



<http://portaildoc.univ-lyon1.fr>

Creative commons : Paternité - Pas d'Utilisation Commerciale -
Pas de Modification 2.0 France (CC BY-NC-ND 2.0)



<http://creativecommons.org/licenses/by-nc-nd/2.0/fr>



N° de mémoire 2312

Mémoire d'Orthophonie

présenté pour l'obtention du

Certificat de capacité d'orthophoniste

Par

ASSELIN Eponine

**Importance du niveau de vocabulaire sur les compétences en
lecture de mots chez les enfants allophones**

Étude transversale multicentrique chez des enfants de niveau primaire

Mémoire dirigé par

BIGNON Matthieu et CASALIS Séverine

Mémoire évalué par

PETIT Nicolas et CAVALLI Eddy

Année académique

2022-2023

INSTITUT DES SCIENCES ET TECHNIQUES DE READAPTATION
DEPARTEMENT ORTHOPHONIE

Directeur ISTR
Pr. Jacques LUAUTÉ

Équipe de direction du département d'orthophonie

Directeur de formation
Solveig CHAPUIS

Coordinateur de cycle 1
Claire GENTIL

Coordinateur de cycle 2
Ségoène CHOPARD

Responsables de l'enseignement clinique
Johanne BOUQUAND
Ségoène CHOPARD
Alice MICHEL-JOMBART

Responsables des travaux de recherche
Mélanie CANAULT
Floriane DELPHIN-COMBE
Claire GENTIL
Nicolas PETIT

Responsable de la formation continue
Johanne BOUQUAND

Responsable du pôle scolarité
Rachel BOUTARD

Secrétariat de scolarité
Audran ARRAMBOURG
Sigolène-Victoria CHEVALIER
Danièle FEDERICI

UNIVERSITE CLAUDE BERNARD LYON 1

Président
Pr. FLEURY Frédéric

Vice-président CFVU
Mme BROCHIER Céline

Vice-président CA
Pr. REVEL Didier

Vice-président CR
M. HONNERAT Jérôme
Délégué de la Commission Recherche Secteur
Santé

Directeur Général des Services
M. ROLLAND Pierre

1 Secteur Santé

U.F.R. de Médecine Lyon Est Doyen
Pr. RODE Gilles

Institut des Sciences Pharmaceutiques et
Biologiques
Pr. DUSSART Claude

U.F.R. de Médecine et de maïeutique
Lyon-Sud Charles Mérieux Doyenne
Pr. PAPAREL Philippe

Institut des Sciences et Techniques de la
Réadaptation (I.S.T.R.)
Pr LUAUTÉ Jacques

U.F.R. d'Odontologie
Pr. MAURIN Jean-Christophe

2 Secteur Sciences et Technologie

U.F.R. Faculté des Sciences
Directeur **M. ANDRIOLETTI Bruno**

Institut des Sciences Financières et
d'Assurance (I.S.F.A.)
Directeur **M. LEBOISNE Nicolas**

U.F.R. Biosciences
Directrice **Mme GIESELER Kathrin**

Observatoire Astronomique de Lyon
Directeur **M. GUIDERDONI Bruno**

U.F.R. de Sciences et Techniques des
Activités Physiques et Sportives
(S.T.A.P.S.)
Directeur **M. BODET Guillaume**

POLYTECH LYON
Directeur **M. PERRIN Emmanuel**

Institut National Supérieure du
Professorat et de l'Éducation (INSPé)
Directeur **M. CHAREYRON Pierre**

Institut Universitaire de Technologie de
Lyon 1 (I.U.T. LYON 1)
Directeur **M. MASSENZIO Mich**

Résumé

Les enfants allophones font face à un défi majeur à leur arrivée en France puisqu'ils doivent apprendre à maîtriser la langue française orale et écrite de façon concomitante. Pour les accompagner dans cette étape, ils sont scolarisés dans des unités d'enseignement spécialisé, les UPE2A (Unités Pédagogiques pour Élèves Allophones Arrivants). Le langage écrit apparaît alors comme la base de tous les futurs apprentissages scolaires et quotidiens. La littérature scientifique ayant peu étudié les compétences nécessaires au bon développement de la lecture dans le cadre de l'allophonie, cette étude a pour objectif principal de déterminer et comparer le poids des différents facteurs cognitifs et langagiers lors de la lecture de mots. L'hypothèse soulevée est que les différences inter-individuelles en vocabulaire chez les enfants allophones auront plus d'impact sur la réussite en lecture que chez les enfants tout-venant. Le niveau de vocabulaire serait alors plus prédictif de la réussite en lecture de mots chez les enfants allophones que chez les enfants monolingues. Pour répondre à cette problématique, les performances de 231 enfants tout-venant ont été comparées à celles de 157 enfants allophones grâce à des épreuves réalisées sur le temps scolaire et évaluant les facteurs prédictifs de la lecture. Les résultats de la présente étude montrent que pour les deux groupes étudiés, tous les facteurs prédictifs testés sont corrélés significativement aux performances en lecture de mots. De plus, le vocabulaire occupe une place plus importante lors de la lecture pour les enfants allophones que pour les enfants tout-venant. Ces résultats sont à nuancer si l'on prend en compte la part de variance expliquée par le modèle des enfants allophones qui est supérieure à celle des enfants tout-venant. En conclusion, ce sont les facteurs phonologiques qui occupent une place prédominante lors de la lecture de mots, indépendamment du statut linguistique.

Mots clés : allophonie – lecture – prédicteur – vocabulaire – enfant – langue seconde

Abstract

Allophone children face a major challenge at their arrival in France, as they must learn to master the French language, both spoken and written, at the same time. To support them in this step, they are enrolled in specialized teaching units, the UPE2A (Unités Pédagogiques pour Élèves Allophones Arrivants). Written language appears to be the basis of all future school and daily learning. Since the scientific literature has not studied the skills necessary for the proper development of reading in the context of allophony, the main objective of this study is to determine and compare the weight of the different cognitive and language factors during word reading. The hypothesis raised is that inter-individual differences in vocabulary among allophone children will have a bigger impact on reading success than among non-allophone children. The level of vocabulary would then be more predictive of word reading success in allophone children than in monolingual children. To address this issue, the performance of 231 non-native children was compared to the one of 157 allophone children using school-based tests that assessed predictive factors of reading. The results of the present study show that for both groups studied, all the predictive factors tested were significantly correlated with word reading performance. In addition, vocabulary is more important in reading for allophone children than for non-allophone children. These results must be qualified if we take into account the proportion of variance explained by the allophone children's model, which is greater than that of the general population. In conclusion, phonological factors are predominant in word reading, regardless of linguistic status.

Key words : allophony - reading - predictor - vocabulary - child - second language

Remerciements

Je tiens tout d'abord à remercier mes deux directeurs de mémoire, Matthieu Bignon et Séverine Casalis. Merci Matthieu particulièrement pour ta disponibilité, tes précieux conseils et tes nombreuses relectures qui m'ont permis de mener à bien ce travail durant deux ans. Je remercie également les instituteurs, institutrices et tous les enfants qui ont donné de leur temps et de leur énergie pour concrétiser ce projet. Travailler avec vous fut très enrichissant, tant sur le plan professionnel que personnel.

Je remercie toutes les maîtres de stages que j'ai côtoyées durant 5 ans, et spécifiquement Pauline et Céline, qui m'ont permis de me projeter en tant que future orthophoniste. Vos pratiques cliniques m'ont profondément inspirée.

Puis je remercie chaleureusement ma famille, ma mère, mon père, Léopold, Hippolyte et mamie. Merci du soutien sans faille que vous m'apportez au quotidien. Maman, merci de me conseiller, de me guider sans jamais t'imposer. J'espère te rendre fière, aujourd'hui et les jours à venir.

Un merci tout particulier à Clotilde, amie du premier jour de cette aventure orthophonique et qui m'accompagne depuis. Tu es un exemple d'intelligence et de courage.

Merci à Ines, pour ton soutien et tes conseils si justes. Depuis ce voyage inlassablement remémoré autour de tisanes tardives, tu as été particulièrement cette année un soleil vers qui me tourner. Le meilleur nous attend.

Merci à Romane, colocataire toujours présente et rassurante. Je te souhaite un merveilleux voyage.

Merci à toi, tu te reconnaîtras, de m'offrir les plus belles conditions pour terminer ce mémoire. J'ai hâte de voir ce que la suite nous réserve.

Et enfin un immense merci à Alice, merci de m'avoir proposé de visiter Lyon pour ne plus jamais me quitter. Tu as donné une autre dimension à ces années lyonnaises. Merci de ton soutien, de tes conseils et de tous ces moments partagés qui dépassent bien évidemment le cadre des études. Ce n'est que le début et nos chemins se recroiseront sans aucun doute.

SOMMAIRE

I	PARTIE THEORIQUE	1
1	Introduction générale.....	1
2	La lecture.....	2
2.1	Le modèle simple de la lecture.....	2
2.2	Les stades et les phases d'apprentissage de la lecture	2
3	Les prédicteurs de la lecture	4
3.1	La conscience phonologique	4
3.2	La mémoire phonologique à court terme	5
3.3	La dénomination rapide automatisée	6
3.4	Le niveau de vocabulaire.....	7
4	L'apprentissage de la lecture dans une seconde langue	8
4.1	Le transfert entre deux langues.....	8
4.2	L'importance du niveau de vocabulaire lors de l'apprentissage de la lecture dans une deuxième langue	8
5	Buts et hypothèses.....	9
II	METHODE	10
1	Population.....	10
2	Matériel.....	11
2.1	A l'attention des familles : courriers d'information, formulaires de consentement et questionnaires	11
2.2	A l'attention des enfants : les épreuves	11
2.2.1	<i>Matrices progressives de Raven (Raven et al., 1998).</i>	12
2.2.2	<i>Analyse phonologique, EVALEC (Sprenger-Charolles et al., 2010).</i>	12
2.2.3	<i>Dénomination rapide automatisée (Plaza & Robert-Jahier, 2006).</i>	13
2.2.4	<i>Répétition de pseudo-mots (Korkman et al., 2012).</i>	13

2.2.5	<i>Désignation d'images, EVIP (Dunn et al., 1993).</i>	13
2.2.6	<i>Lecture de mots, EVALEC (Sprenger-Charolles et al., 2010).</i>	14
3	Procédure	14
4	Gestion des données	15
III	RESULTATS	15
1	Données socio-démographiques de l'échantillon	16
2	Présentation des scores aux différentes épreuves	17
2.1	Prédicteurs cognitifs et langagiers	17
2.2	Épreuve de lecture	18
3	Corrélation entre les variables étudiées	19
3.1	Corrélation des variables socio-démographiques et des scores aux épreuves pour les enfants tout-venant	19
3.2	Corrélation des variables socio-démographiques et des scores aux épreuves pour les enfants allophones	21
4	Régressions linéaires multiples	23
4.1	Régressions linéaires multiples pour l'épreuve de lecture de mots chez les enfants tout-venant	23
4.2	Régressions linéaires multiples pour l'épreuve de lecture de mots chez les enfants allophones	24
IV	DISCUSSION	25
1	Rappel des objectifs de l'étude	25
2	Interprétation et discussion des résultats	25
3	Limites, perspectives et intérêts de l'étude	28
V	CONCLUSION	31
VI	REFERENCES	32

I PARTIE THEORIQUE

1 Introduction générale

Depuis l'application des lois Ferry de 1881 et 1882, la scolarisation est gratuite et obligatoire pour tous les enfants de 3 ans à 16 ans révolus. L'accès à l'école pour les enfants est un enjeu majeur en France, qui fait face à plusieurs crises sociales et politiques importantes dont la question migratoire et de l'asile. Entre 2018 et 2019, la Division de l'Évaluation, de la Prospective et de la Performance (DEPP) recensait 67 909 Élèves Allophones Nouvellement Arrivés (EANA), soit 3 559 de plus qu'en 2017-2018. Les EANA sont des enfants arrivés en France depuis moins de deux ans et dont la langue maternelle est différente du français (Armagnague & Rigoni, 2018). Ces élèves peuvent avoir un parcours scolaire atypique : parmi eux, certains ont connu des ruptures dans leur scolarisation et d'autres n'ont jamais été scolarisés. Actuellement, la scolarisation des EANA est régie par les circulaires de 2012 qui ont restructuré d'anciens dispositifs en Unités Pédagogiques pour Élèves Allophones Arrivants (UPE2A). Scolarisés en UPE2A pour une durée officielle maximale de deux ans, ces enfants bénéficient d'un soutien linguistique tout en étant inscrits dans un cursus ordinaire afin de favoriser leur future insertion. Lors de ces années de scolarisation, l'acquisition du langage écrit sera un défi de taille pour ces élèves car elle constituera le socle de tous les futurs apprentissages scolaires. De plus, des difficultés en langage écrit peuvent compromettre l'accès à l'information, à l'emploi, aux logements ainsi qu'aux soins (Agence Nationale de Lutte Contre l'Illettrisme, 2020). Or les chiffres sont sans appel : au collège, 44% des enfants allophones sont en retard d'un an par rapport à leurs pairs (Brun, 2020). Ce retard pourrait s'expliquer, entre autres, par le défi auquel sont confrontés les enfants allophones : apprendre simultanément la langue française orale et écrite. Contrairement aux enfants allophones, les enfants monolingues ont pu développer le système phonologique de leur langue et acquérir un bon niveau de vocabulaire avant d'apprendre la langue écrite. Les recherches en psycholinguistiques ont démontré que certaines compétences cognitivo-linguistiques, parfois nommées prédicteurs, étaient nécessaires au bon développement du processus d'identification des mots écrits, et plus généralement de la lecture (Labat et al., 2013). Certaines de ces compétences s'appuient sur des connaissances langagières, ce qui pourrait expliquer en partie les difficultés rencontrées par les enfants allophones pour apprendre à lire en langue seconde. Cette étude a pour objectif d'analyser et comparer le poids des prédicteurs cognitivo-linguistiques chez les enfants monolingues et allophones. Une attention particulière sera portée au niveau de vocabulaire.

Dans les sections suivantes, nous tâcherons de définir la lecture ainsi que son développement. Puis nous identifierons précisément les différents prédicteurs de la lecture et nous définirons leur rôle. Enfin, nous nous intéresserons aux spécificités de l'apprentissage de la lecture chez les enfants allophones.

2 La lecture

Nous nous intéresserons tout d'abord aux différentes composantes de la lecture puis nous analyserons les différentes compétences qui sous-tendent le développement du langage écrit.

2.1 Le modèle simple de la lecture

Gough & Tunmer élaborent en 1986 le modèle simple de la lecture. Celle-ci est traduite par une équation à trois composantes : la compréhension écrite (lecture) est égale au produit des capacités d'identification de mots écrits (décodage) et des capacités de compréhension linguistique : $L = D \times LC$. Cela signifie que si le décodage est de mauvaise qualité, il est impossible de comprendre ce qui est lu et inversement, il est impossible de comprendre ce qui est lu sans un bon niveau en langage oral.

Le terme « décodage » est parfois restreint aux capacités de recodage phonologique. Dans le cadre de cette étude, nous l'emploierons dans le sens plus large que lui prêtent les auteurs du modèle simple de la lecture, c'est-à-dire la capacité à faire le lien entre un mot écrit et une forme phonologique, indépendamment du contexte, que ce soit par une procédure orthographique ou grâce aux règles de conversion grapho-phonémique.

La compréhension orale se réfère quant à elle aux différents niveaux de la langue : lexical, syntaxique et discursif.

La présente étude porte sur les capacités de décodage car elles constituent l'étape fondatrice de la compréhension écrite.

