

DEPARTEMENT ORTHOPHONIE
FACULTE DE MEDECINE
Pôle Formation
59045 LILLE CEDEX
Tél : 03 20 62 76 18
departement-orthophonie@univ-lille.fr



MEMOIRE

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

Julie GUILLAUD

soutenu publiquement en juin 2023

Le repérage et la prévention des troubles du langage oral chez l'enfant de 0 à 6 ans Une revue systématique de la littérature

MEMOIRE dirigé par

Sandrine MEJIAS, Maître de conférences, Enseignante au Département d'Orthophonie de Lille,
Université de Lille

Sophie RAVEZ, Orthophoniste, Enseignante au Département d'Orthophonie de Lille,
Université de Lille

Lille – 2023

Remerciements

Je tiens premièrement à remercier mes directrices de mémoire, Sandrine Mejias et Sophie Ravez, pour leur soutien et leur accompagnement tout au long de ce projet, et qui m'ont permis d'aboutir ce travail.

Merci également à tous les maîtres de stage qui m'ont accueillie et partagé leur expérience depuis le début de mes études, et grâce à qui j'ai pu me construire en tant que future professionnelle.

Je remercie aussi ma famille pour leur présence et leurs encouragements, et particulièrement mes parents pour leur soutien inconditionnel durant ces études. Merci à Guillaume pour m'avoir encouragée, soutenue et supportée pendant cette dernière année pleine de stress et de doutes.

Enfin, je remercie toutes mes copines lilloises, ma famille de cœur, avec qui j'ai passé de superbes années remplies de rires, de pleurs et surtout d'amour. Merci à elles pour ce soutien, cette force qui nous ont permis de venir à bout de ces cinq années d'études. Merci particulièrement à Clémentine pour avoir contribué à une partie de la réalisation de ce mémoire et pour le soutien durant cette dernière année.

Résumé :

Les troubles du langage oral constituent une part importante des prises en charge de l'orthophoniste. Il est important de dépister au plus tôt ces troubles afin de proposer une prise en charge adaptée précocement. L'objectif de cette étude est de faire un état des lieux des outils utilisés pour évaluer le langage oral chez les enfants de zéro à six ans, et d'en extraire les tâches spécifiques permettant le dépistage des troubles du langage. Nous avons donc réalisé une revue systématique afin de faire une synthèse de données probantes et récentes pour répondre au critère de pratique fondée sur les preuves. Nos recherches se sont basées sur 5 bases de données (PubMed, WebofScience, Google Scholar, Lillocat et ScienceDirect) qui ont permis d'inclure 42 articles dans la synthèse finale. Ces articles ont été analysés, les données concernant l'évaluation du langage oral et les qualités psychométriques en ont été extraites. Nous avons obtenu des outils évaluant chaque domaine du langage oral en compréhension et en expression et des tâches spécifiques pour chaque domaine, avec des qualités psychométriques acceptables, même si les tailles d'échantillon étaient globalement faibles. Ces résultats permettent d'avoir des données récentes et internationales sur le dépistage des troubles du langage oral.

Mots-clés :

Trouble du langage oral – dépistage – évaluation – préscolaire – revue systématique

Abstract :

Language disorders are an important part of the speech therapist's care. It's important to detect these disorders as early as possible in order to provide an appropriate treatment at an early stage. The aim of this study is to review the tests used to assess language in children aged zero to six years old, and to extract specific tasks that allow the identification of language disorders. Therefore, we conducted a systematic review to synthesize recent evidence to meet the criterion of evidence-based practice. Our search was based on 5 databases (PubMed, WebofScience, Google Scholar, Lillocat, and ScienceDirect), which resulted in 42 articles being included in the final synthesis. These articles were analyzed, and data concerning language assessment and psychometric quality were extracted. We obtained tests assessing each domain of language in comprehension and expression, and their specific tasks, with acceptable psychometric qualities, even if the sample sizes were globally small. These results provide recent and international data on screening for oral language disorders.

Keywords :

Language disorder - screening - assessment – preschool – systematic review

Table des matières

Introduction.....	1
Contexte théorique, buts et objectifs.....	1
1. Le projet LOLEMATH.....	1
2. Le langage oral.....	2
2.1. Le développement du langage.....	2
2.2. Les prérequis au langage oral.....	2
2.3. Les domaines du langage oral.....	3
3. Les troubles du langage oral.....	3
4. Le repérage.....	3
5. La prévention.....	4
6. L'évaluation du langage oral.....	4
6.1. Outils d'évaluation du langage oral.....	4
6.2. Qualités psychométriques.....	4
6.2.1. Validité.....	4
6.2.2. Fidélité.....	5
6.2.3. Sensibilité et spécificité.....	5
7. But du mémoire.....	5
8. Objectifs.....	5
Matériels et méthodes.....	6
1. Guide PRISMA de rédaction.....	6
2. Sources d'information et stratégie de recherche	6
3. Sélection des articles et critères d'éligibilité.....	7
4. Protocole de collecte des données.....	8
Résultats.....	8
1. Sélection des articles.....	8
2. Extraction des données.....	9
2.1. Domaines du langage et tâches spécifiques correspondantes.....	9
2.1.1. Prérequis.....	10
2.1.2. Prosodie.....	10
2.1.3. Articulation.....	10
2.1.4. Phonologie.....	10
2.1.5. Lexique.....	11
2.1.6. Morphosyntaxe, grammaire.....	12
2.1.7. Discours, narration.....	13
2.1.8. Pragmatique.....	13
2.1.9. Résumé des tâches spécifiques d'évaluation du langage oral.....	13
2.2. Évaluation de la qualité.....	14
2.2.1. Date de publication des articles.....	14
2.2.2. Taille des échantillons.....	15
2.2.3. Qualité psychométrique.....	15
Discussion.....	16
1. Synthèse des résultats obtenus.....	16
1.1. Les outils d'évaluation du langage oral pour les enfants de 0 à 6 ans.....	17
1.2. Les tâches spécifiques utilisées dans les tests d'évaluation du langage oral.....	17
2. Qualité et validité des résultats.....	19
2.1. Méthodologie de la revue systématique.....	19
2.2. Biais de l'étude.....	19
2.3. Qualité psychométrique des études.....	20
3. Apport pour la pratique clinique.....	21
4. Limites de l'étude.....	21
Conclusion.....	22

Bibliographie.....	23
Liste des annexes.....	28
Annexe n°1 : Tableau de résultats.....	28
Annexe n°2 : Tableau récapitulatif des outils d'évaluation.....	28

Introduction

Le développement du langage oral se réalise en plusieurs étapes et nécessite l'acquisition de prérequis. L'enfant doit développer des aspects langagiers formels (tels que la phonologie, le lexique, la syntaxe, la prosodie) ainsi que des aspects plutôt interactionnels (par exemple la cognition sociale, l'attention conjointe) (Kern, 2019). Quand les enfants ont des difficultés à mettre en place ces composants, cela peut entraîner des difficultés ou troubles du langage oral. L'orthophonie intervient grandement dans ce domaine. Mais pour pouvoir réaliser une prise en charge, encore faut-il avoir repéré les premiers signes d'appel qui peuvent faire penser aux troubles du langage. Le repérage et la prévention sont indispensables dans le cadre des troubles du langage oral car ils permettent une prise en charge la plus précoce possible afin que ces troubles aient le moins de répercussions sur la vie de l'enfant (Billard, 2004). Pour dépister ces troubles, les éducateurs, enseignants, parents, ou médecins peuvent observer l'enfant, remplir des questionnaires. Pour obtenir un diagnostic plus précis, l'orthophoniste intervient en utilisant des outils standardisés permettant d'évaluer le langage et la communication. Ces outils doivent répondre à des critères de qualité psychométrique tels que la validité, la fidélité, la sensibilité et la spécificité.

Ce mémoire consiste à réaliser une revue systématique de la littérature qui s'inscrit dans le cadre des troubles du langage oral pour faire un état des connaissances sur le repérage et la prévention de ces troubles chez l'enfant de zéro à six ans et en obtenir des données récentes et probantes. Ce mémoire aura pour objectif d'obtenir des outils permettant d'évaluer le langage oral chez les enfants de zéro à six ans, et d'en extraire les tâches spécifiques permettant de dépister les troubles du langage oral. Il s'inscrit dans le projet plus global LOLEMATH ayant plusieurs objectifs. Les données issues de ce mémoire pourront servir au développement de la plateforme informatique à destination des parents, des enseignants et des professionnels de santé.

Premièrement, nous développerons les aspects plus théoriques du sujet en ce qui concerne le projet LOLEMATH, le langage oral et ses troubles, le repérage et la prévention des troubles du langage oral, ainsi que des informations concernant l'évaluation du langage oral comprenant les outils et les qualités psychométriques. Nous appliquerons par la suite la méthodologie rigoureuse et reproductible d'une revue systématique de la littérature afin d'obtenir des outils d'évaluation du langage oral et les tâches qui les composent. Nous présenterons nos résultats et discuterons de ceux-ci.

Contexte théorique, buts et objectifs

1. Le projet LOLEMATH

Ce mémoire s'inscrit dans le projet LOLEMATH qui comporte plusieurs objectifs. Il vise à développer des tests de repérage en langage oral, langage écrit et cognition mathématique afin de les rendre disponibles pour les enseignants et les professionnels de santé. Il vise aussi à mettre à disposition des conseils de stimulation sur une plateforme internet qui sera accessible aux parents, enseignants et professionnels de santé. Enfin, il vise à suivre une cohorte d'enfants de manière longitudinale afin de pouvoir dégager des profils

d'enfants à risque de développer des difficultés en langage oral, langage écrit et/ou cognition mathématique.

Mon mémoire s'axera principalement sur l'aspect du développement du site internet afin de pouvoir y ajouter des repères théoriques à propos du repérage et de la prévention des troubles du langage oral chez l'enfant de zéro à six ans.

2. Le langage oral

2.1. Le développement du langage

Le développement du langage oral peut se diviser en plusieurs étapes : la période prélinguistique et la période linguistique. Dans la première période, le nourrisson va pouvoir être capable de reconnaître les phonèmes de sa langue et commencer à distinguer les premiers mots familiers, sans pouvoir encore les prononcer. C'est une capacité innée. Le début de la période linguistique est marqué par la prononciation des premiers mots, pour ensuite évoluer vers la formation de constructions morphosyntaxiques plus semblables à celles de l'adulte. En ce qui concerne la compréhension, elle évolue tout d'abord davantage au niveau lexical vers deux ans et demi, pour ensuite évoluer au niveau morphosyntaxique vers l'âge de trois ans et demi jusqu'à six ou sept ans (Delahaie, 2004).

L'acquisition du langage oral nécessite des prérequis. Kern (2019) en distingue deux types : les prérequis formels du langage et les prérequis de la cognition sociale. Premièrement, pour développer son langage, l'enfant présente dès son plus jeune âge des capacités de segmentation qui permettront de mettre en place les différents composants du langage tels que la phonologie, le lexique ou encore la syntaxe. Le nourrisson est aussi en mesure de repérer les indices prosodiques dans le langage : les pauses, les groupes rythmiques, la modulation de la hauteur vocale ainsi que les variations d'intensité de la prosodie. En plus de ces composantes formelles, des prérequis de la cognition sociale sont nécessaires. Ceux-ci sont plutôt liés aux actions et aux interactions. Ils concernent le regard, le sourire social, l'attention conjointe, le pointage, ainsi que le jeu symbolique. La mise en place de ces prérequis permet donc un développement harmonieux du langage oral.

