

Cotation des tests francophones de langage oral et écrit chez l'enfant : quelques recommandations à l'usage des orthophonistes

Matthieu BIGNON *, Loïc GAMOT **, Marie-Pierre LEMAITRE ***, Lucie MACCHI ****

* Doctorant, SCALab, UMR 9193, Université de Lille, France

** Orthophoniste, Enseignant, CHU de Lille, CRDTA, Département d'orthophonie, Université de Lille, France

*** Neuropédiatre, CHU de Lille, CRDTA, France

**** Maître de conférences, laboratoire STL UMR CNRS 8163, Université de Lille, France

Auteur de correspondance :
matthieu.bignon@univ-lille.fr

Résumé :

L'évaluation orthophonique du langage oral et écrit de l'enfant suppose l'utilisation de tests normés qui permettent d'obtenir des indications quantitatives et qualitatives sur les compétences des patients. Trois points de vigilance concernant la comparaison des scores des patients à la norme semblent importants à considérer pour réduire la marge d'erreur de mesure lors de l'évaluation. Ils correspondent à des limites possibles présentées par certains tests et concernent les sujets suivants : (1) l'adéquation entre le type de norme utilisé (en âge ou en niveau scolaire) et la modalité langagière évaluée (le langage oral ou écrit), (2) l'adéquation entre l'étendue des tranches de développement (en âge ou en niveau scolaire) et la vitesse de progression des compétences évaluées dans la population de référence, (3) la clarté de la présentation des tranches d'âge dans les manuels ou les logiciels. L'analyse méthodique de 54 tests francophones utilisés par les praticiens pour évaluer le langage oral et/ou écrit des enfants montre qu'une part non négligeable de ces tests présente des imperfections en ce qui concerne ces 3 aspects. Il semble donc important que le clinicien utilise les tests avec prudence, en ayant conscience de leurs éventuelles limites. Des tableaux synoptiques permettent d'identifier les écueils et les solutions possibles, pour chacun de ces tests. Des conseils pratiques sont également proposés afin de contourner ces écueils et d'optimiser l'usage des tests.

Mots clés : test, évaluation, enfant, langage

**Scoring french spoken and written language tests for children :
recommendations for speech and language therapists**

Summary:

Speech and language therapy assessment of the child's spoken and written language involves the use of standardized tests that provide quantitative and qualitative indications of patient competence. Three points of caution regarding the comparison of patients' scores to the standard seem important to consider in order to reduce the margin of measurement error during the evaluation. They correspond to some limits presented by certain tests and concern the following subjects: (1) the adequacy between the type of standard used (in age or grade level) and the language modality assessed (spoken or written language), (2) the adequacy between the extent of the development ranges (in age or grade level) and the rate of progression of the skills assessed in the reference population, (3) the clarity of the presentation of the age ranges in the textbooks or software. A methodological analysis of 54 French-language tests used by practitioners to assess children's spoken and/or written language shows that a non-negligible proportion of these tests have imperfections in these 3 aspects. It therefore seems important that the clinician uses the tests with caution, being aware of their possible limitations. Synoptic tables allow the clinician to identify possible pitfalls and solutions for each of these tests. Practical advice is also provided to overcome these potential pitfalls and optimize the use of tests.

Keywords: test, assessment, children, language

-----INTRODUCTION-----

Réaliser un bilan orthophonique des compétences langagières d'un enfant nécessite d'analyser qualitativement et quantitativement de nombreuses informations, et de recourir à des tests standardisés dont l'utilisation est recommandée (Haute Autorité de Santé, 2017 ; Inserm, 2007). Ces outils d'évaluation permettent d'attester de manière quantifiée l'existence d'un trouble, par des capacités nettement au-dessous du niveau escompté pour l'âge chronologique de l'enfant (American Psychiatric Association, 2013). Ils permettent également de mesurer la sévérité du trouble, d'émettre un pronostic, d'élaborer un projet thérapeutique ou d'évaluer les progrès d'un patient (Garcia et al., 2006 ; Kerr et al., 2003).