2.2 Les stades et les phases d'apprentissage de la lecture

Le développement du décodage a été décrit par Frith en 1986. Dans ce modèle, l'autrice suggère que le décodage suit trois stades distincts. Avant l'apprentissage formel de la lecture, l'enfant reconnaît certains mots grâce aux indices visuels des lettres, sans effectuer de correspondance entre les graphèmes et les phonèmes. A ce stade, l'enfant posséderait un vocabulaire visuel de 40 à 100 mots (Morton, 1989) ; c'est le stade logographique. L'enfant accède ensuite au stade alphabétique (ou phonologique) grâce à l'apprentissage formel de la lecture. L'apprenti lecteur décode phonologiquement un mot grâce aux règles de conversion grapho-phonémique. Les unités visuelles discrètes (graphèmes) sont mises en correspondance avec les unités phonémiques abstraites (phonèmes). Cette procédure est sous-tendue par l'enseignement de l'alphabet et le développement de la conscience phonémique. Enfin, le dernier stade est nommé stade orthographique. L'enfant réussit alors à récupérer en mémoire les mots qu'il a stockés dans son lexique mental.

Ehri (2014) met à jour le modèle de Frith et présente dans son étude plusieurs phases d'apprentissage de la lecture. Il y a tout d'abord la phase pré-alphabétique : les enfants ne connaissent pas le système alphabétique et s'appuient sur des caractéristiques visuelles ou

contextuelles saillantes pour lire les mots. Puis suit la phase alphabétique partielle lorsque les enfants apprennent les noms et les sons des lettres. Ils forment des connexions partielles en reliant les lettres les plus saillantes (par exemple la première ou la dernière lettre) aux sons dans la prononciation. Les connexions ainsi formées sont incomplètes car les enfants ne maîtrisent pas encore la conscience phonémique ainsi que les correspondances grapho-phonémiques. Lorsqu'ils sont capables d'établir des liens complets entre les graphèmes et les phonèmes pour lire des mots, il s'agit de la phase alphabétique complète. Enfin, à mesure que les mots s'accumulent en mémoire et qu'ils peuvent être lus comme des unités individuelles, il convient d'évoquer la phase alphabétique consolidée.

Ehri précise que le développement de la lecture est un processus continu et que la succession des phases ressemblerait plus à des vagues superposées qu'à des étapes distinctes. Pour donner un exemple concret, les enfants dans la phase partielle apprennent à former des connexions grapho-phonémiques pour lire des nouveaux mots courts, mais utilisent encore des indices partiels pour les mots plus longs.

Dans son étude, Ehri (2014) présente également deux stratégies utilisées par les enfants lorsqu'ils apprennent à lire des mots inconnus. Tout d'abord, ils peuvent utiliser leur connaissance du système alphabétique et appliquer les règles de conversion grapho-phonémique. Ensuite, les enfants peuvent se servir de l'analogie : ils vont faire un parallèle entre le mot qu'ils essaient de décoder et un autre mot connu à l'oral. Ce parallèle entre lecture de mot et lexique sera détaillé dans une partie dédiée. Share (1995) affirme que le recodage phonologique fonctionnerait comme un mécanisme d'auto-apprentissage et permettrait à l'enfant de développer de façon autonome son lexique orthographique. Ce sont les réussites répétées lors du décodage qui permettraient d'assimiler les connexions entre la représentation orthographique et le sens du mot.

Les deux stratégies présentées ci-dessus évoquent le modèle à double voie de Coltheart, qui décrit les processus en lecture chez le lecteur expert. Pour identifier un mot écrit, l'enfant recourt à deux voies de lecture (Coltheart et al., 2001). Tout d'abord la voie d'assemblage (ou voie phonologique) active les représentations phonologiques en mémoire grâce à l'application des règles de conversion grapho-phonémique. Cette voie permet le déchiffrement des mots nouveaux et des pseudo-mots. Puis la voie d'adressage (ou voie lexicale ou orthographique), qui permet un accès direct à la représentation orthographique du mot stocké dans le lexique mental du lecteur. Cette voie favorise la lecture de mots réguliers ou irréguliers connus. Elle est ainsi particulièrement sensible à la fréquence du mot rencontré. Il est possible d'évaluer ces deux voies de lecture par des épreuves de lecture de pseudo-mots (utilisation de la voie phonologique) et de lecture de mots. Cette dernière peut contenir des mots réguliers (qui respectent les règles de conversion grapho-phonémique) et irréguliers (qui ne respectent pas les règles de conversion). Lorsque le lecteur procède au décodage d'un

mot, les deux voies de lecture s'activent en parallèle. Selon les propriétés du mot, la réponse est récupérée plus rapidement par l'une ou l'autre voie de lecture. Lorsque le mot est inconnu et régulier, c'est la voie phonologique qui sera utilisée préférentiellement. Si le mot fait partie du stock orthographique du lecteur, alors la voie lexicale sera plus rapide pour le décoder. Pour lire des mots irréguliers, le lecteur a forcément recourt à la voie lexical et à ses connaissances orthographiques s'il veut éviter de commettre des erreurs de régularisation, c'est-à-dire prononcer le mot tel qu'il se lit phonologiquement. Lorsque l'enfant lit, ces deux niveaux de traitement sont toujours utilisés en parallèle et contribuent au décodage.

Nous avons décrit les étapes par lesquelles passe un lecteur pour acquérir des capacités de décodage. A présent, nous allons présenter plus en détail leurs soubassements cognitifs et linguistiques.

3 Les prédictors de la lecture

Les prédictors de la lecture regroupent les connaissances et les compétences précoces nécessaires au bon développement de la lecture, et plus particulièrement aux capacités de décodage. Cette section vise à décrire les principaux prédictors en décodage que sont la conscience phonologique, la mémoire phonologique à court terme, la dénomination rapide automatisée et le niveau de vocabulaire.

3.1 La conscience phonologique

La conscience phonologique se définit comme la capacité à percevoir, découper et manipuler consciemment les unités sonores du langage (phonèmes, syllabes et rimes ; Adams, 1994). La conscience phonologique évolue avant et après l'apprentissage de la lecture, c'est-à-dire que l'enfant va dans un premier temps développer une conscience implicite sur des tâches d'identification (compter, identifier ou discriminer des unités de langage) puis cette conscience deviendra explicite avec des tâches de manipulation (inverser, supprimer ou segmenter des unités langagières). Le traitement des unités linguistiques va lui aussi évoluer, passant du mot, à la syllabe, à la rime et enfin au phonème (Liberman et al., 1974).

La conscience phonologique se décompose en trois parties : la conscience phonémique (capacité à extraire les composants phonémiques d'une syllabe), la conscience syllabique (capacité à découper et manipuler les syllabes d'un mot) et la conscience de la rime. Pour tester la conscience phonémique, les tâches les plus couramment utilisées sont l'identification et/ou la suppression du phonème initial ainsi que la localisation phonémique. Pour évaluer la conscience syllabique, les tâches utilisées sont celles d'identification, d'inversion ou encore de fusion syllabique. De plus, il est important de noter que la conscience phonologique est étroitement liée aux compétences mnésiques, et plus précisément à la mémoire de travail car

l'enfant doit coupler la tâche phonologique à un travail de maintien et de manipulation de l'information verbale en mémoire (Demont & Botzung, 2003).

C'est la conscience phonémique (conscience des phonèmes) qui reflète le plus la qualité des représentations phonologiques, qui sont elles-mêmes déterminantes pour associer une lettre à un son (Melby-Lervåg et al., 2012). La conscience phonémique nécessite un apprentissage formel et se développe grâce à l'expérience de l'écrit et la rencontre avec le système d'écriture alphabétique. C'est donc lors de l'apprentissage de la lecture que la conscience des phonèmes va se développer (Morais et al., 1979). Inversement, une conscience phonémique déficitaire ne permet pas l'accès au système alphabétique ni aux règles de conversion grapho-phonémique.

Plusieurs études ont démontré que la conscience phonémique était un corrélât puissant des différences interindividuelles dans la capacité de décodage (Hulme & Snowling, 2013). Elle est reliée plus fortement à la réussite en lecture que les tâches portant sur des unités plus larges comme les syllabes (Melby-Lervåg et al., 2012). Dans une étude de comparaison entre bons et mauvais lecteurs, Catts et al. (2016) ont démontré que les apprentis mauvais lecteurs avaient généralement un déficit au niveau de la conscience phonémique.

3.2 La mémoire phonologique à court terme

La mémoire phonologique à court terme fait référence à la capacité de maintenir consciemment et activement en mémoire des informations verbales et visuo-spatiales (Majerus, 2014). C'est une composante de la mémoire de travail, définie comme un système permettant à la fois de mémoriser et de manipuler des informations à court terme (Fanuel et al., 2021). Nous utiliserons dans cet écrit les termes de « mémoire phonologique » car ils englobent les deux mémoires définies précédemment.

Les modèles actuels présentent la mémoire phonologique sous trois composantes : les interactions avec la mémoire à long terme pour stocker temporairement des informations (c'est la mémoire des représentations phonologiques et des représentations lexico-sémantiques), les processus permettant de stocker l'ordre d'apparition des informations (le traitement sériel des informations) ainsi que les mécanismes attentionnels (l'orientation et le focus attentionnel ; Majerus, 2014). C'est la composante du traitement sériel des informations qui se distingue comme le prédicteur le plus important des capacités d'apprentissage pour tous les domaines qui nécessitent le traitement et l'acquisition d'informations séquentielles, dont fait partie l'activité de lecture (Abrahamse et al., 2017).

En outre, le lien entre lecture de mots et mémoire phonologique peut s'expliquer ainsi : la mémoire phonologique permet au lecteur de coupler une lettre à un son, de déterminer la syllabe constituée par plusieurs sons, de maintenir en mémoire les différentes syllabes pour enfin les fusionner et produire oralement le mot écrit (Peng et al., 2018).

Il est possible de mesurer la mémoire phonologique grâce à des tâches de répétition de pseudo-mots car elles nécessitent la rétention d'informations phonologiques. En effet, les pseudo-mots font uniquement appel au stock phonologique de l'individu qui ne peut activer aucune représentation lexicale en mémoire. Plusieurs études ont démontré que la tâche de répétition de pseudo-mots était souvent échouée par les enfants dyslexiques (Melby-Lervåg & Lervåg, 2012).

3.3 La dénomination rapide automatisée

Lors d'une tâche de dénomination rapide automatisée, le participant doit dénommer le plus justement et rapidement possible des séries de lettres, chiffres, couleurs, objets ou symboles qui se répètent (Georgiou et al., 2013). Cette tâche met en jeu des processus similaires à ceux de la lecture (l'attention aux stimuli, les processus visuels, l'intégration des informations visuelles à des représentations stockées en mémoire, les processus lexicaux avec l'accès et la récupération des codes phonologiques ainsi que l'output d'articulation ; Araújo et al., 2011). Cependant il n'y a pas encore de consensus quant au processus qui expliquerait le mieux ce lien.

D'un côté, la dénomination rapide évalue les capacités de récupération d'une forme auditive en mémoire (tout comme les suites de lettres permettent de récupérer un mot stocké auditivement en mémoire). D'un autre côté, pour certains auteurs, ce serait plutôt la capacité à reconnaître visuellement un item qui expliquerait la relation entre la lecture et la dénomination rapide automatisée (Araújo et al., 2011). Bowers & Newby-Clark (2002) font également le lien entre la vitesse de dénomination et les représentations orthographiques. Si le traitement des lettres est ralenti, les lettres ne pourront pas être regroupées en un modèle commun qui pourra servir de base à la construction d'un nouveau modèle orthographique. En outre, Castel et al. apportent en 2008 la preuve physiologique du lien entre dénomination rapide automatisée et lecture lorsqu'ils constatent que les mêmes circuits neuronaux sont activés lors d'une tâche de lecture de mots et de dénomination rapide automatisée.

Ces différentes études tendent à se rejoindre sur un point : la performance en dénomination rapide automatisée est un puissant prédicteur de la capacité de lecture (Landerl et al., 2019). Elle permet également de distinguer de façon fiable les bons lecteurs des lecteurs dyslexiques, qui rencontrent des difficultés marquées lors de l'épreuve de dénomination rapide automatisée (Wolf et al., 2000). De plus, les performances en dénomination de stimuli alphanumériques (lettres et chiffres) seraient plus fortement corrélées aux capacités de décodage que la dénomination de couleurs ou d'objets (Norton & Wolf, 2012).

3.4 Le niveau de vocabulaire

Le niveau de vocabulaire peut être analysé sous deux angles : le vocabulaire réceptif (en compréhension) et le vocabulaire expressif (en production). Le vocabulaire en réception peut être évalué par une tâche de désignation d'images sur présentation orale d'un mot tandis que le vocabulaire en expression est généralement évalué par une tâche de dénomination orale d'images. Plusieurs études ont démontré que les connaissances en vocabulaire réceptif et expressif favorisent les compétences en décodage (Wise et al., 2007).

Notons que le mot lu peut être décrit avec trois composantes : la composante phonologique (la forme auditive du mot), la composante orthographique (la représentation orthographique du mot en mémoire) et la composante sémantique (la signification du mot ; Perfetti & Hart, 2001). Il existe un lien entre la représentation phonologique stockée (la forme auditive du mot) et sa représentation orthographique. De ce fait, les enfants ayant un niveau de vocabulaire restreint peuvent présenter des difficultés à identifier des mots écrits parce qu'ils ne disposent pas de représentations phonologiques stables et intériorisées à faire correspondre avec les mots écrits (Wise et al., 2007). En ce sens, il est plus simple d'identifier un nouveau mot écrit lorsque sa forme auditive est déjà en mémoire (Hulme et al., 2012).

Cela s'explique également par le fait que la connaissance orale d'un mot viendra soutenir le processus d'auto-correction des productions erronées (Tunmer & Chapman, 2012). Le lecteur produira différentes prononciations du mot cible jusqu'à faire correspondre une prononciation avec un mot ancré dans sa mémoire lexicale. Lorsque le lecteur est exposé à la forme phonologique d'un mot qu'il apprend à lire, cela augmente considérablement les performances en lecture de ce mot (Duff & Hulme, 2012).

Enfin, l'information sémantique peut fonctionner comme un mécanisme compensatoire lorsque l'individu ne parvient pas à retrouver la forme phonologique correcte du mot, notamment pour les mots dont les correspondances grapho-phonémiques sont incongruentes (par exemple dans le mot « yacht » ; Ricketts et al., 2007). De cette façon, il est plus facile d'apprendre à lire de nouveaux mots lorsque des informations sur le sens de ces mots sont connues (McKay et al., 2008).

Les différents prédicteurs de la lecture présentés ci-dessus jouent tous un rôle lors de l'apprentissage de la lecture et plus particulièrement lors du décodage. Il est intéressant de les évaluer le plus tôt possible afin de dépister les enfants à risque de développer un futur trouble de la lecture. Nous allons maintenant nous intéresser aux spécificités de l'apprentissage de la lecture chez les enfants allophones.

4 L'apprentissage de la lecture dans une seconde langue

Il existe peu d'études ayant concentré leurs recherches sur le développement du langage écrit chez les enfants allophones. Nous allons dès à présent nous intéresser aux particularités du développement du langage écrit dans une seconde langue.

4.1 Le transfert entre deux langues

Apprendre à lire dans une deuxième langue implique des ajustements et une adaptation aux exigences de deux systèmes langagiers différents. Geva (2009) présente deux cadres de référence concernant l'apprentissage de la lecture dans une deuxième langue.

Tout d'abord le cadre central (ou universel) indique que les mêmes habiletés de traitement (conscience phonémique, rapidité du traitement de l'information et processus visuels) sont utilisées lorsque les enfants apprennent à lire en L1 (langue première, maternelle) et en L2 (langue seconde). Ces habiletés sont donc transférées entre les deux langues et les performances de traitement dans une langue sont de ce fait reliées aux habiletés en lecture dans l'autre langue. En ce sens, des corrélations positives et significatives ont été constatées entre les habiletés de la L1 et de la L2, notamment celles appliquées à la reconnaissance des mots écrits et le décodage de pseudo-mots. De plus, les sources de différences individuelles associées au développement de la lecture dans la L1 comme la conscience phonologique, la dénomination rapide et la mémoire phonologique sont également sources de différences individuelles dans la L2.