2.2. Les prérequis au langage oral

Le développement du langage oral nécessite au préalable l'acquisition de pré-requis. Différentes compétences sont nécessaires afin de mettre en place le langage, telles que l'attention conjointe, le tour de rôle, l'imitation ou encore le pointage, comme énoncé précédemment. L'attention conjointe correspond au fait de porter son attention ensemble à un objet ou une situation. Elle permettrait ensuite, selon les auteurs, de développer les compétences lexicales (Soubeyrand et al., 2007). Le tour de rôle, quant à lui, est le processus permettant la communication entre deux personnes en respectant la prise de parole chacun son tour, en faisant varier différents paramètres. C'est un élément indispensable à acquérir afin d'avoir les codes des règles de communication. L'imitation est la capacité d'imiter les mouvements de la face de son interlocuteur, ainsi que ses gestes, et sa parole. Soubeyrand et al. (2007) expliquent que l'imitation permet une forme de communication dès la naissance. En ce qui concerne le pointage, il s'agit d'un système de communication précédant le système verbal, et permettant ensuite d'associer le geste à la parole. On peut l'observer tôt dans le développement de l'enfant, vers l'âge de douze mois environ. Il permet de désigner des

objets, faire une demande, partager avec l'adulte et également d'avoir un échange (Morgenstern et al., 2008).

2.3. Les domaines du langage oral

Le langage oral est une capacité complexe qui se divise en plusieurs domaines : la prosodie, l'articulation, la phonologie, le lexique, la morphosyntaxe, la pragmatique et la narration. Ces domaines s'expriment sur deux versants : expressif et réceptif (Colas et al., 2021). La prosodie d'une langue regroupe les règles de prononciation, qui incluent l'intonation, l'accentuation, le rythme, le débit et les pauses. Ces éléments se caractérisent par des variations de fréquence, de hauteur, d'intensité et/ou de durée (Soubeyrand et al., 2007). L'articulation (aussi parfois appelée phonétisme) est la prononciation isolée des sons de la parole, d'un point de vue moteur (Colas et al., 2021). La phonologie est une branche de la linguistique qui se concentre sur l'étude de l'organisation des sons de la parole afin de former des phrases. En ce qui concerne le langage oral, la phonologie est l'enchaînement des sons de la parole (Colas et al., 2021). Le lexique est le répertoire de vocabulaire acquis par l'enfant. Les premiers mots sont prononcés par l'enfant entre onze et quatorze mois. Entre seize et vingt mois environ, le lexique augmente fortement, on appelle cela l'explosion lexicale (Plaza, 2014). La morphosyntaxe est l'association des mots pour former des phrases. Vers trente mois, l'enfant peut associer correctement la plupart des phrases (Plaza, 2014). La pragmatique est l'utilisation du langage de manière communicative, en prenant en considération à la fois le contexte dans lequel la communication a lieu, mais aussi les connaissances de l'interlocuteur (Colas et al., 2021). Enfin, la narration apparaît chez l'enfant vers l'âge de cinq ans. Il s'agit de la capacité de produire un discours. L'enfant doit utiliser le langage de manière non contextuelle en évoquant des personnages et événements absents, tout en prenant en compte les connaissances de son interlocuteur. Il doit également être capable de marquer la relation entre des phrases successives (Plaza, 2014).

3. Les troubles du langage oral

Les troubles du langage oral sont des troubles affectant la communication. Dans la terminologie, on distingue les troubles des sons de la parole, les troubles du langage oral associés à une condition biomédicale (par exemple une déficience auditive, une déficience intellectuelle, un syndrome génétique), et les troubles développementaux du langage. Ces troubles peuvent affecter la phonologie, la syntaxe, le lexique, la pragmatique, le discours ainsi que l'apprentissage verbal et la mémoire (Bishop et al., 2017). Les difficultés et troubles du langage peuvent souvent être le premier signe d'alerte d'une autre pathologie (par exemple une déficience intellectuelle, un trouble du spectre de l'autisme). Il est donc d'autant plus important de repérer les enfants dont le langage démarre difficilement pour mettre rapidement en place une prise en charge adaptée (Kern, 2019).

4. Le repérage

Le repérage est très présent dans le domaine de l'orthophonie. Dans le cadre du langage oral, il peut être initié par les enseignants, les parents, ou le médecin de famille. Il semble donc nécessaire que tous ces acteurs soient sensibilisés aux points d'alerte des

troubles du langage oral afin de faire intervenir au plus vite les professionnels spécifiques auprès de l'enfant. De nombreux outils de repérage des troubles du langage oral existent tels que des questionnaires (par exemple le DPL 3) ou des tests rapides (comme l'ERTL 4 ; Dellatolas, 2005).

5. La prévention

D'après la définition de l'Organisation mondiale de la santé (OMS, 1948), "la prévention est l'ensemble des mesures visant à éviter ou réduire le nombre et la gravité des maladies, des accidents et des handicaps". Elle distingue trois types de prévention : primaire, secondaire et tertiaire, qui vont de pair avec les différents états dans lesquels les patients peuvent se retrouver dans une maladie. La prévention consistera donc à mettre en œuvre des moyens pour éviter l'apparition de maladies, appliquer des traitements quand elles font surface, et réduire les complications (Flageolet, 2008). La prévention intervient largement en orthophonie, notamment dans le domaine des troubles du langage oral.

6. L'évaluation du langage oral

6.1. Outils d'évaluation du langage oral

Il est crucial de repérer précocement les troubles du langage oral, ce qui peut être accompli à l'aide de divers outils. Les éducateurs, enseignants, parents ou médecins peuvent effectuer des observations ou remplir des questionnaires pour détecter les difficultés ou troubles du langage. Cependant, pour obtenir un diagnostic plus précis, il est nécessaire de consulter un spécialiste du langage (orthophoniste) qui utilise des outils plus spécifiques et normés pour évaluer le langage et la communication. La pose du diagnostic nécessite un avis pluridisciplinaire en consultant d'autres professionnels du médical et du paramédical afin d'avoir une vision plus globale du patient et de ses capacités (Kern, 2019).

6.2. Qualités psychométriques

La psychométrie est une branche de la psychologie qui met l'accent sur les méthodes de recherche, et qui a pour objectifs principaux d'étudier les mesures en psychologie et de développer des méthodes pour évaluer les caractéristiques individuelles (Bernaud, 2014). Ces mesures sont également exploitables pour les outils d'évaluation en orthophonie. C'est grâce à celles-ci que l'on peut évaluer la qualité d'un outil. L'élaboration de la qualité psychométrique se divise en plusieurs parties : la validité, la fidélité, la sensibilité et la spécificité. C'est à l'aide de ces mesures que l'on va évaluer la qualité des instruments d'évaluation du langage oral (Rondal, 1997).

6.2.1. Validité

La validité d'un test se rapporte à la corrélation entre ce qu'il mesure et ce qu'il est censé mesurer, en fonction des objectifs d'évaluation posés par les auteurs du test (Bernaud, 2014 ; Rondal, 1997). Il existe plusieurs aspects de la validité. Premièrement, la validité interne (ou validité de contenu) concerne les items d'une épreuve. Ceux-ci doivent être représentatifs et pertinents pour évaluer les compétences requises (Beech, 1994; Rondal, 1997). La validité prédictive consiste à évaluer la relation entre les résultats du test et la

performance des individus dans des tâches de la vie quotidienne qui utilisent les mêmes compétences (Rondal, 1997). Ensuite, la validité concourante consiste à évaluer la relation entre les résultats d'un test et ceux d'autres épreuves déjà validées qui mesurent les mêmes compétences (Beech, 1994). Pour terminer, la validité théorique se rapporte aux modèles théoriques de capacités cognitives et la validité de construction, quant à elle, vise à démontrer l'acceptabilité théorique d'un test (Bernaud, 2007; Rondal, 1997).

6.2.2. Fidélité

La fidélité (aussi appelée fiabilité dans de nombreuses études) correspond à la stabilité des résultats obtenus à travers une épreuve (Rondal, 1997). Pour qu'un outil soit considéré comme fiable, les résultats doivent être cohérents, quels que soient le moment de passation (fidélité test-retest), l'évaluateur (fidélité inter-juge ou inter-évaluateur) et l'extrait de l'épreuve utilisé (cohérence interne) (Bernaud, 2007).

6.2.3. Sensibilité et spécificité

La sensibilité d'un outil concerne sa capacité à discriminer des différences fines entre les sujets et à identifier les individus présentant un trouble. En ce qui concerne la sensibilité d'un test, cela correspond à sa capacité à exclure la présence d'un trouble chez les individus qui ne sont pas atteints (Plante & Vance, 1994).

7. But du mémoire

Le but de ce mémoire est de réaliser une revue systématique de la littérature afin d'avoir des données probantes sur le repérage et la prévention des troubles du langage oral chez l'enfant de zéro à six ans. Il consistera plus précisément à faire un état des lieux actuel des outils et des tâches spécifiques d'évaluation du langage oral en France et à l'étranger.

La revue systématique de la littérature permet d'évaluer le niveau de qualité et d'ampleur des preuves existantes sur une question ou un sujet d'intérêt selon une méthodologie standardisée et reproductible. La revue systématique identifie toutes les preuves publiées et non publiées sur un sujet en particulier dans le but d'avoir une vue d'ensemble objective et transparente sur ce sujet (Hemingway & Brereton, 2009).

Avant de réaliser cette revue, j'ai vérifié si mon projet de recherche correspondait bien à une revue systématique. Tout d'abord, j'ai recherché dans les bases de données si des revues systématiques avaient déjà été réalisées sur le sujet ciblé (Google Scholar, Cairn, Cochrane). D'après ces recherches, aucune revue récente sur ce sujet précis n'a été publiée. Il est donc d'autant plus intéressant et pertinent de réaliser une revue systématique sur ce thème afin de faire la synthèse des recherches et connaissances publiées dans la littérature scientifique.

Publier cette revue permettra aux personnes qui souhaitent s'informer sur ce sujet d'avoir une synthèse qualitative des connaissances issues de la littérature scientifique en excluant les problèmes de fiabilité. Cette revue serait surtout à destination des orthophonistes afin qu'ils aient une analyse globale des références sur le sujet et qu'ils puissent répondre au critère de pratique fondée sur les preuves (*Evidence Based-Pratice*, EBP), mais aussi aux autres professionnels de santé ainsi qu'aux enseignants.

8. Objectifs

La revue systématique a été réalisée pour répondre aux questions suivantes, se basant sur le modèle PICO (*Population/Problem, Intervention/Exposure, Comparison, Outcome*).

- a) Chez les enfants de 0 à 6 ans, quels outils peut-on utiliser pour évaluer les capacités en langage oral ?
- b) Quelles sont les tâches spécifiques utilisées dans les tests d'évaluation du langage oral ?

Matériels et méthodes

1. Guide PRISMA de rédaction

La rédaction de la revue systématique a été réalisée avec l'aide du guide PRISMA (Prisma 2020 Checklist). Celui-ci a permis de nous aider à réaliser les étapes nécessaires afin de rédiger une revue systématique de la littérature avec une liste contrôle. La majeure partie des points du guide a pu être abordée dans la rédaction de cette revue systématique.

