

MÉMOIRE

En vue de l'obtention du
Certificat de Capacité d'Orthophoniste
présenté par
Léa VERMES

soutenu publiquement en juin 2021

**Création d'épreuves lexicales
visant à tester les effets de la Lecture Dialogue auprès
d'enfants présentant une déficience intellectuelle**

MÉMOIRE dirigé par
Bruno FACON, Professeur des universités, Université de Lille, Villeneuve d'Ascq
Lucie MACCHI, Maître de conférences, Université de Lille, Lille

Mémoire réalisé dans le cadre du Parcours Recherche

Lille – 2021

Remerciements

Je souhaite d'abord remercier Mme MACCHI, co-directrice de ce mémoire, grâce à qui je peux dire que celui-ci est à la hauteur de mes attentes. Je lui suis reconnaissante pour sa disponibilité, son écoute, ses conseils avisés et toujours bienveillants, sa réactivité et ses encouragements. J'ai été ravie et honorée de pouvoir profiter de son expertise grâce à cette expérience inédite que représente le mémoire de fin d'études. Je remercie également M. FACON, qui a codirigé ce mémoire, pour son temps précieux et pour le partage de savoirs et d'anecdotes agrémentés d'une touche d'humour dont il a le secret. Je m'estime chanceuse d'avoir été si bien accompagnée durant ces deux années. Enfin, l'équipe de choc ne serait pas complète sans Amélie REMY-NERIS, que je remercie sincèrement pour m'avoir permis de participer à cette belle étude menée dans le cadre de son doctorat, pour nos nombreux échanges instructifs et la confiance qu'elle m'a accordée.

Je suis extrêmement reconnaissante aux cinq jeunes participants que j'ai eu le plaisir de rencontrer et d'accompagner dans le cadre de ce mémoire. J'espère sincèrement que cette expérience leur aura apporté autant qu'à moi. Je remercie leurs familles et les professionnels des Instituts Médico-Educatifs m'ayant accueillie pour avoir rendu cela possible.

Je remercie aussi toutes les orthophonistes ayant croisé mon chemin au cours des cinq dernières années. Elles ont participé à façonner ma vision de ce métier incroyable et bientôt mien. Elles m'ont partagé leur passion et leur goût du soin, qui m'ont animée durant toute la partie pratique de ce mémoire.

Enfin, je remercie tout particulièrement mes parents pour m'avoir offert la possibilité de me réaliser au fil de ce cursus, qui se conclut par cet écrit, et ce, dans les meilleures conditions qui soient. Je remercie également tous mes proches qui de près ou de loin, en étant convaincus de ma réussite, m'ont portée jusqu'ici. Le parcours a été extraordinaire à vos côtés.

Résumé :

La lecture dialogue est une méthode de stimulation langagière efficace auprès de diverses populations. Nous souhaitons, grâce à notre étude, être en mesure de tester son efficacité auprès d'enfants présentant une déficience intellectuelle, car cela n'a été que peu étudié. Afin de récolter des données précises concernant d'éventuels progrès lexicaux chez les participants, nous avons créé des épreuves spécifiques à notre étude, car elles sont reconnues plus sensibles que les tests standardisés. Selon nos hypothèses, les épreuves lexicales spécifiques sont en mesure d'objectiver des progrès significatifs pour les enfants entraînés en vocabulaire, contrairement aux épreuves standardisées. Lors du pré- et du post-test, les deux types d'épreuves ont été présentés à 22 enfants. Entre-temps, plusieurs séances individuelles de lecture dialogue ont été réalisées par des étudiantes en orthophonie. Les résultats tendent à confirmer notre hypothèse : les épreuves lexicales spécifiques ont objectivé une différence significative entre le pré- et le post-test pour les enfants entraînés en vocabulaire. Afin d'attester la validité des épreuves créées, différents questionnaires ont été diffusés. Les résultats des questionnaires d'accord sur l'image et sur le nom apportent des preuves en faveur de la validité des épreuves créées. Grâce à ces épreuves, nous avons pu constater que des séances individuelles de lecture dialogue permettent à des enfants d'un âge intellectuel environ égal à quatre ans d'acquérir, en réception comme en expression, des mots entraînés. Les résultats plutôt satisfaisants d'accord sur l'image et sur le nom sont encourageants, mais devront être complétés pour parfaire l'analyse quantitative de l'étude.

Mots-clés :

Épreuves spécifiques, lexique, lecture dialogue, déficience intellectuelle, validité

Abstract :

Dialogic reading is an effective method of language stimulation for various populations. We hope, through our study, to be able to test its effectiveness on children with intellectual disability, because it has been poorly studied. In order to collect accurate data regarding potential lexical progress for the participants, we have created specific tests for our study, because they are known to be more sensitive than standardized tests. According to our hypotheses, the specific lexical tests are able to objectify significant progress for children trained in vocabulary, unlike the standardized tests. During the pre- and post-test, both types of tests were presented to 22 children. In the interval, several individual dialogic reading sessions were conducted by speech-language pathology students. The results tend to confirm our hypothesis: the specific lexical tests showed a significant difference between the pre- and post-test for children trained in vocabulary. In order to attest the validity of the created tests, different questionnaires were used. The results of the image and name agreement questionnaires provide evidence for the validity of the created tests. Thanks to these tests, we have been able to observe that individual sessions of dialogic reading allow children with an intellectual age of approximately four years to acquire trained words in reception and expression. The rather good results of image and name agreements are encouraging, but will have to be completed to improve the quantitative analysis of the study.

Keywords :

Specific tests, lexicon, dialogic reading, intellectual disability, validity

Table des matières

Introduction	3
Contexte théorique, buts et hypothèses	5
.1. Lecture Dialogue et déficience intellectuelle	5
.2. Évaluer l'efficacité des interventions	5
2.1. Dans le domaine du vocabulaire	5
2.2. Dans le domaine de la LD	6
2.3. Recommandations dans le domaine de la LD	6
.3. Les particularités de la déficience intellectuelle	7
3.1. Le développement langagier	7
3.2. L'apprentissage dans les interventions langagières	7
.4. Les objectifs de la création d'épreuves spécifiques	8
4.1. Évaluer précisément les effets de la Lecture Dialogue chez les enfants présentant une déficience intellectuelle	8
4.2. Comparer les résultats des épreuves spécifiques et standardisées	9
4.3. Démontrer la validité interne des épreuves	9
.5. Hypothèses	10
5.1. Amélioration significative des scores aux épreuves lexicales spécifiques dans le Groupe Vocabulaire	10
5.2. Absence d'effet significatif aux tests standardisés	10
Méthode	11
.1. Création des épreuves	11
1.1. Choix des épreuves	11
1.2. Contenu des épreuves	11
1.2.1. Sélection des mots cibles	11
1.2.2. Sélection des distracteurs et positionnement sur les planches	12
1.2.3. Sélection des illustrations	13
1.3. Validité des épreuves	13
1.3.1. Analyse des variables	13
1.3.2. Accord sur l'image	13
1.3.3. Accord sur le nom	14
1.3.4. Distance sémantique	14
1.4. Fidélité des épreuves	14
.2. Entraînement	15
2.1. Population	15
2.2. Matériel	15
2.2.1. Pour le pré- et le post-test	15
2.2.2. Pour l'entraînement	15

2.3. Procédure	16
Résultats	16
Discussion	18
Conclusion	21
Bibliographie.....	22
Liste des annexes	28
Annexe 1. Variables psycholinguistiques des items cibles et de leurs distracteurs.....	28
Annexe 2. Premier item exemple et item 1 de l'épreuve spécifique de lexique en production	28
Annexe 3. Premier item exemple et item 1 de l'épreuve spécifique de lexique en réception.....	28
Annexe 4. Liste des albums sélectionnés et des mots cibles associés.....	28
Annexe 5. Résultats des questionnaires d'accord sur l'image et d'accord sur le nom	28
Annexe 6. Comparaison des fréquences d'occurrence des mots ayant obtenu un accord sur le nom suffisant et insuffisant	28

Introduction

Lors de la conception d'une étude expérimentale, les auteurs doivent choisir les mesures qui permettront d'évaluer les effets de leur intervention. De ce choix dépendent les résultats et les conclusions. Dans la littérature, il est fréquent que les auteurs optent pour des tests standardisés. Ces mesures étalonnées permettent de déterminer le niveau langagier global d'un individu et de le situer par rapport à une norme en comparant ses scores avec ceux d'une population de référence. Cependant, toute intervention n'a pas obligatoirement un effet sur les résultats à ces tests en raison de leur caractère très général. Lorsque les interventions sont ciblées et/ou brèves la sensibilité des tests standardisés est réduite. Dans le cadre de ces interventions, l'utilisation de mesures moins globales, capables de révéler des progrès réels, mais circonscrits, serait intéressante. Les mesures spécifiques pourraient alors être plus adaptées, simplement parce qu'elles concernent directement les apprentissages ciblés lors de l'intervention. Pourtant, malgré leurs avantages reconnus dans la littérature, ces mesures ne sont pas systématiquement utilisées. Cela est peut-être dû au manque de données concernant leur élaboration ou aux preuves parfois modestes de leur validité. C'est dans le souci de pallier ces manques que nous nous intéresserons, dans le présent mémoire, aux épreuves créées dans le cadre d'une étude expérimentale dont le but est d'objectiver l'efficacité de la Lecture Dialogue (LD) sur les acquisitions lexicales d'enfants présentant une déficience intellectuelle (DI).

La LD est une méthode d'intervention langagière structurée décrite par Whitehurst et ses collègues en 1988. Elle consiste en une lecture d'albums au cours de laquelle les enfants s'investissent activement (Lonigan & Whitehurst, 1998) grâce au lecteur adulte qui encourage, évalue, reformule et répète les productions orales des enfants (Grygas Coogle et al., 2018) et, en outre, renforce leur implication à l'aide de questions incitatives. Bien que l'efficacité de cette méthode soit largement validée par des études menées auprès de différentes populations, elle reste encore à démontrer auprès d'enfants présentant une DI.

Dans le cadre de notre étude, des enfants avec une DI ont été répartis en deux groupes nommés Groupe Vocabulaire (GV) et Groupe Syntaxe (GS) selon le domaine langagier travaillé au cours des séances. Nous supposons que les progrès réalisés en vocabulaire, à l'issue de notre intervention, seront significativement plus importants au sein du GV que du GS. Afin de le démontrer, le GS sera le groupe contrôle du GV, lors des analyses statistiques. Dans le GV, entre septembre et novembre 2020, 30 mots ont été entraînés au cours de trois à onze sessions de LD ($Mdn = 7$) d'une durée minimale de vingt minutes chacune. L'apprentissage de mots nouveaux, en dehors des mots cibles, étant limité, nous supposons que les tests standardisés ne permettront pas de mettre en évidence l'efficacité de la LD. Selon nous, seules des épreuves spécifiquement élaborées pour l'étude seront en mesure d'y parvenir, car elles permettent d'évaluer la maîtrise des 30 mots cibles travaillés en expression et en réception lors des séances d'entraînement.