L'autre cadre de référence prend en compte les différences typologiques entre les langues. Celles-ci peuvent tout d'abord se différencier sur la manière de coder les unités orales (système logographique, syllabique ou alphabétique). Elles peuvent également présenter des différences en termes de régularité dans les règles de correspondance grapho-phonémique. Par exemple, si la L2 possède un système d'écriture plus régulier que la L1, alors les enfants réussiront mieux à lire les mots dans la L2 que dans la L1. Pour les enfants qui rencontrent des difficultés en lecture et qui sont confrontés à différents systèmes orthographiques, ils auront des résultats communs aux deux langues (par exemple lors du décodage de pseudo-mots) mais également des résultats différents, en lien avec la typologie de la langue et le système d'écriture. Le développement du langage écrit dans une seconde langue implique donc l'apprentissage de nouvelles règles grapho-phonologiques ainsi qu'une adaptation de la part de l'enfant à un nouveau système d'écriture potentiel.

4.2 L'importance du niveau de vocabulaire lors de l'apprentissage de la lecture dans une deuxième langue

Selon Yeung et al. (2013), le vocabulaire expressif en L1 serait directement lié à la lecture de mots en L1 et en L2 au cours des premières années. Les travaux de Liu et al (2017) viennent confirmer l'importance du vocabulaire lors d'une tâche de décodage. Ils ont démontré

dans une étude longitudinale portant sur des enfants chinois de Hong Kong apprenant l'anglais en langue seconde que lorsque l'âge, le QI non verbal, la conscience phonologique de l'anglais, la connaissance des lettres et la lecture des caractères chinois étaient contrôlés, le niveau initial de vocabulaire expressif en chinois permettait de prédire la lecture de mots en anglais 15 mois plus tard. De plus, le taux de croissance du vocabulaire expressif au cours de ces 15 mois prédisait les performances en lecture de mots en anglais.

Les études présentées précédemment apportent des informations concernant les enfants bilingues, qui apprennent à lire dans une langue différente de leur langue maternelle. Les enfants allophones récemment arrivés en France se trouvent également dans la situation où ils ne maîtrisent pas la langue orale de leur pays d'accueil. Ils ne peuvent donc pas s'appuyer sur celle-ci lors du processus de décodage. Concernant spécifiquement le niveau de vocabulaire, même si l'écart se réduit au fil du temps, les compétences en vocabulaire des enfants allophones dans la langue française restent très en dessous de celles de leurs pairs monolingues (Farnia & Geva, 2011). Cela s'explique par le fait que les enfants allophones disposent de peu de représentations phonologiques stockées en mémoire, conséquence directe du manque d'exposition à la langue française. Ils se retrouvent donc en difficulté lorsqu'il s'agit de faire des liens avec les représentations orthographiques des mots. Le processus d'auto-correction des productions se retrouve également affaibli par le manque de connaissance orale des mots.

5 Buts et hypothèses

Cette étude a pour objectif d'explorer le développement des compétences en décodage chez les enfants allophones scolarisés en écoles élémentaires françaises grâce à l'analyse de l'impact des différents prédicteurs de la lecture. Plus spécifiquement, elle vise à évaluer le poids de ces différents prédicteurs sur l'acquisition du décodage, le but étant d'identifier ceux qui ont le plus d'importance pour les enfants allophones.

Nous nous attarderons spécifiquement sur le lien entre le niveau de vocabulaire et les compétences en décodage chez les enfants allophones. Contrairement aux enfants tout-venant qui disposent d'un niveau de vocabulaire suffisant lorsqu'ils apprennent à lire et pour lesquels les différences inter-individuelles en terme de vocabulaire ont peu d'importance, les enfants allophones n'ont pas acquis le vocabulaire de base et présentent des difficultés à reconnaître des mots écrits très simples.

Nous émettons l'hypothèse que les différences inter-individuelles en vocabulaire chez les enfants allophones auront plus d'impact sur la réussite en lecture que chez les enfants monolingues. En d'autres termes, le niveau de vocabulaire serait plus prédictif de la réussite en lecture de mots chez les enfants allophones que chez les enfants monolingues.

Cette question de recherche est née du constat qu'il existe très peu d'études concernant les enfants allophones, qui sont pourtant de plus en plus nombreux en France. Étant donné leur faible temps d'exposition à la langue d'apprentissage, les outils en matière d'évaluation langagière sont peu voire pas adaptés. Travailler sur les précurseurs de la lecture permettrait également d'identifier des profils d'enfants ayant des compétences faibles dans les acquisitions clés préalables à la lecture. Un repérage précoce de ces enfants à risque de développer des futures difficultés en lecture permettra une prise en soin orthophonique anticipée et plus efficiente. Cette étude pourra également servir d'outil pour orienter les pédagogies des enseignants UPE2A sur les apprentissages les plus déterminants dans le développement de la lecture chez les enfants allophones.

II METHODE

Ce projet de recherche est inclus au sein du projet TANMALL (Tâche d'Apprentissage de Nouveaux Mots chez des enfants ALLOphones), mené par Monsieur Bignon, doctorant en psychologie cognitive. Il s'agit d'une étude transversale observationnelle multicentrique qui a reçu un avis favorable de la part du comité d'éthique de Lille (référence : 2021-466-S90).

1 Population

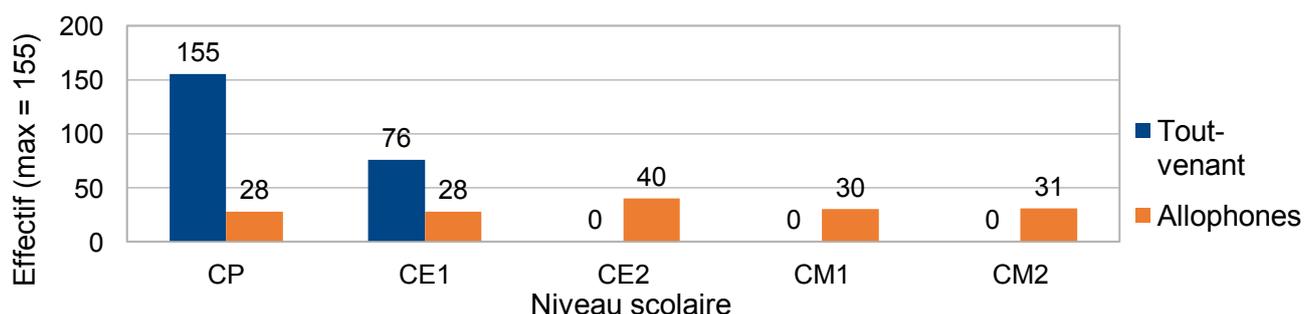
Dans cette étude, nous avons inclus 157 enfants allophones suivis en UPE2A (Unité Pédagogique pour Élèves Allophones Arrivants) et scolarisés depuis plus de 6 mois et moins de 3 ans en France. Une durée minimale de scolarisation a été imposée afin de nous assurer que ces enfants aient acquis les compétences minimales en français pour communiquer avec les examinateurs. Puisque les enfants allophones de niveau CP étaient trop peu nombreux, nous avons sélectionné des enfants allophones qui avaient commencé à apprendre à lire même s'ils étaient scolarisés dans des classes de niveaux plus élevés.

Un effectif de 231 enfants tout-venant monolingues de niveau CP et CE1 a également été inclus dans l'étude afin de constituer un groupe contrôle.

La figure 1 détaille la répartition des participants par niveau scolaire et par groupe.

Figure 1.

Répartition des participants par niveau scolaire et par groupe (tout-venant et allophones)



Nous avons fait le choix ne pas sélectionner les enfants allophones en fonction de leurs langues maternelles, mais celles-ci sont tout de même répertoriées en annexe A (langues maternelles principales) et en annexe B (langues maternelles secondaires).

2 Matériel

2.1 A l'attention des familles : courriers d'information, formulaires de consentement et questionnaires

Des courriers d'information (cf. annexe C) et des formulaires de consentement (cf. annexe D) ont été adressés aux parents ou aux tuteurs légaux des enfants. Dans le même temps, des questionnaires (cf. annexe E) ont été adressés aux parents ou tuteurs légaux des enfants afin de recueillir des informations complémentaires essentielles à l'analyse des données telles que :

- La date de naissance de l'enfant
- Le sexe de l'enfant
- La date approximative d'arrivée en France et de première scolarisation en France, le niveau de classe
- Des informations concernant la scolarisation antérieure (niveau atteint, apprentissage ou non de la lecture)
- Le développement langagier (âge des premiers mots)
- Les antécédents familiaux de difficultés langagières
- La ou les langues parlée(s) à la maison
- Les performances générales des enfants et des parents dans ces langues
- Le niveau socio-culturel : nombre d'années d'études et dernier diplôme obtenu par les parents

Tous les documents ont été donnés en même temps aux familles pour éviter d'avoir à les solliciter à plusieurs reprises. Ce sont les enseignants qui se sont chargés de fournir en mains propres les documents aux familles afin de limiter les difficultés liées à la lecture et à la langue. Lors du remplissage des questionnaires, ce sont les enseignants, en tant qu'interlocuteurs privilégiés, qui ont aidé les familles à répondre aux questions. Ils ont été prévenus au préalable et ont accepté de prendre ce temps pour participer au bon déroulement de l'étude.

2.2 A l'attention des enfants : les épreuves

La passation pour un élève comprenait neuf épreuves, mais seulement six seront analysées dans le cadre de cette étude sur les prédicteurs du langage écrit : les matrices progressives de Raven (Raven et al., 1998), la tâche d'analyse phonologique de l'EVALEC (Sprenger-Charolles et al., 2010), la tâche de dénomination rapide automatisée (Plaza & Robert-Jahier, 2006), la tâche de répétition de pseudo-mots (Korkman et al., 2012), la tâche

de désignation d'images de l'EVIP (Dunn et al., 1993) ainsi que la tâche de lecture de mots de l'EVALEC (Sprenger-Charolles et al., 2010). Les sections suivantes ont pour but de détailler chaque épreuve.

2.2.1 Matrices progressives de Raven : les habiletés non-verbales (Raven et al., 1998).

Cette épreuve permettait de contrôler les habiletés non-verbales des enfants. Lors de cette tâche, les élèves devaient compléter des systèmes d'images abstraits auxquels il manquait un morceau. Ils sélectionnaient alors une proposition parmi six en s'appuyant sur des indices visuo-spatiaux.

L'examineur inscrivait sur le cahier de passation les réponses de l'enfant. La notation était de 1 point par item réussi, donc le score maximal pour cette épreuve était de 36 points. Le temps indicatif pour cette épreuve était de 5 minutes.

2.2.2 EVALEC : analyse phonologique (Sprenger-Charolles et al., 2010).

Cette épreuve permettait d'évaluer la conscience phonologique de l'enfant et ses compétences en matière de suppression syllabique et phonémique. Concrètement, trois tâches ont été proposées à l'enfant :

- Une épreuve d'élimination de syllabe initiale : l'enfant devait retirer la première syllabe d'un pseudo-mot trisyllabique. Cette épreuve comprenait 2 essais et 10 items. La notation était de 1 point par item correctement produit donc le score maximal était de 10 points.
- Une épreuve d'élimination de phonème initial : l'enfant devait retirer le premier phonème d'une syllabe de type consonne/voyelle/consonne. Cette épreuve comprenait 2 essais et 12 items. La notation était de 1 point par item correctement produit donc le score maximal était de 12 points.
- Une épreuve d'élimination de phonème final : l'enfant devait retirer le dernier phonème d'une syllabe de type consonne/consonne/voyelle. Cette épreuve comprenait 2 essais et 12 items. La notation était de 1 point par item correctement produit donc le score maximal était de 12 points.

Avant de commencer chaque tâche, l'examineur présentait à l'enfant des exemples à l'aide de jetons afin de s'assurer de la bonne compréhension de la consigne. Dans un premier temps, l'examineur manipulait les jetons et oralisait la suppression syllabique ou phonémique puis demandait à l'enfant de faire de même avec un deuxième exemple. Enfin, l'examineur retirait les jetons et proposait à l'enfant un troisième exemple. En plus de ces exemples, 2 items d'essais étaient proposés au début de chaque tâche.

Les réponses de l'enfant étaient enregistrées. La notation était informatique et manuelle : l'ordinateur enregistrait les scores temps pour chaque item et l'examineur retranscrivait les

réponses de l'enfant sur le fichier de passation. Un point était accordé par réponse correcte et les auto-corrrections étaient acceptées. Le score maximal pour cette épreuve était de 34 points. Le temps indicatif total de l'épreuve était de 5 minutes.

2.2.3 Dénomination rapide automatisée (Plaza & Robert-Jahier, 2006).

Pour cette épreuve, il s'agissait de dénommer le plus rapidement possible des items présentés en lignes. Deux conditions étaient présentées à l'enfant : 4 chiffres (1, 2, 7, 9) et 4 lettres (A, B, C, U). Pour chaque condition, les stimuli étaient répétés dans un ordre aléatoire sur 8 lignes avec 6 items par ligne.

Un entraînement sur les 4 items était proposé par l'examineur avant chaque tâche afin de s'assurer que les enfants étaient capables de les dénommer. L'épreuve était enregistrée et chronométrée. La notation correspondait au temps mis par l'enfant pour lire tous les chiffres puis toutes les lettres. Le temps indicatif pour cette épreuve était de 3 minutes.

2.2.4 Répétition de pseudo-mots (Korkman et al., 2012).

Cette épreuve permettait de tester la mémoire à court terme de l'enfant qui devait répéter des pseudo-mots énoncés par l'ordinateur. L'épreuve comprenait 31 items de longueur croissante. Les réponses de l'enfant étaient enregistrées puis retranscrites sur le fichier de passation par l'examineur. Les items ont été créés à partir de sons et de structures syllabiques presque universels afin d'éviter toute difficulté liée à la connaissance des phonèmes spécifiques au français.

Un point était accordé par pseudo-mots correctement répété et les auto-corrrections étaient acceptées. Le score maximal pour cette épreuve était de 31 points. Le temps indicatif pour cette épreuve était de 2 minutes.

2.2.5 EVIP : désignation d'images et lexique passif (Dunn et al., 1993).

Cette épreuve permettait de tester le stock lexical passif de l'enfant. Classiquement, l'EVIP est composée de 170 items classés par difficulté croissante, mais seulement 50 items ont été retenus pour cette étude afin de limiter le temps de passation. Le temps indicatif pour cette épreuve était de 5 minutes.

L'item était dénommé oralement par l'ordinateur et l'enfant devait alors désigner une image parmi quatre présentées sur l'écran. Pour ce faire, des étiquettes numérotées de 1 à 4 ont été collées sur le clavier d'ordinateur et l'enfant devait appuyer sur la touche correspondant au numéro de l'image qu'il souhaitait désigner. Le temps de réponse était limité à 10 secondes pour chaque item. L'épreuve s'arrêtait automatiquement lorsque 5 erreurs consécutives étaient produites et les réponses suivantes étaient considérées comme fausses. La notation

était réalisée par le logiciel. La notation était de 1 point par item réussi, donc le score maximal pour cette épreuve était de 50 points.

2.2.6 EVALEC : lecture de mots (Sprengr-Charolles et al., 2010).

Cette épreuve permettait d'évaluer les procédures de décodage. L'enfant devait lire à voix haute 48 mots de complexité croissante (mots réguliers, digraphes, mots comprenant des graphies contextuelles et mots irréguliers) qui apparaissaient successivement sur l'écran d'ordinateur. Il n'y avait pas de contrainte de temps pour cette épreuve.

La notation était manuelle car l'examineur retranscrivait les réponses de l'enfant sur le fichier de passation. Un point était attribué par réponse correcte, les auto-corrrections étaient acceptées et signalées sur le fichier de passation. Le score maximal pour cette épreuve était de 48 points. Le temps indicatif de l'épreuve était de 10 minutes.

Après avoir détaillé les différentes épreuves sélectionnées dans le cadre de cette étude, nous allons décrire la procédure employée pour effectuer les passations.