2. Sources d'information et stratégie de recherche

La recherche d'articles a été menée sur cinq bases de données (PubMed, Webofscience, Google Scholar, Lillocat, ScienceDirect). Les mots-clés utilisés ont été établis grâce au logiciel MeSH en amont afin qu'ils soient les mêmes dans toutes les bases de données. Les connecteurs "OR" et "AND" ont été utilisés afin de préciser la recherche. Le tableau suivant expose les formules de recherche utilisées pour chaque base de données.

Tableau 1 : Équations de recherche selon les bases de données sélectionnées

Bases de données	Équation de recherche
PubMed	("language disorder" OR "language impairment" OR "specific language impairment" OR "developmental language disorder" OR "speech disorder" OR "language delay" OR "late talkers") AND ("assessment" OR "test" OR "screening" OR "diagnosis") AND ("preschool" OR "kindergarten" OR "nursery school") + 2010-2023 Filters applied: Newborn: birth-1 month, Infant: birth-23 months, Infant: 1-23 months, Preschool Child: 2-5 years. Date de recherche : 27/01/2023
WebOfScience	Topic : ("language disorder" OR "language impairment" OR "specific language impairment" OR "developmental language disorder" OR "speech disorder" OR "language delay" OR "late talkers") AND ("assessment" OR "test" OR "screening" OR "diagnosis") AND ("preschool" OR "kindergarten" OR "nursery school") + Publication year : 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023 Date de recherche : 27/01/2023
Google Scholar	("language disorder" or "developmental language disorder" or "language impairment" or "specific language impairment" or "speech disorder" or "language delay") AND ("test" or "assessment" or "screening") and ("preschool" or "kindergarten" or "nursery school") + 2010-2023 Date de recherche : 27/01/2023
Lillocat	Sujet contient ("language disorder" OR "language impairment" OR "specific language impairment" OR "developmental language disorder" OR "speech disorder" OR "language delay" OR "late talkers") AND ("assessment" OR "test" OR "screening" OR "diagnostic") AND ("preschool" OR "kindergarten" OR "nursery school") Date de publication : du 01/01/2010 au 31/12/2023 Date de recherche :30/01/2023
ScienceDirect	("language disorder" OR "language impairment" OR "language delay") AND (assessment OR screening OR test OR diagnostic) AND (preschool OR kindergarten OR "nursery school") + diagnostic dans les mots-clés + 2010-2023 Date de recherche : 27/01/2023

3. Sélection des articles et critères d'éligibilité

La sélection des études a été réalisée en double aveugle grâce à l'utilisation d'un tableur partagé (GoogleSheet). Les doublons ont été gardés en un unique exemplaire. Les titres et résumés ont d'abord été examinés. Puis les textes complets des articles jugés potentiellement pertinents ont été lus afin de vérifier qu'ils correspondaient bien à l'étude.

Les critères de sélection étaient les suivants :

1. Études portant sur l'élaboration d'un test d'évaluation du langage oral
2. Participants : enfants de 0 à 6 ans
3. Articles de 2010 à 2023
4. Langue de publication : français, anglais
5. Études incluant un groupe contrôle (enfants au développement typique).

Les critères d'exclusion étaient les suivants :

1. Études portant sur des personnes ayant des troubles associés aux troubles du langage (Troubles du Spectre de l'Autisme, trouble sensoriel, déficience intellectuelle, syndromes, paralysie cérébrale...)
2. Études portant sur l'évaluation d'un autre domaine que le langage oral (langage écrit, mathématique ...)
3. Études dans un contexte de bilinguisme
4. Tests ayant de mauvaises qualités psychométriques.

4. Protocole de collecte des données

Les informations de chaque étude et les conclusions principales ont été extraites grâce à un tableau (Annexe 1). Les données suivantes ont été analysées : nom du test, domaine qu'il évalue, but de l'étude, nombre et âge des participants, langue, détail de l'évaluation, tâches de l'évaluation, test de comparaison, résultats, conclusion, limites.

Résultats

1. Sélection des articles

Le diagramme de flux ci-après (Figure 1) expose la procédure de recherche. Il présente la sélection des articles issus des bases de données avec application des critères d'inclusion et d'exclusion. La formule de recherche a permis initialement de récolter 2693 articles. Après suppression des doublons, 2099 articles ont été évalués par lecture du titre. L'éligibilité par les titres a été réalisée en double aveugle avec Clémentine B., également étudiante en M2 en orthophonie. Après cette étape, 274 études ont été jugées pertinentes et ont été analysées en lisant le résumé. Cette phase a permis d'exclure 219 articles qui étaient hors sujet ou rencontraient des critères d'exclusion. Enfin, 55 articles ont été lus intégralement, ce qui a permis d'obtenir la sélection de 42 articles qui entrent dans la synthèse finale.

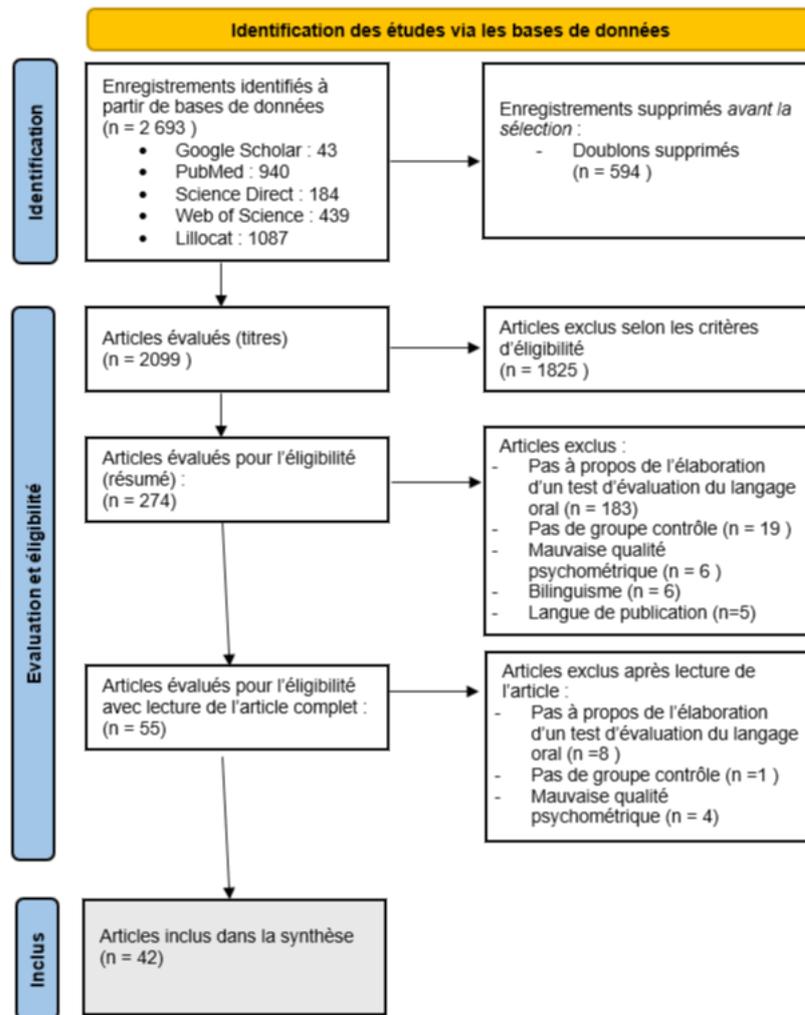


Figure 1 : Digramme de flux de la sélection des articles

Parmi les études incluses, 9 études sont des créations de tests, on compte 24 études qui évaluent la qualité psychométrique du test, 3 études sont des traductions de tests existants, 3 articles consistent à évaluer les performances d'enfants au développement typique par rapport à des enfants souffrant de troubles développementaux du langage oral grâce à un test, et enfin, 3 études sont une standardisation du test pour obtenir des normes.

2. Extraction des données

Le tableau ayant servi à l'extraction des données est disponible en Annexe 1. Il y comporte une synthétisation des résultats obtenus.

2.1. Domaines du langage et tâches spécifiques correspondantes

Chaque outil d'évaluation portait sur un ou plusieurs domaines du langage oral. Afin de pouvoir évaluer ces domaines, il fallait des tâches spécifiques. Les études sélectionnées, ayant un niveau de preuve et de fiabilité suffisant, ont permis d'obtenir quelles tâches pouvaient être utilisées pour les différents domaines.

2.1.1. Prérequis

Une seule des études sélectionnées rendait compte de l'évaluation des prérequis. L'étude de Lüke et al. (2017) avait pour objectif d'établir si l'évaluation de la communication préverbale (notamment ici le pointage) permettait de détecter les troubles du langage oral. La tâche de l'outil d'évaluation consistait à susciter le pointage impératif et le pointage déclaratif de jeunes enfants (à 12 mois, et à 24 mois). Les résultats montraient que le pointage avec l'index était une étape importante dans le développement de l'enfant et que c'était un moyen efficace d'évaluer l'enfant préverbal.

2.1.2. Prosodie

L'étude de Gibbon & Smyth (2013) étudiait le domaine de la prosodie chez les enfants âgés de quatre ans à quatre ans et onze mois, avec l'outil PEPS-C (*Profiling Elements of Prosody in Speech-Communication*). Cet outil d'évaluation comportait douze sous-tests qui évaluaient les quatre fonctions de communication de la prosodie : l'expression des émotions, la délimitation des unités linguistiques dans le discours, la closure des énoncés, les accents sur les éléments linguistiques. En réception, la prosodie était évaluée avec une tâche de discrimination auditive, en expression, elle était évaluée avec une tâche d'imitation.

2.1.3. Articulation

La sélection des articles avait permis d'inclure trois études évaluant l'articulation. Le DEMSS (*Dynamic Evaluation of Motor Speech Skill*) était un outil composé de neuf sous-tests permettant d'évaluer l'habileté motrice de la parole (Strand et al., 2013). Cet outil se composait de quatre types de scores : la précision articulatoire globale, la précision des voyelles, la précision prosodique et la cohérence, à l'aide d'une tâche d'imitation d'une série de mots, avec répétition si nécessaire. Le deuxième outil évaluant l'articulation (CAI : *Computer Articulation Instrument*) comportait quant à lui une tâche d'imitation de non-mots (van Haften et al., 2019). Cette étude incluait des participants âgés de deux à sept ans. Les tâches de répétition de mots et de répétition de non-mots étaient également utilisées dans l'étude de Granocchio et al. (2021). Il s'agissait d'une adaptation en italien du protocole Robbins et Klee. Cet outil comportait deux autres tâches : les praxies et l'observation des structures anatomiques des voies vocales. Ces épreuves permettaient également de rendre compte des compétences de production motrice des enfants âgés de deux ans et six mois à six ans et onze mois.

2.1.4. Phonologie

La phonologie pouvait être évaluée sur les versants réceptif et expressif. Les études permettaient, ensemble, de représenter des participants âgés de deux ans à dix ans et onze mois. L'étude de Rakhlin et al. (2021) établissait les propriétés psychométriques d'un sous-test de l'ALEF (*Arabic Language : Evaluation of Function*) évaluant la phonologie en réception. La tâche utilisée était une tâche de discrimination de pseudomots.

