milieu de chaque cours. Les deux séquences didactiques diffèrent essentiellement par leur approche implicite ou explicite de l'enseignement des représentations phonétiques et phonologiques.

Dans la présente étude, la question de recherche cherche à établir quelle didactique est la plus efficace entre deux didactiques, l'une explicite et l'autre implicite, ciblées sur l'apprentissage des représentations phonétiques, phonologiques et du vocabulaire réceptif. Certains auteurs plaident pour des effets différents en fonction de la compétence observée ou de la difficulté de l'objet d'enseignement / apprentissage (Ellis, N., 2006b). Ma prédiction était que j'observerai une supériorité de l'enseignement explicite car les deux phonèmes observés sont deux phonèmes difficiles.

Cette prédiction a été en partie seulement confirmée par les résultats de cette étude. La didactique explicite a été significativement plus favorable au développement de la maîtrise des correspondances phonème-graphème, au progrès des élèves les plus faibles et à l'apprentissage des représentations phonétiques et phonologiques liées au phonème le plus difficile (i.e., le phonème le /ʒ/-j). Cependant, aucune différence significative entre les deux séquences didactiques n'a été observée dans l'évolution des compétences de discrimination, de prononciation, de maintien en mémoire de travail et de vocabulaire, pour la moyenne des deux phonèmes ou pour le phonème /ʃ/-on. Ces résultats sont intéressants pour l'enseignement des représentations phonétiques et phonologiques, du vocabulaire et des correspondances phonème-graphème en contexte scolaire.

L'absence de différence entre les deux séquences didactiques pour le phonème /ʃ/-on, alors qu'il est connu pour générer des difficultés importantes, indiquent qu'un enseignement implicite ciblé est un outil efficace pour faciliter l'apprentissage des représentations phonétiques et phonologiques qui lui sont associées. Un tel outil permettrait de pallier le manque de formation des enseignants et à leur inquiétude lorsqu'il s'agit d'enseigner ces représentations. Si les méthodes communicatives étaient un peu moins ignorantes de ces contraintes phonologiques et un peu plus structurées autour de ces apprentissages, l'apprentissage de ces représentations phonétiques et phonologiques indispensables au processus communicatif serait peut-être facilité pour la plupart des élèves.

Le progrès du vocabulaire réceptif est identique pour les deux séquences didactiques alors que la valeur sémantique n'a jamais été enseignée dans la séquence didactique explicite pour les représentations phonétiques et phonologiques. Le sens a pu être acquis pendant les périodes de travail avec l'enseignante. L'important n'est pas de savoir si réellement la valeur sémantique est acquise mais bien de souligner combien les représentations phonologiques sont importantes pour l'acquisition du vocabulaire. L'augmentation de leur niveau de maîtrise a permis de libérer suffisamment d'attention pour qu'une représentation multimodale (même incomplète) soit

établie. Il est important de repenser notre perception des représentations phonétiques et phonologiques et de considérer la charge cognitive importante que génère leur traitement cognitif quand elles sont peu ou pas maîtrisées. Leur enseignement devrait être réintroduit dans toutes les salles de classe au moins pour les deux phonèmes étudiés dans la présente étude.

Peu d'auteurs s'intéressent aux correspondances phonème-graphème alors qu'elles sont au cœur des interférences cross-linguistiques quand la L1 et la LE partagent le même système d'écriture comme dans la présente étude. Comme annoncé par Ellis (N., 2006b), seul un enseignement explicite de ces correspondances a pu permettre de dépasser les difficultés cross-linguistiques. Cet enseignement devrait être intégré dans les manuels scolaires au moins pour les deux phonèmes de la présente étude.

Cette recherche a confirmé pour le contexte scolaire l'indispensable recours à un enseignement explicite pour surmonter les interférences cross-linguistiques (Ellis, N., 2006b), pour l'apprentissage des représentations les plus difficiles et pour les élèves les plus faibles. Cependant, l'enseignement implicite ciblé a été totalement oublié pour les représentations phonétiques et phonologiques et son potentiel semble vraiment prometteur. Pour les deux phonèmes de la présente étude, il a été possible de faire remarquer à l'apprenant, sans qu'il en ait conscience, les particularités phonétiques et phonologiques de la LE. Des études confirmant ce potentiel seraient bienvenues.

Cette étude présente quatre biais et limites qui devront être explorés dans des études complémentaires.

Premièrement, concernant l'observation d'une influence de l'enseignement explicite sur les procédures et routines de traitement cognitif, cette étude n'a pas inclus de tâche productive qui aurait permis d'observer ou non cette influence dans les deux modes réceptif et productif.

Deuxièmement, les deux séquences didactiques, bien que conçues pour être le plus similaire possible, présentent néanmoins des différences au niveau des compétences travaillées. Il ne peut pas être exclu que certaines différences de performances soient induites par un travail légèrement différent de la compétence visée. Cependant, la supériorité de l'enseignement explicite pour les élèves les plus faibles ou le phonème le plus difficile ont confirmé un effet différent des deux séquences didactiques.

Troisièmement, la prise en compte de la moyenne des deux phonèmes a indiqué que les séquences didactiques induisent un progrès similaire. En effet, pour le phonème /ʒ/-on la séquence implicite induit un progrès supérieur, et inversement pour le phonème /z/-j la séquence explicite induit un progrès supérieur. Le choix des phonèmes a donc influencé les résultats et les résultats de cette étude devraient être confirmés avec d'autres phonèmes du FLE, pour d'autres interférences cross-linguistiques.

Quatrièmement, l'expérimentatrice a enseigné les deux séquences didactiques. Un même enseignement mené par l'enseignante régulière de la classe aurait-il eu le même impact ?

Plusieurs points ont besoin d'être vérifiés et plus d'études en contexte scolaire sont nécessaires. En particulier, l'exploration des modes réceptif /productif d'une même compétence ou l'enseignement d'autres phonèmes du FLE devraient être explorés dans de nouvelles études. Cependant, cette thèse a mis en évidence l'intérêt de l'enseignement implicite ciblé pour les représentations phonétiques et phonologiques, l'importance des représentations phonétiques et phonologiques pour l'apprentissage du vocabulaire réceptif et la nécessité d'introduire un enseignement explicite des correspondances phonème-graphème lorsque la LE et la L1 partagent le même système d'écriture. Elle a également permis d'ouvrir les possibilités de modélisation de l'apprentissage des représentations phonétiques et phonologiques en LE en intégrant les différents types de représentations, les modes d'enseignement explicite et implicite, la charge cognitive et la représentation phonémique dans un même modèle cohérent basé sur l'acquisition d'exemplaires.

Annexes

Toutes les annexes sont accessibles avec le lien ci-dessous.

https://osf.io/c6es3/?view_only=7d7033049cab49d8a5d759960d323345

Abréviations

C séquence didactique implicite ciblée communicative actionnelle

CAF Complexity Accuracy Fluency theory

CPG correspondances phonème-graphème

CV séquence composée d'une consonne suivie par une voyelle

e.g. par exemple

FLE français langue étrangère

E séquence didactique explicite

i.e. c'est-à-dire

ISI intervalle de silence entre deux éléments d'un stimulus

L1 langue première

LE langue étrangère

p probabilité

p. page

T0 pré-test

T1 post-test immédiat

T2 post-test à 3 mois

vs. Versus

Séries de la tâche de mesure du maintien en mémoire de travail

PIT inclut des phonèmes communs L1-LE et des phonèmes LE inconnus travaillés

PInT inclut des phonèmes communs L1-LE et des phonèmes LE inconnus NON travaillés

PC inclut des phonèmes communs L1-LE

SA inclut des phonèmes communs L1-LE et a été réalisé en condition de suppression articulaire

Correspondances des niveaux de scolarité entre les systèmes scolaires suisses et français

7H = CM1 ; 8H = CM2 ; 9H = 6eme

Bibliographie

- Abdallah, F. (2010). *The role of phonological memory in L2 acquisition in adults at different proficiency levels*. Thèse de Doctorat. <https://corpus.ulaval.ca/jspui/handle/20.500.11794/28443>
- Abel, C. (2018). L'enseignement et l'évaluation de la prononciation en classe de FLE et l'approche par compétences / l'approche actionnelle—Opposition ou synergie ? *Travaux de didactique du français langue étrangère*, 72(51).
- Abreu, P. de, & Gathercole, S.E. (2012). Executive and phonological processes in second-language acquisition. *Journal of Educational Psychology*, 104, 974-986.
- Adlof, S. M., & Patten, H. (2017). Nonword Repetition and Vocabulary Knowledge as Predictors of Children's Phonological and Semantic Word Learning. *Journal of Speech, Language, and Hearing Research*, 60(3), 682-693. https://doi.org/10.1044/2016_JSLHR-L-15-0441
- Amelot, A. (2004). *Etude aérodynamique, fibroscopique, acoustique et perceptive des voyelles nasales du français*. Thèse de Doctorat, Université de la Sorbonne nouvelle - Paris III. <https://tel.archives-ouvertes.fr/tel-00315998>
- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, 111(4), 1036-1060. <https://doi.org/10.1037/0033-295X.111.4.1036>
- Antoniou, M., & Wong, P. C. M. (2015). Poor phonetic perceivers are affected by cognitive load when resolving talker variability. *The Journal of the Acoustical Society of America*, 138(2), 571-574. <https://doi.org/10.1121/1.4923362>
- Atkinson, A. L., Allen, R. J., Baddeley, A., Hitch, G., & Waterman, A. (2020). Can valuable information be prioritized in verbal working memory? *Journal of Experimental Psychology. Learning, Memory, and Cognition*. <https://doi.org/10.1037/xlm0000979>
- Atkinson, A. L., Waterman, A. H., & Allen, R. J. (2019). Can Children Prioritize More Valuable Information in Working Memory ? An Exploration Into the Effects of Motivation and Memory Load. *Developmental psychology*, 55(5), 967-980.
- Atmaca, E. (2017). Un aperçu général des méthodes de FLE : De la méthode audio-Orale jusqu'à l'approche communicative. *International Journal of Languages' Education and Teaching*, 5(4), 367-377. <https://doi.org/10.18298/ijlet.2451>
- Baars, B. J. (1997). Some essential differences between consciousness and attention, perception, and working memory. *Consciousness and Cognition*, 6(2-3), 363-371. <https://doi.org/10.1006/ccog.1997.0307>
- Baayen, R. H. (2008). *Analyzing Linguistic Data : A Practical Introduction to Statistics using R*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511801686>
- Baddeley, A. (2000). The episodic buffer : A new component of working memory? *Trends in Cognitive Sciences*, 417-423.
- Baddeley, A. (2003). Working memory and language : An overview. *Journal of communication disorders*, 36, 189-208. [https://doi.org/10.1016/S0021-9924\(03\)00019-4](https://doi.org/10.1016/S0021-9924(03)00019-4)