3 Procédure

Dans un premier temps, les Inspecteurs de l'Éducation Nationale de dix circonscriptions lyonnaises ont été contactés afin d'obtenir leur accord pour effectuer l'étude sur leur territoire. L'expérimentation ne pouvait avoir lieu sans la signature d'une convention (cf. annexe F) et d'une demande d'autorisation (cf. annexe G) pour chaque circonscription. Lorsque le dossier était complété et signé par toutes les parties, il convenait de prendre contact avec les enseignants UPE2A de la circonscription concernée pour leur présenter l'étude et leur demander s'ils souhaitaient y participer. Les enseignants qui acceptaient recevaient alors par voie postale les lettres d'information, les formulaires de consentement ainsi que les questionnaires à faire remplir aux familles ou aux tuteurs légaux des enfants allophones. Lorsque les documents étaient signés, l'examineur et l'enseignant choisissaient une ou plusieurs dates pour effectuer les passations sur le temps scolaire. Il était important que l'examineur et l'enfant se trouvent dans une salle calme et isolée. Deux passations de 30 minutes étaient nécessaires pour faire passer l'ensemble du test. Afin de remercier les enfants pour leur participation, un autocollant leur était offert. De plus, une fois l'étude terminée, les enseignants ayant participé au projet bénéficieront d'une présentation des résultats de l'étude.

Concernant les épreuves, l'enfant a réalisé dans l'ordre suivant l'épreuve des matrices progressives de Raven, la tâche de désignation d'images, l'épreuve de dénomination rapide automatisée, la tâche d'analyse phonologique, la tâche de répétition de pseudo-mots puis la lecture de mots.

4 Gestion des données

Les différentes épreuves ont été administrées grâce au logiciel PsychoPy (Peirce et al., 2019) sur l'ordinateur. Ce programme lançait les épreuves et enregistrait les réponses des élèves en temps réel.

Un code d'anonymisation a été attribué à tous les élèves. Ce code a été reporté sur le fichier de passation et avant le début de chaque épreuve sur le logiciel PsychoPy. Les participants étaient anonymisés selon la procédure suivante : numéro affecté au terrain d'expérimentation, âge en mois à la date de création du code et numéro aléatoire entre 0 et 9 pour distinguer les enfants qui pourraient avoir un code équivalent.

Pour certaines tâches, un dictaphone était utilisé pour enregistrer les productions orales des enfants et faciliter la future transcription des réponses. Ces enregistrements ont été anonymisés selon le code de l'enfant concerné et sauvegardés le jour-même dans un fichier crypté via le logiciel VeraCrypt. Les enregistrements étaient ensuite supprimés du dictaphone avant de quitter le terrain d'expérimentation.

Toutes les réponses des élèves ont ensuite été reportées et codées sur un tableur par l'examineur. Pour les épreuves de conscience phonologique, de répétition de pseudo-mots et de lecture de mots, une cotation souple a été pratiquée. En effet, certains contrastes phonologiques pouvaient poser des difficultés durables aux enfants allophones et ne reflétaient pas leurs capacités de recodage phonologique. Certaines déviations phonologiques ont alors été acceptées (la liste des déformations phonologiques autorisées se trouve en annexe H). Cette cotation permettait d'analyser uniquement les compétences à évaluer et diminuait dans le même temps la marge d'erreur due aux perceptions auditives de l'examineur.

III RESULTATS

Dans un premier temps nous présenterons les caractéristiques de notre échantillon. Les données concernant l'âge et la scolarisation des deux groupes d'enfants (tout-venant et allophones) ont été recueillies grâce aux questionnaires remplis par les parents ou les tuteurs légaux des enfants. Puis nous détaillerons les performances des deux groupes aux différentes épreuves. Dans un troisième temps, nous nous intéresserons aux analyses corrélationnelles que nous avons effectuées afin d'analyser quels sont les prédicteurs corrélés à la lecture de mots. Enfin, nous réaliserons des régressions linéaires multiples en intégrant dans le modèle les variables corrélées de façon significative avec la lecture de mots afin d'obtenir la contribution unique de chaque variable sur les compétences en décodage.

1 Données socio-démographiques de l'échantillon

Pour chaque variable, nous avons calculé la moyenne et l'écart-type ainsi que reporté la donnée minimale et maximale de chaque groupe. Pour comparer les deux groupes, un test t de Welch a été effectué afin de tester l'hypothèse d'égalité des moyennes des deux groupes. Enfin, le d de Cohen permet de calculer la taille de l'effet.

Les deux groupes que nous cherchons à comparer sont composés de 231 enfants pour le groupe « tout-venant » et de 157 enfants pour le groupe « allophones » (cf. tableau 1.) Nous relevons une différence significative entre les deux groupes lorsque nous comparons les moyennes d'âge en mois lors de la passation ($p < .001$, d de Cohen = -1.82). La moyenne des enfants tout-venant est significativement inférieure à celle des enfants allophones ($M = 86.52$ pour les enfants tout-venant et $M = 113.11$ pour les enfants allophones ; cf. tableau 1.). Ces données sont congruentes avec le niveau scolaire des enfants tout-venant ($M = 1.33$, $ET = 0.47$) qui est inférieur à celui des enfants allophones ($M = 3.05$, $ET = 1.37$). Notons que l'âge moyen lors de la scolarisation en France pour les enfants allophones est de 98.42 mois ($ET = 20.81$; cf. tableau 1.). La figure 2 rend compte du niveau scolaire de première scolarisation en France des enfants allophones. Nous remarquons que la majorité des enfants allophones ont été scolarisés pour la première fois en classe de CE1, CE2, CM1 et dans une plus petite proportion en classe de CP (cf. figure 2). C'est pourquoi nous comparons les deux groupes non pas en fonction de leur niveau scolaire mais plutôt à partir du moment où ils entrent dans l'apprentissage de la lecture.

Tableau 1.

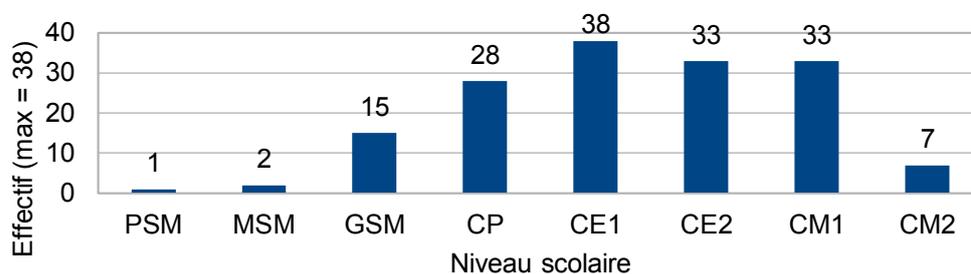
Présentation des données concernant l'âge et la scolarisation de l'échantillon

Variables	Tout-venant (N=231)		Allophones (N=157)		Différences entre les groupes			
	Moyenne (écart-type)	Min-Max	Moyenne (écart-type)	Min-Max	t*	df	p	d de Cohen
Âge (lors de la passation)	86.52 (6.33)	77-110	113.11 (19.69)	75-191	-16.36	178.07	<.001	-1.82
Nombre de mois en primaire (exposition explicite à l'écrit)	12.22 (5.13)	5-31	13.34 (6.88)	6-31	-1.74	270.04	0.08	-0.18
Niveau scolaire numérique	1.33 (0.47)	1-2	3.05 (1.37)	1-5	–	–	–	–
Temps de séjour en France	–	–	16.24 (8.06)	6-38	–	–	–	–
Temps de scolarisation en France	–	–	14.68 (7.26)	6-37	–	–	–	–
Âge lors de la scolarisation en France	–	–	98.42 (20.81)	47-181	–	–	–	–

*test t de Wesch car les variances n'étaient pas homogènes

Figure 2.

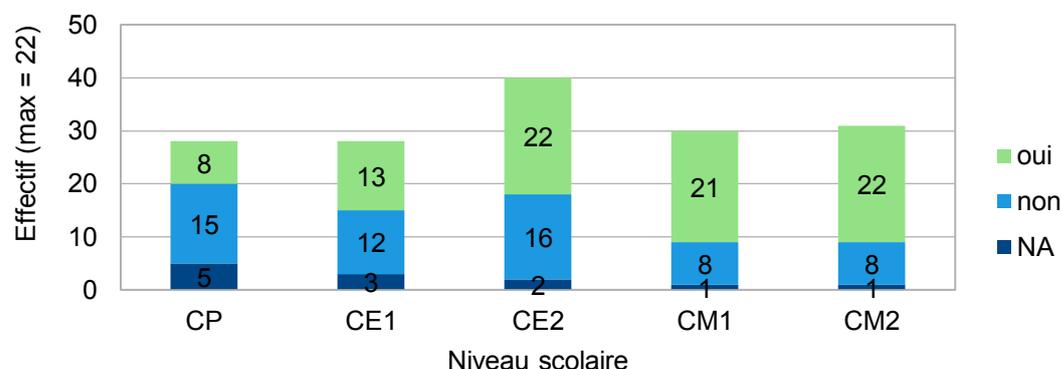
Répartition des enfants allophones par niveau scolaire de première scolarisation en France



Nous remarquons que les groupes ne se différencient pas sur le nombre de mois de scolarisation en primaire, qui représente approximativement une durée d'exposition explicite à l'écrit ($p = 0.08$; cf. tableau 1.). Si l'on s'intéresse plus particulièrement à l'apprentissage du langage écrit, sur 28 élèves scolarisés en classe de CP, 15 n'ont pas commencé à apprendre à lire avant d'arriver en France contre 8 ayant déjà commencé l'apprentissage de la lecture. A partir du CE1, plus de la moitié des enfants allophones ont au moins commencé à apprendre à lire avant d'arriver en France (cf. figure 3.).

Figure 3.

Répartition des enfants allophones par niveau scolaire ayant au moins commencé à apprendre à lire avant l'arrivée en France



2 Présentation des scores aux différentes épreuves

2.1 Prédicteurs cognitifs et langagiers

Le tableau 2 ci-dessous présente les résultats des groupes d'enfants tout-venant et allophones aux épreuves des matrices de Raven, de conscience phonologique, de dénomination rapide automatisée, de répétition de pseudo-mots et de vocabulaire.

Le groupe des enfants tout-venant obtient des performances supérieures à celui des enfants allophones sur l'ensemble des épreuves. C'est pour l'épreuve de vocabulaire que la taille de

l'effet est la plus importantes ($p < .001$, d de Cohen = 1.80 ; cf. tableau 2.). Le groupe des enfants allophones présente donc des performances nettement inférieures à celles des enfants tout-venant pour cette tâche. Notons que les deux groupes ne se différencient pas sur l'épreuve de dénomination rapide automatisée ($p = 0.74$; cf. tableau 2.).

Lorsque l'on compare les deux cotations (stricte et souple) pour les tâches de conscience phonologique et de répétition de pseudo-mots, la taille de l'effet (d de Cohen) diminue. Celle-ci passe de 0.64 (cotation stricte) à 0.51 (cotation souple) pour l'épreuve de conscience phonologique et de 0.56 (cotation stricte) à 0.22 (cotation souple) pour l'épreuve de répétition de pseudo-mots (cf. tableau 2.).

Tableau 2.

Scores aux différentes épreuves des prédicteurs de la lecture et comparaison entre les groupes

Variables	Tout-venant (N=231)		Allophones (N=157)		Différences entre les groupes			
	Moyenne (écart-type)	Min-Max	Moyenne (écart-type)	Min-Max	t*	df	p	d de Cohen
Matrices de Raven	24.76 (5.33)	7-36	22.97 (6.18)	6-34	2.95	301.30	<.001	0.31
Conscience phonologique (stricte)	26.50 (5.64)	3-34	21.80 (8.68)	0-34	5.98	244.60	<.001	0.64
Conscience phonologique (souple)	26.73 (5.59)	4-34	22.94 (8.78)	0-34	4.79	241.33	<.001	0.51
Dénomination rapide (temps moyen en sec.)	32.97 (8.06)	18,5-79	33.31 (10.91)	17,5-82	-0.34	268.14	0.74	-0.04
Répétition de pseudo-mots (stricte)	26.59 (3.10)	13-31	24.22 (5.08)	4-31	5.22	234.90	<.001	0.56
Répétition de pseudo-mots (souple)	26.94 (3.01)	14-31	26.15 (4.18)	4-31	2.04	262.93	0.04	0.22
Vocabulaire	30.90 (6.41)	13-46	18.09 (7.76)	0-35	17.11	291.54	<.001	1.80

* test t de Welch car les variances n'étaient pas homogènes.

2.2 Épreuve de lecture

Pour la tâche de lecture de mots, les performances des enfants tout-venant sont significativement supérieures aux enfants allophones ($M = 39.69$ pour les enfants tout-venant et $M = 25.45$ pour les enfants allophones en cotation stricte). La taille de l'effet est très importante ($p < .001$, d de Cohen supérieur à 1). Lorsque la cotation est souple, la taille de l'effet diminue, passant de 1.42 à 1.25 (cf. tableau 3.).

Les écarts de score sont toujours plus importants pour le groupe des enfants allophones : le minimum de 0 est atteint, l'écart-type est de 12.20 lorsque la cotation est stricte et de 12.02

lorsque la cotation est souple. Le score maximal des enfants tout-venant n'est jamais atteint par les enfants allophones (cf. tableau 3.).

Tableau 3.

Scores à l'épreuve de lecture de mots et comparaison entre les groupes

Variables	Tout-venant (N=231)		Allophones (N=157)		Différences entre les groupes			
	Moyenne (écart-type)	Min-Max	Moyenne (écart-type)	Min-Max	t*	df	p	d de Cohen
Lecture de mots (stricte)	39.69 (7.21)	11-48	25.45 (12.20)	0-46	13.15	230.03	<.001	1.42
Lecture de mots (souple)	40.54 (6.74)	14-48	28.34 (12.02)	0-47	11.55	222.90	<.001	1.25

* test t de Welch car les variances n'étaient pas homogènes.

Nous avons présenté les différents résultats obtenus par les deux groupes d'enfants lors des différentes épreuves. Nous allons maintenant observer les relations entre les prédicteurs et les performances en lecture.

3 Corrélations entre les variables étudiées

Pour tester l'existence d'une relation linéaire entre deux variables continues, nous avons réalisé des corrélations de Pearson. Selon Winter et al. (2016), lorsque le coefficient de corrélation (r) est égal à 0, cela montre une corrélation nulle. Si $r = 0.4$ la corrélation est forte et si $r = 0.8$ la corrélation entre deux variables est très forte.

Nous n'avons conservé que les scores avec la cotation souple pour nous assurer que les relations observées entre les variables ne tenaient pas compte des déformations phonologiques liées aux langues maternelles des enfants allophones.

Nous allons dans un premier temps analyser la matrice de corrélation pour les enfants tout-venant puis nous analyserons la matrice de corrélation des enfants allophones. Nous tâcherons d'explicitier les corrélations existantes entre l'épreuve de lecture de mots et celles évaluant les différents prédicteurs cognitifs et linguistiques. Puis nous procéderons de la même façon avec l'épreuve de vocabulaire.

3.1 Corrélations des variables socio-démographiques et des scores aux épreuves pour les enfants tout-venant

Nous remarquons dans un premier temps que la lecture de mots est significativement et fortement corrélée aux épreuves de conscience phonologique ($r = 0.54$, $p < .001$), de

répétition de pseudo-mots ($r = 0.45, p < .001$) et de dénomination rapide automatisée ($r = -0.44, p < .001$). Elle est corrélée plus faiblement aux matrices de Raven ($r = 0.37, p < .001$) et à l'épreuve de vocabulaire ($r = 0.35, p < .001$; cf. tableau 4.).

Concernant les variables socio-démographiques, l'épreuve de lecture de mots est très faiblement corrélée à l'âge lors de la passation ($r = 0.16, p < .05$), au nombre de mois d'exposition explicite à l'écrit ($r = 0.17, p < .01$) et au niveau scolaire ($r = 0.14, p < .05$; cf. tableau 4.).