Sur le versant de la production, plusieurs études incluses évaluaient la phonologie. L'ALEF (Rakhlin et al., 2021) contenait une tâche d'articulation de mots pour évaluer ce domaine. Plusieurs études indiquaient que la tâche de dénomination d'image permettait également de coter la phonologie en production. Parmi ces études on retrouvait les outils suivants : LogoFoVa (Clausen & Fox-Boyer, 2022), Kurdish Speech Test (Syadar et al., 2019), MCELP-CS (*Mandarin Clinical Evaluation of Language for Preschooler's Core*

Scale ; Wu et al., 2020) et le CAI (van Haaften et al., 2019). Ce dernier outil contenait également d'autres tâches telles que l'imitation de non-mots, la répétition de mots et de non-mots et le taux de répétition maximum (diadococinésies). D'autres méthodes, telle que le calcul du Pourcentage de Consonnes Correctes (PCC) permettait d'évaluer l'intelligibilité, et donc la phonologie en production, sur une parole spontanée de l'enfant (Lagerberg et al., 2014). Enfin, des questionnaires parentaux pouvaient également être utilisés pour faire un compte rendu des capacités de l'enfant, comme par exemple l'ICS (*Intelligibility Context Scale ; McLeod et al., 2015*) ou le SIOLD (*Screening Identification of preschool Oral Language Difficulties ; Puglisi et al., 2020*).

D'autres outils permettaient d'évaluer la phonologie sur les deux versants (expressif et réceptif). Ces tests utilisaient tous la même tâche : la répétition de non-mots, qui intègre des processus d'entrée (réceptif) et de sortie (expressif). Parmi ces tests nous retrouvons : le GAPS (*Grammar And Phonology Screening ; van der Lely et al., 2011 ; Nash et al., 2011*), l'étude de Lavesson et al. (2018), l'étude de Kapalkova et al. (2013), le TR-NWR (*Turkish NonWord Repetition Test ; Topbaş et al., 2014*) ainsi que le Danish SI-3 (Bleses et al., 2010).

2.1.5. Lexique

Afin d'évaluer le lexique en réception, deux tâches ressortaient principalement : la désignation d'image sur mot entendu, et la reconnaissance des parties du corps. Nous pouvions retrouver la tâche de désignation d'image dans dix études. Ces études incluaient les tests suivants : une adaptation en grec du ROWPVT-II (Okalidou et al., 2011), l'ALEF (Rakhlin et al., 2021), le PING (*Picture Naming Game ; Bello et al., 2012*), le QUILS (*Quick Interactive Language Screener ; Pace et al., 2022*), l'étude de Lavesson et al. (2018), Logometro (Antoniou et al., 2022), le MCELP-CS (Wu et al., 2020) ainsi que le PDAS (*Preschool Developmental Assessment Scale ; Wong et al., 2012*). Deux études portaient sur la tâche de reconnaissance des parties du corps (Romski et al., 2018; Visser-Bochane, Luinge, et al., 2021) avec les outils suivants : MSEL (*Mullen*) et le *Dutch well child language screening protocol*.

A propos de l'évaluation du lexique en production, les recherches avaient permis de trouver huit études, utilisant toutes la tâche de dénomination d'images présente dans les outils suivants : ALEF, PING, PPVT, Logometro, MCELP-CS, Danish SI-3, PDAS, KHLA (Rakhlin et al., 2021 ; Bello et al., 2012 ; Katzenberger & Meilijson, 2014 ; Spaulding et al., 2013 ; Antoniou et al., 2022 ; Wu et al., 2020 ; Bleses et al., 2010 ; Wong et al., 2012).

D'autres études permettaient d'évaluer le lexique sur les deux versants (expressif et réceptif). Parmi celles-ci, quatre études comprenaient un questionnaire parental permettant d'indiquer quels mots sont compris et produits par l'enfant. On retrouve cette tâche dans les outils suivants : SIOLD (Puglisi et al., 2020), CDI-III (*Communicative Development Inventory*) (Kas et al., 2022), ELS (*Early Language Scale*) (Visser-Bochane, van der Schans, et al., 2021), SDDS (Smolík & Bytešníková, 2021). Les trois autres études permettant d'évaluer le lexique en réception et en expression contenaient toutes des tâches différentes. L'étude de Dias et al. (2020) comprenait une observation de l'enfant par l'expérimentateur afin d'avoir un aperçu de son lexique. Ensuite, une tâche d'organisation lexicale et de maîtrise de mots permettait d'évaluer l'accès lexical de l'enfant, c'est-à-dire la capacité de retrouver les mots en mémoire (Lavesson et al., 2018). Pour terminer, l'outil Logometro contenait une tâche consistant à donner la définition de mots (Antoniou et al., 2022).

Les études incluses pour évaluer le lexique sur chacun des versants, permettent de représenter des enfants de un an et sept mois à six ans et neuf mois. Le détail de l'âge des participants de chaque étude est disponible en Annexe 1.

2.1.6. Morphosyntaxe, grammaire

D'après les études sélectionnées, deux tâches permettaient d'évaluer la morphosyntaxe en réception. Tout d'abord, huit études contenaient la tâche de désignation d'image sur phrase entendue, qui consistait à pointer une image parmi plusieurs correspondant à la phrase que l'examineur (ou l'outil informatique) énonce. Nous pouvions trouver cette tâche dans les études qui suivent : ALEF (Rakhlin et al., 2021), LOGiK-S (*Logopädie im Kindergarten – Screening*) (Holzinger et al., 2022), l'étude de Lavesson et al. (2018), PSCT (*Persian Syntax Comprehension Test*) (Mohamadi et al., 2019), Logometro (Antoniou et al., 2022), MCELP-CS (Wu et al., 2020), SlaM (*Short Language Measure*) (Matov et al., 2020) et PDAS (Wong et al., 2012). Puis, dans des outils tels que le Token Test (Schmoeger et al., 2020), MSEL (Romski et al., 2018) et le Danish SI-3 (Bleses et al., 2010), on retrouvait la tâche d'exécution de commandes ou de manipulation d'objets selon une phrase entendue.

En ce qui concernait la morphosyntaxe en production, nous retrouvions diverses tâches. D'une part, on observait dans l'ALEF (Rakhlin et al., 2021), LOGiK-S (Holzinger et al., 2022) ainsi que PDAS (Wong et al., 2012) une tâche d'achèvement de phrases qui consistait pour l'enfant à terminer un énoncé amorcé par l'examineur. D'autre part, des études incluaient une tâche de description d'image. Cette tâche était incluse dans des outils tels que le QUILS (Pace et al., 2022), GLI (*Grammatically and utterance Length Instrument*) (Castilla-Earls & Fulcher-Rood, 2018), PGR (Pourcentage de Réponses Grammaticales) (Eisenberg & Guo, 2018) ainsi que le T-PEC (*Test di Produzione Elicitata di Clitici*) (Crocetti et al., 2021). Ensuite, les outils MCELP-CS (Wu et al., 2020) et SlaM (Matov et al., 2020) incluaient une tâche de répétition de phrases. Dans le MSEL (échelle de Mullen), la morphosyntaxe en production était évaluée à l'aide d'une tâche de réponse à des questions (Romski et al., 2018). Et pour terminer, le *Dutch well child screening protocol* évaluait la production de phrases de deux mots par des enfants de deux ans (Visser-Bochane, Luinge, et al., 2021).

Certaines études permettaient d'évaluer la grammaire sur le versant expressif et aussi sur le versant réceptif. Premièrement, dans le GAPS (van der Lely et al., 2011 ; Nash et al., 2011), on retrouve la tâche de répétition de phrases, qui, d'après les auteurs, incluait des processus d'entrée (réceptif) et de sortie (expressif). Ensuite, le KHLA (Katzenberger & Meilijson, 2014) proposait une tâche de questions sur des images pour évaluer la morphosyntaxe (compréhension des questions, et production d'une réponse). Dans leur étude, Dias et al. (2020) évaluaient la précision diagnostique d'un outil qui permettait d'observer les capacités linguistiques de l'enfant également sur le domaine de la morphosyntaxe. Pour terminer, trois outils étaient des questionnaires parentaux incluant des questions sur la compréhension et la production de phrases : SIOLD (Puglisi et al., 2020), CDI-III (Kas et al., 2022) ainsi que l'ELS (Visser-Bochane, van der Schans, et al., 2021). Les résultats portant sur l'évaluation de la morphosyntaxe ont porté sur des enfants âgés de un an et neuf mois à dix ans et onze mois. Les informations détaillées sur l'âge des participants de chaque étude peuvent être également consultées dans l'Annexe 1.

2.1.7. Discours, narration

Les études évaluant le discours contenaient toutes la tâche de narration d'histoire en image. Elles incluaient des participants âgés de deux ans et six mois à dix huit ans. Les six tests suivants incluaient cette tâche : ALEF (Rakhlin et al., 2021), NAP-2 (*Narrative Assessment Protocol-2*) (Bowles et al., 2020), KHLA (Katzenberger & Meilijson, 2014), MCELP-CS (Wu et al., 2020), l'évaluation dynamique de Romero et al. (2021) et le PDAS (Wong et al., 2012). Ce dernier outil contenait également une tâche de questions sur texte pour évaluer la compréhension de la narration.

2.1.8. Pragmatique

Trois études rendaient compte de l'évaluation de la pragmatique, sur une tranche d'âge de deux ans et un mois à dix ans et onze mois. Chacune d'entre elles contenait des tâches différentes pour l'évaluer. Tout d'abord, l'APLSI (Version arabe du *Pragmatic Language Skill Inventory*) permettait de faire l'inventaire des compétences de l'enfant concernant l'interaction de classe, l'interaction sociale et l'interaction personnelle (Alduais et al., 2022). L'ALEF (Rakhlin et al., 2021) cité précédemment utilisait une tâche de réponse à une question sur une situation. On évaluait alors si l'enfant parvenait à s'adapter à la situation. Enfin, l'étude de Khodeir et al. (2017), ayant pour but de standardiser l'EAPLT (*Egyptian Arabic Pragmatic Language Test*), contenait les tâches suivantes : une conversation, un entretien semi-dirigé et la narration d'une histoire en contexte.

2.1.9. Résumé des tâches spécifiques d'évaluation du langage oral

Le tableau suivant résume les tâches utilisées dans les études sélectionnées en fonction des domaines du langage oral.

Tableau 2 : Synthèse des tâches utilisées dans l'évaluation du langage oral

Domaines	Versant réceptif	Versant expressif
Prérequis		Susciter le pointage impératif et déclaratif
Prosodie	Discrimination auditive	Imitation des émotions Closure d'énoncés
Articulation		Répétition de mots Répétition de non-mots Praxies Observation des structures anatomiques vocales
Phonologie	Discrimination de pseudomots Répétition de non-mots	Articulation de mots Dénomination d'images Répétition de mots Répétition de non-mots Taux de répétition maximum (diadococinésies) Pourcentage de Consonnes Correctes Questionnaire parental
Lexique	Désignation d'images sur mot entendu Reconnaissance des parties du corps Questionnaire parental Observation de l'enfant Définition de mots Organisation lexicale	Dénomination d'images Questionnaire parental Observation de l'enfant Définition de mots Organisation lexicale
Morphosyntaxe	Désignation d'images sur phrase entendue Exécution de commande Répétition de phrases Questions sur images Questionnaire parental Observation de l'enfant	Achèvement de phrases Description d'une scène Répétition de phrases Réponse aux questions Production de phrases Répétition de phrases Questions sur images Questionnaire parental Observation de l'enfant
Narration	Questions sur images	Narration d'histoires sur images
Pragmatique		Inventaire de compétences pragmatiques (observation) Questions sur une situation Conversation Entretien semi-dirigé Narration d'une histoire en contexte

2.2. Évaluation de la qualité

Afin d'évaluer la qualité des outils d'évaluation du langage, nous nous étions basé sur différents aspects : la date de publication des articles, la taille de l'échantillon dans les études, et la qualité psychométrique des tests.