Nous exposerons, pour commencer, des données relatives à la recherche sur l'efficacité de la LD dans le champ de la DI. Nous présenterons également une revue de la littérature synthétisant, dans un premier temps, les connaissances disponibles concernant l'évaluation des interventions dans les domaines du vocabulaire et de la LD puis, dans un second temps, les caractéristiques du développement langagier des enfants DI et des interventions destinées à en assurer la promotion. Ensuite, après avoir mis en évidence les intérêts des épreuves spécifiques, le processus de création de ces épreuves sera détaillé afin de permettre leur reproduction ou leur adaptation dans de futures études. Leur validité sera évaluée, notamment grâce à des questionnaires d'accord sur le nom et sur l'image dont les réponses seront traitées et analysées quantitativement. Finalement, nous préciserons, selon les premiers résultats de cette étude, les enseignements que ces épreuves permettent de tirer.

Ces résultats seront discutés. Suivra une conclusion générale destinée à synthétiser le travail réalisé et à ouvrir des perspectives.

Contexte théorique, buts et hypothèses

.1. Lecture Dialogue et déficience intellectuelle

De nombreuses études tendent à montrer que la LD influence positivement le langage d'enfants de milieux défavorisés (Burgoyne et al., 2018; Dickinson & Porche, 2011; Korat et al., 2013; Lonigan & Whitehurst, 1998; Mendelsohn et al., 2001; Valdez-Menchaca & Whitehurst, 1992; Van Kleeck et al., 2006; Wasik & Bond, 2001; Whitehurst et al., 1994; Zevenbergen & Whitehurst, 2003), à risque de difficultés scolaires (Lonigan et al., 2008) ou présentant un retard langagier. Cette méthode s'avère donc efficace auprès de populations assez variées. En revanche, peu de travaux ont, jusqu'à présent, été menés auprès d'enfants présentant une DI.

Or, certains principes de la LD correspondent à des pratiques très favorables au développement langagier, même en cas de DI. On peut mentionner l'ajustement du contenu et de la forme des énoncés de l'adulte au niveau langagier de l'enfant ainsi que l'utilisation de reformulations verbales qui, pour être efficaces auprès des enfants déficients intellectuels, doivent être plus fréquentes que chez les enfants au développement typique (Proctor-Williams & Fey, 2007). Les quelques études sur la LD menées dans le domaine de la DI ne portent que sur un faible nombre de participants, ce qui constitue un problème sur le plan méthodologique. Cependant, toutes ont abouti à des résultats prometteurs (Dale et al., 1996). La plupart ont été réalisées auprès d'enfants présentant des troubles du spectre autistique (Bellon et al., 2000; Fleury & Schwartz, 2017; Grygas Coogle et al., 2018). D'autres pathologies comme le syndrome de Rett (Koppenhaver et al., 2001; Skotko et al., 2004) ou l'infirmité motrice cérébrale (Liboiron & Soto, 2006) ont également fait l'objet d'études dont les résultats se sont avérés positifs. C'est dans la continuité de ces données empiriques que s'inscrit la présente étude visant à évaluer l'efficacité de la LD auprès d'enfants présentant une DI.

.2. Évaluer l'efficacité des interventions

2.1. Dans le domaine du vocabulaire

Une part non négligeable des auteurs étudiant les effets de méthodes d'enrichissement du vocabulaire utilise exclusivement des tests standardisés afin de quantifier les effets de leurs interventions sur le vocabulaire de jeunes enfants. C'est notamment le cas dans 32 des 67 (48%) études incluses dans la méta-analyse de Marulis et Neuman (2010). Ces mesures visent à évaluer le niveau lexical global des enfants. De ce fait, elles ne contiennent pas ou peu de mots entraînés lors de l'intervention (Sénéchal, 2010).

Les épreuves spécifiques, à l'inverse, permettent de tester uniquement le vocabulaire ciblé par l'intervention (Marulis & Neuman, 2013) et sont donc aptes à mesurer avec plus de sensibilité les effets du traitement (Rosenshine, 2010). Ces mesures ont été recommandées par des chercheurs dans le domaine du vocabulaire (Biemiller, 2006; Sénéchal & Cornell, 1993). Elles sont jugées plus pertinentes dans la quantification des progrès. Malgré cela, elles ne sont utilisées que dans 52% des études recensées par Marulis et Neuman (2010). Parmi les 35 études en question, 16 les associent à des tests standardisés.

2.2. Dans le domaine de la LD

De nombreux auteurs ont cherché à démontrer les effets de la LD sur le vocabulaire d'enfants d'âges et de milieux différents. Pour cela, la majorité a utilisé des tests standardisés. Les résultats obtenus grâce à ces tests sont très hétérogènes. Certaines études ont mis en évidence des effets significatifs sur les deux versants du langage (Marulis & Neuman, 2013) ou uniquement sur le versant expressif (Karrass & Braungart-Rieker, 2005; Lonigan, 1993; Lonigan & Whitehurst, 1998; Whitehurst et al., 1988). En revanche, d'autres études n'ont objectivé aucun effet significatif, que ce soit en réception ou en expression (Crain-Thoreson & Dale, 1999; DeBaryshe et al., 1992). Cette discordance dans les résultats s'explique en partie par des différences méthodologiques entre les études. Cependant, elle s'observe également lors de la comparaison de deux études aux conditions d'intervention et d'évaluation similaires (Arnold et al., 1994; Whitehurst et al., 1988). En effet, bien que l'âge, le niveau langagier et le milieu socioculturel des enfants ainsi que la durée et la fréquence de l'intervention en LD soient sensiblement similaires entre ces deux études, les effets obtenus sont différents. L'étude de Whitehurst et ses collègues (1988) révèle des différences significatives entre le groupe contrôle et le groupe expérimental aux deux épreuves expressives (à l'ITPA-VE¹, $p = .005$; à l'EOWPVT², $p = .009$) et un effet non significatif à l'épreuve réceptive ($p = .065$ au PPVT-R³). En revanche, dans l'étude d'Arnold et ses collègues (1994), toutes les épreuves lexicales ont permis d'objectiver une différence significative entre les deux groupes, l'ITPA-VE ($p = .02$), l'EOWPVT ($p = .01$) et le PPVT-R ($p = .03$). Il est à noter que la taille de l'échantillon est supérieure dans l'étude d'Arnold et ses collègues par rapport à celle de Whitehurst et ses collègues (64 participants contre 29). Cette relative inconsistance des effets pose question quant à la fiabilité des conclusions basées exclusivement sur les résultats obtenus grâce à des tests standardisés. En outre, dans plusieurs études, les épreuves spécifiques ont permis d'obtenir des résultats plus francs que ceux enregistrés avec des épreuves standardisées (Elleman et al., 2009; Hargrave & Sénéchal, 2000; Marulis & Neuman, 2010, 2013; Sénéchal et al., 1995; Whitehurst et al., 1994).

En résumé, les tests standardisés, largement utilisés dans la littérature, possèdent à la fois des avantages et des inconvénients (Arnold et al., 1994). Néanmoins, leur caractère très général explique, selon certains auteurs, l'absence apparente de progrès suite à une intervention en LD (Hargrave & Sénéchal, 2000). Par conséquent, nous avons recherché si des préconisations explicites, qui permettraient d'identifier de manière éclairée les mesures adéquates à chaque étude, étaient formulées dans la littérature.

2.3. Recommandations dans le domaine de la LD

Le choix des mesures doit se faire au regard des hypothèses émises. En l'occurrence, la formulation d'hypothèses précises est recommandée afin de mieux rendre compte des effets d'une intervention sur les différents domaines langagiers (Scarborough & Dobrich, 1994). Les mesures spécifiquement créées dans le cadre d'une intervention ont donc un grand intérêt, car elles permettent d'évaluer précisément l'efficacité de celle-ci (Ebbels, 2017).

De plus, Rosenshine (2010) recommande l'utilisation de mesures spécifiques dans de nombreux domaines, langagiers et autres, puisqu'elles donnent systématiquement lieu à des tailles

¹ Illinois Test of Psycholinguistic Ability – Verbal Expression

² Expressive One-Word Picture Vocabulary Test

³ L'Échelle de Vocabulaire en Image Peabody (E.V.I.P) est une adaptation française de ce test.

d'effet plus importantes que les mesures standardisées. Cependant, ces dernières restent pertinentes pour déterminer le niveau langagier de l'enfant, le comparer à une norme (Pollard-Durodola et al., 2011) et rendre compte de la validité consécutive d'une intervention en témoignant de l'évolution des capacités langagières générales (Marulis & Neuman, 2013). En effet, les mesures spécifiques ne peuvent témoigner d'une éventuelle généralisation des apprentissages dans le domaine du vocabulaire.

En conclusion, la combinaison des deux types de mesures (standardisées et spécifiques) semble être le meilleur choix. C'est d'ailleurs ce que conseille le National Reading Panel (2000) dans sa méta-analyse narrative. En effet, les tests standardisés seuls ne permettraient pas de conclure, car l'absence de résultats significatifs à ces tests ne signifie pas nécessairement une inefficacité de l'intervention (Ebbels, 2017). D'un autre côté, les conclusions tirées sur la base de mesures spécifiques sont souvent remises en question à cause du manque de confiance en leur validité (Elleman et al., 2009). Ce choix semble également être le plus adapté compte tenu des caractéristiques de la population étudiée, thématique que nous aborderons en détail dans la section suivante.

3. Les particularités de la déficience intellectuelle

3.1. Le développement langagier

Le développement cognitif précède et conditionne le développement langagier, ce qui explique le retard langagier invariablement observé chez les enfants déficients intellectuels. La maîtrise d'un code comme le langage nécessite des capacités cognitives suffisantes. Réciproquement, le langage permet l'acquisition de concepts et de représentations (INSERM, 2016). Or, la DI affecte les mécanismes d'apprentissage et limite l'acquisition du code linguistique, qui sera plus lente et restera inachevée, c'est-à-dire en deçà du niveau attendu chez un individu typique et ce, même à l'âge adulte.

Le caractère systématique des difficultés langagières peut également s'expliquer par la qualité plus modeste des apports langagiers fournis aux enfants. Cela s'explique d'une part, par le lien établi entre DI et pauvreté (INSERM, 2016). En effet, les apports langagiers sont reconnus comme suboptimaux dans les milieux sociaux défavorisés. D'autre part, parce que l'adulte adapte inconsciemment la complexité de son discours au niveau intellectuel supposé de l'enfant (Siegel, 1963a, 1963b; Siegel & Donovan, 1964; Siegel & Harkins, 1963). Cet ajustement de l'individu le situe dès lors dans la zone d'autonomie de l'enfant, dans laquelle les concepts abordés sont acquis. Or, c'est en dépassant légèrement le niveau langagier et cognitif de l'enfant, en se situant dans sa zone proximale de développement, que l'adulte lui permet de progresser.