- Baddeley, A. (2012). Working Memory: Theories, Models, and Controversies. *Annual Review of Psychology*, 63(1), 1-29. <https://doi.org/10.1146/annurev-psych-120710-100422>
- Baddeley, A., & Hitch, G. (1974). Working Memory. In G.H. Bower (Ed.), *Psychology of Learning and Motivation*, Vol. 8, (pp. 47-89). Academic Press. [https://doi.org/10.1016/S0079-7421\(08\)60452-1](https://doi.org/10.1016/S0079-7421(08)60452-1)
- Baddeley, A., Papagno, C., & Vallar, G. (1988). When long-term learning depends on short-term storage. *Journal of Memory and Language*, 27(5), 586-595. [https://doi.org/10.1016/0749-596X\(88\)90028-9](https://doi.org/10.1016/0749-596X(88)90028-9)
- Baddeley, A., & Hitch, G. (1993). The recency effect: Implicit learning with explicit retrieval? *Memory & Cognition*, 21(2), 146-155. <https://doi.org/10.3758/BF03202726>
- Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology. General*, 133(1), 83-100. <https://doi.org/10.1037/0096-3445.133.1.83>
- Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 33(3), 570-585. <https://doi.org/10.1037/0278-7393.33.3.570>
- Barrouillet, P., & Camos, V. (2012). As Time Goes By: Temporal Constraints in Working Memory. *Current Directions in Psychological Science*, 21(6), 413-419. <https://doi.org/10.1177/0963721412459513>
- Barrouillet, P., & Camos, V. (2015). *Working memory: Loss and reconstruction*. Hove, UK : Psychology Press.
- Barth, B.-M. (2013). *L'apprentissage de l'abstraction*. Paris: Retz.
- Bartolotti, J., Bradley, K., Hernandez, A. E., & Marian, V. (2017). Neural signatures of second language learning and control. *Neuropsychologia*, 98, 130-138. <https://doi.org/10.1016/j.neuropsychologia.2016.04.007>
- Bassetti, B. (2017). Orthography affects second language speech: Double letters and geminate production in English. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 43(11), 1835-1842. <https://doi.org/10.1037/xlm0000417>
- Bassetti, B., Escudero, P., & Hayes-Harb, R. (2015). Second language phonology at the interface between acoustic and orthographic input. *Applied Psycholinguistics*, 36(1), 1-6.
- Bassetti, B., Mairano, P., Masterson, J., & Cerni, T. (2020). Effects of Orthographic Forms on Second Language Speech Production and Phonological Awareness, With Consideration of Speaker-Level Predictors. *Language Learning*, 70(4), 1218-1256. <https://doi.org/10.1111/lang.12423>
- Bédard, P., Audet, A.-M., Drouin, P., Roy, J.-P., Rivard, J., & Tremblay, P. (2017). SyllabO+ : A new tool to study sublexical phenomena in spoken Quebec French. *Behavior Research Methods*, 49(5), 1852-1863. <https://doi.org/10.3758/s13428-016-0829-7>
- Beddor, P., Krakow, R., & Goldstein, L. (1986). Perceptual constraints and phonological change: A study of nasal vowel height. *Phonology*, 3, 197-217. <https://doi.org/10.1017/S0952675700000646>

- Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception : Commonalities and complementarities. Dans O.S. Bohn & M. J. Munro (Eds.), *Language Learning & Language Teaching*, Vol. 17 (pp. 13-34). Amsterdam : John Benjamins Publishing Company. <https://doi.org/10.1075/llt.17.07bes>
- Bhide, A., Ortega-Llebaria, M., Fraundorf, S., & Perfetti, C. (2020). The contribution of orthographic input, phonological skills, and rise time discrimination to the learning of non-native phonemic contrasts. *Applied Psycholinguistics*, *41*(3), 481-516. <https://doi.org/10.1017/S0142716419000511>
- Billières, M. (2005). Codage phonologique et boucle articulatoire en mémoire de travail. Un support pour la facilitation de l'accès à l'oral et à la lecture pour des publics débutants en français langue étrangère. *Corela. Cognition, représentation, langage, HS-1*, Article HS-1. <https://doi.org/10.4000/corela.1110>
- Bion, R., Escudero, P., Rauber, A., & Baptista, B. (2006). Category formation and the role of spectral quality in the perception and production of English front vowels. *Interspeech 2006*, 1270.
- Boersma, P., & Weenink, D. (2018). *Praat: Doing phonetics by computer [Computer program]* (Version 6.0.37) [Computer software]. <http://www.praat.org/>
- Bolker, B. (2015). *Bbolker/mixedmodels-misc*. @bates_fitting_2015. <https://github.com/bbolker/mixedmodels-misc>
- Bolker, B., Brooks, M., Clark, C., Geange, S., Poulsen, J., Stevens, M.H., & White, J.-S. (2009). Generalized linear mixed models: A practical guide for ecology and evolution. *Trends in Ecology & Evolution*, *24*(3), 127-135. <https://doi.org/10.1016/j.tree.2008.10.008>
- Bonneau, A. (2015). Realizations of French voiced fricatives by German learners as a function of speaker level and prosodic boundaries. *18th International Congress of Phonetic Sciences, ICPHS 2015*, 5. <https://hal.inria.fr/hal-01186062>
- Bosse, M.-L., & Pacton, S. (2006). Comment l'enfant produit-il l'orthographe des mots ? Dans P. Gentaz & E. Dessus (Eds.), *Apprendre et enseigner à l'école* (pp. 43-58). Paris: Dunod. <https://hal.archives-ouvertes.fr/hal-00826018>
- Bouchard, K., Mesgarani, N., Johnson, K., & Chang, E. F. (2013). Functional organization of human sensorimotor cortex for speech articulation. *Nature*, *495*(7441), 327.
- Bowers, J. S., Kazanina, N., & Andermane, N. (2016). Spoken word identification involves accessing position invariant phoneme representations. *Journal of Memory and Language*, *87*, 71-83. <https://doi.org/10.1016/j.jml.2015.11.002>
- Brown, C. (2000). The Interrelation between Speech Perception and Phonological Acquisition from Infant to Adult. *Second Language Acquisition and Linguistic Theory*, *1*, 4-64.
- Buchsbaum, B., Baldo, J., Okada, K., Berman, K., Dronkers, N., D'Esposito, M., & Hickok, G. (2011). Conduction aphasia, sensory-motor integration, and phonological short-term memory—An aggregate analysis of lesion and fMRI data. *Brain and Language*, *119*(3), 119-128. <https://doi.org/10.1016/j.bandl.2010.12.001>
- Bunting, M., & Engle, R. (2015). Foreword. Dans Z. Wen, M. Borges Mota, & A. McNeill, (Eds.), *Working Memory in Second Language Acquisition and Processing* (pp. XVii). Bristol, UK: Multilingual Matters.

- Bürki, A. (2018). Variation in the speech signal as a window into the cognitive architecture of language production. *Psychonomic Bulletin & Review*, 25(6), 1973-2004. <https://doi.org/10.3758/s13423-017-1423-4>
- Bürki, A., Ernestus, M., & Frauenfelder, U. (2010). Is there only one “fenêtre” in the production lexicon? On-line evidence on the nature of phonological representations of pronunciation variants for French schwa words. *Journal of Memory and Language*, 62, 421-437. <https://doi.org/10.1016/j.jml.2010.01.002>
- Burnham, K., & Anderson, D. (2002). *Model selection and multimodel inference: A practical information—Theoretic approach*. NY: Springer. https://doi.org/10.1007/978-0-387-22456-5_8
- Bybee, J. (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change*, 14(3), 261-290. <https://doi.org/10.1017/S0954394502143018>
- Camos, V., Lagner, P., & Barrouillet, P. (2009). Two maintenance mechanisms of verbal information in working memory. *Journal of Memory and Language*, 61(3), 457-469. <https://doi.org/10.1016/j.jml.2009.06.002>
- Camos, V., Johnson, M., Loaiza, V., Portrat, S., Souza, A., & Vergauwe, E. (2018). What is attentional refreshing in working memory? *Annals of the New York Academy of Sciences*, 1424(1), 19-32. <https://doi.org/10.1111/nyas.13616>
- Camos, V., Mora, G., & Barrouillet, P. (2013). Phonological similarity effect in complex span task. *Quarterly Journal of Experimental Psychology (2006)*, 66(10), 1927-1950. <https://doi.org/10.1080/17470218.2013.768275>
- Carroll, J.M., & Snowling, M.J. (2004). Language and phonological skills in children at high risk of reading difficulties. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 45(3), 631-640. <https://doi.org/10.1111/j.1469-7610.2004.00252.x>
- Case, R., Kurland, D.M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*, 33(3), 386-404. [https://doi.org/10.1016/0022-0965\(82\)90054-6](https://doi.org/10.1016/0022-0965(82)90054-6)
- Catach, N. (1995). La variation graphique et les rectifications de l’orthographe française. *Langue française*, 108(1). https://www.persee.fr/issue/lfr_0023-8368_1995_num_108_1
- Celce-Murcia, M., Dornyei, Z., & Thurrell, S. (1995). Communicative Competence : A Pedagogically Motivated Model with Content Specifications. *Issues in Applied Linguistics*, 6(2). <https://escholarship.org/uc/item/2928w4zj>
- Chen, Y. (2021). Comparing incidental vocabulary learning from reading-only and reading-while-listening. *System*, 97, 102442. <https://doi.org/10.1016/j.system.2020.102442>
- Chevrot, J.-P., Dugua, C., & Fayol, M. (2009). Liaison acquisition, word segmentation and construction in French : A usage-based account. *Journal of Child Language*, 36(3), 557-596. <https://doi.org/10.1017/S0305000908009124>
- Clements, G.N. (2003). Feature Economy in Sound Systems. *Phonology*, 20(3), 287-333.
- Clements, G.N. (2006). Feature Organization. Dans *Encyclopedia of Language & Linguistics* (pp. 433-440). Amsterdam: Elsevier. <https://doi.org/10.1016/B0-08-044854-2/00055-9>