2.2.1. Date de publication des articles

Afin d'avoir des publications récentes, nous avons inséré dans nos critères d'inclusion que les études devaient avoir été publiées entre 2010 et 2023. La synthèse finale incluait 42 articles, parmi ceux-ci 14 ont été publiés entre 2010 et 2015 inclus, et 28 articles publiés entre 2017 et 2022 inclus. Nous pouvions compter le plus d'articles parus en 2021, avec 7 articles,

et on pouvait également compter 6 articles publiés en 2020 et 2022. Cette revue systématique se base donc sur une littérature scientifique récente.

2.2.2. Taille des échantillons

Premièrement, la taille des échantillons de chaque étude a été analysée. Seulement douze des études sélectionnées incluaient cent participants par sous-groupe ou plus. La moitié des études se basait sur des groupes de participants compris entre onze et quarante participants. On pouvait retrouver quatre études qui comprenaient très peu de participants (moins de dix) et qui devaient réaliser des études supplémentaires afin d'élargir le niveau de preuve à une population plus large. On pouvait ensuite retrouver trois articles qui comprenaient entre 41 et 71 participants. Enfin, un article comptait entre 71 et 99 participants, ce qui reste un peu faible pour obtenir un très bon niveau de qualité. Le tableau suivant permet d'avoir une vue synthétique des échantillons dans chaque étude. Afin de compléter ce tableau, la moyenne de la répartition des participants par tranche d'âge de chaque étude a été prise en compte (Tableau 3).

Tableau 3 : Taille des échantillons des études sélectionnées

Taille de l'échantillon	Nombre d'articles	Références
< 10	4	(Lagerberg et al., 2014; Romero et al., 2021; Spaulding et al., 2013; van der Lely et al., 2011)
11 < n < 40	21	(Bello et al., 2012; Castilla-Earls & Fulcher-Rood, 2018; Clausen & Fox-Boyer, 2022; Crocetti et al., 2021; Eisenberg & Guo, 2018; Gibbon & Smyth, 2013; Granocchio et al., 2021; Kapalkova et al., 2013; Khodeir et al., 2017; Matov et al., 2020; Nash et al., 2011; Okalidou et al., 2011; Pace et al., 2022; Rakhlin et al., 2021; Romski et al., 2018; Schmoeger et al., 2020; Smolík & Bytešníková, 2021; Strand et al., 2013; Syadar et al., 2019; Topbaş et al., 2014; Visser-Bochane, van der Schans, et al., 2021)
41 < n < 71	3	(Alduais et al., 2022; Bowles et al., 2020; Lüke et al., 2017)
71 < n < 99	1	(Wu et al., 2020)
≥ 100	12	(Antoniou et al., 2022; Bleses et al., 2010; Dias et al., 2020; Holzinger et al., 2022; Kas et al., 2022; Katzenberger & Meilijson, 2014; McLeod et al., 2015; Mohamadi et al., 2019; Puglisi et al., 2020; van Haaften et al., 2019; Visser-Bochane, Luinge, et al., 2021; Wong et al., 2012)

2.2.3. Qualité psychométrique

Les critères d'exclusion avaient permis, dès la sélection d'articles, d'écarter les études qui présentaient de mauvaises qualités psychométriques et qui n'étaient donc pas pertinentes pour notre synthèse. Pour chaque article inclus, j'ai donc vérifié qu'ils contenaient bien au moins une des qualités psychométriques nécessaires à l'élaboration d'un test de langage oral : validité, fidélité (ou fiabilité), sensibilité, spécificité. En ce qui concerne la validité, 27 articles contenaient cette qualité. Plus précisément, j'ai analysé seize études qui avaient une validité

globale, j'ai retrouvé un article ayant un critère de validité de contenu, deux articles ayant un critère de validité prédictive, dix articles ayant un critère de validité concourante (aussi appelée coefficient de corrélation dans de nombreuses études) et enfin quatre articles ayant un critère de validité de construction. Au niveau de la fidélité, j'ai pu intégrer 22 articles présentant cette qualité. La fidélité test-retest était observée dans trois études, j'ai pu analyser la fidélité inter-juges dans quatre études, le critère de cohérence interne était présent dans cinq études, et quinze études présentaient des critères de fidélité globale. Pour terminer, onze études étaient sensibles et treize étaient analysées comme ayant une bonne spécificité. Le tableau suivant (Tableau 4) résumait quelles études contenaient les différentes qualités psychométriques.

Tableau 4 : Résumé des qualités psychométriques des études

Qualité psychométrique	Nombre d'articles	Références
Validité	27	(Alduais et al., 2022; Antoniou et al., 2022; Bello et al., 2012; Bleses et al., 2010; Bowles et al., 2020; Castilla-Earls & Fulcher-Rood, 2018; Eisenberg & Guo, 2018; Kas et al., 2022; Khodeir et al., 2017; Lagerberg et al., 2014; Lavesson et al., 2018; Lüke et al., 2017; Matov et al., 2020; Mohamadi et al., 2019; Nash et al., 2011; Okalidou et al., 2011; Smolík & Bytešniková, 2021; Spaulding et al., 2013; Strand et al., 2013; Syadar et al., 2019; Topbaş et al., 2014; van der Lely et al., 2011; van Haaften et al., 2019; Visser-Bochane, Luinge, et al., 2021; Visser-Bochane, van der Schans, et al., 2021; Wong et al., 2012; Wu et al., 2020)
Fidélité	22	(Antoniou et al., 2022; Bello et al., 2012; Bleses et al., 2010; Clausen & Fox-Boyer, 2022; Crocetti et al., 2021; Gibbon & Smyth, 2013; Granocchio et al., 2021; Kapalkova et al., 2013; Katzenberger & Meilijson, 2014; Khodeir et al., 2017; Lagerberg et al., 2014; Lavesson et al., 2018; Matov et al., 2020; McLeod et al., 2015; Rakhlin et al., 2021; Strand et al., 2013; Syadar et al., 2019; Topbaş et al., 2014; van der Lely et al., 2011; van Haaften et al., 2019; Wong et al., 2012; Wu et al., 2020)
Sensibilité	11	(Crocetti et al., 2021; Dias et al., 2020; Holzinger et al., 2022; Katzenberger & Meilijson, 2014; Lavesson et al., 2018; McLeod et al., 2015; Pace et al., 2022; Puglisi et al., 2020; Romero et al., 2021; Topbaş et al., 2014; Visser-Bochane, Luinge, et al., 2021)
Spécificité	13	(Crocetti et al., 2021; Dias et al., 2020; Holzinger et al., 2022; Lavesson et al., 2018; Lüke et al., 2017; McLeod et al., 2015; Nash et al., 2011; Pace et al., 2022; Puglisi et al., 2020; Romero et al., 2021; Topbaş et al., 2014; Visser-Bochane, Luinge, et al., 2021; Visser-Bochane, van der Schans, et al., 2021)

Discussion

1. Synthèse des résultats obtenus

Ce mémoire avait pour but de réaliser une revue systématique afin d'avoir des données probantes et récentes sur l'évaluation du langage oral. Les objectifs de cette revue systématique de la littérature étaient de récolter des outils d'évaluation du langage oral pour les enfants de zéro à six ans afin d'en extraire les tâches spécifiques et sensibles au dépistage des troubles du langage oral. Nous avons analysé des articles trouvés dans six bases de

données en appliquant les critères d'inclusion et d'exclusion. L'analyse s'est divisée en plusieurs étapes : la lecture des titres s'est réalisée en double aveugle, ensuite nous avons lu les résumés des articles sélectionnés, et pour finir, les études éligibles ont été lues intégralement. Ces étapes nous ont permis d'inclure dans notre synthèse 42 études ayant des qualités psychométriques suffisantes, malgré de nombreux articles dont la taille de l'échantillon reste trop faible.

1.1. Les outils d'évaluation du langage oral pour les enfants de 0 à 6 ans

Un des objectifs de cette revue systématique était d'obtenir des outils d'évaluation du langage oral récents pour les enfants de zéro à six ans. La recherche bibliographique nous a permis de rassembler les études traitant de l'élaboration, de la validation ou de la normalisation d'outils d'évaluation du langage oral. Ces études étaient publiées entre 2010 et 2023 afin d'avoir des outils récents et étant en lien avec les données scientifiques actuelles. Nous avons donc pu collecter 42 outils conçus en France ou dans d'autres pays du monde. Les tests découverts ont permis d'évaluer l'ensemble des compétences du langage oral, sur les deux versants. Le tableau disponible en Annexe 2 résume tous les outils d'évaluation du langage oral que nous avons intégrés à l'étude.

1.2. Les tâches spécifiques utilisées dans les tests d'évaluation du langage oral

Le second objectif était de récolter les tâches spécifiques issues des tests d'évaluation du langage qui permettent de dépister les troubles du langage oral. Afin d'évaluer le pointage, qui est un élément important pour le développement du langage oral, et faisant partie des prérequis, les résultats indiquent qu'il est possible de susciter le pointage impératif et déclaratif afin de voir si les enfants parviennent à le faire. L'étude ayant évalué ce domaine a permis de montrer qu'il est possible d'évaluer les éventuels troubles du langage avant même que l'enfant ne parle (Lüke et al., 2017). L'évaluation précoce du langage permet ensuite de prévenir et repérer les troubles le plus tôt possible, afin qu'une prise en charge adaptée soit rapidement mise en place. Cependant, les autres pré-requis du langage oral n'étaient pas évalués dans les études sélectionnées. Ce point sera plus largement abordé dans la partie « Limites » de la discussion.

Pour évaluer la prosodie, deux tâches étaient proposées pour l'expression (l'imitation des émotions et la closure d'énoncés) et une tâche pour la réception (la discrimination auditive). Ces tâches provenaient toutes du test PEPS-C (Gibbon & Smyth, 2013) qui était basé sur des bases théoriques fiables. L'évaluation de la communication à travers les indices prosodiques semble donc importante, même si elle reste peu proposée en clinique. L'étude n'était cependant basée que sur des participants âgés de quatre ans à quatre ans et onze mois.

Les résultats indiquent ensuite que les tâches de répétition de mots et la répétition de non-mots permettent d'évaluer la prononciation, d'un point de vue moteur (articulation) et également d'un point de vue phonologique. Ces tâches étaient incluses dans plusieurs études. Pour évaluer l'articulation, les praxies ou l'observation des structures anatomiques vocales sont également utilisées et validées. D'autres épreuves sont largement utilisées dans l'évaluation de la phonologie en production : l'articulation de mots, la dénomination d'images, le taux de répétition maximum (diadococinésies), le pourcentage de consonnes correctes. Au niveau réceptif, la tâche la plus répandue est la discrimination de pseudomots.

Pour les domaines de l'articulation et de la phonologie, des enfants de deux ans à dix ans et onze mois étaient inclus. L'évaluation de ces domaines est donc possible quand le langage commence à se stabiliser et que l'enfant peut prononcer des mots et des phrases courtes.