Au niveau lexical, les enfants présentant une DI disposent de compétences approximativement similaires à celles d'enfants typiques de même âge mental. Globalement, les processus d'acquisition des mots et de leurs significations sont les mêmes (Pruthi, 2013). Seules diffèrent la concrétude des mots, qui a tendance à favoriser l'acquisition langagière des enfants présentant une DI, et l'acquisition de mots issus de classes grammaticales diversifiées, qui est plus tardive et moins marquée (Rondal, 1975).

3.2. L'apprentissage dans les interventions langagières

Les interventions langagières conduites auprès de personnes DI sont fondamentales, car leurs difficultés langagières sont systématiques. Ces interventions visent une meilleure communication

pour permettre à ces personnes d'exprimer leurs besoins, leurs choix, leur inconfort ou encore leur douleur. Le rapport de l'INSERM (2016) répertorie plusieurs modalités d'interventions langagières menées auprès des enfants avec une DI et rassure en rejetant l'idée que la DI pourrait limiter leur efficacité (Cole et al., 1990; Fey et al., 1994). En effet, si les interventions influencent le développement langagier des enfants typiques, alors elles profiteront également aux enfants présentant une DI puisque leur développement est assez comparable à celui des enfants typiques. Les seules adaptations potentiellement nécessaires lors des interventions langagières et notamment concernant la LD, seraient un respect strict des modalités de mise en œuvre ainsi qu'une augmentation de l'intensité (nombre de sessions par unité de temps, ex. deux sessions par semaine) et de la durée totale des interventions.

.4. Les objectifs de la création d'épreuves spécifiques

4.1. Évaluer précisément les effets de la Lecture Dialogue chez les enfants présentant une déficience intellectuelle

Les épreuves spécifiques, créées par les auteurs des études, permettent de tester la connaissance des mots ciblés lors de l'intervention. Elles sont généralement plus sensibles que les épreuves standardisées et permettent également une analyse qualitative plus fine des performances. Il est aussi possible, grâce à ces épreuves, d'explorer les liens entre le pourcentage de réussite à un item (ou à un groupe d'items) et diverses variables psycholinguistiques, par exemple la classe grammaticale, l'âge moyen d'acquisition ou la fréquence d'occurrence de l'item (ou du groupe d'items). Les conclusions permises grâce à ces épreuves devraient s'avérer plus précises en ce qui concerne l'efficacité de la LD pour les personnes présentant une DI. La reprise et l'adaptation de ces données pourraient sans doute permettre d'augmenter les connaissances sur l'acquisition du vocabulaire dans de futures études conduites dans le champ de la DI.

Qui plus est, le choix de créer des épreuves peut également être motivé par les particularités cognitives de la population testée. Les études incluant des enfants présentant une DI, dont le vocabulaire est limité, nécessitent des adaptations dans la conception et l'exécution de l'intervention (Pollard-Durodola et al., 2011), mais également dans l'évaluation des effets de l'intervention. Chez ces participants, les effets sont potentiellement moindres et/ou plus lents à apparaître par rapport aux enfants au développement typique. Il revient donc aux chercheurs de sélectionner des mesures suffisamment sensibles aux progrès réalisés (Dunning et al., 1994). Or, les mesures spécifiques sont plus à même de détecter des progrès ciblés et limités que les tests standardisés.

D'un point de vue plus général, parmi les études interventionnelles en LD recensées dans le champ du handicap, aucun des auteurs n'a opté pour l'utilisation exclusive de tests standardisés. Les auteurs ont utilisé, en y associant ou non des mesures standardisées, le codage et l'analyse de corpus (Bellon et al., 2000; Koppenhaver et al., 2001; Liboiron & Soto, 2006; McGinty et al., 2012; Pile et al., 2010; Skotko et al., 2004), des mesures spécifiques (Fleury & Schwartz, 2017; Grygas Coogle et al., 2018; Stephenson, 2009) et/ou des questionnaires parentaux (Pile et al., 2010). La volonté de ces auteurs, et aussi la nôtre, est de se mettre dans la situation la plus favorable pour détecter les effets de l'intervention. Dans notre étude, nous combinons épreuves spécifiques, tests standardisés et analyses de corpus.

4.2. Comparer les résultats des épreuves spécifiques et standardisées

La création de mesures spécifiques permet également de comparer leurs résultats avec ceux obtenus à l'aide de tests standardisés. Si tous les résultats aux tests standardisés et aux épreuves spécifiques sont significatifs, cela supposerait que l'intervention a permis l'apprentissage des mots cibles ainsi que la généralisation des apprentissages à d'autres mots. Dans le cadre de notre étude, ce cas de figure n'est pas envisagé puisque 26 des 30 mots cibles sont des radicaux ⁴, des items spécifiques, non dérivés par l'adjonction d'un affixe. Or pour ces mots, aucune règle ne peut être extraite ni généralisée. Seuls quatre mots cibles peuvent être considérés comme construits : épuisette, poulailler, pansement et monocycle.

En revanche, si seules les mesures spécifiques apportent des résultats significatifs, cela signifierait que l'augmentation du vocabulaire est possible, mais se limite, pour une intervention de 3 à 11 semaines ($Mdn = 7$), aux mots entraînés lors des sessions de LD. Dans ce cas, l'utilisation exclusive de tests standardisés aurait entraîné des conclusions lacunaires ne portant que sur l'absence d'apprentissage de mots de vocabulaire non entraînés. Enfin, si aucun des résultats n'est significatif, cela signifierait que l'intervention proposée en LD est inefficace (sur le vocabulaire à la fois entraîné et non entraîné). Quoi qu'il en soit, les comparaisons entre les résultats obtenus sont pertinentes (Rosenshine, 2010) et les conclusions associées également. Ces conclusions, si elles étaient confirmées lors d'études futures, fourniraient des arguments pour considérer ou non la LD comme un moyen d'enrichir le lexique des enfants présentant une DI.

4.3. Démontrer la validité interne des épreuves

En créant les épreuves spécifiques de manière structurée, en vérifiant et en contrôlant de nombreux paramètres, nous souhaitons démontrer leur validité, c'est-à-dire leur capacité à tester ce qu'elles sont supposées tester. Plus spécifiquement nous cherchons à vérifier leur validité interne, ou validité de contenu. Celle-ci atteste l'exhaustivité et la représentativité des items par rapport au domaine testé. Pour la justifier, nous aborderons, dans un premier temps, le choix des items cibles, l'analyse approfondie de ceux-ci ainsi que leur nombre. Dans un second temps, nous évoquerons les questionnaires d'accords sur le nom et sur l'image, sur la distance sémantique, sur l'âge d'acquisition, et les différentes variables psycholinguistiques qui ont été contrôlées.

Une validité interne satisfaisante conférerait aux résultats obtenus un caractère plus probant, nous permettant d'enrichir les connaissances actuelles dans le domaine de la DI et des interventions langagières. Cela légitimerait également l'utilisation éclairée de nos conclusions, basées sur ces épreuves, par la communauté scientifique. Actuellement, le manque de confiance en la validité des épreuves spécifiques peut constituer un frein à leur utilisation. Dans le cadre de notre étude, nous choisissons de les exploiter pour leurs atouts en tentant, néanmoins, d'en limiter les faiblesses.

⁴ Selon le dictionnaire de l'Académie française en ligne, « élément commun à plusieurs mots d'une même famille, [...] que l'on peut isoler en retranchant les différents préfixes, suffixes, infixes et désinences » (s.d.).

.5. Hypothèses

5.1. Amélioration significative des scores aux épreuves lexicales spécifiques dans le Groupe Vocabulaire

En accord avec les données de la littérature concernant l'efficacité de la LD, il est raisonnable de considérer que cette intervention langagière est adaptée à des enfants aux besoins spécifiques (Lonigan et al., 2008). Nous faisons donc l'hypothèse qu'elle sera efficace auprès des enfants présentant une DI et que les épreuves spécifiques seront en mesure de le confirmer.

Pour obtenir des résultats significatifs, la durée de l'intervention doit être suffisante. Puisque de tels résultats ont pu être obtenus après seulement 5 sessions de 20 minutes de LD réparties sur 2 semaines (soit 100 minutes d'interaction) auprès d'enfants issus de milieux défavorisés (Korat et al., 2013), nous espérons en observer également suite à 18 sessions de 20 minutes (soit 360 minutes d'interaction minimum) auprès d'enfants atteints de DI. La durée de notre entraînement a été raccourcie en raison du contexte sanitaire, celle-ci étant finalement comprise entre 60 et 220 minutes environ. Nous maintenons néanmoins que cette durée devrait permettre d'objectiver des résultats significatifs aux épreuves lexicales spécifiques pour le GV.

De plus, dans la présente étude, l'âge mental des participants des deux groupes est d'environ quatre ans. Or, en cas de DI, il existe une corrélation entre l'âge mental et le niveau langagier des enfants. Puisque l'effet maximal de la LD est considéré atteint aux environs de deux-trois ans chez l'enfant au développement typique (INSERM, 2016), nous pensons que les participants de notre étude sont susceptibles de progresser significativement sur le plan langagier grâce à cette intervention. Cela va dans le sens des résultats prometteurs obtenus lors des rares travaux portant sur la LD dans le cadre de la DI. Nous émettons donc l'hypothèse selon laquelle les scores aux épreuves lexicales spécifiques lors du post-test seront significativement supérieurs à ceux obtenus lors du prétest au sein du GV.

5.2. Absence d'effet significatif aux tests standardisés

Nous faisons l'hypothèse que les épreuves standardisées ne permettront pas d'objectiver de différence significative entre le pré- et le post-test, pour deux raisons. La première raison est que, dans les études ayant démontré un effet grâce à ces épreuves, les intervenants adultes sont fréquemment les parents, susceptibles de généraliser les principes de la méthode aux activités de la vie quotidienne (Burgoyne et al., 2018), ou les enseignants, pouvant diffuser ces principes à des domaines d'apprentissage plus généraux (Wasik & Bond, 2001). Dans notre intervention, les étudiantes en orthophonie ne rencontrent les enfants qu'à l'occasion des sessions de LD, nous pouvons donc nous attendre à une absence ou une très faible application des principes de la LD dans d'autres contextes.

La seconde raison est qu'une intervention plus longue permettrait aux participants d'être confrontés à un vocabulaire plus étendu et donc, potentiellement, à des mots inclus dans les tests standardisés. Cependant, cela semble peu probable à l'issue des trois à onze sessions que nous avons été en mesure de conduire dans le cadre de notre étude.