Cette recherche porte sur les tests standardisés permettant d'examiner les difficultés langagières orales et/ou écrites des enfants. En ce qui concerne l'oral, les troubles peuvent affecter le langage ou les sons de la parole. Les troubles du langage concernent des individus dont les difficultés langagières sont susceptibles de persister jusqu'au milieu de l'enfance et au-delà, avec un impact fonctionnel significatif sur leurs interactions sociales quotidiennes ou leurs apprentissages scolaires (Bishop et al., 2017). La catégorie des troubles du langage comprend des troubles liés à une condition biomédicale spécifique (ex. paralysie cérébrale) et des troubles sans condition biomédicale connue : les troubles développementaux du langage. Les troubles des sons de la parole se caractérisent par un trouble articulatoire ou phonologique isolé généralement de pronostic favorable (Beitchman et al., 1996). Concernant la modalité écrite, les difficultés relèvent des troubles des apprentissages tels que décrits dans le DSM-5 (Diagnostic and Statistical Manual of Mental Disorders, American Psychiatric Association, 2013, 2015). De tels troubles, dits neurodéveloppementaux, se définissent par la persistance (depuis au moins 6 mois) d'au moins une des manifestations suivantes, malgré la mise en place de mesures de soutien : une lenteur ou un manque d'exactitude en lecture de mots, des difficultés de compréhension en lecture, des difficultés d'orthographe et des difficultés d'expression écrite. Ces difficultés sont évaluées au moyen de tests standardisés dont les résultats doivent être déficitaires, c'est-à-dire, selon le DSM-5, inférieurs à -1.5 ET ou au percentile 7 par rapport à la population de référence. De plus, un retentissement doit être observé sur la réussite scolaire ou les activités quotidiennes. Enfin, les difficultés doivent débiter au cours de la scolarité et ne pas être le seul fait de facteurs environnementaux ou d'autres troubles ou handicaps (Launay, 2018).

Notre étude porte sur plusieurs qualités des tests de langage oral ou écrit destinés à l'enfant. Les qualités des tests concernent tant le matériel et les éventuels rapports générés par ordinateur que les caractéristiques psychométriques des tests, c'est-à-dire les caractéristiques liées aux échelles de mesures (Evers et al., 2013 ; Laveault & Grégoire, 2014). Différentes caractéristiques psychométriques peuvent être distinguées : la fidélité (précision des mesures d'un test), la validité (faculté d'un test à mesurer ce qu'il prétend mesurer) et les normes (Evers et al., 2013 ; Laveault & Grégoire, 2014). Ces propriétés psychométriques doivent être satisfaisantes pour permettre une évaluation clinique de qualité (Hogan, 2017). C'est pourquoi des chercheurs ont analysé les qualités psychométriques des tests de langage oral ou écrit. Par exemple, McCauley et Swisher (1984) ont examiné, pour 30 tests américains de langage oral chez l'enfant, la présence de 10 critères psychométriques. Seuls 20 % des tests satisfont à la moitié de ces critères. Les informations nécessaires pour évaluer chaque critère sont souvent absentes des manuels. Une autre étude a permis l'analyse de 16 critères psychométriques pour 31 outils d'évaluation du langage oral et écrit, utilisés par des orthophonistes auprès d'enfants francophones canadiens (Bouchard et al., 2009). Plus de la moitié des 31 outils étudiés n'apporte pas de preuves de validité ni de fidélité (Bouchard et al., 2009). De même, Leclercq

et Veys (2014) ont étudié 13 critères psychométriques de 5 batteries à destination des orthophonistes de France et de Belgique pour évaluer le langage oral des enfants. Les conclusions sont similaires : ces batteries réunissent entre 3 et 8 critères psychométriques sur les 13 nécessaires à un outil de qualité. Plus récemment, en France, le comité directeur de l'Union Nationale pour le Développement de la Recherche et de l'Évaluation en Orthophonie (UNADREO) a proposé une grille d'évaluation à l'aide de laquelle il se propose d'étudier la qualité des tests en orthophonie (Hilaire-Debove, 2017 ; Rousseau, 2016). La Fédération Européenne des Associations de Psychologues a également proposé un formulaire détaillé pour l'examen des tests utilisés dans le domaine de la psychologie et de l'éducation (Evers et al., 2013). En somme, bien qu'une réelle réflexion soit menée par les chercheurs et les cliniciens sur la qualité des tests, dans l'ensemble, peu d'outils d'évaluation du langage oral et écrit pour les enfants répondent actuellement aux propriétés psychométriques attendues pour une pratique de qualité.