En ce qui concerne le lexique, le stock lexical actif est toujours évalué par une tâche de dénomination d'images ; huit études en prouvent la validité. Le stock lexical passif est lui souvent évalué par une tâche de désignation d'images sur mot entendu ou une tâche de reconnaissance des parties du corps. Il est important également d'évaluer ce domaine du langage car il peut lui aussi être la cause de difficultés ou d'un trouble du langage oral. Nos résultats permettent de balayer les enfants d'une tranche d'âge d'un an et sept mois à six ans et neuf mois. Des tâches telles que l'observation de l'enfant ou un questionnaire parental permettent d'aborder l'évaluation de ce domaine quand l'enfant n'a que peu ou pas de langage afin de repérer d'une manière qualitative s'il pourrait présenter des troubles du langage oral. Par exemple si un enfant de deux ans ne prononce régulièrement qu'une dizaine de mots, on pourra conclure qu'il présente des difficultés de langage, car en effet, vers l'âge de vingt mois, se produit l'explosion lexicale (Plaza, 2014). Les questionnaires parentaux et l'observation de l'enfant, dans les résultats, sont également utilisés pour rendre compte des capacités en phonologie ainsi qu'en morphosyntaxe. Le lexique, notamment du point de vue sémantique, c'est-à-dire à propos du sens des mots, est également évalué avec des tâches de définition de mots et d'organisation lexicale.

Au niveau morphosyntaxique, l'étude de (Visser-Bochane, Luinge, et al., 2021) indique que les enfants de deux ans sont capables de produire des phrases de deux mots. Dans cette étude, cette capacité est évaluée d'une manière qualitative. Dans d'autres études, pour des enfants plus âgés, capables de produire des phrases plus longues, les résultats indiquent l'utilisation de tâches d'achèvement de phrases, de description d'une scène imagée, de répétition de phrases et de réponses à des questions. Les tâches d'achèvement, de description et de répétition étaient incluses dans plusieurs outils. Sur le versant réceptif de la morphosyntaxe, deux tâches étaient utilisées dans les onze études évaluant ce domaine : la désignation d'images sur phrase entendue et l'exécution de commandes.

Les outils intégrant l'évaluation de la narration contenaient tous la tâche de narration d'histoire sur image pour l'expression, et la tâche de questions sur texte pour la réception. Cette capacité apparaît vers l'âge de cinq ans. En pratique clinique, c'est rarement le premier signe apparent d'un trouble du langage oral (Gillon, 2017). Les études incluaient des participants âgés de deux ans et six mois à dix huit ans, mais l'évaluation spécifique de la narration portait sur des enfants à partir de quatre ans. En effet les études ayant des participants à partir de deux ans contenaient d'autres tâches évaluant les autres domaines.

Pour terminer, diverses tâches permettent d'évaluer la pragmatique chez des enfants de deux ans et un mois à dix ans et onze mois. Un inventaire des compétences pragmatiques par l'observation permet au praticien d'analyser le patient dans différents contextes de communication, avec différentes personnes (parents, enseignants, pairs) pouvant aider à repérer les difficultés de l'individu à s'adapter aux différentes situations sociales (Alduais et al., 2022). Ensuite, l'utilisation de questionnaires, ou un entretien semi-dirigé permettent au clinicien de collecter des informations sur les compétences du patient. La pragmatique peut également être évaluée grâce à une tâche de narration pour observer comment la personne utilise et organise son langage.

L'étude de tous ces articles a permis de nous rendre compte que parmi les résultats, beaucoup d'outils d'évaluation du langage utilisaient une même tâche pour évaluer le même

domaine du langage oral. Cela nous confirme que les données sont probantes car les tâches sont semblables et que les données psychométriques de tous les tests sont acceptables. Les études ont permis d'évaluer avec des outils standardisés des participants âgés de deux à sept ans. Des enfants de zéro à deux ans n'auraient pas pu être inclus, car en effet avant deux ans, les enfants ont une grande variabilité dans leur développement linguistique, ce qui rend difficile l'administration d'évaluations normées (Owens, 2016).

2. Qualité et validité des résultats

Nous avons évalué la qualité et la validité des résultats à travers plusieurs aspects. Tout d'abord, la méthodologie rigoureuse de la revue systématique, les possibles biais de sélection inclus dans la revue, et la qualité psychométrique des études.

2.1. Méthodologie de la revue systématique

Une revue systématique repose sur une méthodologie rigoureuse et reproductible. Les différentes étapes de réalisation d'une revue systématique ont donc été suivies. La première étape a permis de formuler la question de recherche PICO (*Patient, Intervention, Comparison, Outcome*) et de rédiger le protocole de recherche.

La seconde étape a consisté à rechercher les études dans les bases de données grâce aux formulations détaillées dans le tableau 1 dans la partie « Matériels et méthodes ». Nous avons choisi de traiter des articles publiés en français ainsi qu'en anglais afin d'obtenir un maximum de données. En effet, la langue anglaise permet une ouverture plus large à la littérature internationale et permet également d'avoir davantage d'études récentes. Les formulations de recherche bibliographique ont donc toutes été rédigées en anglais. Cela a peut-être pu écarter des articles uniquement en français, mais cela nous a permis d'obtenir des données probantes à l'international. Nous avons inclus dans cette étude des articles publiés entre 2010 et 2023 afin de garantir des données issues d'une littérature scientifique récente. En effet, les résultats indiquent que la majorité des articles a été publiée entre 2020 et 2023, et nous retrouvons la moitié des articles qui a été publiée après 2018.

En troisième étape, nous avons sélectionné les références grâce aux critères d'inclusion et d'exclusion. Les données des articles sélectionnés ont ensuite été extraites à l'aide d'un tableau puis la qualité des études a été analysée. Toutes ces données ont été synthétisées pour être incluses dans ce mémoire. Pour terminer, la revue systématique a été rédigée grâce au guide de rédaction PRISMA, détaillant les éléments nécessaires à la rédaction d'une revue.

2.2. Biais de l'étude

Cette revue systématique peut contenir des biais de sélection des articles. En effet, des critères d'inclusion et d'exclusion ont été établis au préalable afin de pouvoir sélectionner les études. Ces critères nous ont permis d'obtenir des résultats homogènes. De plus, ce biais est également applicable pour la sélection des mots-clés. Ceux-ci ont été sélectionnés par lecture d'articles, de revues systématiques, en prenant en compte les mots-clés les plus fréquents. Il est tout de même probable que des mots-clés qui auraient été pertinents n'aient pas été inclus dans les formules de recherche bibliographique. La sélection par lecture des titres a été réalisée en double aveugle ce qui permet de réduire le risque de biais (Borel et al., 2022). La suite de la sélection n'a, quant à elle, pas pu être réalisée en double aveugle, ce qui peut augmenter le risque de biais. Par ailleurs, lors de la lecture de résumés, il est possible que

certains résumés ne soient pas suffisamment informatifs et pertinents ce qui nous a conduit à les exclure.

Les articles sélectionnés pouvaient également contenir eux-mêmes des biais ou des limites. Ces biais pouvaient concerner la sélection des participants ; par exemple la parité homme/femme n'était pas toujours respectée, ou certains niveaux socio-économiques n'étaient pas forcément représentés, ce qui pourrait causer des biais dans les résultats. La méthode appliquée pour construire les outils d'évaluation dans les études pouvait également constituer un biais. Cependant, ces biais ne devraient pas altérer la qualité de cette revue, d'une part parce qu'elle inclut un nombre suffisant d'études, et également parce que l'on retrouve des résultats similaires dans les différentes études. En effet, la même tâche permet dans plusieurs études d'évaluer le même domaine du langage oral. Par exemple, la dénomination d'image permet dans huit études d'évaluer le lexique en production.

2.3. Qualité psychométrique des études

Afin de rassembler des données probantes et fiables nous avons évalué la qualité psychométrique des études. Pour cela, nous avons récolté les données concernant la validité, la fiabilité, la sensibilité et la spécificité des outils d'évaluation inclus dans les études. Nous avons également analysé la taille des échantillons de chaque étude.

En ce qui concerne les tailles d'échantillons de participants, il pouvait y avoir des variations considérables au sein d'une même épreuve. Par exemple, l'étude d'Alduais et al. (2022) incluait 14 participants pour les enfants de 4 ans et 148 participants âgés de 7 ans. D'autre part, certaines études ne détaillaient pas la répartition exacte des participants par groupe (Strand et al., 2013). Il a donc fallu faire des moyennes afin d'analyser équitablement les différents groupes. Nous retrouvons dans les résultats uniquement douze études qui comprennent des groupes d'échantillonnage de plus de cent participants. En effet, pour avoir une qualité acceptable et fiable, les études devaient inclure au minimum cent participants par tranche d'âge (Bernaud, 2014). Peu d'études incluses dans la revue ont donc d'excellentes qualités psychométriques en ce qui concerne la taille de l'échantillon.

Nous avons également analysé les qualités psychométriques propres aux outils d'évaluation. Chacun d'entre eux présentait au moins une des qualités psychométriques. La propriété psychométrique qui a le plus été soulevée est la validité. Pour rappel, cela concerne la corrélation entre ce que le test mesure et ce qu'il devrait mesurer (Bernaud, 2014). Nous retrouvons donc 27 études présentant une validité acceptable. Les résultats nous ont permis également de souligner que 22 études indiquent de bons critères de fidélité. Les outils issus de ces études comprennent donc des résultats aux épreuves qui sont stables. La sensibilité, quant à elle, a été relevée dans onze outils d'évaluation qui ont de bonnes capacités à discriminer les différences entre les participants afin d'identifier ceux présentant des troubles. Enfin, treize études comprenaient une bonne spécificité. Ces tests avaient donc la capacité d'exclure la présence d'un trouble chez les participants n'ayant pas de déficit (Plante & Vance, 1994). Cependant, certains outils (MSEL, Danish SI-3) étaient prometteurs au niveau de la qualité psychométrique, mais les études sur la standardisation des tests ou l'étude de la validité n'ont pas encore été établis. Les études choisies répondent donc globalement aux critères de qualité, ce qui permet d'obtenir un bon niveau de preuve de cette revue systématique.

3. Apport pour la pratique clinique

Les troubles du langage oral sont une part importante des prises en charge des orthophonistes. Ces difficultés ont parfois du mal à être détectées par les parents, les médecins ou les enseignants. Nous savons qu'il est important que les enfants soient pris en charge le plus précocement possible afin de diminuer les conséquences. Ce mémoire s'inscrit dans le projet LOLEMATH qui a pour objectifs de développer des tests de repérage, de suivre des cohortes d'enfants de manière longitudinale, et enfin de mettre à disposition une plateforme internet donnant des repères théoriques et des conseils de stimulation aux parents, enseignants et professionnels de santé. Cette revue systématique de la littérature nous a permis d'obtenir des données probantes et récentes sur l'évaluation du langage oral. Nous avons pu observer que le même type d'épreuves est proposé en France et à l'étranger. Ces résultats pourront donc servir pour le développement du site internet afin d'avoir des données théoriques fiables, qui sera à destination des parents, des enseignants et des professionnels de santé.

4. Limites de l'étude

Cette revue systématique présente cependant quelques limites. Premièrement, nous avons observé des limites liées aux outils d'évaluation. Nous avons décidé d'inclure les études uniquement publiées entre 2010 et 2023 afin d'avoir des données récentes. Cependant, nous nous sommes rendu compte que beaucoup de tests, notamment les outils utilisés en France pour les enfants âgés de zéro à six ans, ont été publiés avant 2010. Par exemple, parmi les outils les plus fréquents et récents en clinique en France, on retrouve EVALO 2-6 (Coquet et al., 2009), Exalang 3-6 (Helloin & Thibault, 2006), l'ELO (Khomsi, 2001), l'EVIP (Dunn & Theriault-Whalen, 1993). Ces tests sont tous publiés avant 2010. D'autre part, pour les enfants de zéro à trois ans, de nombreux tests consistent seulement en une observation des prérequis et ne possèdent donc pas spécialement de qualités psychométriques. Les enfants sont évalués en fonction des repères développementaux liés à l'âge.