Bien que ces tests soient peu susceptibles d'objectiver des différences significatives au sein du GV entre le pré- et le post-test, ils permettront de faire le lien entre le niveau langagier des enfants, qu'ils permettront d'évaluer, et les capacités d'apprentissage de ces derniers, comme cela a déjà été réalisé chez les enfants tout-venant (Sénéchal et al., 1995).

Méthode

1. Création des épreuves

1.1. Choix des épreuves

Une épreuve de désignation d'une image parmi quatre sur entrée verbale orale, et une épreuve de dénomination d'images ont été créées. Elles permettent de tester respectivement la compréhension et la production des trente mots ciblés lors de l'entraînement (cf. annexe 1). L'épreuve de définition n'est pas apparue comme pertinente dans le cadre de notre étude, car la capacité à définir est acquise à un stade plus avancé de connaissance du mot que nous ne considérons pas atteindre en neuf semaines. De plus, en réception, il n'est pas nécessaire d'être capable de définir un mot pour pouvoir le comprendre, notamment parce que la compréhension en situation écologique peut s'appuyer sur le contexte (Winick et al., 2009). Il en est de même pour l'épreuve en expression, laquelle ne sollicite pas l'utilisation d'unités lexicales en situation de communication (Thibaut et al., 2001). En outre, il s'agit d'une tâche métalinguistique, c'est-à-dire une tâche qui nécessite une réflexion de l'individu sur sa langue. Elle est donc cognitivement complexe et particulièrement difficile pour des enfants avec une DI.

Les épreuves de dénomination et de désignation nous permettent de différencier l'étape de compréhension de l'étape d'expression, supposée ultérieure (Sénéchal, 1997). La réalisation des deux épreuves permet donc, contrairement à l'utilisation exclusive de l'une ou l'autre, une évaluation plus fine du degré de maîtrise des mots entraînés. Par ailleurs, la dénomination permet de simuler de manière écologique les processus engagés lors de la récupération des mots en mémoire sémantique (Piérart, 2005). Enfin, les tâches proposées à l'enfant lors de ces épreuves sont similaires à celles des épreuves standardisées sélectionnées (désignation d'une image parmi quatre sur entrée verbale orale et dénomination d'un mot sur entrée visuelle). Grâce à cela, la comparaison des scores aux différentes épreuves est possible (Rosenshine, 2010) malgré de légères différences, notamment concernant les illustrations, exclusivement en noir et blanc dans les épreuves standardisées, généralement en couleurs pour les épreuves spécifiques.

1.2. Contenu des épreuves

1.2.1. Sélection des mots cibles

Les items testés dans ces épreuves correspondent aux trente mots entraînés au sein du GV. Ce nombre tient à trois objectifs. D'abord, celui d'élaborer des épreuves suffisamment sensibles et fiables, ce qui nécessite un nombre non négligeable d'items. Ensuite, celui de limiter la durée de l'évaluation et, ainsi, d'éviter trop de lassitude de la part des participants. Enfin, permettre un nombre suffisant de présentations de chaque item lors des séances d'entraînement, afin d'augmenter la probabilité de progrès de la part des enfants.

Les mots ont été sélectionnés au sein d'albums traitant de thèmes spécifiques (ex. les chevaliers) afin d'être peu connus, voire inconnus, des enfants. Un second avantage des albums thématiques est qu'ils favorisent implicitement la mise en réseau sémantique, par exemple le mot « trompette » sera implicitement relié aux autres noms d'instruments d'un album sur la musique. Les albums sélectionnés devaient présenter des caractéristiques précises : relater une histoire, contenir

des illustrations simples (avec peu de distracteurs visuels), en couleurs, attractives et en lien avec leur contenu. Par ailleurs, les albums ne devaient pas nécessiter de manipulations lors de la lecture, mais devaient contenir chacun quatre à cinq mots complexes (qui pouvaient être ciblés par l'entraînement), entourés de mots simples et rappelés chacun sur au moins deux pages différentes du même album. Enfin, ils devaient être propices aux questionnements des enfants. Afin de respecter au mieux l'ensemble de ces critères, et dans un souci d'homogénéisation du contenu des séances, les scripts des dix-huit albums sélectionnés ont été réécrits par la doctorante responsable de l'étude. La réécriture visait une présentation répétée de chaque mot cible et une narration des histoires fidèle aux illustrations dans le but de faciliter le processus d'apprentissage (Hargrave & Sénéchal, 2000). Les structures syntaxiques des phrases ont été simplifiées et la structure narrative a été rendue accessible et prévisible, afin de favoriser la participation des enfants.

Ensuite, pour être sélectionnés, les mots cibles devaient être "imageables", ne pas posséder d'hyperonymes plus fréquents (ex. « gorille » risquant d'être dénommé « singe »), ni être polysémiques (ex. « baguette » : variété de pain, instrument du magicien), car l'image mentale du mot serait potentiellement variable d'un participant à l'autre, ce qui biaiserait les résultats au test de désignation d'images. Il fallait également qu'ils soient définissables sans utiliser (ou en utilisant peu fréquemment) la négation, des propositions relatives ou des pronoms sujets il/elle (ciblés par l'entraînement syntaxique). La fréquence d'occurrence des mots cibles devait être suffisamment faible pour que les mots aient peu de chance d'être connus des participants. Finalement, trois mots cibles ont été sélectionnés au sein de chaque album, en se basant également sur des résultats préliminaires d'accords sur le nom et sur l'image, définis aux paragraphes 1.3.2 et 1.3.3.

1.2.2. Sélection des distracteurs et positionnement sur les planches

L'épreuve de désignation contient, pour chaque mot cible, trois distracteurs destinés à évaluer la précision des représentations sémantiques des enfants. Les distracteurs ne sont pas travaillés au cours de l'entraînement. Un des distracteurs est proche sémantiquement du mot cible (ex. « poivron » et « aubergine »). Le 2^e distracteur est éloigné sémantiquement (ex. « épouvantail » et « aubergine »). De manière équilibrée et aléatoire, les distracteurs proches et éloignés sémantiquement sont alternativement présents ou absents des images de l'album du mot cible. La distance sémantique entre deux mots est un paramètre assez subjectif, initialement fondé sur le jugement de la doctorante et de quelques personnes sondées lors de la sélection des distracteurs. Un questionnaire sur la distance sémantique a par la suite été réalisé afin d'objectiver ce paramètre (cf. paragraphe 1.3.4.). Enfin, le 3^e distracteur est un autre mot cible, sans lien sémantique, issu d'un album différent (ex. « caserne » et « aubergine »). L'objectif de ce dernier distracteur est d'éviter que l'enfant ne pointe par élimination une illustration fréquemment désignée par l'adulte lors de l'entraînement.

L'absence de distracteurs phonologiques au sein de cette épreuve, pourtant souvent présents dans les épreuves standardisées, s'explique par deux raisons. Premièrement, dans le cadre de la DI, la phonologie est un domaine langagier relativement préservé (INSERM, 2016). Il nous a donc semblé plus pertinent d'examiner finement le degré de précision sémantique des mots cibles plutôt que leurs représentations phonologiques. Deuxièmement, nous avons unifié les fréquences d'occurrence des mots cibles et des distracteurs afin d'éviter les désignations par élimination. Cette homogénéisation aurait fait perdre tout intérêt aux distracteurs phonologiques. En effet, les mots cibles devant être peu fréquents pour ne pas être connus des enfants, il est peu probable que les participants disposent en mémoire des représentations phonologiques des mots cibles (avant l'intervention) et des distracteurs (avant et après l'intervention). Les distracteurs phonologiques

n'auraient alors pas été en mesure d'assurer leur fonction première.

Enfin, la localisation du mot cible et de chaque type de distracteur sur les planches a été effectuée de manière aléatoire en randomisant toutes les possibilités de placements possibles. Suite à cette randomisation, nous avons également vérifié que le mot cible n'était pas situé au même endroit sur deux planches successives pour ne pas induire de persévération chez les participants.

1.2.3. Sélection des illustrations

Les illustrations des épreuves spécifiques sont majoritairement des dessins en couleurs, bien que les épreuves standardisées soient exclusivement illustrées en noir et blanc. Cela nous a semblé important compte tenu de l'âge chronologique et du niveau de développement intellectuel des participants (compris entre 4 et 6,83 ans), mais également dans un souci de réalisme afin de minimiser les confusions visuelles susceptibles de biaiser les résultats. De plus, la couleur est l'un des traits sémantiques principaux pour certains mots cibles comme « aubergine » ou « jonquille ».

Des questionnaires préliminaires d'accord sur le nom et sur l'image (cf. paragraphes 1.3.2 et 1.3.3) ont été diffusés à un nombre limité d'individus (entre 15 et 25) afin d'observer une tendance pour chaque mot/illustration et d'effectuer des modifications lorsque la médiane de l'accord était très basse ou lorsque plusieurs répondants étaient gênés par les mêmes aspects.

Enfin, les illustrations des mots cibles dans les épreuves spécifiques sont différentes de celles des albums (Pollard-Durodola et al., 2011) dans le but d'éviter des réponses fondées uniquement sur une reconnaissance visuelle. Nous souhaitons tester la connaissance du concept en nous assurant que l'enfant ait étendu la signification du mot à d'autres représentations du référent.

1.3. Validité des épreuves

1.3.1. Analyse des variables

Les caractéristiques psycholinguistiques des items cibles et des distracteurs ont été répertoriées (cf. annexe 1). Cette démarche vise à contrôler les variables recensées (ex. longueur des mots, fréquence d'occurrence, structure syllabique, nombre de phonèmes) sur les résultats obtenus aux épreuves. La longueur des mots et leur fréquence ont été déterminées grâce à LEXIQUE© et Manulex©. Il est également pertinent d'étudier l'impact de l'âge d'acquisition des mots sur leur apprentissage. Pour cela, il est d'usage d'utiliser l'âge subjectif d'acquisition (âge auquel les individus estiment avoir acquis un mot), lequel s'avère significativement corrélé à l'âge réel d'acquisition (Chalard et al., 2003). Pour le déterminer, un questionnaire sera diffusé à un grand nombre de personnes par le biais du logiciel « Google Form ». Ce moyen de diffusion a été sélectionné pour chacun de nos questionnaires en raison du peu de données personnelles récoltées (âge et sexe uniquement). L'élaboration de ce questionnaire a suscité des réflexions méthodologiques concernant l'information à cibler (âge de production ou âge de compréhension des mots) et la formulation appropriée qui n'induirait aucun biais de réponse. Ces questionnements restent en suspens, les âges d'acquisition ne pourront donc pas être présentés.