- Clements, G.N., & Hallé, P.A. (2010). "Phonetic bases of distinctive features": Introduction. *Journal of Phonetics*, 38(1), 3-9. <https://doi.org/10.1016/j.wocn.2010.01.004>
- Clumeck, H. (1976). Patterns of soft palate movements in six languages. *Journal of Phonetics*, 337-351.
- Coady, J., & Evans, J. (2008). Uses and interpretations of non-word repetition tasks in children with and without specific language impairments (SLI). *International journal of language & communication disorders*, 43(1), 1-40. <https://doi.org/10.1080/13682820601116485>
- Cook, S.V., & Gor, K. (2015). Lexical access in L2 : Representational deficit or processing constraint? *The Mental Lexicon*, 10(2), 247-270. <https://doi.org/10.1075/ml.10.2.04coo>
- Cook, S.V., Pandža, N., Lancaster, A., & Gor, K. (2016). Fuzzy Nonnative Phonolexical Representations Lead to Fuzzy Form-to-Meaning Mappings. *Frontiers in Psychology*, 7, 1345. <https://doi.org/10.3389/fpsyg.2016.01345>
- Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in Brain Research*, 169, 323-338. [https://doi.org/10.1016/S0079-6123\(07\)00020-9](https://doi.org/10.1016/S0079-6123(07)00020-9)
- Cowan, N. (2015). Second-language use, theories of working memory, and the Vennian mind. Dans Z. Wen, M.B. Mota, A. McNeill (Eds.), *Working memory in second language acquisition and processing*. Bristol, UK : Multilingual Matters.
- Curtin, S., Byers-Heinlein, K., & Werker, J.F. (2011). Bilingual beginnings as a lens for theory development: PRIMIR in focus. *Journal of Phonetics*, 39(4), 492-504. <https://doi.org/10.1016/j.wocn.2010.12.002>
- Curtin, S., Hufnagle, D., Mulak, K., & Escudero, P., (2017). Speech Perception: Development. Dans Module in *Neuroscience and Biobehavioral psychology*, (pp. 1-7). Amsterdam: Elsevier Science. <https://doi.org/10.1016/B978-0-12-809324-5.01904-0>
- Curtin, S., & Werker, J.F. (2007). Perceptual foundations of phonological development. Dans M. Gareth Gaskell, G. Altmann, P. Bloom, A. Caramazza, & P. Levelt (Eds.), *The Oxford Handbook of Psycholinguistics*, (pp. 577-600). Oxford University Press.
- Cutler, A. (2015). Representation of second language phonology. *Applied Psycholinguistics*, 36(1), 115-128. <https://doi.org/10.1017/S0142716414000459>
- Cutler, A., Weber, A., & Otake, T. (2006). Asymmetric mapping from phonetic to lexical representations in second-language listening. *Journal of Phonetics*, 34(2), 269-284. <https://doi.org/10.1016/j.wocn.2005.06.002>
- Darcy, I., Daidone, D., & Kojima, C., (2013). Asymmetric lexical access and fuzzy lexical representations in second language learners. *The Mental Lexicon*, 372-420.
- Darcy, I., Dekydtspotter, L., Sprouse, R., Glover, J., Kaden, C., McGuire, M., & Scott, J. (2012). Direct mapping of acoustics to phonology: On the lexical encoding of front rounded vowels in L1 English– L2 French acquisition. *Second Language Research*, 28(1), 5-40. <https://doi.org/10.1177/0267658311423455>
- Darcy, I., & Kügler, F. (2007). Phonological context effects for voicing and devoicing in French. *Proceedings of the 16th International Congress of Phonetic Sciences*, 1257-1260.

https://www.academia.edu/17791433/Phonological_context_effects_for_voicing_and_devoicing_in_French

- DeCarlo, L. T. (1998). Signal Detection Theory and Generalized Linear Models. *Psychological Methods*, 3(2), 186-205.
- DeCarlo, L. T. (2010). On the statistical and theoretical basis of signal detection theory and extensions: Unequal variance, random coefficient, and mixture models. *Journal of Mathematical Psychology*, 54(3), 304-313. <https://doi.org/10.1016/j.jmp.2010.01.001>
- DeCarlo, L. T. (2012). On a signal detection approach to m-alternative forced choice with bias, with maximum likelihood and Bayesian approaches to estimation. *Journal of Mathematical Psychology*, 56(3), 196-207. <https://doi.org/10.1016/j.jmp.2012.02.004>
- Dehaene-Lambertz, G., Dupoux, E., & Gout, A. (2000). Electrophysiological correlates of phonological processing: A cross-linguistic study. *Journal of Cognitive Neuroscience*, 12(4), 635-647. <https://doi.org/10.1162/089892900562390>
- Dekeyser, R. (2003). Explicit and Implicit Learning. Dans C. Doughty & M.H. Longs (Eds.), *The Handbook of Second language Acquisition* (pp. 313-348). Oxford: Blackwell. <http://dx.doi.org/10.1002/9780470756492.ch11>
- DeKeyser, R. (2016). Of moving tardets and chameleons ; Why the Concept of Difficulty Is So Hard to Pin Down. *Studies in Second Language Acquisition*, 38(2), 353-363. <https://doi.org/10.1017/S0272263116000024>
- DeKeyser, R. (2005). What Makes Learning Second-Language Grammar Difficult? A Review of Issues. *Language Learning*, 55(S1), 1-25. <https://doi.org/10.1111/j.0023-8333.2005.00294.x>
- Dell, F. (2004). *Les règles et les sons. Introduction à la phonologie générative*. Paris: Hermann.
- Delvaux, V. (2003). *Contrôle et connaissance phonétique : Les voyelles nasales du français*. Thèse de Doctorat, Université libre de Bruxelles.
- Delvaux, V. (2009). Perception du contraste de nasalité vocalique en français. *Journal of French Language Studies*, 19(1), 25-59. <https://doi.org/10.1017/S0959269508003566>
- Demont, E., & Botzung, A. (2003). Contribution de la conscience phonologique et de la mémoire de travail aux difficultés en lecture : Étude auprès d'enfants dyslexiques et apprentis lecteurs. *L'année psychologique*, 103(3), 377-409. <https://doi.org/10.3406/psy.2003.29642>
- Denhovska, N., Serratrice, L., & Payne, J. (2016). Acquisition of Second Language Grammar Under Incidental Learning Conditions: The Role of Frequency and Working Memory: Frequency and Implicit Learning of L2 Grammar. *Language Learning*, 66(1), 159-190. <https://doi.org/10.1111/lang.12142>
- Derwing, T., & Munro, M. (1997). Accent, intelligibility and comprehensibility: Evidence from Four L1s. *Studies in Second Language Acquisition*, 19(1), 1-16. <https://doi.org/10.1017/S0272263197001010>
- Derwing, T., & Munro, M. (2005). Second Language Accent and Pronunciation Teaching: A Research-Based Approach. *TESOL Quarterly*, 39(3), 379-397. <https://doi.org/10.2307/3588486>

- Derwing, T., & Munro, M. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching*, 42(4), 476-490.
<https://doi.org/10.1017/S026144480800551X>
- Detey, S. (2016). Enseignement de la prononciation et correction phonétique : Principes essentiels. Dans S. Detey, I. Racine, Y. Kawaguchi & J. Eychenne (Eds.), *La prononciation du français dans le monde. Du natif à l'apprenant*, Ch. 36 (pp. 226-235). Paris : CLE International.
- Detey, S., Fontan, L., & Pellegrini, T. (2016). Traitement de la prononciation en langue étrangère : Approches didactiques, méthodes automatiques et enjeux pour l'apprentissage. *Traitement Automatique des Langues*, 57(3), 15-39.
- Detey, S., & Racine, I. (2015). The InterPhonology of Contemporary French (IPFC): An international corpus-based L2 phonology research programme. *Book of extended Abstracts, Phonetic Learner Corpora*, 1-4.
- Dherbey-Chapuis, N. (2016). *Influence des voyelles nasales sur l'acquisition lexicale dans l'apprentissage du français L2 par des enfants de 9 à 11 ans*. Thèse de Master. Toulouse Le Mirail - Octogone Lordat Laboratory.
- Dherbey-Chapuis, N., & Berthele, R. (2020). L'apprentissage des correspondances phonème - graphème en français langue étrangère. *Actes Du 7e Congrès Mondial de Linguistique Française*, 18. <https://doi.org/10.1051/shsconf/20207809011>
- Di Liberto, G., Wong, D., Melnik, G., & de Cheveigné, A. (2019). Low-frequency cortical responses to natural speech reflect probabilistic phonotactics. *NeuroImage*, 196, 237-247. <https://doi.org/10.1016/j.neuroimage.2019.04.037>
- Dupoux, E. (2010). *Phonological Deafnesses : A Summary*.
<http://www.lscp.net/persons/dupoux/#listpub>
- Dupoux, E., Pallier, C., Sebastián-Gallés, N., & Mehler, J. (1997). A Destressing “Deafness” in French? *Journal of Memory and Language*, 36(3), 406-421.
<https://doi.org/10.1006/jmla.1996.2500>
- Dupoux, E., Pallier, C., Kakehi, K., & Mehler, J. (2001). New evidence for prelexical phonological processing in word recognition. *Language and Cognitive Processes*, 16(5-6), 491-505. <https://doi.org/10.1080/01690960143000191>
- Dupoux, E., Sebastián-Gallés, N., Navarrete, E., & Peperkamp, S. (2008). Persistent stress ‘deafness’: The case of French learners of Spanish. *Cognition*, 106(2), 682-706.
<https://doi.org/10.1016/j.cognition.2007.04.001>
- Elgort, I., Brysbaert, M., Stevens, M., & Assche, E. (2018). Contextual word learning during reading in a second language: An eye-movement study. *Studies in Second Language Acquisition*, 40(2), 341-366. <https://doi.org/10.1017/S0272263117000109>
- Ellis, N.C. (2005). At the interface : Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, 27(02).
<https://doi.org/10.1017/S027226310505014X>
- Ellis, N.C. (2006a). Selective Attention and Transfer Phenomena in L2 Acquisition: Contingency, Cue Competition, Salience, Interference, Overshadowing, Blocking, and Perceptual Learning. *Applied Linguistics*, 27(2), 164-194.
<https://doi.org/10.1093/applin/aml015>

- Ellis, N.C. (2006b). Cognitive perspectives on SLA : The associative-cognitive CREED. *AILA Review*, 19, 100-121. <https://doi.org/10.1075/aila.19.08ell>
- Ellis, N.C. (2015). Implicit and Explicit Language Learning : Their dynamic interface and complexity. Dans P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 3-23). Amsterdam: John Benjamins.
- Ellis, N.C., & Beaton, A. (1993). Psycholinguistic Determinants of Foreign Language Vocabulary Learning. *Language Learning*, 43(4), 559-617.
- Ellis, N.C., & Wulff, S. (2012). Second Language Acquisition. Dans E. Dabrowska and D. Divjak (Eds.), *Handbook of Cognitive Linguistics*, Introduction. Berlin: De Gruyter Mouton.
- Ellis, R. (2009). Implicit and Explicit Learning, Knowledge and Instruction. Dans H. Reinders, R. Erlam, J. Philp, S. Loewen & C. Elder (Eds.), *Implicit and Explicit Knowledge in Second Language Learning, Testing and Teaching*, (pp. 3-26). Bristol, UK : Multilingual Matters. <https://doi.org/10.21832/9781847691767-003>
- Ellis, R. (2015). *Understanding Second Language Acquisition* (Second Edition). Oxford University Press.
- Engle, R., Kane, M., & Tuholski, S. (1999). Individual Differences in Working Memory Capacity and What They Tell Us About Controlled Attention, General Fluid Intelligence, and Functions of the Prefrontal Cortex. Dans A. Miyake & S. Priti (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 102-134). Cambridge University Press. <https://doi.org/10.1017/CBO9781139174909.007>
- Ernestus, M., & Mak, W. (2004). Distinctive phonological features differ in relevance for both spoken and written word recognition. *Brain and Language*, 90, 378-392. [https://doi.org/10.1016/S0093-934X\(03\)00449-8](https://doi.org/10.1016/S0093-934X(03)00449-8)
- Escudero, P. (2005). Linguistic Perception and Second Language Acquisition: Explaining the attainment of optimal phonological categorization. Thèse de Doctorat. <http://dspace.library.uu.nl/handle/1874/7349>
- Escudero, P. (2009). The linguistic perception of SIMILAR L2 sounds. Dans P. Boersma & S. Hamann (Eds.), *Phonology in perception*, (pp. 151-190). Berlin: Mouton de Gruyter.
- Escudero, P., Hayes-Harb, R., & Mitterer, H. (2008). Novel second-language words and asymmetric lexical access. *Journal of Phonetics*, 36(2), 345-360. <https://doi.org/10.1016/j.wocn.2007.11.002>
- Escudero, P., Simon, E., & Mulak, K. (2014). Learning words in a new language: Orthography doesn't always help. *Bilingualism: Language and Cognition*, 17(2), 384-395. <https://doi.org/10.1017/S1366728913000436>
- Flege, J.E. (1999). The relation between production and perception. *ICPhS*, 2, 1273-1276.
- Flege, J.E. (1995). Second language speech learning: Theory, Findings, and Problems. Dans W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research*, (pp. 233-272). York: Timonium.