Lorsque nous analysons l'épreuve de vocabulaire, nous constatons qu'elle est significativement et faiblement corrélée aux matrices de Raven ($r = 0.35, p < .001$), à l'épreuve de conscience phonologique ($r = 0.29, p < .001$) et à l'épreuve de répétition de pseudo-mots ($r = 0.20, p < .01$). L'épreuve de vocabulaire est très faiblement corrélée à la tâche de dénomination rapide automatisée ($r = -0.15, p < .05$; cf. tableau 4.).

Tableau 4.

Matrice de corrélation des variables socio-démographiques et des scores aux différentes épreuves du groupe d'enfants tout-venant

	1	2	3	4	5	6	7	8
1. Âge (lors de la passation)	—							
2. Nombre de mois en primaire (exposition explicite à l'écrit)	0.83***	—						
3. Niveau scolaire numérique	0.78***	0.94***	—					
4. Matrices de Raven	0.17**	0.12	0.09	—				
5. Conscience phonologique (souple)	-0.03	-0.07	-0.06	0.39***	—			
6. Dénomination rapide (temps en sec.)	-0.22***	-0.18**	-0.21**	-0.22***	-0.27***	—		
7. Répétition de pseudo-mots (souple)	0.07	0.06	0.06	0.16*	0.39***	-0.30***	—	
8. Vocabulaire	0.11	0.03	-0.01	0.35***	0.29***	-0.15*	0.20**	—
9. Lecture de mots (souple)	0.16*	0.17**	0.14*	0.37***	0.54***	-0.44***	0.45***	0.36***

* $p < .05$, ** $p < .01$, *** $p < .001$

3.2 Corrélations des variables socio-démographiques et des scores aux épreuves pour les enfants allophones

Tout d'abord, nous remarquons que chez les enfants allophones, l'épreuve de lecture de mots est significativement et fortement corrélée à toutes les épreuves évaluant les prédicteurs cognitifs et linguistiques de la lecture, c'est-à-dire la conscience phonologique ($r = 0.71, p < .001$), la dénomination rapide automatisée ($r = 0.67, p < .001$), les matrices de Raven ($r = 0.55, p < .001$), la répétition de pseudo-mots ($r = 0.44, p < .001$) et le vocabulaire ($r = 0.40, p < .001$; cf. tableau 5.).

Si l'on s'intéresse aux variables socio-démographiques, l'épreuve de lecture de mots est significativement et fortement corrélée au niveau scolaire ($r = 0.45, p < .001$) et au fait d'avoir appris ou non à lire antérieurement ($r = 0.41, p < .001$). Elle est faiblement corrélée au fait d'avoir été scolarisé antérieurement ($r = 0.35, p < .001$), à l'âge lors de la passation ($r = 0.24, p < .01$) et à l'âge lors de la scolarisation en France ($r = 0.23, p < .01$; cf. tableau 5.).

Concernant l'épreuve de vocabulaire, elle est significativement et faiblement corrélée aux matrices de Raven ($r = 0.29, p < .001$), à l'épreuve de dénomination rapide automatisée ($r = 0.25, p < .01$) et à la conscience phonologique ($r = 0.22, p < .01$; cf. tableau 5.).

Si l'on s'intéresse aux variables socio-démographiques, l'épreuve de vocabulaire est significativement et faiblement corrélée au nombre de mois d'exposition explicite à l'écrit ($r = 0.30, p < .001$), au temps de scolarisation en France ($r = 0.24, p < .01$), au temps de séjour en France ($r = 0.22, p < .01$), au niveau scolaire ($r = 0.22, p < .01$) et à l'âge lors de la passation ($r = 0.19, p < .05$; cf. tableau 5.).

Nous avons remarqué pour les deux groupes d'enfants que tous les prédicteurs de la lecture testés étaient corrélés significativement à la lecture de mots et qu'ils étaient également tous corrélés à l'épreuve de vocabulaire, excepté la mémoire phonologique testée avec l'épreuve de répétition de pseudo-mots chez les enfants allophones. Les différents prédicteurs peuvent donc potentiellement représenter des facteurs de confusion si nous nous intéressons au lien entre vocabulaire et lecture de mots. Nous allons donc effectuer des régressions linéaires multiples afin de tester la contribution unique de chacun de ces facteurs sur les performances en lecture de mots.

Tableau 5.

Matrice de corrélation des variables socio-démographiques et des scores aux différentes épreuves du groupe d'enfants allophones

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Âge (lors de la passation)	—												
2. Nombre de mois en primaire (exposition explicite à l'écrit)	0.26**	—											
3. Temps de scolarisation en France	0.03	0.83***	—										
4. Temps de séjour en France	0.04	0.76***	0.93***	—									
5. Âge lors de la scolarisation en France	0.93***	0.18*	-0.03	-0.01	—								
6. Niveau scolaire numérique	0.70***	0.31***	0.05	0.07	0.60***	—							
7. Scolarisation antérieure	0.24**	-0.03	-0.22**	-0.28***	0.22**	0.32***	—						
8. Apprentissage de la lecture antérieur à l'arrivée en France	0.15	-0.03	-0.14	-0.11	0.18*	0.26**	0.50***	—					
9. Matrices de Raven	0.12	0.11	0.01	0.005	0.09	0.32***	0.32***	0.41***	—				
10. Conscience phonologique (souple)	0.14	0.10	0.02	0.02	0.13	0.35***	0.39***	0.36***	0.46***	—			
11. Dénomination rapide (temps en sec.)	-0.34***	-0.15	0.09	0.12	-0.32***	-0.52***	-0.45***	-0.33***	-0.38***	-0.54***	—		
12. Répétition de pseudo-mots (souple)	0.17*	0.11	-0.003	-0.03	0.18*	0.32***	0.42***	0.32***	0.35***	0.54***	-0.32***	—	
13. Vocabulaire	0.19*	0.30***	0.24**	0.22**	0.13	0.22**	0.05	0.12	0.29***	0.22**	-0.25**	0.16	—
14. Lecture de mots (souple)	0.24**	0.15	0.04	0.04	0.23**	0.45***	0.35***	0.41***	0.55***	0.71***	-0.67***	0.44***	0.40***

* p <.05, ** p <.01, *** p <. 001

4 Régressions linéaires multiples

Pour pouvoir comparer les deux modèles, les données socio-démographiques propres aux enfants allophones (temps de séjour en France, temps de scolarisation en France, âge lors de la scolarisation en France, scolarisation antérieure et apprentissage de la lecture antérieure à l'arrivée en France) ne sont pas prises en compte.

4.1 Régressions linéaires multiples pour l'épreuve de lecture de mots chez les enfants tout-venant.

Le modèle des enfants tout-venant (cf. tableau 6.) explique 49% de variance en lecture de mots. Les différents prédicteurs testés présentent tous une contribution unique et significative à la réussite en lecture de mots sauf l'épreuve des matrices de Raven. Par ordre d'importance, les facteurs explicatifs sont la conscience phonologique ($p < .001$, $\Delta R^2 = 7.84\%$), la dénomination rapide automatisée ($p < .001$, $\Delta R^2 = 4.41\%$), la répétition de pseudo-mots ($p < .001$, $\Delta R^2 = 3.61\%$), le vocabulaire ($p < .001$, $\Delta R^2 = 1.69\%$) et l'âge lors de la passation ($p = 0.02$, $\Delta R^2 = 1.21\%$). Nous observons que le vocabulaire occupe une place marginale dans la réussite en lecture de mots pour le groupe d'enfants tout-venant (cf. tableau 6.).

Tableau 6.

Modèle de régressions linéaires multiples des performances en lecture de mots chez les enfants tout-venant

$$F(7,222) = 29.95, p < .001, R^2 = .49$$

Variables	NS	95 % CI		S	t	p	ΔR^2
		Bas	Haut				
(Intercept)	21.22	5.42	37.01		2.65	8.70e-3	
Nombre de mois en primaire (exposition explicite à l'écrit)	-0.10	-0.28	0.08	-0.10	-1.09	0.28	0.25%
Âge (lors de la passation)	0.27	0.04	0.49	0.20	2.35	0.02	1.21%
Matrices de Raven	0.12	-0.01	0.26	0.10	1.77	0.08	0.81%
Conscience phonologique (souple)	0.40	0.26	0.54	0.33	5.72	< .001	7.84%
Dénomination rapide (temps en sec.)	-0.19	-0.28	-0.10	-0.23	-4.35	< .001	4.41%
Répétition de pseudo-mots (souple)	0.47	0.23	0.71	0.21	3.88	< .001	3.61%
Vocabulaire	0.16	0.05	0.27	0.15	2.78	< .001	1.69%

Note. NS = coefficient de pente non standardisé ; S = coefficient de pente standardisé.

4.2 Régressions linéaires multiples pour l'épreuve de lecture de mots chez les enfants allophones

Le modèle des enfants allophones (cf. tableau 7.) explique 68% de la variance en lecture de mots. Les différents prédicteurs testés présentent tous une contribution unique et significative à la réussite en lecture de mots sauf l'épreuve de répétition de pseudo-mots, qui teste la mémoire phonologique. Par ordre d'importance, les facteurs explicatifs sont la conscience phonologique ($p < .001$, $\Delta R^2 = 7.84\%$), la dénomination rapide automatisée ($p < .001$, $\Delta R^2 = 6.25\%$), le vocabulaire ($p < .001$, $\Delta R^2 = 2.56\%$) et les matrices de Raven ($p < .001$, $\Delta R^2 = 2.56\%$; cf. tableau 7.).

Tableau 7.

Modèle de régressions linéaires multiples des performances en lecture de mots chez les enfants allophones

$$F(7, 149) = 45.37, p < .001, R^2 = .68$$

Variables	NS	95 % CI		S	t	p	ΔR^2
		Bas	Haut				
(Intercept)	11.65	-0.91	24.20		1.83	0.07	
Nombre de mois en primaire (exposition explicite à l'écrit)	-0.03	-0.20	0.14	-0.02	-0.34	0.73	0.04%
Âge (lors de la passation)	0.01	-0.05	0.07	0.02	0.43	0.67	0.04%
Matrices de Raven	0.37	0.16	0.58	0.19	3.52	< .001	2.56%
Conscience phonologique (souple)	0.54	0.36	0.71	0.39	6.08	< .001	7.84%
Dénomination rapide (temps en sec.)	-0.36	-0.49	-0.23	-0.32	-5.47	< .001	6.25%
Répétition de pseudo-mots (souple)	0.07	-0.25	0.38	0.02	0.41	0.68	0.04%
Vocabulaire	0.27	0.11	0.42	0.17	3.40	< .001	2.56%

Note. NS = coefficient de pente non standardisé ; S = coefficient de pente standardisé.

Nous avons relevé la contribution unique de chaque variable corrélée à la réussite en lecture de mots chez les enfants tout-venant et allophones. Nous allons par la suite discuter des résultats obtenus lors de cette étude.

IV DISCUSSION

Après un bref rappel des objectifs de l'étude, nous discuterons et analyserons les différents résultats relevés dans le but de répondre à notre hypothèse. Nous évoquerons ensuite les limites ainsi que les perspectives offertes par ce travail de recherche.

1 Rappel des objectifs de l'étude

Lors de cette étude, nous avons cherché à déterminer et comparer le poids des différents prédicteurs cognitifs et linguistiques de la lecture sur les compétences en décodage (lecture de mots) entre un groupe d'enfants tout-venant et allophones. Pour cela, nous avons réalisé une observation basée sur six tâches visant à tester les différents prédicteurs de la lecture : une épreuve de conscience phonologique, une tâche de dénomination rapide automatisée, une tâche de répétition de pseudo-mots, une épreuve de désignation d'images pour tester le vocabulaire passif, une épreuve de lecture de mots ainsi que les matrices de Raven pour contrôler les habiletés non-verbales. Puis nous avons réalisé des analyses de corrélations simples et de régressions linéaires pour examiner et comparer la contribution de toutes les variables testées lors de la tâche de lecture de mots. Nous avons émis l'hypothèse que le vocabulaire expliquait une part de variance plus importante lors de la lecture de mots chez les enfants allophones que chez les enfants tout-venant.

2 Interprétation et discussion des résultats

Dans un premier temps, les analyses des scores aux différentes épreuves ont permis de relever que les performances des enfants allophones sont toujours inférieures à celles des enfants tout-venant, quelle que soit la tâche réalisée (cf. tableaux 2 et 3). Ce premier constat n'est pas étonnant compte tenu du parcours scolaire atypique des enfants allophones. En considérant que les enfants débutent l'apprentissage de la lecture en classe de CP, sur 157 enfants allophones interrogés, 111 ont été scolarisés pour la première fois en France dans un niveau de classe égal ou supérieur au CE1 (cf. figure 2). Les enfants scolarisés en France tardivement n'ont donc pas bénéficié d'un apprentissage explicite et intensif de la lecture comme cela est le cas en classe de CP. Or la littérature fait état que l'acquisition d'une langue est plus aisée lorsqu'elle s'effectue tôt dans le développement (Birdsong, 2018). Il est également plus simple d'acquérir les contrastes phonologiques fins d'une langue dans la première année de développement (Gervain & Mehler, 2010). Nous retrouvons également une grande différence d'âge lors de la passation entre les enfants allophones et tout-venant. En moyenne, les enfants tout-venant ont 86.52 mois (soit environ 7 ans) et les enfants allophones ont en moyenne 113.11 mois (soit environ 9 ans ; cf. tableau 1). Cet écart d'âge s'explique par le fait que nous avons inclus des enfants allophones de tous niveaux à partir du CP, à condition qu'ils soient sur le territoire français depuis plus de 6 mois et moins de 3 ans. En effet, le niveau dans lequel sont scolarisés les enfants allophones n'est pas forcément le reflet de leurs

compétences comme cela pourrait être le cas pour les enfants tout-venant, exclusivement de niveau scolaire CP ou CE1.

Comme précisé ci-dessus, les performances des enfants tout-venant sont toujours supérieures à celles des enfants allophones. Le test t de Welch permet de comparer significativement les moyennes des deux groupes. Cette comparaison est toujours en faveur du groupe d'enfants tout-venant, excepté pour la tâche de dénomination rapide automatisée ($p = 0.74$ et d de Cohen = -0.04 ; cf. tableau 2). Que cette tâche évalue les capacités de récupération d'une forme auditive en mémoire ou la capacité à reconnaître visuellement un item (Araújo et al., 2011), elle semble indépendante de la langue parlée par le locuteur, à condition qu'il sache dénommer les items. Il paraît donc logique que la différence entre les moyennes des deux groupes ne soit pas significative pour cette tâche.

Nous avons ensuite cherché à déterminer quelles variables étaient corrélées significativement aux performances en lecture de mots. Tous les prédicteurs de la lecture testés sont corrélés significativement aux performances en décodage pour les deux groupes (cf. tableaux 4 et 5). Nous relevons de façon indifférenciée pour les deux groupes que la conscience phonologique est la variable la plus fortement corrélée à la lecture de mots ($r = 0.54$, $p < .001$ pour les tout-venant et $r = 0.71$, $p < .001$ pour les allophones) et que le vocabulaire est la variable la moins corrélée ($r = 0.36$, $p < .001$ pour les tout-venant et $r = 0.40$, $p < .001$ pour les allophones ; cf. tableau 4 et 5). Nous avons également constaté que l'épreuve de vocabulaire était significativement corrélée à toutes les épreuves des prédicteurs cognitifs et langagiers pour les enfants tout-venant (cf. tableau 4). Nous effectuons le même constat pour le groupe d'enfants allophones, excepté pour l'épreuve de répétition de pseudo-mots ($p = 0.16$; cf. tableau 5). Comme le vocabulaire est corrélé à la lecture de mots, mais également aux autres facteurs prédictifs qui sont eux-mêmes corrélés à lecture de mots, nous ne pouvons pas conclure à une relation pure entre le niveau de vocabulaire et les performances en lecture de mots.