1.3.2. Accord sur l'image

L'accord sur l'image permet de confirmer que l'image proposée est conforme à l'image mentale du mot pour un échantillon d'individus supposé être représentatif de la population générale

(Snodgrass & Vanderwart, 1980). Celui-ci a été calculé pour tous les items cibles. Pour l'obtenir, nous avons diffusé un questionnaire à un grand nombre de personnes qui devaient, après avoir lu le mot écrit puis vu l'image correspondante, indiquer sur une échelle de Likert à quel point l'image leur semblait représenter le mot en question. Cette échelle proposait cinq choix de réponses allant de 1 (pas du tout d'accord) à 5 (tout à fait d'accord). Un accord sur l'image moyen inférieur à 4,3 (Bragard et al., 2010) est considéré comme insuffisant. Le questionnaire a également été diffusé grâce au logiciel « Google Form ».

1.3.3. Accord sur le nom

L'accord sur le nom permet de vérifier que l'image proposée active chez le participant la représentation lexicale attendue. Il a été calculé pour les illustrations des items cibles et des distracteurs. Pour cela, un questionnaire a été diffusé à grande échelle via le logiciel « Google Form ». Les répondants devaient indiquer le premier mot auquel ils pensaient en voyant chaque image. Si la réponse correspondait au mot souhaité, alors l'item était coté un, sinon il était coté zéro. Une seconde cotation, plus souple a été réalisée pour les distracteurs. Celle-ci évalue l'activation du concept correct et cotait également un pour les synonymes et les définitions correctes du concept. En effet, les distracteurs étant sémantiques, et non phonologiques, une représentation correcte du concept suffit. Un accord sur le nom moyen inférieur à 0,95 (Bragard et al., 2010) est considéré comme insuffisant. Les questionnaires d'accord sur le nom et sur l'image n'ont pas été soumis aux mêmes individus pour éviter tout biais. Lors de la diffusion des deux questionnaires, nous avons tenté d'équilibrer l'âge ainsi que le ratio hommes/femmes des répondants.

1.3.4. Distance sémantique

La distance sémantique entre l'item cible et ses distracteurs sera évaluée grâce à un questionnaire. Une échelle de Likert allant de 1 à 7 permettra aux répondants d'indiquer si les deux images proposées représentent des concepts très proches ou au contraire très éloignées sémantiquement. Un questionnement concernant l'ordre de l'échelle a retardé l'élaboration et la diffusion du questionnaire. Nous n'avons pas obtenu de données chiffrées avant la soutenance du présent mémoire (présentation prévue dans le cadre de la thèse d'Amélie Rémy-Néris). Nous espérons que les distracteurs non reliés obtiendront 6 ou 7 si l'échelle concerne la distance sémantique, 1 ou 2 si elle concerne la proximité sémantique.

1.4. Fidélité des épreuves

La formule 20 de Kuder Richardson, KR-20, sera utilisée afin de mesurer la fidélité des épreuves lexicales créées. Il s'agit d'un cas particulier du coefficient alpha de Cronbach s'appliquant uniquement aux items dichotomiques (Kuder & Richardson, 1937). Il permettra d'estimer le degré de cohérence de l'ensemble des items en se basant sur les réponses des participants aux épreuves lexicales spécifiques. Celui-ci doit être supérieur ou égal à 0,70 pour être considéré comme acceptable (Nunally, 1978). En raison du nombre limité de participants inclus dans l'étude en juin 2021, celui-ci n'aurait pas été pertinent. Il sera calculé ultérieurement, lorsque le nombre de participants sera suffisant.

.2. Entraînement

2.1. Population

Les 22 participants de l'étude ont été recrutés par les étudiantes en orthophonie avec l'aide des professionnels de leurs IME respectifs. Des critères d'inclusion et d'exclusion avaient préalablement été définis afin de constituer un échantillon homogène. Les participants devaient avoir un âge de développement approximativement égal à quatre ans, comprendre des consignes orales simples et être capables de communiquer oralement grâce à des énoncés de quelques mots. Les troubles sensoriels, les troubles du spectre autistique avérés, les troubles massifs du comportement et/ou de l'attention et les situations de polyhandicap étant incompatibles avec l'entraînement en LD ont constitué des critères d'exclusion. Nous avons par la suite obtenu le consentement oral des participants ainsi que le consentement parental écrit.

Les participants inclus dans l'étude lors de la première session d'entraînement (septembre à novembre 2020) étaient âgés de 7,09 à 11,75 ans. Nous avons suivi 10 filles et 12 garçons, dont les étiologies à l'origine de la déficience intellectuelle étaient diverses (ex. trisomie 21, syndrome de Williams, X Fragile) et parfois inconnues. J'ai personnellement suivi 2 participants durant cette période. Les prétests, post-tests et séances d'entraînement ont été réalisés au sein de 9 IME dans 6 départements français : 75, 64, 78, 92, 59, 31.

2.2. Matériel

2.2.1. Pour le pré- et le post-test

Lors du prétest nous avons évalué le développement cognitif des participants grâce au subtest des matrices de l'échelle non verbale d'intelligence de Wechsler (WNV, Wechsler et Naglieri, 2009) et l'épreuve de mémoire auditivo-verbale de la Nouvelle Échelle Métrique de l'Intelligence (NEMI-2, Cognet, 2006). Le lexique et la morphosyntaxe ont été évalués en réception et en production grâce à des épreuves spécifiques et standardisées. Les quatre épreuves standardisées sélectionnées pour l'étude sont l'Épreuve de Compréhension Syntaxico-Sémantique (E.CO.S.SE, Lecocq, 1996) pour la morphosyntaxe en réception, le Test de Closure Grammaticale Révisé (TCG-R, Deltour, 1998) pour la morphosyntaxe en production, l'Échelle de Vocabulaire en Images Peabody (EVIP, Dunn et Dunn, 1993) pour la compréhension lexicale et l'épreuve de Lexique en Production de l'ELO (Khomsî, 2001).

Toutes ces épreuves ont de nouveau été proposées lors du post-test à l'exception du subtest de mémoire auditivo-verbale de la NEMI-2 et de raisonnement non verbal de la WNV. Le post-test a été avancé de plusieurs semaines en raison d'un confinement national nous amenant à suspendre les séances d'entraînement. La doctorante n'ayant pu se déplacer au sein des IME dans ce contexte, celui-ci n'a pas pu être effectué en simple aveugle.

2.2.2. Pour l'entraînement

Les séances de LD du GV avaient pour support matériel neuf albums comprenant chacun trois à quatre mots cibles (cf. annexe 4). À chaque séance, trois albums étaient proposés à l'enfant qui, par conséquent, était exposé à neuf mots cibles minimum. Chaque mot était répété de nombreuses fois au cours de la lecture de l'album grâce aux scripts réécrits par la doctorante et aux interactions

langagières inhérentes à la LD. Les séances se déroulaient au sein des IME accueillant les participants, dans de petites salles mises à disposition. Les séances étaient enregistrées grâce à un téléphone portable ou un enregistreur vocal. Des mesures sanitaires strictes ont été respectées en raison de la situation sanitaire. Le matériel était plastifié et scrupuleusement désinfecté entre chaque utilisation, le port du masque était permanent et du gel hydroalcoolique était utilisé conformément aux recommandations en vigueur.

2.3. Procédure

Cette étude a, dans un premier temps, été validée par le comité d'éthique de l'Université de Lille, suite à quoi nous avons pu commencer à contacter des structures accueillant des enfants présentant une DI, recueillir les consentements des différents acteurs et finalement entamer les prétests. Les épreuves standardisées et spécifiques ont été administrées individuellement aux enfants par des étudiantes en cinquième année d'orthophonie lors du pré- et du post-test.

Lors du prétest, chaque session débutait par une épreuve de pointage, car ces épreuves sont moins intimidantes pour l'enfant qui n'est pas obligé de prendre la parole. Cela permet de le mettre en confiance au contact d'un examinateur peu connu, voire inconnu (Hargrave & Sénéchal, 2000). Afin d'éviter toute interférence entre les épreuves, deux épreuves permettant d'évaluer un même aspect du langage ou un même versant ne se suivaient pas directement. Les deux épreuves lexicales spécifiques ont été proposées lors de deux journées d'évaluation distinctes. L'épreuve spécifique de production a été proposée avant l'épreuve de réception afin d'éviter toute surestimation de la capacité de dénomination pouvant être engendrée par une activation préalable des représentations lexicales. Enfin, les épreuves spécifiques et standardisées ont été alternées afin d'éviter tout effet de fatigue ou de lassitude sur l'une ou l'autre mesure.

Lors du post-test, nous avons été contraints de limiter le nombre de critères à respecter au profit d'un classement des épreuves par ordre de priorité. En effet, la situation sanitaire instable rendait l'accès aux IME très incertain. Pour cela, nous avons préféré proposer toutes les épreuves spécifiques (lexicales et syntaxiques) avant les épreuves standardisées. L'objectif étant, en cas d'arrêt soudain des passations causé par la fermeture des établissements, d'avoir en notre possession les premiers résultats nous permettant de confirmer ou non une partie de nos hypothèses. Comme lors du prétest, et pour les mêmes motifs, l'épreuve lexicale spécifique en production a été proposée avant l'épreuve en réception. En revanche, cela n'a pas été le cas pour les épreuves syntaxiques spécifiques. Enfin, les épreuves standardisées visant à évaluer les compétences réceptives ont été proposées avant celles qui concernaient la production. Que ce soit en compréhension ou en expression, le lexique a été testé avant la syntaxe.

Résultats

Validité des épreuves lexicales spécifiques

Les données relatives à la validité des épreuves lexicales spécifiques sont présentées en Annexe 5. Le questionnaire d'accord sur l'image a été complété par 69 répondants d'un âge moyen de 25,78 ans et dont 87% étaient des femmes. Parmi les 30 items cibles, 24 ont obtenu un accord moyen sur l'image satisfaisant, c'est-à-dire supérieur ou égal à 4,3 sur 5 (Bragard et al., 2010). Pour les 6 items restants, la médiane des résultats est toujours égale à 4. L'accord sur l'image minimal

moyen est de 3,67 pour le mot « tuyau ». Le questionnaire d'accord sur le nom a été administré à 49 répondants d'âge moyen égal à 39,69 ans et dont 55% sont des femmes. La moyenne des accords sur le nom des mots cibles est de 92% et la médiane de 94%.

Tableau 1. Valeurs de l'accord sur le nom (X).