Dans cette recherche, nous nous intéressons à différents aspects d'une caractéristique psychométrique spécifique : les normes. Les normes sont les valeurs d'une population de référence auxquelles les scores d'un individu testé peuvent être comparés (Laveault & Grégoire, 2014). Plus précisément, nous nous intéressons à l'étape de comparaison des scores à la norme, lors de l'évaluation des performances d'un enfant. Nous analyserons les tests de langage oral et écrit utilisés par les orthophonistes, au regard de 3 points de vigilance, relevés au cours de notre pratique clinique. Ces aspects sont importants à prendre en compte pour réduire la marge d'erreur de mesure lors de l'évaluation. Il s'agit de : (1) l'adéquation entre les types de norme utilisés (en âge ou en niveau scolaire) et la modalité langagière évaluée (orale ou écrite), (2) l'adéquation entre l'étendue des tranches de développement (en âge ou en niveau scolaire) et la vitesse de progression des compétences évaluées dans la population de référence, (3) la clarté de la présentation des tranches d'âge dans les tableaux d'étalonnage ou les logiciels des tests. L'objectif est à la fois d'effectuer, sur ces aspects, une revue critique d'un grand nombre de tests utilisés par les cliniciens, mais aussi de fournir aux professionnels des conseils pratiques pour contourner certains écueils éventuels de cotation et optimiser l'usage des tests. Dans la première partie de cet article, nous décrirons chacun de ces points de vigilance. Par la suite, nous détaillerons notre méthodologie d'analyse en vue d'étudier ces points de façon systématique dans les tests francophones actuellement disponibles. La présentation des résultats sera suivie d'une discussion. Avant toute chose, notons que par facilité de communication, les épreuves, batteries ou questionnaires seront tous appelés « tests », dans la suite de cet article.

1. Adéquation entre les types de normes et la modalité langagière évaluée

L'interprétation des scores de l'individu à un test est fortement influencée par le groupe de référence utilisé et par la pertinence de la comparaison entre l'individu testé et le groupe de référence (Hogan, 2017 ; Hutchinson, 1996). Lors de la comparaison des scores d'un patient à la norme, il est donc important de choisir le groupe de référence adéquat. La Commission Internationale des Tests formalise cette recommandation dans son guide d'utilisation des tests de la manière suivante (Vrignaud et al., 2003) : « Les utilisateurs de tests compétents devraient utiliser des normes ou des groupes de référence appropriés lorsqu'ils sont disponibles » (p. 23) et « s'assurer que des conclusions invalides ne sont pas tirées de la comparaison de scores avec des normes inadaptées aux personnes testées » (p. 22).

Les compétences langagières évoluant au cours de la vie, les groupes de référence sont généralement répartis suivant des tranches de développement tenant compte soit de l'âge des personnes testées, soit de leur niveau scolaire. Suivant la modalité d'évaluation, orale ou écrite, l'une ou l'autre possibilité permet de fournir des normes plus ou moins appropriées. Les

paragraphes qui suivent détaillent ces aspects.

a. Les tests de langage oral

Contrairement à l'apprentissage du langage écrit, l'apprentissage du langage oral chez l'enfant au développement typique ne nécessite pas d'enseignement explicite. Avant tout, pour apprendre à parler, l'enfant doit être exposé à une langue maternelle (Lieberman, 1992). L'évolution aurait doté le cerveau de la capacité à apprendre naturellement une langue, dès lors que l'enfant est immergé dans un environnement linguistique (Dehaene, 2019). Tous les environnements humains qui offrent à l'enfant l'occasion de communiquer favorisent son développement langagier (Hoff, 2006), en particulier le milieu familial, bien avant l'entrée à l'école. Le développement du langage oral dépend donc fortement du temps d'exposition général à la langue maternelle, qui s'avère lui-même étroitement lié à l'âge de l'enfant. Pour les tests de langage oral, les normes construites en fonction de l'âge semblent donc plus adaptées que les normes construites en fonction du niveau scolaire.

Pourtant, certains tests de langage oral présentent des normes scolaires, par exemple l'ELO (Khomsi, 2001). Un tel type d'étalonnage ne paraît pas tout à fait adapté, puisque l'âge de l'enfant peut être différent de l'âge moyen des enfants du groupe de référence correspondant à son niveau scolaire, suivant son mois de naissance et la période de l'année scolaire au cours de laquelle les données normatives ont été recueillies. Ce type d'étalonnage peut également induire le clinicien en erreur si l'enfant a redoublé une classe. Il semble en effet plus opportun de comparer ses performances à celles des enfants de la classe qu'il aurait fréquentée s'il n'avait pas redoublé, plutôt qu'à celles des enfants de la classe qu'il fréquente en réalité.