Pour répondre à notre hypothèse, nous avons besoin de mettre en lumière la contribution unique de chaque facteur aux performances en décodage. Les matrices de régressions linéaires ont permis de mettre en évidence la contribution unique de quatre variables lors de l'épreuve de lecture de mots chez les enfants tout-venant (cf. tableau 6). Il s'agit de la conscience phonologique, de la dénomination rapide automatisée, de la répétition de pseudo-mots et du vocabulaire. Nous retrouvons trois de ces variables dans le modèle des enfants allophones : la conscience phonologique, la dénomination rapide automatisée et le vocabulaire (cf. tableau 7). Nous constatons également que le vocabulaire a une contribution unique plus forte chez les enfants allophones que chez les enfants tout-venant pour la lecture de mots ($p < .001$, $\Delta R^2 = 1.69\%$ pour les enfants tout-venant et $p < .001$, $\Delta R^2 = 2.56\%$ pour les enfants allophones), ce qui viendrait valider notre hypothèse de départ. Cela peut s'expliquer

par le fait que les enfants tout-venant apprentis lecteurs possèdent déjà un bagage lexical oral important en français au début de l'apprentissage de la lecture (Lervåg & Aukrust, 2010). L'impact du vocabulaire serait probablement sous-estimé pour ces enfants car la tâche de lecture proposée dans cette étude n'exige pas de connaître beaucoup de mots comparé au nombre de mots connus par les enfants tout-venant. Les enfants allophones, peu exposés au français oral, présenteraient quant à eux des différences interindividuelles importantes sur ces mêmes mots.

Cependant cette différence reste relative car nous relevons seulement un écart de 0.87 point concernant le pourcentage de variance entre les enfants allophones et tout-venant lorsque nous prenons en compte la contribution unique du niveau de vocabulaire lors de l'épreuve de décodage. Nous cherchons alors à comprendre pourquoi le vocabulaire explique une si faible part de variance. Une hypothèse serait de prendre en compte le parcours de vie des enfants allophones. On peut supposer que le parcours migratoire de ces enfants vienne impacter leurs futurs apprentissages dans le pays d'accueil. Les enfants ayant vécu un événement traumatisant comme une guerre, nœud de la situation politique mondiale actuelle, sont peut-être moins disponibles aux apprentissages par rapport aux enfants au parcours de vie plus commun. Par ailleurs, certaines familles n'ont pas de visibilité quant à leur avenir sur le territoire d'accueil ou espèrent retourner un jour dans leur pays d'origine, ce qui peut signer un manque d'investissement pour la langue du pays d'accueil (Bensekhar-Bennabi, 2010). Si la langue française n'est pas ou peu investie au sein du foyer familial, alors le vocabulaire des enfants ne se développera pas ou peu et ce déficit pourrait expliquer la faible part de variance expliquée par le modèle.

Une nuance est également à apporter lorsque nous prêtons attention à la variance expliquée de chaque modèle. Pour les enfants tout-venant, le modèle explique 49% de la variance en lecture de mots, ce qui est inférieur à celui des enfants allophones qui explique 68% de la variance en lecture de mots (cf. tableaux 6 et 7). Le pouvoir prédictif des facteurs cognitifs et linguistiques sur la réussite en lecture est donc supérieur chez les enfants allophones par rapport aux enfants tout-venant. Cette différence de 19 points est non négligeable et vient nuancer la validation de notre hypothèse de départ. Certes, le vocabulaire à lui seul explique une part plus importante de variance chez les enfants allophones que chez les enfants tout-venant, mais si nous considérons que tous les facteurs prédictifs expliquent mieux la réussite en lecture chez les enfants allophones, alors il paraît logique que les pourcentages de variance soient plus importants pour les enfants allophones que les enfants tout-venant. Cet écart de variance expliquée entre les deux modèles pourrait venir du fait que les enfants allophones sont moins avancés dans le processus d'apprentissage de la lecture et accorderaient donc une part plus importante aux facteurs cognitifs et langagiers que les enfants tout-venant. L'apprentissage récent de la langue écrite et le faible niveau des enfants

allophones pour la tâche de lecture pourraient expliquer la part prégnante des prédicteurs cognitifs et langagiers en lecture de mots.

Si nous nous intéressons aux pourcentages de variance des différents facteurs prédictifs, nous remarquons que c'est la conscience phonologique qui contribue le plus aux performances en lecture de mots des deux modèles ($p < .001$, $\Delta R^2 = 7.84\%$; cf. tableaux 6 et 7). C'est donc la stratégie de recodage phonologique qui est prioritaire lors de la lecture de mots pour les deux groupes testés. Ce constat est congruent avec plusieurs études qui ont démontré que la contribution de la conscience phonologique dans la réussite en lecture était plus importante au début de l'apprentissage de la lecture (Poulsen & Elbro, 2018). Puisque la conscience phonémique se développe grâce à l'expérience de l'écrit, il ne semble pas étonnant qu'elle occupe une part de variance majeure lors de l'apprentissage de la lecture (Morais et al., 1979).

De plus, nous remarquons que chez les enfants allophones, la tâche de répétition de pseudo-mots, qui met en jeu la mémoire phonologique, ne contribue pas au modèle prédictif de la lecture, malgré un coefficient de corrélation élevé et significatif ($r = 0.44$ et $p < .01$; cf. tableaux 5 et 7). Pourtant plusieurs auteurs citent bien la mémoire phonologique comme compétence cognitive de base prédisant les compétences en lecture de mots (Melby-Lervåg & Lervåg, 2012). Il se pourrait alors que d'autres variables soient venues biaiser les résultats lors de cette tâche. En ce sens, nous constatons que la répétition de pseudo-mots est fortement corrélée à la conscience phonologique ($r = 0.54$, $p < .001$; cf. tableau 5). Nous pouvons donc supposer que la tâche de conscience phonologique capte une part importante de la variance liée à la mémoire phonologique chez les enfants allophones. La mémoire phonologique serait fortement sollicitée lors de la tâche de conscience phonologique car les items peuvent être éloignés de la langue maternelle des enfants allophones.

En outre, nous remarquons que certaines variables socio-démographiques sont corrélées significativement aux capacités de décodage chez les enfants allophones comme l'âge de la scolarisation en France ($r = 0.23$, $p < .01$), la scolarisation antérieure ($r = 0.35$, $p < .001$) et l'apprentissage de la lecture antérieur à l'arrivée en France ($r = 0.41$, $p < .001$; cf. tableau 5). Cela signifie que les enfants allophones qui ont été scolarisés avant leur arrivée en France et qui ont appris à lire dans leur langue maternelle ont de meilleures performances en lecture de mots. L'apprentissage de la lecture dans la langue maternelle viendrait alors soutenir les apprentissages dans la langue de scolarisation (Geva, 2009).

3 Limites, perspectives et intérêts de l'étude

3.1 Limites de l'étude

Nous relevons plusieurs limites à cette étude. Tout d'abord, malgré les précautions mises en place comme l'utilisation d'un dictaphone pour enregistrer les réponses des participants ou

la cotation souple décrite précédemment, la retranscription des réponses lors des passations reste subjective. Certaines différences phonologiques seront plus ou moins finement perçues selon l'examineur. Cela entraîne une cotation et donc des résultats différents.

Concernant l'épreuve de vocabulaire spécifiquement, nous relevons une limite due au hasard. Pour rappel, 50 items sont proposés et le participant doit faire un choix entre quatre réponses. Au bout de 5 réponses fausses, le test s'arrête automatiquement et toutes les réponses suivantes sont considérées comme fausses. La moyenne pour l'épreuve de vocabulaire des enfants allophones est de 18.09 contre 30.90 pour les enfants tout-venant (cf. tableau 2) mais cela sous-estime les scores que les enfants allophones auraient eu sans critère d'arrêt. Afin de contrôler le critère de hasard, il faudrait modifier la notation du test en considérant qu'après le critère d'arrêt, une réponse sur quatre serait correcte. De cette façon, il serait possible de comparer les scores des participants à un score de hasard. Une analyse complémentaire grâce au test de student et au calcul de la p-value viendrait enrichir cette nouvelle notation.

Il est important de noter que les écoles accueillant des dispositifs UPE2A sont majoritairement des écoles notifiées « Réseau d'Éducation Prioritaire » (REP ou REP+). Sur 157 enfants allophones, 93 sont scolarisés en REP/REP+ alors que sur 231 enfants tout-venant, ils sont uniquement 47 (cf. annexe I). Nous pouvons alors nous demander s'il existe une différence entre les performances des enfants scolarisés en REP/REP+ et ceux scolarisés en écoles ordinaires. Des analyses complémentaires ont été menées concernant les performances en lecture de mots en fonction du statut linguistique et des zones REP/REP+ (cf. annexe J). Nous remarquons qu'indépendamment du statut linguistique, les performances des enfants scolarisés en écoles ordinaires ne sont pas supérieures à celles des enfants scolarisés en écoles classées REP/REP+. Cette limite n'est donc pas vérifiée pour notre échantillon.

Une autre limite concerne l'échantillonnage du groupe d'enfants allophones. Comme précisé précédemment, ce groupe comporte des enfants scolarisés entre la classe de CP et de CM2, contrairement au groupe d'enfants tout-venant scolarisés uniquement en classe de CP et CE1. Même si les caractéristiques des enfants allophones ont été prises en compte grâce aux questionnaires (par exemples le temps de séjour en France, l'âge lors de la scolarisation en France ou le nombre de mois d'exposition explicite à l'écrit), toutes les données ont été analysées ensemble et les écarts d'âges entre les groupes ainsi que les différentes variables liées au développement des enfants ont pu fausser les résultats.

Nous relevons également une limite concernant le plan d'étude mené. En effet, une étude avec un plan transversal, comme c'est le cas pour ce projet de recherche, ne permet pas de déterminer le sens de la relation entre deux facteurs. Il est donc impossible de déterminer avec certitude si la connaissance des mots à l'oral est un support à la lecture de ces mots ou si ce

sont les compétences en langage écrit qui viennent soutenir l'acquisition du vocabulaire (Duff et al., 2015).

3.2 Perspectives de l'étude

Afin de contourner ces limites, plusieurs mesures pourraient être mises en place. Tout d'abord un appariement plus rigoureux permettrait une comparaison plus juste entre les groupes. Il pourrait être intéressant de n'inclure que des enfants scolarisés en écoles REP ou REP+ dans les deux groupes.

De plus, nous pouvons imaginer qu'un plan d'étude longitudinal testerait les enfants allophones à leur arrivée en France et avant le début de l'apprentissage formel de la lecture. Ce type de plan pourrait permettre de statuer sur le sens de la relation entre les différents prédicteurs cognitifs et langagiers et la réussite en lecture de mots.

Enfin, il est important de développer et mettre en valeur le modèle de cotation alternatif utilisé dans cette étude. Celui-ci autorise certaines déviations phonologiques et permet d'augmenter les scores des enfants allophones tout en n'ayant que très peu d'impact sur celui des enfants tout-venant. Les résultats obtenus grâce à la cotation souple restent cohérents et permettent d'analyser uniquement les compétences évaluées en évitant que les particularités phonologiques des enfants allophones ne viennent fausser les résultats.

3.3 Intérêts de l'étude

La présente étude tente de pallier, à son échelle, un manque de travaux scientifiques concernant la population allophone, pourtant de plus en plus nombreuse en France. Peu de données sont disponibles pour améliorer la prise en soin de ces enfants, plus à risque de développer du retard lors des différents apprentissages que les enfants tout-venant.

Ce mémoire a pour but premier d'améliorer la compréhension des processus mis en jeu lors de l'apprentissage de la lecture chez les enfants allophones, afin de pouvoir ensuite agir de façon plus efficace face à leurs difficultés.

Cette étude met également en valeur l'importance des précurseurs de la lecture sur le développement de celle-ci. Porter une attention particulière aux prédicteurs cognitifs et langagiers permet d'identifier des profils d'enfants ayant des compétences faibles aux acquisitions clés préalables à l'activité de lecture. Un repérage précoce des enfants à risque de développer des difficultés en lecture permettra alors une prise en soin orthophonique plus rapide et efficiente.

Ce projet de recherche pourra également informer les enseignants UPE2A quant aux compétences les plus déterminantes dans le développement de la lecture chez les enfants allophones.

V CONCLUSION

Les élèves allophones sont des enfants aux parcours de vie et scolaire atypiques. Ils nécessitent un accompagnement particulier et sont de ce fait inclus pendant les deux premières années qui suivent leur arrivée en France dans des unités scolaires spécifiques, les UPE2A. Ces élèves se trouvent alors dans une position délicate car ils doivent entrer dans le processus d'apprentissage de la lecture alors qu'ils ont, pour la plupart, une mauvaise maîtrise de la langue orale française.

La présente étude cherche à déterminer sur quelles compétences précoces s'appuient préférentiellement les enfants allophones lorsqu'ils apprennent à lire. Nous avons émis l'hypothèse que le niveau en vocabulaire des enfants allophones serait plus corrélé à la réussite en lecture de mots que pour les enfants tout-venant.

Nous avons donc testé et comparé deux groupes d'enfants (allophones et tout-venant) sur des épreuves évaluant les prédicteurs cognitifs et langagiers de la lecture. Nous avons déterminé que pour les deux groupes, tous les prédicteurs de la lecture étaient corrélés aux performances en lecture de mots. Quant au niveau de vocabulaire, il est lui aussi corrélé à tous les prédicteurs de la lecture, excepté pour l'épreuve de répétition de pseudo-mots chez les enfants allophones. L'analyse des modèles de régressions linéaires a permis de mettre en évidence quatre facteurs qui contribuent de façon unique à la réussite en lecture de mots chez les enfants allophones : la conscience phonologique, la dénomination rapide automatisée, les habiletés non verbales et le vocabulaire. Nous relevons également que le niveau de vocabulaire explique une part de variance plus importante lors de la lecture de mots pour les enfants allophones que pour les enfants tout-venant. Les résultats de cette étude vont donc le sens d'une contribution unique et significative du vocabulaire sur les compétences en lecture de mots plus élevée chez les enfants allophones que chez les enfants tout-venant. Il convient néanmoins de préciser que la part de variance expliquée par le modèle des enfants allophones est supérieure à celle des enfants tout-venant, ce qui nous amène à nuancer ces résultats. In fine, ce sont bien les facteurs phonologiques qui prédominent lors de l'apprentissage de la lecture, y compris pour les enfants allophones.

La littérature scientifique a très peu décrit le statut linguistique spécifique de l'allophonie, souvent compris dans une catégorie plus large de « langue minoritaire ». Ce projet de recherche a mis en lumière l'importance des différents prédicteurs cognitifs et linguistiques lors de l'apprentissage de la lecture chez les enfants allophones. Concernant la pratique clinique de l'orthophonie, cette étude a permis d'aborder les différents mécanismes développés lors de l'apprentissage de la lecture pour ces enfants au statut langagier particulier.

VI REFERENCES

- Abrahamse, E., van Dijck, J.-P., Majerus, S., & Fias, W. (2017). Finding the answer in space : The mental whiteboard hypothesis on serial order in working memory. *Frontiers in Human Neuroscience*, 8, 1-7. <https://doi.org/10.3389/fnhum.2014.00932>
- Adams, M. J. (1994). *Beginning to Read : Thinking and Learning about Print*. MIT Press.
- Agence Nationale de Lutte Contre l'Illettrisme. (2020, octobre). Les causes de l'illettrisme. www.anlci.gouv.fr.
- Araújo, S., Inácio, F., Francisco, A., Faísca, L., Petersson, K. M., & Reis, A. (2011). Component Processes Subservicing Rapid Automatized Naming in Dyslexic and Non-dyslexic Readers. *Dyslexia*, 17(3), 242-255. <https://doi.org/10.1002/dys.433>
- Armagnague, M., & Rigoni, I. (2018). Éditorial : Expériences scolaires des mineurs migrants. *Revue européenne des migrations internationales*, 34(4), 7-11.
- Bensekhar-Bennabi, M. (2010). La bilinguïté des enfants de migrants face aux enjeux de la transmission familiale. *Enfances & Psy*, 47(2), 55. <https://doi.org/10.3917/ep.047.0055>
- Birdsong, D. (2018). Plasticity, Variability and Age in Second Language Acquisition and Bilingualism. *Frontiers in Psychology*, 9. <https://www.frontiersin.org/article/10.3389/fpsyg.2018.00081>
- Bowers, P. G., & Newby-Clark, E. (2002). The role of naming speed within a model of reading acquisition. *Reading and Writing*, 15(1), 109-126. <https://doi.org/10.1023/A:1013820421199>
- Brun, L (2020). Note d'information n°20.39. *Ministère de l'Éducation nationale, de la Jeunesse et des Sports, DEPP*.
- Castel, C., Pech-Georgel, C., George, F., & Ziegler, J. C. (2008). Lien entre dénomination rapide et lecture chez les enfants dyslexiques. *L'Année psychologique*, 108(03), 395. <https://doi.org/10.4074/S0003503308003011>

Catts, H. W., Nielsen, D. C., Bridges, M. S., & Liu, Y.-S. (2016). Early Identification of Reading Comprehension Difficulties. *Journal of Learning Disabilities*, 49(5), 451-465. <https://doi.org/10.1177/0022219414556121>

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A Dual Route Cascaded Model of Visual Word Recognition and Reading Aloud. *Psychological Review*, 108(1), 204-256.