	Nombre de mots cibles	Nombre de distracteurs
X = 100%	4	14
95% ≤ X < 100%	9	19
90% ≤ X < 95%	6	7
85% ≤ X < 90%	8	8
X < 85%	3	12
Total	30	60

Parmi les 30 mots cibles, 13 ont obtenu un accord sur le nom suffisant, c'est-à-dire supérieur ou égal à 95%. L'accord sur le nom minimum pour les mots cibles est de 73%, pour le mot « cymbales ». En ce qui concerne les 60 distracteurs, nous différencions deux modalités de cotations : l'une mesurant l'accord sur le nom et l'autre, plus souple, attestant la reconnaissance du concept. Pour cette seconde cotation, les synonymes (ex. becqueter pour picorer) et les définitions correctes du mot attendu (ex. appareil pour prendre la tension pour tensiomètre) ont été cotés 1. La moyenne des accords sur le nom des distracteurs est de 90% et la médiane de 96%. L'accord sur le nom est satisfaisant pour 33 distracteurs, il est au minimum de 39% pour « désinfectant ». Concernant la cotation souple, la moyenne est de 95% et la médiane de 98%. La reconnaissance du concept est supérieure ou égale à 95% pour 44 distracteurs. Le minimum est de 65% pour « trapèze », et le maximum de 100% est atteint pour 25 mots (ex. muguet).

Influence des séances de lecture dialogue

Les scores bruts des enfants présentant une DI au pré- et au post-test ont été collectés et analysés quantitativement. Nous avons calculé la valeur p du test non paramétrique U de Mann-Whitney comparant les médianes des groupes Vocabulaire et Syntaxe ainsi que la valeur p du test non paramétrique W de Wilcoxon comparant les scores des participants d'un même groupe au prétest et au post-test. Ces calculs ont été réalisés pour les épreuves spécifiques et standardisées (cf. Tableau 1). En effet, l'hypothèse formulée précédemment mentionnait que seules les épreuves lexicales spécifiques seraient en mesure d'objectiver un effet significatif de l'intervention en LD auprès des enfants entraînés en vocabulaire.

Tableau 2. Différences entre le pré- et le post-test, d'après les résultats du test non paramétrique W de Wilcoxon (valeurs de p).

			Groupe Vocabulaire	Groupe Syntaxe
Épreuves spécifiques	Lexique	Réception	.004	.018
		Production	.010	.309
	Morphosyntaxe	Réception	.051	.006
		Production	.127	.014
Épreuves standardisées	Lexique	Réception (EVIP B)	.964	.014
		Production (ELO – LexP)	.023	.120
	Morphosyntaxe	Réception (E.CO.S.SE)	.142	.006
		Production (TCG-R)	.606	.135

Note. Les valeurs en gras correspondent à des $p < .05$.

Concernant les épreuves lexicales spécifiques, les différences de scores bruts entre le pré- et le post-test au sein du GV sont significatives en réception (progrès médian de 6 points bruts) comme en expression (progrès médian de 3 points bruts). En revanche, aucune des deux épreuves spécifiques évaluant la morphosyntaxe n'a objectivé de différences significatives au sein du GV entre le pré- et le post-test. Les épreuves standardisées évaluant la morphosyntaxe ainsi que celle évaluant le lexique en réception n'ont pas objectivé de différence significative au sein du GV entre le pré- et le post-test. En revanche, à l'épreuve standardisée de l'ELO LexP, évaluant le lexique en production, les scores du GV étaient significativement supérieurs lors du post-test par rapport au prétest.

Au sein du GS, trois des quatre épreuves spécifiques ont mis en évidence une différence significative entre le pré- et le post-test : les épreuves évaluant la morphosyntaxe, en réception comme en expression, ainsi que l'épreuve évaluant le lexique en réception. Aux épreuves standardisées les scores du GS sont significativement plus élevés, lors du post-test par rapport au prétest, aux deux épreuves réceptives : en lexique et en morphosyntaxe.

Les différences entre les deux groupes (GV et GS) ne se sont jamais révélées significatives, d'après les calculs du test non paramétrique U de Mann-Whitney. Les deux groupes ne se distinguent donc jamais significativement l'un de l'autre, ni au prétest ni au post-test dans aucune des huit épreuves proposées.

Discussion

Dans le cadre de cette étude, nous avons créé des épreuves lexicales spécifiques afin d'objectiver des progrès significatifs en vocabulaire suite à des séances de LD réalisées auprès d'enfants présentant une DI. Leur élaboration a été le fruit d'un long processus, visant à contrôler de nombreux paramètres. Des questionnaires ont été diffusés ou sont en cours d'élaboration afin de démontrer la validité de contenu des épreuves lexicales créées.

Nous avons pu analyser les résultats de deux questionnaires : l'un relatif à l'accord sur l'image des trente mots cibles et l'autre concernant l'accord sur le nom des mots cibles et de leurs distracteurs. Le premier questionnaire objective un accord sur l'image satisfaisant pour 24 des 30 items cibles, selon le seuil fixé à 4,3 par Bragard et al. (2010). Cela représente 80% des mots cibles. Concernant

les 20% restants, il convient de rappeler que les mots cibles ont été sélectionnés pour leur faible fréquence d'utilisation. Or, nous supposons que des mots peu fréquents activent des représentations moins précises chez les répondants que des mots plus courants. De plus, notre questionnaire présentait des représentations dessinées, là où Bragard et al. (2010) proposaient des photographies. Il est possible qu'une représentation dessinée soit implicitement plus éloignée de la représentation mentale du concept chez les répondants qu'une représentation photographiée. En somme, les résultats d'accord sur l'image confortent la validité des épreuves lexicales créées.

Concernant le questionnaire d'accord sur le nom, seuls 13 mots cibles sur 30, soit 43%, ont obtenu un accord sur le nom suffisant, selon le seuil fixé à 95% par Bragard et al. (2010). Pour certains items, le pourcentage d'accord sur le nom est limité du fait d'une orientation ambiguë de flèches ayant entraîné des dénominations telles que « mur » pour « caserne » ou « toit » pour « chapiteau ». Parmi les 14 items pour lesquels une flèche a été utilisée pour désigner le concept ciblé, 9 ont obtenu un accord sur le nom insuffisant (4 mots cibles sur 6 et 5 distracteurs sur 8). Ce biais n'est pas attendu lors de l'évaluation des participants puisque l'examineur présent peut préciser l'orientation de la flèche grâce à un pointage. Une analyse des erreurs de dénomination pour chacun de ces items permettra de confirmer l'absence de ce biais. Le pourcentage de mots cibles ayant un accord sur le nom satisfaisant est faible, bien que les résultats soient globalement encourageants. Afin d'expliquer cela, une analyse qualitative approfondie des réponses au questionnaire serait pertinente.

Concernant les 60 distracteurs, la reconnaissance du concept est au minimum de 95% pour 44 d'entre eux (73%). Ce chiffre est très positif et signifie qu'une grande majorité des répondants fait référence au concept souhaité. Le nombre de répondants connaissant le mot exact attendu est moindre puisque seulement 55% des distracteurs ont un accord sur le nom supérieur ou égal à 95%. Cependant, cette référence au terme exact est accessoire dans le cadre de nos épreuves puisque les distracteurs ont été sélectionnés pour leur proximité sémantique avec les mots cibles, donc, seule la référence au concept visé importe. Les résultats de reconnaissance conceptuelle des distracteurs étayent la validité de contenu de l'épreuve créée évaluant la compréhension des items cibles. Cette épreuve semble donc bien en mesure de tester ce qu'elle est supposée tester.

Nous avons examiné qualitativement les données récoltées grâce au questionnaire d'accord sur le nom. Nous avons ainsi observé que la moyenne et la médiane des fréquences d'occurrence des mots ayant obtenu un accord sur le nom insuffisant sont plus faibles que celles des mots ayant obtenu un accord sur le nom suffisant (cf. Annexe 6). Cela se vérifie pour les mots cibles comme pour les distracteurs. Ces comparaisons ont été réalisées selon les données chiffrées issues de deux bases de données, Lexique3© et Manulex©. La fréquence d'occurrence des mots semble donc influencer sur le pourcentage d'accord sur le nom. Or, pour cette étude nous avons spécifiquement sélectionné des mots cibles avec une fréquence d'occurrence faible afin que ceux-ci ne soient pas connus ou soient peu connus des participants avant l'entraînement. Les distracteurs ayant été appariés en fréquence avec les mots cibles, ils sont également peu fréquents. Cela peut donc expliquer le faible pourcentage de mots ayant obtenu un accord sur le nom suffisant. Des analyses statistiques supplémentaires permettraient de déterminer si ces différences, de moyennes et de médianes, sont significatives. Il serait également intéressant d'étudier l'influence de l'âge d'acquisition sur les deux types d'accords, sur le nom et sur l'image. Ces données justifient la faiblesse des résultats d'accord sur le nom qui auraient pu remettre en cause la validité des épreuves lexicales spécifiques.

Il convient à présent de confronter les résultats obtenus aux hypothèses émises précédemment. D'abord, nous posons l'hypothèse selon laquelle les scores obtenus par le GV aux épreuves lexicales

spécifiques lors du post-test seraient significativement supérieurs à ceux obtenus lors du prétest. Les résultats confirment cette hypothèse. La progression médiane des scores bruts est deux fois supérieure en réception ($Mdn = 6$ points) par rapport à l'expression ($Mdn = 3$ points). Cela suggère que les mots sont d'abord compris avant d'être produits. Toutefois, nous n'avions pas anticipé l'amélioration significative également constatée au sein du GS à l'épreuve spécifique de lexique en réception. Cette dernière est difficilement explicable, car ce groupe n'est pas supposé avoir été exposé aux trente mots cibles en dehors du prétest. Cela semble signifier qu'un entraînement spécifique en morphosyntaxe grâce à la LD permet d'améliorer significativement les compétences lexicales. Néanmoins, nous notons qu'à cette épreuve, la valeur de p du GV ($p = .004$) est inférieure à celle du GS ($p = .018$). Ainsi, la différence des scores entre le pré et le post-test semblerait plus importante au sein du GV, par rapport au GS, et la probabilité que celle-ci soit due au hasard est moindre. Cela semble indiquer qu'un entraînement en LD ciblé sur le vocabulaire permet un apprentissage lexical plus important qu'un entraînement morphosyntaxique s'appuyant sur la même méthode.

De façon similaire, notre seconde hypothèse a été en partie confirmée. En effet, nous supposions une absence d'effets significatifs aux épreuves standardisées entre le pré- et le post-test au sein du GV. Cela s'est vérifié pour trois épreuves sur les quatre réalisées. La différence significative objectivée à l'épreuve lexicale en production de l'ELO au sein du GV ($p = .023$) est inattendue. Plusieurs facteurs explicatifs sont envisageables : la stimulation langagière inhérente aux séances de LD, le développement spontané du vocabulaire, l'effet placebo, l'effet d'un traitement préalable et/ou concomitant ou encore l'effet de pratique du test. Nous ne disposons pas actuellement de données suffisantes afin de privilégier l'une de ces explications.