Pour une utilisation appropriée des tests, il est donc judicieux de choisir le groupe de niveau scolaire dont la moyenne d'âge s'approche le plus de l'âge du patient. Certains manuels de tests indiquent explicitement l'âge moyen des enfants des groupes de référence de chaque niveau scolaire. Par exemple, dans le manuel de l'ELO (Khomsi, 2001), il est mentionné que l'âge moyen des enfants de moyenne section de maternelle est de 4 ; 3 ans et que celui des enfants de grande section de maternelle est de 5 ; 3 ans. Ainsi, si l'on évalue les performances d'un enfant de 5 ; 4 ans scolarisé en moyenne section, l'on comparera ses résultats préférentiellement à ceux du groupe de référence de grande section plutôt qu'à ceux de moyenne section.

b. Les tests de langage écrit

L'expertise en lecture et en orthographe ne s'acquiert pas par une simple exposition à l'écrit. En effet, contrairement au langage oral, le cerveau humain ne dispose pas à la naissance de réseaux spécifiques pour apprendre à lire. L'écriture est une invention culturelle récente dans l'histoire humaine (Dehaene, 2007) qui nécessite un enseignement explicite, dispensé essentiellement à l'école à partir du CP (Ecalte & Magnan, 2002). Cet apprentissage est assez laborieux ; il s'étend sur plusieurs années et nécessite de longues heures de pratique (Van Rompaey et al., 2004).

Dans les systèmes d'écriture alphabétique, l'acquisition de la lecture se produit en 2 temps. Tout d'abord, l'enfant applique des procédés de conversion graphème-phonème et de fusion phonémique, de manière à produire la forme sonore du mot écrit. Puis il met en relation cette forme sonore avec le mot oral connu. Dans un second temps, l'enfant devient capable de traiter plusieurs lettres simultanément et de les mettre directement en relation avec le sens et la forme phonologique du mot écrit. Cela correspond à l'automatisation de la lecture. En réalité, ces

2 temps ne sont pas parfaitement distincts et se recouvrent au cours du développement. De plus, la procédure de conversion graphème-phonème reste fonctionnelle et utile à un stade avancé du développement, notamment pour lire de nouveaux mots (Dehaene, 2019).

L'enseignement du langage écrit dépend des programmes scolaires dont les objectifs d'apprentissage sont échelonnés sur plusieurs niveaux scolaires. Par exemple, l'apprentissage des correspondances graphèmes-phonèmes occupe une grande partie du programme scolaire en CP. Ce temps passé à découvrir les bases du décodage n'est pas utilisé pour apprendre d'autres notions (ex. en grammaire). De ce fait, l'on peut s'attendre à ce qu'un enfant de CP ayant redoublé (c'est-à-dire ayant l'âge des enfants de niveau CE1) ait acquis moins de connaissances en langage écrit à la fin de son deuxième CP qu'un enfant en fin de CE1. De même, les conjugaisons font l'objet d'enseignements spécifiques échelonnés tout au long des classes en école élémentaire (du CP au CM2). Ainsi, en ce qui concerne les conjugaisons, l'on peut s'attendre à ce qu'un enfant ayant redoublé le CE1 ait moins de connaissances langagières écrites qu'un enfant de CE2. C'est pourquoi il semble opportun d'utiliser des normes scolaires pour le langage écrit et non des normes d'âge, afin de comparer les performances des patients avec celles d'enfants exposés à des enseignements similaires. Toutefois, certains tests de langage écrit sont étalonnés par âge comme l'ANALEC (Inizan, 1998). Ce mode d'étalonnage ne paraît pas idéal, parce que les enfants d'un âge donné ne sont pas nécessairement de même niveau scolaire. Par exemple, selon son mois de naissance, un enfant de 9 ans peut tout autant être en CE2 qu'en CM1 (même s'il n'a ni redoublé ni sauté de classe). De ce fait, dans la population de référence d'un test de langage écrit normé par âge, un même groupe d'âge peut comprendre des enfants de différents niveaux scolaires.

De plus, si l'orthophoniste compare les performances de son patient au groupe d'âge qui correspond à l'âge de l'enfant, il peut arriver qu'il les compare à un groupe d'âge d'un niveau scolaire différent de l'enfant, ce qui peut entraîner un biais dans l'évaluation de ses compétences. Par exemple, un patient né en début d'année civile peut avoir 8 ans en CE1, alors que c'est environ l'âge moyen observé en CE2. Si ses performances sont comparées au groupe d'âge de 8 ans, donc au groupe d'enfants majoritairement en CE2, il est possible qu'elles soient sous-estimées.