Ensuite, nous avons également repéré des limites au sein des articles qui ont été sélectionnés. D'une part, comme indiqué dans la partie « Discussion », beaucoup d'études ont peu de participants pour convenir à une bonne qualité psychométrique. Nous avons décidé de tout de même les inclure dans la revue systématique pour ne pas avoir un nombre trop faible d'articles qui aurait pu avoir une conséquence sur les résultats obtenus. Les études ayant tout de même de bonnes qualités psychométriques concernant la validité, la fiabilité, la sensibilité et la spécificité, ont été incluses afin d'avoir de plus larges résultats.

Parmi les études sélectionnées beaucoup contenaient des outils portant sur l'évaluation du lexique, de la phonologie et de la morphosyntaxe, mais peu d'études portaient sur la prosodie, l'articulation, les prérequis, la narration et la pragmatique. Les tâches utilisées n'apparaissent parfois qu'une seule fois ce qui peut nous amener à nous questionner sur leur validité, même si les qualités psychométriques étaient acceptables. Nos résultats regroupent également peu d'outils normés pour les enfants âgés de moins de trois ans, d'où le peu d'informations récoltées sur l'évaluation des prérequis au langage. En effet, les tranches d'âge n'étaient pas toujours toutes représentées dans les études. Les compétences pouvant varier d'un âge à l'autre, nous n'avons donc pas pu récolter les preuves nécessaires dans toutes les tranches d'âge pour tous les domaines. Par exemple, pour l'évaluation de la prosodie, seuls les enfants de quatre ans à quatre ans et onze mois sont représentés.

Pour terminer, dans certains outils, la même tâche était utilisée pour évaluer l'articulation et la phonologie (répétition de mots et répétition de non-mots). En clinique, il faudrait alors être bien capable de discerner les compétences phonologiques et les compétences motrices afin d'avoir la meilleure analyse et donc pouvoir proposer la prise en charge la plus adaptée.

Conclusion

Ce mémoire avait pour but de réaliser une revue systématique de la littérature afin d'obtenir des données probantes sur l'évaluation du langage oral chez les enfants de zéro à six ans. Les objectifs étaient de trouver les outils que l'on peut utiliser pour évaluer les capacités en langage et ainsi d'en extraire les tâches spécifiques permettant de détecter les troubles du langage oral.

Pour ce faire, la méthode rigoureuse et reproductible d'une revue systématique a été réalisée. La recherche bibliographique s'est concentrée sur cinq bases de données (PubMed, WebofScience, Google Scholar, Lillocat et ScienceDirect) avec des mots-clés préalablement définis ainsi que des critères d'inclusion et d'exclusion. Ces critères ont ensuite servi à la sélection des articles, pour ne compter plus que 42 articles inclus dans la synthèse finale. Nous avons veillé à ce que les articles aient de bonnes qualités psychométriques (validité, fidélité, sensibilité et spécificité) afin d'obtenir des données probantes. L'extraction des données des études s'est faite avec un tableau, ce qui a permis de mettre en avant plus facilement les informations dont nous avons besoin.

Nous pouvons alors conclure qu'entre zéro et deux ans, l'évaluation du langage oral se fait principalement par une observation de l'enfant, des prérequis au langage qu'il a ou non acquis, par une analyse qualitative. Avant l'âge de deux ans, il n'existe pas, ou peu, d'outils d'évaluation standardisés. A partir de deux ans, les compétences linguistiques interindividuelles deviennent plus stables et il est donc possible d'établir des normes. Nous avons donc obtenu diverses tâches pouvant évaluer le langage sous tous ses domaines et versants. La prosodie peut être évaluée en analysant les fonctions de la communication grâce à des tâches de discrimination auditive pour la réception et d'imitation pour la production. Au niveau articulatoire et phonologique, les tâches de répétition de mot ou de non-mots sont les plus fréquentes en production. En réception phonologique, la tâche de discrimination phonologique est observée. Pour le lexique, deux tâches sont largement ressorties : la désignation d'images pour la réception, et la dénomination d'images pour la production. La morphosyntaxe est évaluée par diverses tâches consistant à évaluer la compréhension de l'enfant (désignation d'images sur phrases entendues), ou sa capacité à produire des phrases. La narration est évaluable à partir de l'âge de cinq ans, en production (narration d'histoires) et en réception (questions sur texte). Enfin, la pragmatique est évaluée plutôt de manière qualitative en faisant un inventaire des compétences de l'enfant. Des questionnaires parentaux sont aussi utilisables pour évaluer les capacités langagières des enfants de zéro à six ans.

Ce mémoire permet donc d'avoir des données récentes sur les outils et tâches d'évaluation du langage oral, afin que les difficultés et troubles puissent être détectés le plus précocement possible. Les données de cette revue systématique pourront être reprises afin de compléter la construction de la plateforme informatique du projet LOLEMATH, qui sera à destination des parents, des enseignants et des professionnels de santé.

Bibliographie

- Alduais, A., Qasem, F., Alfadda, H., Alfadda, N., & AlAmri, L. (2022). Arabic Validation of the Pragmatic Language Skills Inventory to Assess Pragmatic Language Development in Preschoolers with and without Pragmatic Language Impairment. *CHILDREN-BASEL*, 9(6). <https://doi.org/10.3390/children9060809>
- Antonioni, F., Ralli, A. M., Mouzaki, A., Diamanti, V., & Papaioannou, S. (2022). Logometro (R): The psychometric properties of a norm-referenced digital battery for language assessment of Greek-speaking 4-7 years old children. *FRONTIERS IN PSYCHOLOGY*, 13. <https://doi.org/10.3389/fpsyg.2022.900600>
- Beech, J. R. (1994). Tests mode d'emploi : Guide de psychométrie. In *Tests, mode d'emploi : Guide de psychométrie*. ECPA, les Ed. du Centre de psychologie appliquée.
- Bello, A., Giannantoni, P., Pettenati, P., Stefanini, S., & Caselli, M. C. (2012). Assessing lexicon : Validation and developmental data of the Picture Naming Game (PiNG), a new picture naming task for toddlers. *International Journal of Language & Communication Disorders*, 47(5), 589-602. <https://doi.org/10.1111/j.1460-6984.2012.00168.x>
- Bernaudo, J.-L. (2007). Introduction à la psychométrie. In *Introduction à la psychométrie*. Dunod.
- Bernaudo, J.-L. (2014). *Méthodes de tests et questionnaires en psychologie*. Dunod. <https://univ.scholarvox.com/book/88920722>
- Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & and the CATALISE-2 consortium. (2017). Phase 2 of CATALISE : A multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology and Psychiatry*, 58(10), 1068-1080. <https://doi.org/10.1111/jcpp.12721>
- Bleses, D., Vach, W., Jørgensen, R. N., & Worm, T. (2010). The internal validity and acceptability of the Danish SI-3 : A language-screening instrument for 3-year-olds. *Journal of Speech, Language, and Hearing Research*, 53(2), 490-507. [https://doi.org/10.1044/1092-4388\(2009/08-0132\)](https://doi.org/10.1044/1092-4388(2009/08-0132))
- Borel, S., Gatignol, P., Gros, A., & Tran, T. M. (2022). *Manuel de recherche en orthophonie : Formation initiale et continue, Toutes les UE Recherche*. De Boeck Supérieur.
- Bowles, R. P., Justice, L. M., Khan, K. S., Piasta, S. B., Skibbe, L. E., & Foster, T. D. (2020). Development of the Narrative Assessment Protocol-2 : A Tool for Examining Young Children's Narrative Skill. *Language, Speech & Hearing Services in Schools*, 51(2), 390-404. https://doi.org/10.1044/2019_LSHSS-19-00038
- Castilla-Earls, A., & Fulcher-Rood, K. (2018). Convergent and Divergent Validity of the Grammaticality and Utterance Length Instrument. *JOURNAL OF SPEECH LANGUAGE AND HEARING RESEARCH*, 61(1), 120-129. https://doi.org/10.1044/2017_JSLHR-L-17-0152
- Clausen, M. C., & Fox-Boyer, A. V. (2022). Diagnostic validity, accuracy and inter-rater reliability of a phonological assessment for Danish-speaking children. *Journal of Communication Disorders*, 95, 106168. <https://doi.org/10.1016/j.jcomdis.2021.106168>

- Colas, P., Ruiz, S., & Delteil, F. (2021). Dépistage des troubles du langage oral chez l'enfant et leur classification. *Journal de Pédiatrie et de Puériculture*, 34(2), 68-77. <https://doi.org/10.1016/j.jpp.2021.02.002>
- Coquet, F., Ferrand, P., & Roustit, J. (2009). *EVALO 2-6*. OrthoEdition.
- Crocetti, P., Fancelli, S., Colpizzi, I., Suozzi, A., Crocetti, E., Borgogni, E., & Gagliardi, G. (2021). T-PEC: a novel test for the elicited production of clitic pronouns in Italian. Preliminary data. *CLINICAL LINGUISTICS & PHONETICS*, 35(7), 636-662. <https://doi.org/10.1080/02699206.2020.1818129>
- Delahaie, M. (s. d.). *L'évolution du langage chez l'enfant, de la difficulté au trouble* (Vincent Fournier). Inpes.
- Dellatolas, L. V. G. (2005). *Recommandations sur les outils de Repérage, Dépistage et Diagnostic pour les Enfants atteints d'un Trouble Spécifique du Langage*. 129.
- Dias, D. C., Rondon-Melo, S., & Molini-Avejonas, D. R. (2020). Sensitivity and specificity of a low-cost screening protocol for identifying children at risk for language disorders. *Clinics*, 75, e1426. <https://doi.org/10.6061/clinics/2020/e1426>
- Dunn, L., & Theriault-Whalen, C. (1993). *Echelle de Vocabulaire en Images Peabody (EVIP)*. Pearson.
- Eisenberg, S. L., & Guo, L.-Y. (2018). Percent Grammatical Responses as a General Outcome Measure : Initial Validity. *Language, Speech & Hearing Services in Schools*, 49(1), 98-107. https://doi.org/10.1044/2017_LSHSS-16-0070
- Flageolet, A. (2008). *Annexe 1 : La prévention : Définition, notions générales sur l'approche française et comparaisons internationales, Les disparités territoriales des politiques de prévention sanitaire. Ministère de la Santé et des Solidarités*. <https://solidaritesante.gouv.fr/IMG/pdf/annexes.pdf>
- Gibbon, F. E., & Smyth, H. (2013). Preschool children's performance on Profiling Elements of Prosody in Speech-Communication (PEPS-C). *Clinical Linguistics & Phonetics*, 27(6-7), 428-434. <https://doi.org/10.3109/02699206.2012.741184>
- Gillon, G. T. (2017). *Phonological awareness : From research to practice*. Guilford Publications.
- Granocchio, E., Gazzola, S., Scopelliti, M. R., Criscuoli, L., Airaghi, G., Sarti, D., & Magazù, S. (2021). Evaluation of oro-phonatory development and articulatory diadochokinesis in a sample of Italian children using the protocol of Robbins & Klee. *Journal of Communication Disorders*, 91, 106101. <https://doi.org/10.1016/j.jcomdis.2021.106101>
- Helloin, M.-C., & Thibault, M.-P. (2006). *Exalang 3-6*. Motus.
- Hemingway, P., & Brereton, N. (2009). *What is a systematic review?* 8.
- Holzinger, D., Weber, C., & Diendorfer, B. (2022). Development and Validation of a Language Screening for Implementation in Pre-School Settings. *Frontiers in Public Health*, 10, 866598. <https://doi.org/10.3389/fpubh.2022.866598>
- Kapalkova, S., Polisenka, K., & Vicenova, Z. (2013). Non-word repetition performance in Slovak-speaking children with and without SLI: novel scoring methods. *INTERNATIONAL*