- Flege, J.E. (2003). Assessing constraints on second-language segmental production and perception. *Phonetics and phonology in language comprehension and production: Differences and similarities*, 6, 319-355.
- Flege, J.E., Schirru, C., & McKay, I. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication*, 40(4), 467-491. [https://doi.org/10.1016/S0167-6393\(02\)00128-0](https://doi.org/10.1016/S0167-6393(02)00128-0)
- Flege, J.E. (2016, juin). The role of phonetic category formation in second language speech acquisition. *8th International Conference on second language speech*. New Sounds, Aarhus University, Denmark.
- Flege, J.E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, 25(4), 437-470. <https://doi.org/10.1006/jpho.1997.0052>
- Fleischer, J. & Schmid, S. (2006). Zurich German. *Journal of the International Phonetic Association*, 36(2), 243-255. <https://doi.org/10.1017/S0025100306002441>
- Fourtassi, A., Schatz, T., Varadarajan, B., & Dupoux, E. (2014). Exploring the Relative Role of Bottom-up and Top-down Information in Phoneme Learning. *Proceedings of the 52nd Annual meeting of the association for computational linguistics (Short papers)*, 1-6. <http://www.anthology.aclweb.org/P/P14/P14-2001.pdf>
- Foy, J. G., & Mann, V. (2001). Does strength of phonological representations predict phonological awareness in preschool children? *Applied Psycholinguistics*, 22(3), 301-325. <https://doi.org/10.1017/S0142716401003022>
- Friesen, D., Jared, D., & Haigh, C. (2014). Phonological processing dynamics in bilingual word naming. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 68, 179-193. <https://doi.org/10.1037/cep0000026>
- Fuchs, S., Brunner, J., & Busler, A. (2007). Temporal and spatial aspects concerning the realizations of the voicing contrast in German alveolar and postalveolar fricatives. *Advances in Speech Language Pathology*, 9(1), 90-100. <https://doi.org/10.1080/14417040601094315>
- Gaonac'h, D. (2000). La lecture en langue étrangère : Un tour d'horizon d'une problématique de psychologie cognitive. *Acquisition et interaction en langue étrangère*. <https://doi.org/10.4000/aile.970>
- Gaonac'h, D., & Larigauderie, P. (2000). *Mémoire et fonctionnement cognitif: La mémoire de travail*. Paris: Armand Colin.
- Garlock, V., Walley, A., & Metsala, J. (2001). Age-of-Acquisition, Word Frequency, and Neighborhood Density Effects on Spoken Word Recognition by Children and Adults. *Journal of Memory and Language*, 45(3), 468-492. <https://doi.org/10.1006/jmla.2000.2784>
- Gatbonton, E., Trofimovich, P., & Segalowitz, N. (2011). Ethnic Group Affiliation and Patterns of Development of a Phonological Variable. *The Modern Language Journal*, 95(2), 188-204. <https://doi.org/10.1111/j.1540-4781.2011.01177.x>
- Gathercole, S.E., Frankish, C. R., Pickering, S. J., & Peaker, S. (1999). Phonotactic influences on short-term memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 25(1), 84-95. <https://doi.org/10.1037//0278-7393.25.1.84>

- Gathercole, S.E., Willis, C., Baddeley, A., & Emslie, H. (1994). The Children's Test of Nonword Repetition : A test of phonological working memory. *Memory (Hove, England)*, 2(2), 103-127. <https://doi.org/10.1080/09658219408258940>
- Gathercole, S.E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, 27, 513-543.
- Gathercole, S.E., Dunning, D., Holmes, J., & Norris, D. (2019). Working memory training involves learning new skills. *Journal of Memory and Language*, 105, 19-42. <https://doi.org/10.1016/j.jml.2018.10.003>
- Gathercole, S.E., Willis, C., Emslie, H., & Baddeley, A. (1992). Phonological memory and vocabulary development during the early school years : A longitudinal study. *Developmental Psychology*, 28(5), 887-898. <https://doi.org/10.1037/0012-1649.28.5.887>
- Giannakopoulou, A., Brown, H., Clayards, M., & Wonnacott, E. (2017). High or Low? Comparing High-and Low-Variability Phonetic Training in Adult and Child Second Language Learners. High or Low? Comparing High-and Low-Variability Phonetic Training in Adult and Child Second Language Learners. *PeerJ*, 5, 3209. <https://doi.org/10.7717/peerj.3209>
- Gineste, M., & Le Ny, J.-F. (2002). *Psychologie cognitive du langage : De la reconnaissance à la compréhension*. Dunod.
- Glanz, O., Derix, J., Kaur, R., Schulze-Bonhage, A., Auer, P., Aertsen, A., & Ball, T. (2018). Real-life speech production and perception have a shared premotor-cortical substrate. *Scientific Reports*, 8(1), 8898. <https://doi.org/10.1038/s41598-018-26801-x>
- Goad, H., & White, L. (2008). Prosodic structure and the representation of L2 functional morphology : A nativist approach. *Lingua*, 118(4), 577-594. <https://doi.org/10.1016/j.lingua.2007.01.008>
- Granena, G., & Yilmaz, Y. (2019). Corrective Feedback and the Role of Implicit Sequence-Learning Ability in L2 Online Performance. *Language Learning*, 69(S1), 127-156.
- Green, D.M., & Swets, J.A. (1966). *Signal detection theory and psychophysics*. Oxford: John Wiley.
- Gunnarsson-Largy, C., Dherbey, N., & Largy, P. (2019). How do L2 learners and L1 writers differ in their reliance on working memory during the formulation subprocess? *Reading and Writing*, 32(8), 2083-2110. <https://doi.org/10.1007/s11145-019-09941-y>
- Gupta, P. (2005). Primacy and recency in nonword repetition. *Memory (Hove, England)*, 13(3-4), 318-324. <https://doi.org/10.1080/09658210344000350>
- Gupta, P., McWhinney, B., Feldman, H. M., & Sacco, K. (2003). Phonological memory and vocabulary learning in children with focal lesions. *Brain and language*, 87(2), 241-252.
- Haas, W., & Hove, I. (2009). Die Standardaussprache in der deutschsprachigen Schweiz. Dans Krach, E.-M. et al. (Eds.), *Deutsches Aussprachewörterbuch* (pp. 259-277). Berlin: De Gruyter. https://serval.unil.ch/notice/serval:BIB_E777852E6AE2
- Han, Z. (2013). Forty years later: Updating the Fossilization Hypothesis. *Language Teaching*, 46(2), 133-171. <https://doi.org/10.1017/S0261444812000511>

- Hancin-Bhatt, B. (1994). Segment transfer: A consequence of a dynamic system. *Second Language Research*, *10*(3), 241-269. <https://doi.org/10.1177/026765839401000304>
- Havard, W. N., Chevrot, J.-P., & Besacier, L. (2019). Word Recognition, Competition, and Activation in a Model of Visually Grounded Speech. *Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL) Hong Kong, Chine*, 339-348. <https://doi.org/10.18653/v1/K19-1032>
- Havy, M., Bouchon, C., & Nazzi, T. (2016). Phonetic processing when learning words: The case of bilingual infants. *International Journal of Behavioral Development*, *40*(1), 41-52. <https://doi.org/10.1177/0165025415570646>
- Hawkins, E., Astle, D.E., & Rastle, K. (2015). Semantic advantage for learning new phonological form representations. *Journal of Cognitive Neuroscience*, *27*(4), 775-786.
- Hayes-Harb, R., & Masuda, K. (2008). Development of the ability to lexically encode novel second language phonemic contrasts. *Second Language Research*, *24*(1), 5-33.
- Hayes-Harb, R., Nicol, J., & Barker, J. (2010). Learning the Phonological Forms of New Words: Effects of Orthographic and Auditory Input. *Language and Speech*, *53*(3), 367-381. <https://doi.org/10.1177/0023830910371460>
- Heald, S., & Nusbaum, H.C. (2014). Speech perception as an active cognitive process. *Frontiers in Systems Neuroscience*, *8*(35). <https://doi.org/10.3389/fnsys.2014.00035>
- Herman, A., Houde, J., Vinogradov, S., & Nagarajan, S. (2013). Parsing the Phonological Loop: Activation Timing in the Dorsal Speech Stream Determines Accuracy in Speech Reproduction. *The Journal of Neuroscience*, *33*(13), 5439-5453. <https://doi.org/10.1523/JNEUROSCI.1472-12.2013>
- Hickok, G. (2012). The cortical organization of speech processing: Feedback control and predictive coding the context of a dual-stream model. *Journal of Communication Disorders*, *45*(6), 393-402. <https://doi.org/10.1016/j.jcomdis.2012.06.004>
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews. Neuroscience*, *8*(5), 393-402. <https://doi.org/10.1038/nrn2113>
- Hickok, G., Rogalsky, C., Chen, R., Herskovits, E., Townsley, S., & Hillis, A. (2014). Partially overlapping sensorimotor networks underlie speech praxis and verbal short-term memory: Evidence from apraxia of speech following acute stroke. *Frontiers in Human Neuroscience*, *8*. <https://doi.org/10.3389/fnhum.2014.00649>
- Hisagi, M., Shafer, V., Strange, W., & Sussman, E. (2015). Neural measures of a Japanese consonant length discrimination by Japanese and American English listeners: Effects of attention. *Brain Research*, *1626*, 218-231. <https://doi.org/10.1016/j.brainres.2015.06.001>
- Hisagi, M., & Strange, W. (2011). Perception of Japanese Temporally-cued Contrasts by American English Listeners. *Language and Speech*, *54*(2), 241-264. <https://doi.org/10.1177/0023830910397499>
- Hofmann, M., Stenneken, P., Conrad, M., & Jacobs, A. (2007). Sublexical frequency measures for orthographic and phonological units in German. *Behavior Research Methods*, *39*(3), 620-629. <https://doi.org/10.3758/BF03193034>