Pour comparer les performances langagières écrites du patient à celles du groupe de référence de niveau scolaire le plus proche, nous suggérons aux orthophonistes de garder à l'esprit les correspondances précises généralement observées entre les niveaux scolaires et les tranches d'âge. Par exemple, les performances d'un enfant scolarisé en septembre de CP seront comparées préférentiellement à celles d'enfants de 6 ans, âge correspondant en moyenne à l'âge des enfants de début de CP ; tandis qu'en juin de CP, elles seront préférentiellement comparées à celles des enfants de 7 ans. Pour les enfants ayant redoublé, la même règle s'applique : il est possible de comparer leurs scores en langage écrit à ceux du groupe d'enfants correspondant à leur niveau scolaire. Par exemple un enfant de 8 ans à qui l'on administre un test au début de sa deuxième année de CE1 verra ses scores comparés au groupe d'enfants de 7 ans, âge qui correspond à l'âge moyen des enfants en début de CE1, et non au groupe de référence de 8 ans, même si cela correspond à son âge chronologique. Ces réflexions peuvent aider le clinicien à sélectionner le groupe de référence le plus approprié, tout en gardant à l'esprit que la mesure de l'écart à la norme reste à interpréter avec prudence.

Dans cet article, nos analyses permettront de déterminer la proportion de tests avec des normes appropriées à chaque modalité langagière évaluée. De plus, concernant les tests de langage oral mentionnant uniquement des normes scolaires, nous déterminerons la proportion de ceux

indiquant la moyenne d'âge des participants du groupe de référence par niveau scolaire ainsi que les pages auxquelles ces informations apparaissent, afin d'en faciliter la recherche par les cliniciens. Concernant les tests de langage écrit mentionnant uniquement des normes en âge, nous avons précédemment présenté comment utiliser au mieux les normes du groupe d'âge dont le niveau scolaire est le plus proche de celui du patient. De ce fait, aucune analyse plus approfondie ne sera menée sur ce sujet.

2. Adéquation entre l'étendue des tranches de développement et la vitesse de progression des compétences évaluées dans la population générale

Le deuxième type d'écueil possible se situe au niveau de la sensibilité des épreuves (au sens psychométrique du terme), c'est-à-dire au niveau de leur « capacité à rendre compte du maximum de différences interindividuelles » (Perdrix, 2018, p. 60). La sensibilité dépend notamment de l'étendue des tranches de développement (en âge ou en niveau scolaire) en fonction de la vitesse d'évolution des compétences testées sur la période considérée. En effet, si les tranches sont trop étroites ou si le développement de la compétence ciblée n'est pas suffisamment important au sein de la population au cours de la période en question, les scores des groupes de référence issus des tranches contiguës ne présentent pas nécessairement de différence significative. Dans ce cas, il n'y a pas d'intérêt à créer des sous-groupes d'étalonnage pour ces niveaux de développement (Hogan, 2017). Inversement, si les tranches sont trop larges ou si le développement de la compétence ciblée est très rapide sur la période considérée, le groupe de référence risque d'inclure des enfants dont les niveaux langagiers sont très différents. De ce fait, le test peut perdre de sa sensibilité. C'est de ce cas de figure que nous traitons ci-dessous. Nous développons la question de l'adéquation entre l'étendue des tranches de développement et la vitesse de progression des compétences évaluées pour les tests de langage oral puis écrit.

a. Les tests de langage oral

Au cours des premières années de vie, le langage oral se développe rapidement. Aux alentours de ses 5 ans, l'enfant maîtrise l'essentiel du système de sa langue maternelle (Bassano, 2007 ; Inserm, 2007).¹ Il a acquis les principaux aspects de l'organisation phonologique, possède un vocabulaire de plusieurs milliers de mots et maîtrise les principales structures morphosyntaxiques (Dehaene, 2007). À cet âge, « le langage de l'enfant devient très proche de celui de l'adulte » (Parisse & Maillart, 2004, p. 28). Puisque les compétences langagières orales de l'enfant progressent rapidement entre sa naissance et ses 5 ans, il paraît nécessaire, pour cette période, d'employer des tests dont les tranches d'âge d'étalonnage sont d'étendue relativement limitée. C'est le cas, par exemple, dans les batteries CLÉA (Pasquet et al., 2014) et BILO Petits (Khomsi & Khomsi, 2009) qui disposent de tranches d'âge allant de 6 mois en 6 mois pour les enfants de 5 ans ou moins. Leurs auteurs ont mis en évidence, par des tests statistiques, des différences significatives entre les scores obtenus par les enfants de groupes de tranches d'âge contiguës jusqu'à 5 ans, ce qui démontre l'intérêt d'une telle finesse.