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.

Duff, D., Tomblin, J. B., & Catts, H. (2015). The Influence of Reading on Vocabulary Growth : A Case for a Matthew Effect. *Journal of Speech, Language, and Hearing Research : JSLHR*, 58(3), 853-864. https://doi.org/10.1044/2015_JSLHR-L-13-0310

Duff, F. J., & Hulme, C. (2012). The Role of Children's Phonological and Semantic Knowledge in Learning to Read Words. *Scientific Studies of Reading*, 16(6), 504-525. <https://doi.org/10.1080/10888438.2011.598199>

Dunn, L. M., Dunn, L. M., & Theriault-Whalen, C. (1993). *EVIP: échelle de vocabulaire en image Peabody*. Pearson Canada Assessment.

Ehri, L. C. (2014). Orthographic Mapping in the Acquisition of Sight Word Reading, Spelling Memory, and Vocabulary Learning. *Scientific Studies of Reading*, 18(1), 5-21. <https://doi.org/10.1080/10888438.2013.819356>

Fanuel, L., Jarjat, G., Labaronne, M., Hot, P., Tillmann, B., Portrat, S., & Plancher, G. (2021). Attention et maintien en mémoire de travail : Une revue des données neurophysiologiques. *Revue de neuropsychologie*, 13(3), 171-180.

Farnia, F., & Geva, E. (2011). Cognitive correlates of vocabulary growth in English language learners. *Applied Psycholinguistics*, 32(4), 711-738. <https://doi.org/10.1017/S0142716411000038>

Frith, U. (1986). A developmental framework for developmental dyslexia. *Annals of Dyslexia*, 36(1), 67-81. <https://doi.org/10.1007/BF02648022>

Georgiou, G. K., Parrila, R., Cui, Y., & Papadopoulos, T. C. (2013). Why is rapid automatized naming related to reading ? . *Journal of experimental child psychology*, 115(1), 218-225.

Gervain, J., & Mehler, J. (2010). Speech Perception and Language Acquisition in the First Year of Life. *Annual Review of Psychology*, 61(1), 191-218. <https://doi.org/10.1146/annurev.psych.093008.100408>

Geva, E. (2009). Apprendre à lire dans une langue seconde : Recherche, implications et recommandations pour les services. *Encyclopédie sur le développement des jeunes enfants*, 1-12.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, Reading, and Reading Disability. *Remedial and Special Education*, 7(1), 6-10. <https://doi.org/10.1177/074193258600700104>

Hulme, C., Bowyer-Crane, C., Carroll, J. M., Duff, F. J., & Snowling, M. J. (2012). The Causal Role of Phoneme Awareness and Letter-Sound Knowledge in Learning to Read : Combining Intervention Studies With Mediation Analyses. *Psychological Science*, 23(6), 572-577. <https://doi.org/10.1177/0956797611435921>

Hulme, C., & Snowling, M. J. (2013). Learning to Read : What We Know and What We Need to Understand Better. *Child Development Perspectives*, 7(1), 1-5. <https://doi.org/10.1111/cdep.12005>

Korkman, M., Kirk, U., & Kemp, S. (2012). *NEPSY-II—Bilan neuropsychologique de l'enfant - 2nde édition*. Pearson.

Labat, H., Farhat, S.-L., Andreu, S., Rocher, T., Cros, L., Magnan, A., & Ecalle, J. (2013). Évaluation des connaissances précoces prédictives de l'apprentissage de la lecture en grande section de maternelle. *Revue française de pédagogie. Recherches en éducation*, 184, 41-54. <https://doi.org/10.4000/rfp.4219>

Landerl, K., Freudenthaler, H. H., Heene, M., De Jong, P. F., Desrochers, A., Manolitsis, G., Parrila, R., & Georgiou, G. K. (2019). Phonological Awareness and Rapid Automatized Naming as Longitudinal Predictors of Reading in Five Alphabetic Orthographies with Varying Degrees of Consistency. *Scientific Studies of Reading*, 23(3), 220-234. <https://doi.org/10.1080/10888438.2018.1510936>

Lervåg, A., & Aukrust, V. G. (2010). Vocabulary knowledge is a critical determinant of the difference in reading comprehension growth between first and second language learners : Growth in L1 and L2 reading comprehension. *Journal of Child Psychology and Psychiatry*, 51(5), 612-620. <https://doi.org/10.1111/j.1469-7610.2009.02185.x>

Liberman, I. Y., Shankweiler, D., Fischer, F. W., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology*, 18(2), 201-212. [https://doi.org/10.1016/0022-0965\(74\)90101-5](https://doi.org/10.1016/0022-0965(74)90101-5)

Liu, Y., Yeung, S. S., Lin, D., & Wong, R. K. S. (2017). English expressive vocabulary growth and its unique role in predicting English word reading : A longitudinal study involving Hong Kong Chinese ESL children. *Contemporary Educational Psychology*, 49, 195-202. <https://doi.org/10.1016/j.cedpsych.2017.02.001>

Majerus, S. (2014). L'évaluation de la mémoire à court terme. *Traité de neuropsychologie clinique*, 167-177.

McKay, A., Davis, C., Savage, G., & Castles, A. (2008). Semantic involvement in reading aloud : Evidence from a nonword training study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(6), 1495-1517. <https://doi.org/10.1037/a0013357>

Melby-Lervåg, M., & Lervåg, A. (2012). Oral Language Skills Moderate Nonword Repetition Skills in Children With Dyslexia : A Meta-Analysis of the Role of Nonword Repetition Skills in Dyslexia. *Scientific Studies of Reading*, 16(1), 1-34. <https://doi.org/10.1080/10888438.2010.537715>

Melby-Lervåg, M., Lyster, S.-A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read : A meta-analytic review. *Psychological Bulletin*, 138(2), 322-352. <https://doi.org/10.1037/a0026744>

Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition*, 7(4), 323-331. [https://doi.org/10.1016/0010-0277\(79\)90020-9](https://doi.org/10.1016/0010-0277(79)90020-9)

Morton, J. (1989). An information-processing account of reading acquisition. In *From reading to neurons* (p. 43-66). The MIT Press.

Norton, E. S., & Wolf, M. (2012). Rapid automatized naming (RAN) and reading fluency: Implications for understanding and treatment of reading disabilities. *Annual review of psychology*, 63, 427- 452.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Richard Hochenberger, Sogo, H., ... Jonas Kristoffer Lindelov. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51(1), 195–203. <https://doi.org/10.3758/s13428-018-01193-y>

Peng, P., Barnes, M., Wang, C., Wang, W., Li, S., Swanson, H. L., Dardick, W., & Tao, S. (2018). A meta-analysis on the relation between reading and working memory. *Psychological Bulletin*, 144(1), 48-76. <https://doi.org/10.1037/bul0000124>

Perfetti, C. A., & Hart, L. (2001). The lexical basis of comprehension skill. In *On the consequences of meaning selection : Perspectives on resolving lexical ambiguity* (p. 67-86). American Psychological Association. <https://doi.org/10.1037/10459-004>

Plaza, M., & Robert-Jahier, A.-M. (2006). DRA Enfants—Test de Dénomination Rapide. Adeprio.

Poulsen, M., & Elbro, C. (2018). The short- and long-term predictions of reading accuracy and speed from paired-associate learning. *Journal of Experimental Child Psychology*, 174, 77-89. Scopus. <https://doi.org/10.1016/j.jecp.2018.05.007>

Raven, J. C., Court, J. H., Raven, J., & Marcou, L. (1998). Manuel des Raven Section 2 Progressive matrices couleur (Edition 1998). Editions et Applications Psychologiques.

Ricketts, J., Nation, K., & Bishop, D. V. M. (2007). Vocabulary Is Important for Some, but Not All Reading Skills. *Scientific Studies of Reading*, 11(3), 235-257. <https://doi.org/10.1080/10888430701344306>

Share, D. L. (1995). Phonological recoding and self-teaching : Sine qua non of reading acquisition. *Cognition*, 55(2), 151-226. [https://doi.org/10.1016/0010-0277\(94\)00645-2](https://doi.org/10.1016/0010-0277(94)00645-2)

Sprenger-Charolles, L., Colé, P., Piquard-KipSer, A., & Leloup, G. (2010). *EVALEC*. Ortho Edition.

Tunmer, W. E., & Chapman, J. W. (2012). Does Set for Variability Mediate the Influence of Vocabulary Knowledge on the Development of Word Recognition Skills? *Scientific Studies of Reading*, 16(2), 122-140. <https://doi.org/10.1080/10888438.2010.542527>

Winter, J. C. F., Gosling, S. D., & Potter, J. (2016). Comparing the Pearson and Spearman correlation coefficients across distributions and sample sizes : A tutorial using simulations and empirical data. *Psychological Methods*, 21(3), 273-290. <https://doi.org/10.1037/met0000079>

Wise, J. C., Sevcik, R. A., Morris, R. D., Lovett, M. W., & Wolf, M. (2007). The Relationship Among Receptive and Expressive Vocabulary, Listening Comprehension, Pre-Reading Skills, Word Identification Skills, and Reading Comprehension by Children With Reading Disabilities. *Journal of Speech, Language, and Hearing Research*, 50(4), 1093-1109. [https://doi.org/10.1044/1092-4388\(2007\)076](https://doi.org/10.1044/1092-4388(2007)076)

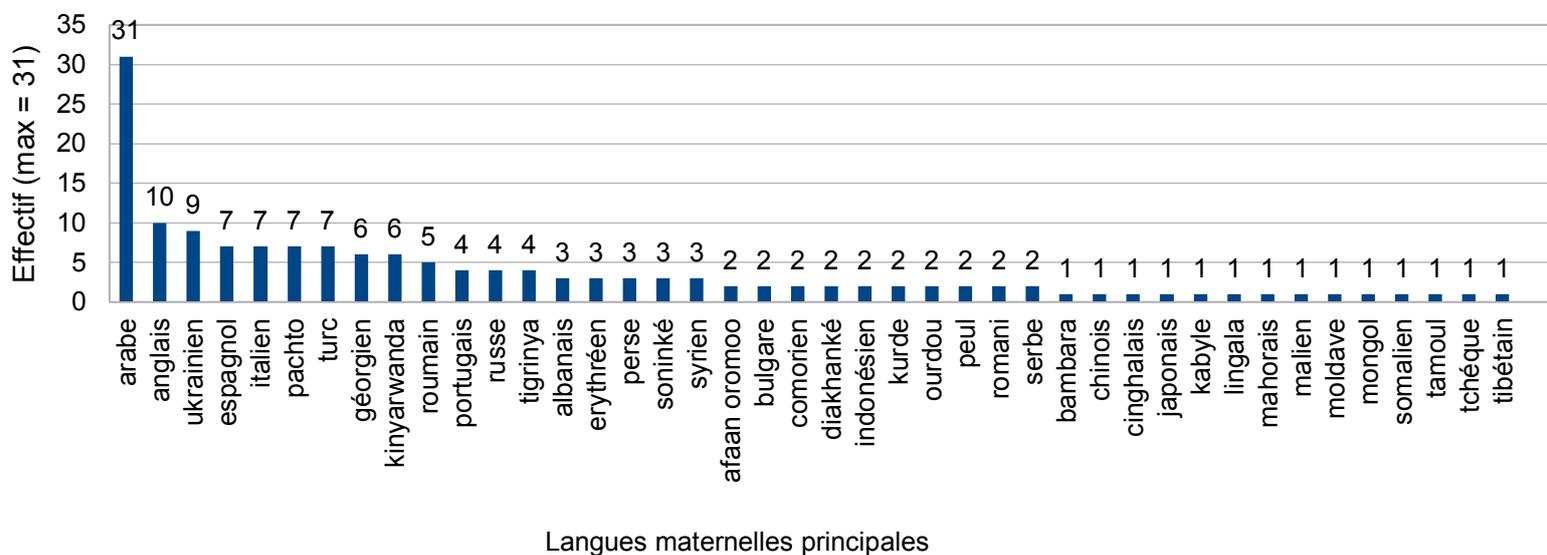
Wolf, M., Bowers, P. G., & Biddle, K. (2000). Naming-Speed Processes, Timing, and Reading : A Conceptual Review. *Journal of Learning Disabilities*, 33(4), 387-407. <https://doi.org/10.1177/002221940003300409>

Yeung, S. S. S., Siegel, L. S., & Chan, C. K. K. (2013). Effects of a phonological awareness program on English reading and spelling among Hong Kong Chinese ESL children. *Reading and Writing*, 26(5), 681-704. <https://doi.org/10.1007/s11145-012-9383-6>

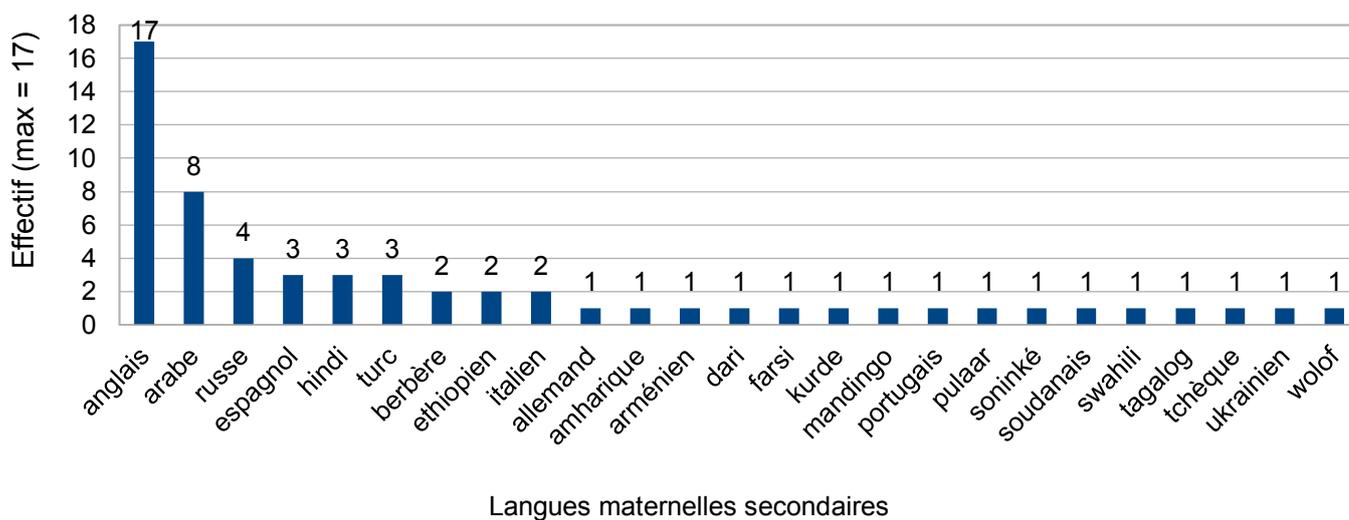
LISTE DES ANNEXES

Annexe A : Nombre d'enfants allophones par langues maternelles principales.....	1
Annexe B : Nombre d'enfants allophones par langues maternelles secondaires	1
Annexe C : Courrier d'information à l'attention des familles	2
Annexe D (2 pages) : Formulaire de consentement de participation libre et éclairée à l'attention des familles.....	3
Annexe E : Questionnaire à l'attention des parents ou tuteurs légaux des enfants allophones.....	5
Annexe F : Convention pour travaux universitaires et de recherche dans les écoles du département du Rhône (exemple de la circonscription de Lyon 7, La Mulatière)	6
Annexe G : Formulaire de demande d'autorisation pour travaux universitaires et de recherche dans les écoles du département du Rhône (exemple de la circonscription de Lyon 7, La Mulatière).....	7
Annexe H : Liste des déformations phonologiques autorisées en cotation souple	8
Annexe I : Répartition de l'effectif en fonction du statut linguistique et des Réseaux d'Éducation Prioritaire (REP et REP+)	8
Annexe J : Moyenne des résultats à l'épreuve de lecture de mots en fonction du statut linguistique et des Réseaux d'Éducation Prioritaire (REP et REP+)	8

Annexe A : Nombre d'enfants allophones par langues maternelles principales



Annexe B : Nombre d'enfants allophones par langues maternelles secondaires



Annexe C : Courrier d'information à l'attention des familles



A l'attention des parents d'élèves

Objet : étude sur l'apprentissage de la lecture

Madame, Monsieur,

Des chercheurs de l'Université de Lille font une étude sur l'apprentissage de la lecture.