- Holt, L., & Lotto, A. (2010). Speech perception as categorization. *Attention, Perception & Psychophysics*, 72(5), 1218-1227. <https://doi.org/10.3758/APP.72.5.1218>
- Housen, A., De Clercq, B., Kuiken, F., & Vedder, I. (2019). Multiple approaches to complexity in second language research. *Second Language Research*, 35(1), 3-21. <https://doi.org/10.1177/0267658318809765>
- Housen, A., & Kuiken, F. (2009). Complexity, Accuracy and Fluency in Second Language Acquisition. *Applied Linguistics*, 30(4), 461-473. <https://doi.org/10.1093/applin/amp048>
- Housen, A., Kuiken, F., & Vedder, I. (2012). *Dimensions of L2 performance and proficiency : Complexity, accuracy and fluency in SLA*. Amsterdam: John Benjamins Publishing Compagny).
- Housen, A., & Pierrard, M. (2005). Instructed Second Language Acquisition: Introduction. Dans A. Housen & P. Pierrard (Eds.), *Investigations in instructed Second Language Acquisition*. (pp. 1-27). Paris: Mouton de Gruyter.
- Housen, A., & Simoens, H. (2016). Introduction: Cognitive perspectives on difficulty and complexity in L2 acquisition. *Studies in Second Language Acquisition*, 38(2), 163-175. <https://doi.org/10.1017/S0272263116000176>
- Huensch, A., & Tremblay, A. (2015). Effects of perceptual phonetic training on the perception and production of second language syllable structure. *Journal of Phonetics*, 52, 105-120.
- Hulme, C., Maughan, S., & Brown, G. D. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language*, 30(6), 685-701. [https://doi.org/10.1016/0749-596X\(91\)90032-F](https://doi.org/10.1016/0749-596X(91)90032-F)
- Hulme, C., Roodenrys, S., Schweickert, R., Brown, G. D., Martin, M., & Stuart, G. (1997). Word-frequency effects on short-term memory tasks: Evidence for a reintegration process in immediate serial recall. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 23(5), 1217-1232.
- Hulme, C., Thomson, N., Muir, C., & Lawrence, A. (1984). Speech rate and the development of short-term memory span. *Journal of Experimental Child Psychology*, 38(2), 241-253. [https://doi.org/10.1016/0022-0965\(84\)90124-3](https://doi.org/10.1016/0022-0965(84)90124-3)
- Hulstijn, J.H. (2013). Incidental learning in second language acquisition. Dans C.A. Chapelle (Ed.), *The encyclopedia of applied linguistics*, (pp. 2632-2640). NY: Wiley-Blackwell. <https://doi.org/10.1002/9781405198431.wbeal0530>
- Hulstijn, J.H., & de Graaff, R. (1994). Under what conditions does explicit knowledge of a second language facilitate the acquisition of implicit knowledge? A research proposal. *AILA Review*, 11, 97-112.
- Isaacs, T. (2014). Assessing pronunciation. Dans A.J. Kunnan (Ed.), *The companion to language assessment*, (pp. 140-155). Hoboken: Wiley-Blackwell.
- Isaacs, T. (2018). Shifting Sands in Second Language Pronunciation Teaching and Assessment Research and Practice. *Language Assessment Quarterly*, 15(3), 273-293. <https://doi.org/10.1080/15434303.2018.1472264>

- Isaacs, T., & Thomson, R. (2013). Rater Experience, Rating Scale Length, and Judgments of L2 Pronunciation: Revisiting Research Conventions. *Language Assessment Quarterly*, *10*, 135-159. <https://doi.org/10.1080/15434303.2013.769545>
- Isaacs, T., & Trofimovich, P. (2010). Falling on Sensitive Ears? The Influence of Musical Ability on Extreme Raters' Judgments of L2 Pronunciation. *TESOL Quarterly*, *44*(2), 375-386. <https://doi.org/10.5054/tq.2010.222214>
- Jacqmin-Gadda, H., Sibillot, S., Proust, C., Molina, J.-M., & Thiébaud, R. (2007). Robustness of the linear mixed model to misspecified error distribution. *Computational Statistics & Data Analysis*, *51*(10), 5142-5154. <https://doi.org/10.1016/j.csda.2006.05.021>
- Jacquemot, C., Pallier, C., LeBihan, D., Dehaene, S., & Dupoux, E. (2003). Phonological Grammar Shapes the Auditory Cortex : A Functional Magnetic Resonance Imaging Study. *Journal of Neuroscience*, *23*(29), 9541-9546. <https://doi.org/10.1523/JNEUROSCI.23-29-09541.2003>
- Jarosz, G., Calamaro, S., & Zentz, J. (2017). Input frequency and the acquisition of syllable structure in Polish. *Language Acquisition*, *24*(4), 361-399. <https://doi.org/10.1080/10489223.2016.1179743>
- Järvikivi, J., Pyykkönen, P., & Niemi, J. (2009). Exploiting degrees of inflectional ambiguity: Stem form and the time course of morphological processing. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *35*(1), 221-237. <https://doi.org/10.1037/a0014355>
- Jatteau, A., Vasilescu, I., Lamel, L., & Adda-Decker, M. (2019). Final devoicing of fricatives in French: Studying variation in large-scale corpora with automatic alignment. *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia* (pp.295-299).
- Jusczyk, P.W. (1999). How infants begin to extract words from speech. *Trends in Cognitive Sciences*, *3*(9), 323-328.
- Kang, H.-S. (2009). *The relative efficacy of explicit and implicit feedback in the learning of a less-commonly-taught foreign language*. *47*(3-4), 303-324. <https://doi.org/10.1515/iral.2009.013>
- Kartushina, N., & Frauenfelder, U. (2014). On the effects of L2 perception and of individual differences in L1 production on L2 pronunciation. *Frontiers in Psychology*, *5*, 1246. <https://doi.org/10.3389/fpsyg.2014.01246>
- Kaushanskaya, M., & Marian, V. (2008). Mapping phonological information from auditory to written modality during foreign vocabulary learning. *Annals of the New York Academy of Sciences*, *1145*, 56-70. <https://doi.org/10.1196/annals.1416.008>
- Kazanina, N., Bowers, J. S., & Idsardi, W. (2018). Phonemes: Lexical access and beyond. *Psychonomic Bulletin & Review*, *25*(2), 560-585. <https://doi.org/10.3758/s13423-017-1362-0>
- Kim, C., Shin, G., & Hur, M. (2016). An Integrative View of Conflict Adaptation and Active Maintenance. *Japanese Psychological Research*, *58*(3), 284-295. <https://doi.org/10.1111/jpr.12113>
- Kim, K.M., & Godfroid, A. (2019). Should We Listen or Read? Modality Effects in Implicit and Explicit Knowledge. *The Modern Language Journal*, *103*(3), 648-664. <https://doi.org/10.1111/modl.12583>

- Kissling, E. (2014). What Predicts the Effectiveness of Foreign-Language Pronunciation Instruction? Investigating the Role of Perception and Other Individual Differences. *Canadian Modern Language Review/ La Revue canadienne des langues vivantes*, 70, 532-558. <https://doi.org/10.3138/cmlr.2161>
- Kissling, E. (2018). Pronunciation Instruction Can Improve L2 Learners' Bottom-Up Processing for Listening. *The Modern Language Journal*, 102(4), 653-675.
- Koller, M. (2016). robustlmm: An R Package for Robust Estimation of Linear Mixed-Effects Models. *Journal of Statistical Software*, 75(1), 1-24. <https://doi.org/10.18637/jss.v075.i06>
- Kowialiewski, B., Lemaire, B., & Portrat, S. (2021). How does semantic knowledge impact working memory maintenance? Computational and behavioral investigations. *Journal of Memory and Language*, 117, 104208. <https://doi.org/10.1016/j.jml.2020.104208>
- Krashen, S. (1982). *Principles and Practice in Second Language Acquisition*. Pergamon Press Inc.
- Krashen, S. (2003). *Explorations in language acquisition and use*. Portsmouth, NH: Heinemann (Elsevier).
- Lachs, L., McMichael, K., & Pisoni, D. (2003). Speech Perception and Implicit Memory: Evidence for Detailed Episodic Encoding of Phonetic Events 1. *Rethinking implicit memory*, 215-235.
- Laeufer, C. (2010). Nasal vowels in French loanwords in German: The effect of linguistic environment. *Folia Linguistica*, 44(1), 53-101. <https://doi.org/10.1515/flin.2010.003>
- Laganaro, M. (2019). Phonetic encoding in utterance production: A review of open issues from 1989 to 2018. *Language, Cognition and Neuroscience*, 34(9), 1193-1201. <https://doi.org/10.1080/23273798.2019.1599128>
- Lahiri, A., & Marslen-Wilson, W. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition*, 38(3), 245-294.
- Laufer, B., & Goldstein, Z. (2004). Testing Vocabulary Knowledge: Size, Strength, and Computer Adaptiveness: Language Learning. *Language Learning*, 54(3), 399-436. <https://doi.org/10.1111/j.0023-8333.2004.00260.x>
- Lee, J., Jang, J., & Plonsky, L. (2015). The Effectiveness of Second Language Pronunciation Instruction: A Meta-Analysis. *Applied Linguistics*, 36(3), 345-366. <https://doi.org/10.1093/applin/amu040>
- Léon, P.R. (1992). *Phonétisme et prononciations du français, avec des travaux pratiques d'application et leurs corrigés*. (1ère Ed.). Nathan Université. <https://elibrary.narr.digital/article/99.125005/vox199310347>
- Levis, J., Sonsaat, S., Link, S., & Barriuso, T. (2016). Native and Nonnative Teachers of L2 Pronunciation: Effects on Learner Performance. *TESOL Quarterly*, 50(4), 894-931. <https://doi.org/10.1002/tesq.272>
- Levis, J., & Wu, A. (2018). Pronunciation—Research Into Practice and Practice Into Research. *The CATESOL Journal*, 30(1), 1-12.
- Levy, E.S. (2009). On the assimilation-discrimination relationship in American English adults' French vowel learning. *The Journal of the Acoustical Society of America*, 126(5), 2670-2682. <https://doi.org/10.1121/1.3224715>

- Li, P., & Farkas, I. (2002). A self-organizing connectionist model of bilingual processing. Dans R. Heredia & J. Altarriba (Eds.), *Bilingual Sentence Processing*, Vol. 134 (pp. 59-85). Amsterdam: Elsevier. [https://doi.org/10.1016/S0166-4115\(02\)80006-1](https://doi.org/10.1016/S0166-4115(02)80006-1)
- Li, P., Farkas, I. & McWhinney, B. (2004). Early lexical development in a self-organizing neural network. *Neural Networks*, 17(8-9), 1345-1362. <https://doi.org/10.1016/j.neunet.2004.07.004>
- Li, S. (2009). The differential effects of implicit and explicit feedback on L2 learners of different proficiency levels. *Applied Language Learning*, 19, 53-79. *Applied Language Learning*, 19(1 & 2), 53-79.
- Li, S. (2010). The Effectiveness of Corrective Feedback in SLA: A Meta-Analysis. *Language Learning*, 60, 309-365. <https://doi.org/10.1111/j.1467-9922.2010.00561.x>
- Liberman, A.M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21(1), 1-36. [https://doi.org/10.1016/0010-0277\(85\)90021-6](https://doi.org/10.1016/0010-0277(85)90021-6)
- Lichtman, K. (2013). Age, ability, and awareness in implicit and explicit second language learning. *LSA Annual Meeting Extended Abstracts*, 4, 26-1-5. <https://doi.org/10.3765/exabs.v0i0.771>
- Linck, J.A., Osthus, P., Koeth, J.T., & Bunting, M.F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review*, 21(4), 861-883. <https://doi.org/10.3758/s13423-013-0565-2>
- Llompart, M., & Reinisch, E. (2017). Articulatory information helps encode lexical contrasts in a second language. *Journal of Experimental Psychology: Human Perception and Performance*, 43(5), 1040-1056. <https://doi.org/10.1037/xhp0000383>
- Loaiza, V., & McCabe, D. (2012). Temporal–contextual processing in working memory: Evidence from delayed cued recall and delayed free recall tests. *Memory & Cognition*, 40(2), 191-203. <https://doi.org/10.3758/s13421-011-0148-2>
- Long, M. (1998). Focus on Form in Task-based Language Teaching. *Working paper in ESL-University of Hawai'i. Language Policy and Pedagogy: Essays in Honor of A. Ronald Walton*, 16(2), 35-49.
- Ludusan, B., & Dupoux, E. (2015). A multilingual study on intensity as a cue for marking prosodic boundaries. *Proceedings of the 18th International Congress of Phonetic Sciences*, 5. ICPhS, Glasgow. <https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0982.pdf>
- Lyster, R. (2018). Roles for Corrective Feedback in Second Language Instruction. Dans C.A. Chapelle, *The Encyclopedia of Applied Linguistics* (pp. 1-6). NY: John Wiley & sons, Ltd. DOI: 10.1002/9781405198431.wbeal1028.pub2
- Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching*, 46(1), 1-40. <https://doi.org/10.1017/S0261444812000365>
- Macmillan, N., & Creelman, C. (2005). *Detection theory: A user's guide*, 2nd ed. Lawrence Erlbaum Associates Publishers.
- Maeda, S. (1993). Acoustics of vowel nasalization and articulatory shifts in French nasal vowels. Dans M. Huffman & R. Krakow (Eds.), *Nasals, Nasalization, and the Velum*,