Nous constatons donc que les progrès significatifs réalisés par les participants sont plus nombreux que ceux attendus a priori. Cette tendance, bien que parfois contradictoire avec les hypothèses émises, est extrêmement positive concernant l'efficacité de la LD. Aussi, une analyse qualitative des erreurs serait tout à fait pertinente et permettrait d'affiner nos conclusions notamment en caractérisant les erreurs réalisées. Si celles du GV sont sémantiquement plus proches des mots cibles que celles du GS aux épreuves lexicales spécifiques, cela traduirait des représentations sémantiques plus précises chez les participants du GV. Les bénéfices des séances de LD sur le lexique seraient alors supérieurs à ceux déjà constatés quantitativement.

Conformément aux données de la littérature présentées au début de ce mémoire, la décision de créer des épreuves lexicales spécifiques dans le cadre de cette étude s'est révélée judicieuse (Hargrave & Sénéchal, 2000; Kirkpatrick, 2003; Pollard-Durodola et al., 2011). En effet, en nous basant uniquement sur les résultats obtenus aux épreuves standardisées, nous aurions pu conclure, à tort, que la LD n'était pas en mesure d'accroître significativement le stock lexical passif d'enfants présentant une DI d'un niveau de développement intellectuel non verbal d'environ quatre ans. Or, les épreuves lexicales spécifiques, plus sensibles, ont objectivé une progression significative du lexique cible sur les deux versants.

Enfin, il est important de souligner le nombre réduit de participants sur lequel se basent ces interprétations et la prudence avec laquelle celles-ci doivent donc être considérées. Pour avoir une idée plus précise des enseignements à tirer de ce programme de recherche, il faudra que puissent être prises en compte les données qui seront collectées par les étudiantes des promotions 2021-2022 et 2022-2023. L'augmentation des effectifs dans les deux groupes permettra de donner aux tests statistiques une plus grande puissance et diminuera les effets aléatoires. Par exemple, il est possible que le « progrès » réalisé par le GS à l'épreuve standardisée de vocabulaire en réception tienne à un effet « test-retest », mais aussi à un aléa statistique inhérent au faible effectif du groupe. De plus, le

nombre de séances d'entraînement en LD a malheureusement dû être limité en raison de la situation sanitaire. Seuls 17 à 61% du temps de LD initialement prévu a pu être atteint. Nous pouvons donc présumer que les résultats sont en deçà de ce qui aurait pu être objectivé suite à l'entraînement initialement planifié.

Conclusion

Ce mémoire s'intègre dans un projet de recherche visant à démontrer l'efficacité de la Lecture Dialogue sur le développement langagier (lexical et morphosyntaxique) d'enfants présentant une déficience intellectuelle. Pour cela, nous avons créé des épreuves lexicales spécifiques à notre étude. Ce choix se justifie par la confrontation des données de la littérature avec les caractéristiques de notre intervention, brève et ciblée, et les particularités développementales, langagières et cognitives, de la population étudiée. Nous supposons que la sensibilité plus importante des épreuves spécifiques dans ce type d'étude, par rapport aux épreuves standardisées, permettra de démontrer des effets significatifs entre le pré- et le post-test pour le groupe entraîné en vocabulaire. Les épreuves spécifiques ont été rigoureusement élaborées, leur validité a été évaluée et leurs résultats au pré- et au post-test ont été confrontés à ceux des épreuves standardisées.

Les résultats satisfaisants d'accord sur le nom et sur l'image des items cibles et des distracteurs composant les épreuves étayent la validité de celles-ci. L'interprétation des résultats aux tests non paramétriques effectués sur les données chiffrées issues des épreuves spécifiques et standardisées gagne donc en crédibilité. Des analyses nous ont permis d'objectiver un bénéfice significatif des séances de LD sur le vocabulaire « cible » chez les enfants avec déficience intellectuelle entraînés en vocabulaire. Grâce aux épreuves créées, nous avons pu mettre en évidence des progrès lexicaux expressifs, qu'objectivent également les épreuves standardisées, mais également réceptifs, ce que les résultats des tests standardisés n'illustrent pas de manière significative.

Notre étude s'intéresse aux enfants d'âge intellectuel environ égal à 4 ans et les données présentées ne concernent qu'un échantillon de 22 participants. Celle-ci se poursuivra durant l'année à venir et la fiabilité des conclusions augmentera conjointement avec la taille de l'échantillon. Cependant, la littérature concernant la lecture dialogue dans le cadre de la déficience intellectuelle reste pauvre. D'autres études, s'intéressant à des publics d'âges intellectuels différents, proposant un plus grand nombre de séances ou une intervention plus longue, seraient extrêmement pertinentes. Nous espérons que les chercheurs intéressés constateront, grâce aux données présentées, tout l'intérêt de recourir aux épreuves spécifiques.

Bibliographie

- Arnold, D. H., Lonigan, C. J., Whitehurst, G. J., & Epstein, J. N. (1994). *Accelerating language development through picture book reading : Replication and extension to a videotape training format*. 86(2), 235-243. <https://doi.org/10.1037/0022-0663.86.2.235>
- Bellon, M. L., Ogletree, B. T., & Harn, W. E. (2000). Repeated storybook reading as a language intervention for children with autism : A case study on the application of scaffolding. *Focus on Autism and Other Developmental Disabilities*, 15(1), 52-58. <https://doi.org/10.1177/108835760001500107>
- Biemiller, A. (2006). Vocabulary development and instruction : A prerequisite for school learning. In *Handbook of Early Literacy Research: Vol. II* (p. 41-51). <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.476.8485&rep=rep1&type=pdf>
- Bragard, A., Schelstraete, M.-A., Collette, E., & Grégoire, J. (2010). Évaluation du manque du mot chez l'enfant : Données développementales récoltées auprès d'enfants francophones de sept à 12 ans. *European Review of Applied Psychology*, 60(2), 113-127. <https://doi.org/10.1016/j.erap.2009.11.003>
- Burgoyne, K., Gardner, R., Whiteley, H., Snowling, M. J., & Hulme, C. (2018). Evaluation of a parent-delivered early language enrichment programme : Evidence from a randomised controlled trial. *Journal of Child Psychology and Psychiatry*, 59(5), 545-555. <https://doi.org/10.1111/jcpp.12819>
- Chalard, M., Bonin, P., Méot, A., Boyer, B., & Fayol, M. (2003). Objective age-of-acquisition (AoA) norms for a set of 230 object names in French : Relationships with psycholinguistic variables, the English data from Morrison et al. (1997), and naming latencies. *European Journal of Cognitive Psychology*, 15(2), 209-245. <https://doi.org/10.1080/095414402440000076>
- Cole, K. N., Dale, P. S., & Mills, P. E. (1990). Defining language delay in young children by cognitive referencing : Are we saying more than we know? *Applied Psycholinguistics*, 11(3), 291-302. <https://doi.org/10.1017/S0142716400008900>
- Crain-Thoreson, C., & Dale, P. S. (1999). Enhancing linguistic performance : Parents and teachers as book reading partners for children with language delays. *Topics in Early Childhood Special Education*, 19(1), 28-39. <https://doi.org/10.1177/027112149901900103>
- Dale, P. S., Crain-Thoreson, C., Notari-Syverson, A., & Cole, K. (1996). Parent-child book reading as an intervention technique for young children with language delays. *Topics in Early Childhood Special Education*, 16(2), 213-235. <https://doi.org/10.1177/027112149601600206>
- DeBaryshe, B. D., Rodarmel, S. L., Daly, B. A., & Huntley, L. (1992). *Shared picture book reading in the home : A language enrichment program for Head Start children*. Conference on Human

Development, Atlanta, GA.

- Dickinson, D. K., & Porche, M. V. (2011). Relation between language experiences in preschool classrooms and children's kindergarten and fourth-grade language and reading abilities: Preschool language experiences and later language and reading. *Child Development*, 82(3), 870-886. <https://doi.org/10.1111/j.1467-8624.2011.01576.x>
- Dunning, D. B., Mason, J. M., & Stewart, J. P. (1994). Reading to preschoolers: A response to Scarborough and Dobrich (1994) and recommendations for future research. *Developmental Review*, 14(3), 324-339. <https://doi.org/10.1006/drev.1994.1012>
- Ebbels, S. H. (2017). Intervention research: Appraising study designs, interpreting findings and creating research in clinical practice. *International Journal of Speech-Language Pathology*, 19(3), 218-231. <https://doi.org/10.1080/17549507.2016.1276215>
- Elleman, A., Lindo, E., Morphy, P., & Compton, D. (2009). The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. *Journal of Research on Educational Effectiveness*, 2(1), 1-44. <https://doi.org/10.1080/19345740802539200>
- Fey, M. E., Long, S. H., & Cleave, P. L. (1994). Reconsideration of IQ criteria in the definition of specific Language Impairment. In R. V. Watkins & M. L. Rice (Eds.), *SLI in children*. (pp. 161-178). Baltimore Paul H. Brookes Publishing.
- Fleury, V. P., & Schwartz, I. S. (2017). A modified dialogic reading intervention for preschool children with autism spectrum disorder. *Topics in Early Childhood Special Education*, 37(1), 16-28. <https://doi.org/10.1177/0271121416637597>
- Grygas Coogle, C., Floyd, K. K., & Rahn, N. L. (2018). Dialogic reading and adapted dialogic reading with preschoolers with autism spectrum disorder. *Journal of Early Intervention*, 40(4), 363-379. <https://doi.org/10.1177/1053815118797887>
- Hargrave, A. C., & Sénéchal, M. (2000). A book reading intervention with preschool children who have limited vocabularies: The benefits of regular reading and dialogic reading. *Early Childhood Research Quarterly*, 15(1), 75-90. [https://doi.org/10.1016/S0885-2006\(99\)00038-1](https://doi.org/10.1016/S0885-2006(99)00038-1)
- INSERM. (2016). *Déficiences intellectuelles*. EDP Sciences.
- Karrass, J., & Braungart-Rieker, J. M. (2005). Effects of shared parent-infant book reading on early language acquisition. *Journal of Applied Developmental Psychology*, 26(2), 133-148. <https://doi.org/10.1016/j.appdev.2004.12.003>
- Kirkpatrick, A. (2003). *A dialogic reading intervention programme for parents and young children*. 7.
- Koppenhaver, D. A., Erickson, K. A., Harris, B., McLellan, J., Skotko, B. G., & Newton, R. A. (2001). Storybook-based communication intervention for girls with Rett syndrome and their mothers.