Votre enfant peut participer à cette étude.

Pour cela, il devra réaliser des exercices de lecture et de langage oral. Ces exercices durent environ une heure et se passeront à l'école.

Votre enfant aura le droit demander d'arrêter les exercices quand il le voudra.

Tous ses résultats seront anonymes.

Parfois, les chercheurs devront enregistrer la voix de votre enfant. Seuls les chercheurs écouteront ces enregistrements. De plus, ces enregistrements seront anonymes et enregistrés dans un dossier crypté.

Si vous le voulez, vous pourrez avoir accès aux résultats de l'étude.

Pour que votre enfant participe, vous devez remplir le « formulaire de consentement » et le « questionnaire ».

Nous vous remercions par avance pour votre participation et restons à votre disposition pour toute question,

Matthieu BIGNON
matthieu.bignon@univ-lille.fr
+33 7 61 11 09 82

Séverine CASALIS
severine.casalis@univ-lille.fr

Sandrine MEJIAS
sandrine.mejias@univ-lille.fr

Annexe D (2 pages) : Formulaire de consentement de participation libre et éclairée à l'attention des familles

CONSETEMENT DE PARTICIPATION LIBRE ET ECLAIREE

Il a été proposé que votre enfant participe à une étude sur **l'apprentissage de la lecture chez les enfants allophones et tout-venant (TANMALL)**

L'investigateur, m'a précisé que je suis libre d'accepter ou de refuser que mon enfant participe.

Afin d'éclairer ma décision, je prends connaissance des informations suivantes et atteste que je les ai bien comprises :

- 1 La participation de mon enfant à cette étude est libre. Il pourra de ce fait interrompre sa participation à tout moment, sans avoir à donner de justification.
- 2 Moi et mon enfant avons été clairement informé(e)s des éléments suivants : But de la Recherche- Méthodologie- Durée de ma participation- Bénéfices attendus- Contraintes- Risques prévisibles, j'ai pris connaissance de la note d'information m'expliquant cette étude.
- 3 Les données transmises à l'expérimentateur sont strictement confidentielles et ne pourront être utilisées pour une publication sans mon accord préalable.
- 4 Les réponses de mon enfant pourront être enregistrées de façon sécurisée lors de la passation expérimentale si je l'autorise.
- 5 L'expérimentateur pourra consulter le niveau de présence de mon enfant à l'école si je l'y autorise.
- 6 L'anonymat de mon enfant sera par ailleurs garanti tant lors de l'analyse des résultats que pour une potentielle publication.
- 7 Les résultats de l'étude pourront m'être communiqués à l'issue des analyses, si je le désire ou si mon enfant le désire.
- 8 Mon consentement et celui de mon enfant ne déchargent en rien le responsable scientifique et l'organisme responsable de la recherche de l'ensemble de leurs responsabilités et moi et mon enfant conservons tous nos droits garantis par la loi.
- 9 Je pourrai être contacté afin de compléter les éventuelles données manquantes des questionnaires.
- 10 J'accepte que les données enregistrées à l'occasion de cette recherche puissent faire l'objet d'un traitement informatisé par l'organisme responsable de la recherche ou pour son compte). Les données recueillies demeureront strictement confidentielles. Je prends note que le droit d'accès prévu par règlement général sur la protection des données s'exerce à tout moment auprès de l'investigateur qui me suit dans le cadre de la recherche et qui connaît mon identité. Je pourrai exercer mon droit de rectification et d'opposition, d'effacement, de limitation et de portabilité auprès de ce même investigateur ou des porteurs du projet pour faire valoir mes droits.

Parent ou tuteur légal de l'enfant : (Nom et Prénom)

Adresse mail : Numéro de téléphone :

Nom et prénom de l'enfant :

Date de naissance de l'enfant :

Compte-tenu des informations qui m'ont été transmises : Cocher les cases appropriées en fonction de votre volonté (OUI / NON)	OUI	NON
1. J'accepte librement et volontairement que mon enfant participe à cette recherche	<input type="checkbox"/>	<input type="checkbox"/>
2. J'accepte que les réponses de mon enfant soient enregistrées de façon sécurisée lors de la passation expérimentale.	<input type="checkbox"/>	<input type="checkbox"/>
3. J'accepte que les données recueillies, préalablement anonymisées, puissent être utilisées par les expérimentateurs dans le cas d'une potentielle publication.	<input type="checkbox"/>	<input type="checkbox"/>
4. Je souhaite que les résultats de l'étude me soient communiqués à l'issue de l'analyse des données.	<input type="checkbox"/>	<input type="checkbox"/>
5. Je consens à ce que l'expérimentateur consulte le niveau de présence de mon enfant à l'école. Ces données resteront strictement confidentielle et seront pseudonymisées.	<input type="checkbox"/>	<input type="checkbox"/>

Date : [] [] [] [] [] [] Date : [] [] [] [] [] []

Signature du représentant légal :

Signature de l'investigateur :

Conformément à la loi Informatique et Libertés du 6 janvier 1978 modifiée et au Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016, vous pouvez exercer votre droit d'accès aux données vous concernant et les faire rectifier ou effacer en contactant le délégué à la protection des données : cil@univ-lille.fr

Annexe E : Questionnaire à l'attention des parents ou tuteurs légaux des enfants allophones

Scolarité

Votre enfant a-t-il été scolarisé dans un autre pays que la France (entourer) ? oui non

Si oui, dans quelles classes (quels niveaux scolaires) ? oui non

Votre enfant a-t-il appris à lire et écrire dans d'autres pays (entourer) ? oui non

Si oui, dans quelle langue (entourer) ? Fr L2 L3

Date d'entrée à l'école en France (mois/année) :

Dans quelle classe votre enfant a été scolarisé pour la première fois en France :

A-t-il redoublé ? oui non Si oui, quelle(s) classe(s) ?

Votre enfant est-il suivi par une UPEZA ? oui non

Depuis quand ? (mois / année)

Parents

Dernier diplôme obtenu : - Père :

- Mère :

Nombre d'années d'études (après le BAC ou équivalent) :

- Père :

- Mère :

Nom de l'enfant :

Prénom de l'enfant :

Ecole :

Classe :

Code de pseudonymisation
(rempli par l'expérimentateur)

.....

Questionnaire parental

Madame, Monsieur,

Merci d'avoir autorisé votre enfant à participer à notre étude. Nous avons besoin de quelques informations complémentaires afin de mener à bien notre étude. Nous vous invitons donc à compléter ce questionnaire si vous le souhaitez. Nous vous assurons que toutes les informations recueillies resteront strictement confidentielles et seront pseudonymisées.

Merci pour votre collaboration.

Informations générales

Votre enfant est un/une : garçon fille Je ne souhaite pas répondre

Votre enfant est-il suivi par un(e) orthophoniste (entourer) ?

Oui avant Oui maintenant Non Je ne souhaite pas répondre

(Si oui, pourquoi :

Date d'arrivée en France (mois/année) :

Quelles langues utilise votre enfant en dehors du français ?

- Langue 2 :
- Langue 3 :

	Niveau			
Compétences en français	1	2	3	4
Compétences dans la Langue 2	1	2	3	4
Compétences dans la Langue 3	1	2	3	4

Entourer le niveau correspondant à chacune des langues dans ce tableau.

1. Il comprend la langue mais il ne la parle pas
2. Il comprend la langue et il peut dire quelques mots
3. Il peut utiliser la langue pour les conversations du quotidien
4. Il est très à l'aise pour communiquer dans cette langue.

Annexe F : Convention pour travaux universitaires et de recherche dans les écoles du département du Rhône (exemple de la circonscription de Lyon 7, La Mulatière)

académie
Lyon
direction des services
départementaux
de l'éducation nationale
Rhône
éducation
nationale



Convention

Travaux universitaires et de recherche dans les écoles du département du Rhône

Nom de l'organisme demandeur : Institut des Sciences et Techniques de la Réadaptation, Université Claude Bernard Lyon 1.

Adresse : 8 avenue Rockefeller, 69008 Lyon.

Tél. : 04 78 77 70 83

Adresse électronique : memoire.orthophonie@univ-lyon1.fr

Nom et prénom du / des intervenant(s) : ASSELIN Eponine

Date et durée de l'intervention : Année scolaire 2022-2023

Renseignement(s) sur l'école / les écoles où se déroule(nt) l'intervention :

Nom : Ecoles avec dispositif UPE2A

Adresse :

Téléphone :

Circonscription concernée : Lyon 7^{ème} / La Mulatière

Adresse : 295 avenue Jean Jaurès 69007 Lyon

Téléphone : 04 78 58 33 90

Intitulé de la problématique : Étude TANMALL sur l'apprentissage de la lecture chez les enfants allophones

Description des objectifs et des modalités des travaux envisagés (thème, mode d'intervention) :

Projet de recherche porté par Matthieu Bignon, doctorant au laboratoire SCALab de l'Université de Lille, qui effectue une thèse sur le développement du langage écrit chez les enfants allophones. Cette recherche est encadrée par Séverine Casalis et Sandrine Mejias, toutes deux enseignantes chercheuses à l'Université de Lille. Par ce courrier, je demande votre autorisation pour intervenir dans les écoles de votre territoire au cours de l'année scolaire 2022-2023 afin de recueillir des données pour une étude sur l'apprentissage de la lecture chez les enfants tout-venant et allophones (étude TANMALL). L'enjeu stratégique de cette étude est de créer des outils permettant de dépister de manière précoce des difficultés en langage oral et en lecture chez les enfants allophones. Cela permettrait de combler un manque dans l'outillage des orthophonistes mais également d'aider les professionnels de l'éducation à proposer des stimulations ou exercices ciblés aux enfants allophones. Pour mener à bien cette recherche, nous souhaitons rencontrer des élèves allophones (sans distinction de langue ou de culture) scolarisés depuis plus de 6 mois en France. Concrètement, des exercices courts de langage oral et de langage écrit leur seront proposés durant deux sessions de 30 minutes dans leur environnement scolaire.

Pièces à fournir :

- Attestation d'assurance responsabilité civile.
- Retour sur projet

A Lyon, le 14/10/2022

L'organisme ou l'établissement demandeur

A Lyon, le 14/10/2022

Le ou les intervenants.

A Lyon, le 31/01/2023

Le directeur académique des services de l'éducation nationale

L'inspecteur d'académie
directeur académique

Philippe CARRIERE

Annexe G : Formulaire de demande d'autorisation pour travaux universitaires et de recherche dans les écoles du département du Rhône (exemple de la circonscription de Lyon 7, La Mulatière)



Direction des services départementaux
de l'éducation nationale
du Rhône

**DEMANDE D'AUTORISATION
POUR TRAVAUX UNIVERSITAIRES ET DE RECHERCHE
DANS LES ECOLES DU DEPARTEMENT DU RHONE**

Pour accord, les formulaires sont à transmettre sous couvert de l'IEN concerné à monsieur l'inspecteur d'académie, à l'adresse suivante : ce.ia69-dos@ac-lyon.fr

Renseignements concernant le déroulement de l'action	
Circonscription	Lyon 7 ^{ème} / La Mulatière
Nom de(s) l'école(s) concernée(s)	Écoles avec dispositif UPE2A
Classe(s) concernée(s) (Nombre/niveau(x))	
Nombre d'élèves concernés	
Mode(s) d'intervention(s) auprès des élèves (Observations dans la classe/Groupes de travail avec les chercheurs ou les étudiants / questionnaire papier et/ou Internet)	Questionnaire papier
Mode(s) d'intervention(s) auprès des enseignants (Questionnaire papier et/ou Internet)	
Durée de l'action	Année scolaire 2022-2023. 2 sessions de 30 minutes.
Renseignement concernant l'établissement demandeur	
Nom de l'établissement supérieur ou de l'organisme demandeur	Laboratoire SCALab UMR CNRS 9193
Adresses de l'établissement supérieur ou de l'organisme	Université de Lille Rue du barreau 59653 Villeneuve d'Ascq
Coordonnées téléphoniques	07 61 11 09 82
Objectifs (Recherche-travaux universitaires)	Thèse et mémoire en orthophonie
Thème des travaux	Développement du langage écrit chez les enfants allophones
Nom de/des Intervenant(s)	Eponine Asselin

Avis de l'IEN : Favorable Défavorable

Date : 25/01/2023

Brigitte PARRINI
Inspectrice de l'Education Nationale
Signature :

Décision de l'inspecteur d'académie : favorable

Date : 31.01.2023

L'inspecteur d'académie
directeur académique
Signature :
Philippe CARRIERE

Annexe H : Liste des déformations phonologiques autorisées en cotation souple

Cible	Productions acceptées
<i>/i/</i>	<i>/i/ /e/ /ɛ/ /y/</i>
<i>/e/</i>	<i>/e/ /ɛ/ /i/ /ø/ /œ/ /a/ /ɛ̃/ /œ̃/</i>
<i>/ɛ/</i>	<i>/ɛ/ /e/ /i/ /ø/ /œ/ /a/ /ɛ̃/ /œ̃/</i>
<i>/a/</i>	<i>/œ/ /ø/ /o/ /ɔ/ /e/ /ɛ/ /ā/ /ē/</i>
<i>/y/</i>	<i>/y/ /i/ /ø/ /œ/ /u/</i>
<i>/ø/</i>	<i>/ø/ /œ/ /y/ /e/ /ɛ/ /a/ /o/ /ɔ/ /œ̃/ /ɛ̃/</i>
<i>/œ/</i>	<i>/œ/ /ø/ /y/ /e/ /ɛ/ /a/ /o/ /ɔ/ /œ̃/ /ɛ̃/</i>
<i>/u/</i>	<i>/u/ /y/ /o/ /ɔ/</i>
<i>/o/</i>	<i>/o/ /ɔ/ /ɔ̃/ /u/ /ø/ /œ/ /a/</i>
<i>/ɔ/</i>	<i>/ɔ/ /o/ /ɔ̃/ /u/ /ø/ /œ/ /a/</i>
<i>/ʁ/</i>	<i>/ʁ/ /r/ /_ʁ/</i>

Annexe I : Répartition de l'effectif en fonction du statut linguistique et des Réseaux d'Éducation Prioritaire (REP et REP+)

	Tout-venant	Allophones	<u>Total Résultat</u>
Ordinaire	184	64	248
REP	0	34	34
REP+	47	59	106
<u>Total Résultat</u>	231	157	388

Annexe J : Moyenne des résultats à l'épreuve de lecture de mots en fonction du statut linguistique et des Réseaux d'Éducation Prioritaire (REP et REP+)

	Tout-venant	Allophones
Ordinaire	40,2	29
REP	/	32,4
REP+	42	25,2