- 5, (pp. 147-167). Academic Press. <https://doi.org/10.1016/B978-0-12-360380-7.50010-7>
- Mair, P., & Wilcox, R. (2020). Robust statistical methods in R using the WRS2 package. *Behavior Research Methods*, 52(2), 464-488. <https://doi.org/10.3758/s13428-019-01246-w>
- Majerus, S. (2010). Les multiples déterminants de la mémoire à court terme verbale : Implications théoriques et évaluatives. *Developpements*, n° 4(1), 5-15.
- Majerus, S., Salmon, E., & Attout, L. (2013). The importance of encoding-related neural dynamics in the prediction of inter-individual differences in verbal working memory performance. *PloS One*, 8(7), e69278. <https://doi.org/10.1371/journal.pone.0069278>
- Majerus, S., & van der Linden, M. (2003). Long-term memory effects on verbal short-term memory : A replication study. *British Journal of Developmental Psychology*, 21(2), 303-310. <https://doi.org/10.1348/026151003765264101>
- Masoura, E.V., & Gathercole, S.E. (1999). Phonological Short-term Memory and Foreign Language Learning. *International Journal of Psychology*, 34(5-6), 383-388. <https://doi.org/10.1080/002075999399738>
- Masoura, E.V., & Gathercole, S.E. (2005). Contrasting contributions of phonological short-term memory and long-term knowledge to vocabulary learning in a foreign language. *Memory*, 13(3-4), 422-429.
- Mattys, S., & Wiget, L. (2011). Effects of cognitive load on speech recognition. *Journal of Memory and Language*, 65(2), 145-160. <https://doi.org/10.1016/j.jml.2011.04.004>
- McAllister, R., Flege, J. E., & Piske, T. (2002). The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics*, 30(2), 229-258. <https://doi.org/10.1006/jpho.2002.0174>
- McClelland, J.L., & Elman, J.L., (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18(1), 1-86. [https://doi.org/10.1016/0010-0285\(86\)90015-0](https://doi.org/10.1016/0010-0285(86)90015-0)
- McKay, I., Flege, J., Piske, T., & Schirru, C. (2001). Category restructuring during second-language (L2) speech acquisition. *The journal of the acoustical society of America*, 110, 516-528. <https://doi.org/10.1121/1.1377287>
- McLeod, S., & Crowe, K. (2018). Children's Consonant Acquisition in 27 Languages : A Cross-Linguistic Review. *American Journal of Speech-Language Pathology*, 27(4), 1546-1571. https://doi.org/10.1044/2018_AJSLP-17-0100
- McLeod, A., Sutton, A., Trudeau, N., & Thordardottir, E. (2011). The acquisition of consonants in Québécois French : A cross-sectional study of pre-school aged children. *International Journal of Speech-Language Pathology*, 13(2), 93-109. <https://doi.org/10.3109/17549507.2011.487543>
- McNicol, D. (1972). Response bias interpretation of the effects of anchors in judgements of lifted weights : A comment on John's paper. *Australian Journal of Psychology*, 24(1), 87-90. <https://doi.org/10.1080/00049537208255789>
- Meara, P., Lightbown, P. M., & Halter, R. H. (1994). The Effect of Cognates on the Applicability of YES/NO Vocabulary Tests. *The Canadian Modern Language Review*, 50(2), 296-311. <https://doi.org/10.3138/cmlr.50.2.296>
- Meara, P. M., & Milton, J. (2003). *X-lex : The Swansea levels test*. Express Publishing.

- Meschyan, G., & Hernandez, A. (2002). Is native-language decoding skill related to second-language learning? *Journal of Educational Psychology, 94*(1), 14-22.
<https://doi.org/10.1037/0022-0663.94.1.14>
- Mesgarani, N., Cheung, C., Johnson, K., & Chang, E.F. (2014). Phonetic feature encoding in human superior temporal gyrus. *Science, 343*(6174), 1006-1010.
- Metsala, J. (2011). Repetition of Less Common Sound Patterns: A Unique Relationship to Young Children's Phonological Awareness and Word Reading. *International Journal of English Linguistics, 1*(2), 3-17. <https://doi.org/10.5539/ijel.v1n2p3>
- Meynadier, Y. (2001). La syllabe phonétique et phonologique : Une introduction. *Travaux Interdisciplinaires du Laboratoire Parole et Langage d'Aix-en-Provence (TIPA), 20*, 91-148.
- Meynadier, Y. (2013). Éléments de phonétique acoustique. In M.A.-D. Noël Nguyen (Ed.), *Méthodes et outils pour l'analyse phonétique des grands corpus oraux* (pp. 25-83). Hermès. <https://hal.archives-ouvertes.fr/hal-01212693>
- Mille Feuilles*. Schulverlag. <https://www.1000feuilles.ch/>
- Miller, G.A., Galanter, E., & Pribram, K.H. (1960). *PLANS and the STRUCTURE of BEHAVIOR*. USA : Holt, Rinehart and Winston, Inc.
- Miwa, K., Dijkstra, T., Bolger, P., & Baayen, R. H. (2014). Reading English with Japanese in mind : Effects of frequency, phonology, and meaning in different-script bilinguals*. *Bilingualism: Language and Cognition, 17*(3), 445-463.
<https://doi.org/10.1017/S1366728913000576>
- Mora, G., & Camos, V. (2013). Two systems of maintenance in verbal working memory : Evidence from the word length effect. *PloS One, 8*(7), e70026.
<https://doi.org/10.1371/journal.pone.0070026>
- Mora, G., & Camos, V. (2015). Dissociating rehearsal and refreshing in the maintenance of verbal information in 8-year-old children. *Frontiers in Psychology, 6*(11).
<https://doi.org/10.3389/fpsyg.2015.00011>
- Munro, M., & Derwing, T. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System, 34*(4), 520-531.
<https://doi.org/10.1016/j.system.2006.09.004>
- Murakami, T., Kell, C., Restle, J., Ugawa, Y., & Ziemann, U. (2015). Left dorsal speech stream components and their contribution to phonological processing. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 35*(4), 1411-1422.
<https://doi.org/10.1523/JNEUROSCI.0246-14.2015>
- Myung, I., Forster, M., & Browne, M.W. (2000). Special Issue on Model Selection. *Journal of Mathematical Psychology, 44*(1-2), 190-204.
[https://www.scirp.org/\(S\(vtj3fa45qm%201ean45vvffcz55\)\)/reference/referencespapers.aspx?referenceid=1542785](https://www.scirp.org/(S(vtj3fa45qm%201ean45vvffcz55))/reference/referencespapers.aspx?referenceid=1542785)
- Näätänen, R., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin, 125*(6), 826-859.
- Nakayama, M., Sears, C.R., Hino, Y., & Lupker, S.J. (2012). Cross-script phonological priming for Japanese-English bilinguals : Evidence for integrated phonological

- representations. *Language and Cognitive Processes*, 27(10), 1563-1583.
<https://doi.org/10.1080/01690965.2011.606669>
- Nation, P. (2001). Planning and Running an Extensive Reading Program. *NUCB journal of language culture and communication*, 3(1), 1-8.
- Nation, P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, 63(1), 59-82. <https://doi.org/10.3138/cmlr.63.1.59>
- Niebuhr, O., & Meunier, C. (2011). The phonetic manifestation of French /s#/#/ and /#s/ sequences in different vowel contexts: On the occurrence and the domain of sibilant assimilation. *Phonetica*, 68(3), 133-160. <https://doi.org/10.1159/000331235>
- Norris, D., McQueen, J., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, 47(2), 204-238. [https://doi.org/10.1016/s0010-0285\(03\)00006-9](https://doi.org/10.1016/s0010-0285(03)00006-9)
- Norris, J., & Ortega, L. (2000). Effectiveness of L2 Instruction: A Research Synthesis and Quantitative Meta-analysis. *Language Learning*, 50(3), 417-528.
<https://doi.org/10.1111/0023-8333.00136>
- Norris, J., & Ortega, L. (2009). Towards an Organic Approach to Investigating CAF in Instructed SLA: The Case of Complexity. *Applied Linguistics*, 30, 555-578.
<https://doi.org/10.1093/applin/amp044>
- O'Dell, F., Read, J., & McCarthy, M. (2000). *Assessing vocabulary*. Cambridge university press.
- Oftinger, A.-L., & Camos, V. (2015). Maintenance Mechanisms in Children's Verbal Working Memory. *Journal of Educational and Developmental Psychology*, 6(1), 16-28.
- Ohala, J. J. (2005). Phonetic explanations for sound patterns. *UC Berkeley Phonology Lab Annual Report. A Figure of Speech: A Festschrift for John Laver*, 23, 269-288.
- Ota, M., Hartsuiker, R., & Haywood, S. (2009). The KEY to the ROCK : Near-homophony in nonnative visual word recognition. *Cognition*, 111(2), 263-269.
<https://doi.org/10.1016/j.cognition.2008.12.007>
- Pajak, B., & Levy, R. (2014). The role of abstraction in non-native speech perception. *Journal of Phonetics*, 46, 147-160. <https://doi.org/10.1016/j.wocn.2014.07.001>
- Pallier, C., Colomé, A., & Sebastián-Gallés, N. (2001). The Influence of Native-Language Phonology on Lexical Access : Exemplar-Based Versus Abstract Lexical Entries. *Psychological Science*, 12(6), 445-449. <https://doi.org/10.1111/1467-9280.00383>
- Papachristou, V. (2011). Explicit vs. Implicit Pronunciation Teaching to Greek children: The case of the acquisition of English vowels. *Selected Papers on Theoretical and Applied Linguistics*, 19, 371-381.
- Paradis, C., & Prunet, J.-F. (2000). Nasal Vowels as Two Segments: Evidence from Borrowings. *Language*, 76(2), 324-357. <https://doi.org/10.1353/lan.2000.0117>
- Paradis, M. (2008). Language and Communication Disorders in Multilinguals. Dans B. Stemmer & H.A. Whitaker (Eds), *Handbook of the Neuroscience of Language*, (pp. 341-349). Amsterdam: Elsevier,
- Peereman, R., Lété, B., & Sprenger-Charolles, L. (2007). Manulex-infra : Distributional characteristics of grapheme—phoneme mappings, and infralexical and lexical units in