- Disability and Rehabilitation*, 23(3-4), 149-159. <https://doi.org/10.1080/09638280150504225>
- Korat, O., Shamir, A., & Heibal, S. (2013). Expanding the boundaries of shared book reading : E-books and printed books in parent-child reading as support for children's language. *First Language*, 33(5), 504-523. <https://doi.org/10.1177/0142723713503148>
- Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika*, 2(3), 151-160. <https://doi.org/10.1007/BF02288391>
- Lété, B., Sprenger-Charolles, L., & Colé, P. (2004). MANULEX: A grade-level lexical database from French elementary school readers. *Behavior Research Methods, Instruments, & Computers*, 36, 156-166. <https://doi.org/10.3917/dbu.didle.2004.01.0241>
- Liboiron, N., & Soto, G. (2006). Shared storybook reading with a student who uses alternative and augmentative communication : A description of scaffolding practices. *Child Language Teaching and Therapy*, 22(1), 69-95. <https://doi.org/10.1191/0265659006ct298oa>
- Lonigan, C. J. (1993). Somebody read me a story : Evaluation of a shared reading program in low-income daycare. *Society for Research in Child Development Abstracts*, 9.
- Lonigan, C. J., Shanahan, T., & Cunningham, A. (2008). *Developing Early Literacy : Report of the National Early Literacy Panel* (p. 153-171). National Institute for Literacy.
- Lonigan, C. J., & Whitehurst, G. J. (1998). Relative efficacy of parent and teacher involvement in a shared-reading intervention for preschool children from low-income backgrounds. *Early Childhood Research Quarterly*, 13(2), 263-290. [https://doi.org/10.1016/S0885-2006\(99\)80038-6](https://doi.org/10.1016/S0885-2006(99)80038-6)
- Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-linguistic easy-access resource for phonological and orthographic neighborhood densities. *PLoS One*, 7(8). <https://doi.org/10.1371/journal.pone.0043230>
- Marulis, L. M., & Neuman, S. B. (2010). The effects of vocabulary intervention on young children's word learning : A meta-analysis. *Review of Educational Research*, 80(3), 300-335. <https://doi.org/10.3102/0034654310377087>
- Marulis, L. M., & Neuman, S. B. (2013). How vocabulary interventions affect young children at risk : A meta-analytic review. *Journal of Research on Educational Effectiveness*, 6(3), 223-262. <https://doi.org/10.1080/19345747.2012.755591>
- McGinty, A. S., Justice, L. M., Zucker, T. A., Gosse, C., & Skibbe, L. E. (2012). Shared-reading dynamics : Mothers' question use and the verbal participation of children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 55(4), 1039-1052. [https://doi.org/10.1044/1092-4388\(2011/10-0298\)](https://doi.org/10.1044/1092-4388(2011/10-0298))
- Mendelsohn, A. L., Mogilner, L. N., Dreyer, B. P., Forman, J. A., Weinstein, S. C., Broderick, M., Cheng, K. J., Magloire, T., Moore, T., & Napier, C. (2001). The impact of a clinic-based

- literacy intervention on language development in inner-city preschool children. *Pediatrics*, 107(1), 130-134. <https://doi.org/10.1542/peds.107.1.130>
- Mohammadian, A., & Dolatabadi, S. M. (2016). The effect of affection on English language learning of children with intellectual disability based on total physical response method of language teaching. *International Journal of English Language and Literature Studies*, 5(2), 92-103. <https://doi.org/10.18488/journal.23/2016.5.2/23.2.92.103>
- National Reading Panel (U.S.). (2000). *Report of the National Reading Panel : Teaching children to read : An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction : Reports of the subgroups*. National Institute of Child Health and Human Development, National Institutes of Health.
- Nunally, J. C. (1978). *Psychometric Theory* (2ème Edition).
- Piérart, B. (2005). L'évaluation du lexique de l'enfant. Un instrument diagnostique : ISADYLE. In *Le langage de l'enfant. Comment l'évaluer?* (p. 65-82). De Boeck Université. <https://doi.org/10.3917/dbu.piera.2005.01.0065>
- Pile, E. J. S., Girolametto, L., Johnson, C. J., Chen, X., & Cleave, P. L. (2010). Shared book reading intervention for children with language impairment : Using parents-as-aides in language intervention. *Canadian Journal of Speech-Language Pathology and Audiology*, 34(2), 14.
- Pollard-Durodola, S. D., Gonzalez, J. E., Simmons, D. C., Kwok, O., Taylor, A. B., Davis, M. J., Kim, M., & Simmons, L. (2011). The effects of an intensive shared book-reading intervention for preschool children at risk for vocabulary delay. *Exceptional Children*, 77(2), 161-183. <https://doi.org/10.1177/001440291107700202>
- Proctor-Williams, K., & Fey, M. E. (2007). Recast density and acquisition of novel irregular past tense verbs. *Journal of Speech, Language, and Hearing Research*, 50(4), 1029-1047. [https://doi.org/10.1044/1092-4388\(2007/072\)](https://doi.org/10.1044/1092-4388(2007/072))
- Pruthi, G. (2013). Language development in children with mental retardation. *Dynamical Psychology*. <http://dynapsyc.org/2013/LanguageDevelopment.pdf>
- Rondal, J. A. (1975). Développement du langage et retard mental : Une revue critique de la littérature en langue anglaise. *L'Année Psychologique*, 75(2), 513-547. <https://doi.org/10.3406/psy.1975.28111>
- Rosenshine, B. (2010, mars). *A comparison of results obtained using standardized tests and researcher-developed tests in intervention studies*. Annual meeting of Society for Research on Educational Effectiveness. <https://pdfs.semanticscholar.org/ab33/c33f730fb60da0d8bba5517cc2ac43d2e10e.pdf>
- Scarborough, H. S., & Dobrich, W. (1994). On the efficacy of reading to preschoolers. *Developmental Review*, 14(3), 245-302. <https://doi.org/10.1006/drev.1994.1010>

- Sénéchal, M. (1997). The differential effect of storybook reading on preschoolers' acquisition of expressive and receptive vocabulary. *Journal of Child Language*, 24(1), 123-138. <https://doi.org/10.1017/S0305000996003005>
- Sénéchal, M. (2010). Reading books to young children : What it does and does not do. In D. Aram & O. Korat (Éds.), *Literacy Development and Enhancement Across Orthographies and Cultures* (Vol. 2, p. 111-122). Springer US. https://doi.org/10.1007/978-1-4419-0834-6_8
- Sénéchal, M., & Cornell, E. (1993). Vocabulary acquisition through shared reading experiences. *Reading Research Quarterly*, 28(4), 360-374. <https://doi.org/10.2307/747933>
- Sénéchal, M., Thomas, E., & Monker, J.-A. (1995). Individual differences in 4-year-old children's acquisition of vocabulary during storybook reading. *Journal of Educational Psychology*, 87(2), 218-229. <https://doi.org/10.1037/0022-0663.87.2.218>
- Siegel, G. M. (1963a). Adult verbal behavior in « play therapy » sessions with retarded children. Part I. *Journal of Speech & Hearing Disorders. Monograph Supplement*, 10, 34-38.
- Siegel, G. M. (1963b). Adult verbal behavior with retarded children labeled as « high » or « low » in verbal ability. *American Journal of Mental Deficiency*, 68(3), 417-424.
- Siegel, G. M., & Donovan, K. E. (1964). Verbal behavior of retarded children in interpersonal assemblies with adults. *American Journal of Mental Deficiency*, 69, 244-252.
- Siegel, G. M., & Harkins, J. P. (1963). Verbal behavior of adults in two conditions with institutionalized retarded children. Part II. *Journal of Speech & Hearing Disorders. Monograph Supplement*, 10, 39-46.
- Skotko, B. G., Koppenhaver, D. A., & Erickson, K. A. (2004). Parent reading behaviors and communication outcomes in girls with Rett syndrome. *Exceptional Children*, 70(2), 145-166. <https://doi.org/10.1177/001440290407000202>
- Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures : Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6(2), 174-215. <https://doi.org/10.1037/0278-7393.6.2.174>
- Stephenson, J. (2009). Picture-book reading as an intervention to teach the use of line drawings for communication with students with severe intellectual disabilities. *Augmentative and Alternative Communication*, 25(3), 202-214. <https://doi.org/10.1080/07434610903031216>
- Thibaut, J.-P., Grégoire, J., & Lion, P. (2001). Sur les difficultés de l'évaluation du lexique : Une perspective développementale. *Glossa*, 76, 52-61.
- Valdez-Menchaca, M. C., & Whitehurst, G. J. (1992). Accelerating language development through picture book reading : A systematic extension to Mexican day care. *Developmental Psychology*, 28(6), 1106-1114. <https://doi.org/10.1037/0012-1649.28.6.1106>

- Van Kleeck, A., Vander Woude, J., & Hammett, L. (2006). Fostering literal and inferential language skills in head start preschoolers with language impairment using scripted book-sharing discussions. *American Journal of Speech-Language Pathology*, 15(1), 85-95. [https://doi.org/10.1044/1058-0360\(2006/009\)](https://doi.org/10.1044/1058-0360(2006/009))
- Wasik, B. A., & Bond, M. A. (2001). Beyond the pages of a book : Interactive book reading and language development in preschool classrooms. *Journal of Educational Psychology*, 93(2), 243-250. <https://doi.org/10.1037/0022-0663.93.2.243>
- Whitehurst, G. J., Arnold, D. S., Epstein, J. N., Angell, A. L., Smith, M., & Fischel, J. E. (1994). A picture book reading intervention in day care and home for children from low-income families. *Developmental Psychology*, 30(5), 679-689. <https://doi.org/10.1037/0012-1649.30.5.679>
- Whitehurst, G. J., Falco, F. L., Lonigan, C. J., Fischel, J. E., DeBaryshe, B. D., Valdez-Menchaca, M. C., & Caulfield, M. (1988). Accelerating language development through picture book reading. *Developmental Psychology*, 24(4), 552-559. <https://doi.org/10.1037/0012-1649.24.4.552>
- Winick, D. M., Avallone, A. P., Smith, C. E., & Crovo, M. (2009). *Reading Framework for the 2009 National Assessment of Educational Progress*. National Assessment of Educational Progress. <https://www.nagb.gov/assets/documents/publications/frameworks/reading/2009-reading-framework.pdf>
- Zevenbergen, A. A., & Whitehurst, G. J. (2003). Dialogic reading : A shared picture book reading intervention for preschoolers. In *On reading books to children : Parents and teachers*. (p. 170-192). Lawrence Erlbaum Associates Publishers.

Liste des annexes

Annexe 1. Variables psycholinguistiques des items cibles et de leurs distracteurs

Annexe 2. Premier item exemple et item 1 de l'épreuve spécifique de lexique en production

Annexe 3. Premier item exemple et item 1 de l'épreuve spécifique de lexique en réception

Annexe 4. Liste des albums sélectionnés et des mots cibles associés

Annexe 5. Résultats des questionnaires d'accord sur l'image et d'accord sur le nom

Annexe 6. Comparaison des fréquences d'occurrence des mots ayant obtenu un accord sur le nom suffisant et insuffisant