- JOURNAL OF LANGUAGE & COMMUNICATION DISORDERS*, 48(1), 78-89. <https://doi.org/10.1111/j.1460-6984.2012.00189.x>
- Kas, B., Jakab, Z., & Lorik, J. (2022). Development and norming of the Hungarian CDI-III: A screening tool for language delay. *INTERNATIONAL JOURNAL OF LANGUAGE & COMMUNICATION DISORDERS*, 57(2), 252-273. <https://doi.org/10.1111/1460-6984.12686>
- Katzenberger, I., & Meilijson, S. (2014). Hebrew language assessment measure for preschool children : A comparison between typically developing children and children with specific language impairment. *LANGUAGE TESTING*, 31(1), 19-38. <https://doi.org/10.1177/0265532213491961>
- Kern, S. (2019). *Le développement du langage chez le jeune enfant* (Deboeck supérieur).
- Khodeir, M. S., Hegazi, M. A., & Saleh, M. M. (2017). Development and Standardization of a Test for Pragmatic Language Skills in Egyptian Arabic : The Egyptian Arabic Pragmatic Language Test (EAPLT). *Folia Phoniatica et Logopaedica: Official Organ of the International Association of Logopedics and Phoniatrics (IALP)*, 69(5-6), 209-218. <https://doi.org/10.1159/000485656>
- Khomsî, A. (2001). *Evaluation du Langage Oral (ELO)*. Editions du Centre de Psychologie Appliqué.
- Lagerberg, T. B., Åsberg, J., Hartelius, L., & Persson, C. (2014). Assessment of intelligibility using children's spontaneous speech : Methodological aspects. *International Journal of Language & Communication Disorders*, 49(2), 228-239. <https://doi.org/10.1111/1460-6984.12067>
- Lavesson, A., Lovden, M., & Hansson, K. (2018). Development of a language screening instrument for Swedish 4-year-olds. *INTERNATIONAL JOURNAL OF LANGUAGE & COMMUNICATION DISORDERS*, 53(3), 605-614. <https://doi.org/10.1111/1460-6984.12374>
- Lüke, C., Grimminger, A., Rohlfing, K. J., Liszkowski, U., & Ritterfeld, U. (2017). In Infants' Hands : Identification of Preverbal Infants at Risk for Primary Language Delay. *Child Development*, 88(2), 484-492. <https://doi.org/10.1111/cdev.12610>
- Matov, J., Mensah, F., Cook, F., Reilly, S., & Dowell, R. (2020). The development and validation of the Short Language Measure (SLaM) : A brief measure of general language ability for children in their first year at school. *International Journal of Language & Communication Disorders*, 55(3), 345-358. <https://doi.org/10.1111/1460-6984.12522>
- McLeod, S., Crowe, K., & Shahaiean, A. (2015). Intelligibility in Context Scale : Normative and Validation Data for English-Speaking Preschoolers. *Language, Speech & Hearing Services in Schools*, 46(3), 266-276. https://doi.org/10.1044/2015_LSHSS-14-0120
- Mohamadi, R., Ahmadi, A., Kazemi, M. D., Minaei, A., & Damarchi, Z. (2019). Development of the Persian syntax comprehension test. *International Journal of Pediatric Otorhinolaryngology*, 124, 22-29. <https://doi.org/10.1016/j.ijporl.2019.05.032>
- Morgenstern, A., Leroy, M., & Mathiot, E. (2008). Le pointage chez l'enfant : Du gestuel au linguistique. *Congrès Mondial de Linguistique Française 2008*, 164. <https://doi.org/10.1051/cmlf08170>

- Nash, H., Leavett, R., & Childs, H. (2011). Evaluating the GAPS test as a screener for language impairment in young children. *INTERNATIONAL JOURNAL OF LANGUAGE & COMMUNICATION DISORDERS*, 46(6), 675-685. <https://doi.org/10.1111/j.1460-6984.2011.00038.x>
- Okalidou, A., Syrika, A., Beckman, M. E., & Edwards, J. R. (2011). Adapting a receptive vocabulary test for preschool-aged Greek-speaking children. *International Journal of Language & Communication Disorders*, 46(1), 95-107. <https://doi.org/10.3109/13682821003671486>
- Owens, R. E. (2016). *Language Development: An Introduction 9e*. Boston: Pearson Education.
- Pace, A., Curran, M., Van Horne, A. O., de Villiers, J., Iglesias, A., Golinkoff, R. M., Wilson, M. S., & Hirsh-Pasek, K. (2022). Classification accuracy of the Quick Interactive Language Screener for preschool children with and without developmental language disorder. *JOURNAL OF COMMUNICATION DISORDERS*, 100. <https://doi.org/10.1016/j.jcomdis.2022.106276>
- Plante, E., & Vance, R. (1994). Selection of preschool language tests : A data-based approach. *Language, Speech, and Hearing Services in Schools*, 25(1), 15-24.
- Plaza, M. (2014). Le développement du langage oral. *Contraste*, 39(1), 99-118. <https://doi.org/10.3917/cont.039.0099>
- Puglisi, M. L., Blasi, H. F., & Snowling, M. J. (2020). Screening for the Identification of Oral Language Difficulties in Brazilian Preschoolers : A Validation Study. *Language, Speech & Hearing Services in Schools*, 51(3), 852-865. https://doi.org/10.1044/2020_LSHSS-19-00083
- Rakhlin, N. V., Aljughaiman, A., & Grigorenko, E. L. (2021). Assessing language development in Arabic : The Arabic language: Evaluation of function (ALEF). *Applied Neuropsychology. Child*, 10(1), 37-52. <https://doi.org/10.1080/21622965.2019.1596113>
- Romero, M. F., Petersen, D. B., Long, A., Arvizu, M. N., & Dechert, C. (2021). The accuracy of a Spanish dynamic assessment of narrative language in identifying developmental language disorder : A cross validation study. *Revista de Logopedia, Foniatria y Audiología*, 41(4), 172-182. <https://doi.org/10.1016/j.rlfa.2021.05.008>
- Romski, M., Bornman, J., White, R., Morwane, R., Tonsing, K., Sevcik, R., & Barton-Hulsey, A. (2018). Adapting and translating the Mullen Scales of Early Learning for the South African context. *South African Journal of Communication Disorders*, 65(1), 1-9. <https://doi.org/10.4102/sajcd.v65i1.571>
- Rondal, J.-A. (1997). L'évaluation du langage. In *L'évaluation du langage*. P. Mardaga.
- Schmoeger, M., Deckert, M., Eisenwort, B., Loader, B., Hofmair, A., Auff, E., & Willinger, U. (2020). Evaluating Part V of the German version of the Token Test as a screening of specific language impairment in preschoolers. *Applied Psycholinguistics*, 41(1), 237-258. <https://doi.org/10.1017/S0142716419000493>
- Smolík, F., & Bytešníková, I. (2021). Validity of the SDDS : A 40-item vocabulary screening tool for 18- to 42-month olds in Czech. *Journal of Communication Disorders*, 93, 106146. <https://doi.org/10.1016/j.jcomdis.2021.106146>

- Soubeyrand, E., Xavier, J., & Guilé, J.-M. (2007). Quand le langage émerge... (Partie 1). *Revue de la littérature sur le développement normal et le risque autistique. Perspectives Psy*, 46(3), 248-255.
- Spaulding, T. J., Hosmer, S., & Schechtman, C. (2013). Investigating the interchangeability and diagnostic utility of the PPVT-III and PPVT-IV for children with and without SLI. *International Journal of Speech-Language Pathology*, 15(5), 453-462. <https://doi.org/10.3109/17549507.2012.762042>
- Strand, E. A., McCauley, R. J., Weigand, S. D., Stoeckel, R. E., & Baas, B. S. (2013). A motor speech assessment for children with severe speech disorders: Reliability and validity evidence. *Journal of Speech, Language, and Hearing Research*, 56(2), 505-520. [https://doi.org/10.1044/1092-4388\(2012\)12-0094](https://doi.org/10.1044/1092-4388(2012)12-0094)
- Syadar, S. F., Zarifian, T., Modarresi, Y., Zahedi, M. S., Ebrahimipour, M., & Biglarian, A. (2019). Kurdish Speech Test: A validation study for children aged 3-5 years. *International Journal of Pediatric Otorhinolaryngology*, 117, 61-66. <https://doi.org/10.1016/j.ijporl.2018.10.009>
- Topbaş, S., Kaçar-Kütükçü, D., & Kopkalli-Yavuz, H. (2014). Performance of children on the Turkish Nonword Repetition Test: Effect of word similarity, word length, and scoring. *Clinical Linguistics & Phonetics*, 28(7-8), 602-616. <https://doi.org/10.3109/02699206.2014.927003>
- van der Lely, H. K. J., Payne, E., & McClelland, A. (2011). An Investigation to Validate the Grammar and Phonology Screening (GAPS) Test to Identify Children with Specific Language Impairment. *PLOS ONE*, 6(7). <https://doi.org/10.1371/journal.pone.0022432>
- van Haaften, L., Diepeveen, S., van den Engel-Hoek, L., Jonker, M., de Swart, B., & Maassen, B. (2019). The Psychometric Evaluation of a Speech Production Test Battery for Children: The Reliability and Validity of the Computer Articulation Instrument. *JOURNAL OF SPEECH LANGUAGE AND HEARING RESEARCH*, 62(7), 2141-2170. https://doi.org/10.1044/2018_JSLHR-S-18-0274
- Visser-Bochane, M. I., Luinge, M., Dieleman, L., van der Schans, C., & Reijneveld, S. (2021). The Dutch well child language screening protocol for 2-year-old children was valid for detecting current and later language problems. *Acta Paediatrica (Oslo, Norway: 1992)*, 110(2), 556-562. <https://doi.org/10.1111/apa.15447>
- Visser-Bochane, M. I., van der Schans, C. P., Krijnen, W. P., Reijneveld, S. A., & Luinge, M. R. (2021). Validation of the Early Language Scale. *European Journal of Pediatrics*, 180(1), 63-71. <https://doi.org/10.1007/s00431-020-03702-8>
- Wong, A. M.-Y., Leung, C., Siu, E. K.-L., & Lam, C. C.-C. (2012). Validating the language domain subtest in a developmental assessment scale for preschool children. *Research in Developmental Disabilities*, 33(5), 1633-1641. <https://doi.org/10.1016/j.ridd.2012.03.002>
- Wu, H., Liu, Q., Yu, B., Zhang, Y., Ren, D., Serdarevic, M., Liang, Z., Wang, Y., Chen, S., Zhang, K., & Chen, S. (2020). Psychometric properties of the mandarin clinical evaluation of language for preschooler's core scale. *Journal of Communication Disorders*, 87, 106000. <https://doi.org/10.1016/j.jcomdis.2020.106000>

Liste des annexes

Annexe n°1 : Tableau de résultats

Annexe n°2 : Tableau récapitulatif des outils d'évaluation