- child-directed written material. *Behavior Research Methods*, 39(3), 579-589.
<https://doi.org/10.3758/BF03193029>
- Peperkamp, S., Pettinato, M., & Dupoux, E. (2002, novembre). Allophonic variation and the acquisition of phoneme categories. *Proceedings of the 27th Annual Boston University Conference on Language Development*, Boston, MA.
- Pérez, A., Carreiras, M., Gillon Dowens, M., & Duñabeitia, J. A. (2015). Differential oscillatory encoding of foreign speech. *Brain and Language*, 147, 51-57.
<https://doi.org/10.1016/j.bandl.2015.05.008>
- Pérez, M. (2014). *Proposition de hiérarchisation des 45 graphèmes de base de l'orthographe du français*. <https://hal-univ-tlse2.archives-ouvertes.fr/hal-00961511>
- Perfetti, C., & Hart, L. (2002). The lexical quality hypothesis. Dans L. Verhoeven, C. Elbro & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189-213). Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/swll.11.14per>
- Petersen, S. (2015). Categorical Nasal Vowel Acquisition in L2 French Learners. *Working papers in Linguistics*, 21(1), Article 25.
<http://repository.upenn.edu/pwpl/vol21/iss1/25/>
- Piaget, J. (1972). *Problèmes de psychologie génétique*. Paris: Denoël-Gonthier.
- Piccardo, E., Goodier, T., & North, B. (2018). *Council of Europe (2018). Common European Framework of Reference for Languages : Learning, Teaching, Assessment. Companion Volume with New Descriptors*. Strasbourg: Council of Europe Publishing. Authors: B. North, E. Piccardo, T. Goodier.
- Piske, T., Flege, J.E., McKay, I., & Meador, D. (2002). The Production of English Vowels by Fluent Early and Late Italian-English Bilinguals. *Phonetica*, 59(1), 49-71.
<https://doi.org/10.1159/000056205>
- Ponniah, J. (2011). The Effectiveness of the Comprehension Hypothesis: A Review on the Current Research on Incidental Vocabulary Acquisition. *Journal on English Language Teaching*, 1(2), 1-4.
- Pulvermüller, F., Huss, M., Kherif, F., Moscoso del Prado M.F., Hauk, O., & Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences of the United States of America*, 103(20), 7865-7870.
<https://doi.org/10.1073/pnas.0509989103>
- R Foundation for Statistical Computing. (2020). *R Core Team (2020). R: A language and environment for statistical computing*. (Version 3-6-3) [Computer software].
<https://www.R-project.org/>
- Racine, I., Detey, S., Bühler, N., Schwab, S., Zay, F., & Kawaguchi, Y. (2010). *The production of French nasal vowels by advanced Japanese and Spanish learners of French : A corpus-based evaluation study*. Dans K. Dziubalska-Kotaczyk, M. Wrembel & M. Kul (Eds.), *Proceedings of New Sounds 2010*, Sixth International Symposium on the acquisition of second language speech, Poznań, 367-372.
- Rafat, Y. (2012). *Orthography-induced Transfer in the Production of Novice Adult English-speaking Learners of Spanish*. Thèse de Doctorat, University of Toronto.
<https://tspace.library.utoronto.ca/handle/1807/31909>

- Raguenes, A. (2011). Mémoire auditivo-verbale chez l'enfant : Entraînement informatisé de l'empan de rythmes, syllabes et mots. *Glossa*, 110, 86-110. <https://www.glossa.fr/index.php/glossa/article/view/571>
- Reetz, H., & Jongman, A. (2020). *Phonetics : Transcription, Production, Acoustics, and Perception*. Hoboken, NJ : Wiley-Blackwell.
- Rentzopoulos, P., & Kokkinakis, G. (1992, octobre). Multilingual Phoneme to grapheme conversion system based on HMM. *Second international conference on spoken language processing*. ICSLP, Banff, Alberta, Canada. https://www.isca-speech.org/archive/icslp_1992/i92_1191.html
- Rentzopoulos, P., & Kokkinakis, G. (1996). Efficient Multilingual Phoneme-to-grapheme Conversion Based on HMM. *Comput. Linguist.*, 22(3), 351-376.
- Ricker, T. (2015). The role of short-term consolidation in memory persistence. *AIMS Neuroscience*, 2(4), 259-279. <https://doi.org/10.3934/Neuroscience.2015.4.259>
- Robinson, P. (2015). The Cognition Hypothesis, second language task demands, and the SSARC model of pedagogic task sequencing. *Domains and Directions in the Development of TBLT*, 87-122.
- Robinson, P., & Gilabert, R. (2007). Task complexity, the Cognition Hypothesis and second language learning and performance. *IRAL - International Review of Applied Linguistics in Language Teaching*, 45, 161-176. <https://doi.org/10.1515/IRAL.2007.007>
- Roodenrys, S., Hulme, C., Alban, J., Ellis, A.W., & Brown, G.D.A. (1994). Effects of word frequency and age of acquisition on short-term memory span. *Memory & Cognition*, 22(6), 695-701. <https://doi.org/10.3758/BF03209254>
- Rose, N., Buchsbaum, B., & Craik, F. (2014). Short-term retention of a single word relies on retrieval from long-term memory when both rehearsal and refreshing are disrupted. *Memory & Cognition*, 42(5), 689-700. <https://doi.org/10.3758/s13421-014-0398-x>
- Rosseel, Y. (2012). lavaan : An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(1), 1-36. <https://doi.org/10.18637/jss.v048.i02>
- Rousse-Malpat, A., & Verspoor, M. (2012). Measuring effectiveness in Focus on Form versus Focus on Meaning. *Dutch Journal of Applied Linguistics*, 263-276.
- Ruellot, V. (2014). Introducing French nasal vowels at the beginner level: A demystification. *Pronunciation in second language learning and teaching*, 151-155.
- Saito, K. (2012). Effects of Instruction on L2 Pronunciation Development : A Synthesis of 15 Quasi-Experimental Intervention Studies. *TESOL Quarterly*, 46(4), 842-854.
- Saito, K. (2021). Effects of Corrective Feedback on Second Language Pronunciation Development. Dans E. Kartchava & H. Nassaji (Eds.), *The Cambridge Handbook of Corrective Feedback in Second Language Learning and Teaching* (pp. 407-428). Cambridge University Press. <https://doi.org/10.1017/9781108589789.020>
- Saito, K., & Lyster, R. (2012). Effects of Form-Focused Instruction and Corrective Feedback on L2 Pronunciation Development of /ɪ/ by Japanese Learners of English : FFI and L2 Pronunciation. *Language Learning*, 62(2), 595-633. <https://doi.org/10.1111/j.1467-9922.2011.00639.x>

- Saito, K., Trofimovich, P., Isaacs, T., & Webb, S. (2017). Re-examining phonological and lexical correlates of second language comprehensibility: The role of rater experience. Dans T. Isaacs & P. Trofimovich (Eds.), *Second language assessment: Interdisciplinary perspectives*. (pp. 141-156). Bristol: Multilingual Matters. <https://www.jstor.org/stable/10.21832/j.ctt1xp3wcc.12>
- Sakai, M., & Moorman, C. (2018). Can perception training improve the production of second language phonemes? A meta-analytic review of 25 years of perception training research. *Applied Psycholinguistics*, 39(1), 187-224. <https://doi.org/10.1017/S0142716417000418>
- Schmid, S. (2009). La prononciation du français par des élèves d'un lycée zurichois. Dans H.R. Nüesch (Ed.), *Galloromanica et Romanica* (pp. 253-267). Basel: Verlag.
- Schmid, S. (2012). The pronunciation of voiced obstruents in L2 French: A preliminary study of Swiss German learners. *Poznań Studies in Contemporary Linguistics*, 48(4), 627-659. <https://doi.org/10.1515/psicl-2012-0028>
- Schmidt, R. (2001). Attention. Dans P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge University Press.
- Schmidt, R. (2012). CH2 Attention, awareness, and individual differences in language learning. Dans W.M. Chan, K.N. Chin, S. Bhatt & I. Walker (Eds.), *Perspectives on Individual Characteristics and Foreign Language Education* (pp. 27-50). Walter de Gruyter.
- Schmitt, N. (2014). Size and Depth of Vocabulary Knowledge: What the Research Shows. *Language Learning*, 64(4), 913-951. <https://doi.org/10.1111/lang.12077>
- Schröger, E. (1997). On the detection of auditory deviations: A pre-attentive activation model. *Psychophysiology*, 34(3), 245-257. <https://doi.org/10.1111/j.14698986.1997.tb02395.x>
- Service, E., & Kohonen, V. (1995). Is the relation between phonological memory and foreign language learning accounted for by vocabulary acquisition? *Applied Psycholinguistics*, 16(2), 155-172. <https://doi.org/10.1017/S0142716400007062>
- Shapiro, K., Raymond, J., & Arnell, K. (1994). Attention to visual pattern information produces the attentional blink in rapid serial visual presentation. *Journal of Experimental Psychology. Human Perception and Performance*, 20(2), 357-371. <https://doi.org/10.1037//0096-1523.20.2.357>
- Sheen, Y., & Ellis, R. (2011). Corrective Feedback in Language Teaching. Dans E. Hinkel (Ed.), *Handbook of Research in Second Language teaching and learning*, 2 (pp. 593-610). Seattle: Routledge, Taylor & Francis Group.
- Shin, J.-A., & Christianson, K. (2012). Structural Priming and Second Language Learning. *Language Learning*, 62(3), 931-964. <https://doi.org/10.1111/j.1467-9922.2011.00657.x>
- Shosted, R., Carignan, C., & Rong, P. (2012). Managing the distinctiveness of phonemic nasal vowels: Articulatory evidence from Hindi. *The Journal of the Acoustical Society of America*, 131(1), 455-465.
- Showalter, C. E. (2020). Russian phono-lexical acquisition and orthographic input. Naïve learners, experienced learners, and interventions. *Studies in Second Language Acquisition*, 42(2), 255-277. <https://doi.org/10.1017/S0272263119000585>

- Showalter, C. E., & Hayes-Harb, R. (2013). Unfamiliar orthographic information and second language word learning : A novel lexicon study. *Second Language Research*, 29(2), 185-200. <https://doi.org/10.1177/0267658313480154>
- Showalter, C. E., & Hayes-Harb, R. (2015). Native English speakers learning Arabic : The influence of novel orthographic information on second language phonological acquisition. *Applied Psycholinguistics*, 36(01), 23-42. <https://doi.org/10.1017/S0142716414000411>
- Skehan, P. (2015). Working memory and second language performance : A commentary. Dans Z. Wen, M. Borges Mota & A. McNeill (Eds.), *Working memory in second language acquisition and processing*. (pp. 189-201). Bristol: Multilinguals matters.
- Smith, E. E., & Jonides, J. (1999). Storage and Executive Processes in the Frontal Lobes. *Science*, 283(5408), 1657-1661. <https://doi.org/10.1126/science.283.5408.1657>
- Snoeren, N., Seguí, J., & Hallé, P. (2008). On the role of regular phonological variation in lexical access: Evidence from voice assimilation in French. *Cognition*, 108(2), 512-521. <https://doi.org/10.1016/j.cognition.2008.02.008>
- Spada, N. (1997). Form-Focussed Instruction and Second Language Acquisition: A Review of Classroom and Laboratory Research. *Language Teaching*, 30(2), 73-87. <https://doi.org/10.1017/S0261444800012799>
- Spada, N., Jessop, L., Tomita, Y., Suzuki, W., & Valeo, A. (2014). Isolated and Integrated form-focused instruction: Effects on different types of L2 knowledge. *Language Teaching Research*, 18(4), 453-473. <https://doi.org/10.1177/1362168813519883>
- Spada, N., & Tomita, Y. (2010). Interactions Between Type of Instruction and Type of Language Feature: A Meta-Analysis: Type of Instruction and Language Feature. *Language Learning*, 60(2), 263-308. <https://doi.org/10.1111/j.1467-9922.2010.00562.x>
- Spinner, P. (2013). Language Production and Reception : A Processability Theory Study. *Language Learning*, 63(4), 704-739. <https://doi.org/10.1111/lang.12022>
- Sprenger-Charolles, L. (2004). Linguistic Processes in Reading and Spelling : The Case of Alphabetic Writing Systems: English, French, German and Spanish. Dans T. Nunes & P. Bryant (Eds.), *Handbook of Children's Literacy* (pp. 43-65). Netherlands; Springer. https://doi.org/10.1007/978-94-017-1731-1_4
- Sprenger-Charolles, L., Siegel, L. S., & Bonnet, P. (1998). Reading and Spelling Acquisition in French: The Role of Phonological Mediation and Orthographic Factors. *Journal of Experimental Child Psychology*, 68(2), 134-165. <https://doi.org/10.1006/jecp.1997.2422>
- Squire, L., & Dede, A. (2015). Conscious and unconscious memory systems. *Cold Spring Harbor Perspectives in Biology*, 7(3), a021667. <https://doi.org/10.1101/cshperspect.a021667>
- Staffeldt, S. (2010). Einführung in die Phonetik, Phonologie und Graphematik des Deutschen. Ein Leitfaden für den akademischen Unterricht. Tübingen: Stauffenburg Verlag.
- Steinschneider, M. (2013). Phonemic Representations and Categories. Dans Y. E. Cohen, A. N. Popper, & R. R. Fay (Eds.), *Neural Correlates of Auditory Cognition*, Vol. 45 (pp. 151-191). New York: Springer. https://doi.org/10.1007/978-1-4614-2350-8_6

- Stemmer, B., & Whitaker, H.A. (2008). *Handbook of the neuroscience of language*. Amsterdam: Elsevier Science.
- Stoel-Gammon, C. (2011). Relationships between lexical and phonological development in young children. *Journal of Child Language*, 38(1), 1-34.
<https://doi.org/10.1017/S0305000910000425>
- Strange, W. (2011). Automatic selective perception (ASP) of first and second language speech : A working model. *Journal of Phonetics*, 39(4), 456-466.
<https://doi.org/10.1016/j.wocn.2010.09.001>
- Suzuki, Y., Nakata, T., & Dekeyser, R. (2019). Optimizing Second Language Practice in the Classroom : Perspectives from Cognitive Psychology. *The Modern Language Journal*, 103(3), 551-561. <https://doi.org/10.1111/modl.12582>
- Swingle, D. (2003). Phonetic Detail in the Developing Lexicon. *Language and Speech*, 46(2-3), 265-294. <https://doi.org/10.1177/00238309030460021001>
- Szewczyk, J., Marecka, M., Chiat, S., & Wodniecka, Z. (2018). Nonword repetition depends on the frequency of sublexical representations at different grain sizes: Evidence from a multi-factorial analysis. *Cognition*, 179, 23-36.
<https://doi.org/10.1016/j.cognition.2018.06.002>
- Tagarelli, K., Mota, M., & Rebuschat, P. (2011). The role of working memory in implicit and explicit language learning. *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*, 33, 2061-2066.
- Tam, H., Jarrold, C., Baddeley, A., & Sabatos-DeVito, M. (2010). The development of memory maintenance : Children's use of phonological rehearsal and attentional refreshment in working memory tasks. *Journal of Experimental Child Psychology*, 107(3), 306-324. <https://doi.org/10.1016/j.jecp.2010.05.006>
- Taylor, R. S., Francis, W. S., Borunda-Vazquez, L., & Carbajal, J. (2019). Mechanisms of word concreteness effects in explicit memory : Does context availability play a role? *Memory & cognition*, 47(1), 169-181. <https://doi.org/10.3758/s13421-018-0857-x>
- Thomson, R. I., & Derwing, T. (2015). The Effectiveness of L2 Pronunciation Instruction : A Narrative Review. *Applied Linguistics*, 36(3), 326-344.
<https://doi.org/10.1093/applin/amu076>
- Tremblay, P., & Bédard, P. (2013). *SyllabO+*. Laboratoire des neurosciences de la parole et de l'audition. <https://speechneurolab.ca/fr/syllabo>
- Trofimovich, P., & Gatbonton, E. (2006). Repetition and Focus on Form in Processing L2 Spanish Words : Implications for Pronunciation Instruction. *The Modern Language Journal*, 90(4), 519-535. <https://doi.org/10.1111/j.1540-4781.2006.00464.x>
- Tse, C.-S., & Altarriba, J. (2009). The word concreteness effect occurs for positive, but not negative, emotion words in immediate serial recall. *British Journal of Psychology*, 100(1), 91-109. <https://doi.org/10.1348/000712608X318617>
- Tyler, M., Best, C., Faber, A., & Levitt, A. (2014). Perceptual assimilation and discrimination of non-native vowel contrasts. *Phonetica*, 71(1), 4-21.
<https://doi.org/10.1159/000356237>
- Ullman, M. T. (2013). The role of declarative and procedural memory in disorders of language. *Linguistic variation*, 13(2), 133-154.

- van Leussen, J.-W., & Escudero, P. (2015). Learning to perceive and recognize a second language : The L2LP model revised. *Frontiers in Psychology*, 6(1000).
<https://doi.org/10.3389/fpsyg.2015.01000>
- Veivo, O., Järvikivi, J., Porretta, V., & Hyönä, J. (2016). Orthographic Activation in L2 Spoken Word Recognition Depends on Proficiency: Evidence from Eye-Tracking. *Frontiers in Psychology*, 7(1120). <https://doi.org/10.3389/fpsyg.2016.01120>
- Ventura, P., Kolinsky, R., Fernandes, S., Querido, L., & Morais, J. (2007). Lexical restructuring in the absence of literacy. *Cognition*, 105, 334-361.
<https://doi.org/10.1016/j.cognition.2006.10.002>
- Vihman, M.M., & Keren-Portnoy, T. (2019). *The Emergence of Phonology. Whole-word approaches and cross-linguistic evidence*. Cambridge University Press.
- Vitevitch, M.S., Luce, P.A., Charles-Luce, J., & Kemmerer, D. (1997). Phonotactics and Syllable Stress : Implications for the Processing of Spoken Nonsense Words. *Language and Speech*, 40(1), 47-62. <https://doi.org/10.1177/002383099704000103>
- Vygotsky, L. (1978). *Mind in society : The development of higher psychological processes*. M. Cole, V. John-Steiner, S. Scribner & E. Souberman, (Eds.). Harvard University Press.
- Walker, I., & Hulme, C. (1999). Concrete words are easier to recall than abstract words : Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(5), 1256-1271.
- Wang, F., Karipidis, I. I., Pleisch, G., Fraga-González, G., & Brem, S. (2020). Development of Print-Speech Integration in the Brain of Beginning Readers With Varying Reading Skills. *Frontiers in Human Neuroscience*, 14(289).
<https://doi.org/10.3389/fnhum.2020.00289>
- Webb, S. (2007). The Effects of Repetition on Vocabulary Knowledge. *Applied Linguistics*, 28(1), 46-65. <https://doi.org/10.1093/applin/aml048>
- Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, 50(1), 1-25. [https://doi.org/10.1016/S0749-596X\(03\)00105-0](https://doi.org/10.1016/S0749-596X(03)00105-0)
- Wen, Z.E., Borges Mota, M., & McNeill, A. (2015). *Working Memory in Second Language Acquisition and Processing*. Bristol, UK : Multilingual Matters.
<https://www.degruyter.com/document/doi/10.21832/9781783093595/html>
- Werker, J.F., & Curtin, S. (2005). PRIMIR : A Developmental Framework of Infant Speech Processing. *Language learning and development*, 1(2), 197-234.
- Werker, J.F., & Logan, J.S. (1985). Cross-language evidence for three factors in speech perception. *Perception & Psychophysics*, 37(1), 35-44.
<https://doi.org/10.3758/BF03207136>
- Werker, J.F., & Yeung, H.H. (2005). Infant speech perception bootstraps word learning. *Trends in Cognitive Sciences*, 9(11), 519-527.
<https://doi.org/10.1016/j.tics.2005.09.003>
- White, K.S., Peperkamp, S., Kirk, C., & Morgan, J. L. (2008). Rapid acquisition of phonological alternations by infants. *Cognition*, 107(1), 238-265.
<https://doi.org/10.1016/j.cognition.2007.11.012>

- Wioland, F. (1991). *Prononcer les mots du français : Des sons et des rythmes* (Hachette). Hachette FLE.
- Wright, D. B., Horry, R., & Skagerberg, E. M. (2009). Functions for traditional and multilevel approaches to signal detection theory. *Behavior Research Methods*, 41(2), 257-267. <https://doi.org/10.3758/BRM.41.2.257>
- Yeung, H.H., & Werker, J.F. (2013). Lip Movements Affect Infants' Audiovisual Speech Perception. *Psychological Science*, 603-612.
- Zatorre, R., & Gandour, J. (2008). Neural specializations for speech and pitch: Moving beyond the dichotomies. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 363(1493), 1087-1104. <https://doi.org/10.1098/rstb.2007.2161>
- Zhang, R., & Yuan, Z. (2020). Examining the effects of explicit pronunciation instruction on the development of L2 pronunciation. *Studies in Second Language Acquisition*, 42(4), 905-918. <https://doi.org/10.1017/S0272263120000121>
- Ziegler, J., Jacobs, A., & Stone, G. (1996). Statistical analysis of the bidirectional inconsistency of spelling and sound in French. *Behavior Research Methods, Instruments, & Computers*, 28(4), 504-515. <https://doi.org/10.3758/BF03200539>
- Ziegler, J., & Goswami, U. (2005). Reading Acquisition, Developmental Dyslexia, and Skilled Reading Across Languages: A Psycholinguistic Grain Size Theory. *Psychological bulletin*, 131, 3-29. <https://doi.org/10.1037/0033-2909.131.1.3>
- Zielinski, B. (2008). The listener: No longer the silent partner in reduced intelligibility. *System*, 36, 69-84. <https://doi.org/10.1016/j.system.2007.11.004>

Financement

Cette thèse a été financée grâce au programme P-9 de Swissuniversities.