Dans cet article, nos analyses nous permettront d'observer quelle proportion de tests de langage oral dispose de tranches d'âge d'étendue inférieure à 1 an pour les enfants de 5 ans ou moins. Puisque l'âge moyen en début de grande section de maternelle est d'environ 5 ans, nous analyserons selon ce même critère les tests de langage oral destinés aux enfants de grande

¹. L'âge charnière généralement retenu est celui de 5 ans, bien qu'il existe de légères variations d'un auteur à l'autre (entre l'âge de 4 et 6 ans).

section de maternelle ou moins.

b. Les tests de langage écrit

Les capacités de reconnaissance et de transcription de mots écrits se développent très rapidement durant les premières années d'école élémentaire, au CP et au CE1 (Sprenger-Charolles et al., 2003). Pour ces niveaux, il semble donc idéal d'utiliser les tests de langage écrit dont les tranches d'étalonnage couvrent une étendue inférieure à 1 an, afin d'évaluer finement les compétences des enfants. Par exemple, le Timé 2 (Ecalte, 2018) présente 2 sessions d'étalonnage pour le CP et le CE1 : une première en septembre-octobre et une seconde en janvier-février. Il existe un effet significatif des sessions d'étalonnage sur les scores des enfants, ce qui justifie le choix de tranches développementales étroites en CP et en CE1.

Pour les autres tests, étalonnés par année scolaire complète, il est possible d'identifier le tableau d'étalonnage le plus adapté en prenant connaissance de la période précise de recueil des données auprès de la population de référence du test, et en la comparant à la période au cours de laquelle le patient est évalué. Par exemple, les données normatives de la BLI (Khomsy, 2003) ont été récoltées entre septembre et décembre pour chacun des niveaux du CE1 au CM2. Ainsi, les performances d'un patient évalué entre septembre et décembre de son année de CE2, pourront être comparées à celles du groupe de référence de CE2. La situation est assez différente pour un patient évalué en juillet de son année de CE2, car 2 mois séparent son évaluation de celle du groupe de CM1 (de juillet à septembre), alors que 7 mois séparent son évaluation de celle du groupe de CE2 (de décembre à juillet). L'on peut donc penser que le profil scolaire de ce patient s'apparente davantage à celui des enfants du groupe de CM1, qu'à ceux de CE2 dont l'utilisation de l'étalonnage serait moins fiable. C'est donc au groupe de CM1 que les performances de ce patient pourront être préférentiellement comparées. Ce raisonnement peut s'appliquer à tous les tests disposant de tranches scolaires d'étendue de 1 an et s'avère fondamental à appliquer au CP et au CE1, puisque les compétences langagières écrites y évoluent rapidement.²

Dans cet article, nous calculerons la proportion de tests de langage écrit disposant de tranches scolaires d'étendue inférieure à 1 an pour les enfants de niveau CP et/ou CE1. Nous analyserons cet aspect également pour les tests normés en âge s'adressant aux enfants de 6 ans et 7 ans, c'est-à-dire aux âges moyens correspondant respectivement aux classes de CP et CE1. Par ailleurs, nous relèverons les périodes de recueil des données normatives des tests présentant des normes scolaires, pour aider les orthophonistes à identifier la norme la plus adaptée à leurs patients, lorsqu'ils utilisent des tests de langage écrit.

². Ce procédé ne s'applique pas aux normes d'âge car un groupe d'âge peut correspondre à des enfants de niveaux scolaires différents. Par exemple, pour un test étalonné de septembre à décembre, la tranche d'âge de 9 ans réunit vraisemblablement des enfants scolarisés en CM1 et en CM2.

3. Les tranches d'âge : la clarté de leur présentation dans les manuels, les tableaux d'étalonnage et les logiciels

Cette section pointe 2 types d'écueils possibles, essentiellement d'ordre pratique, qui peuvent réduire la facilité d'utilisation des tests, voire entraîner des erreurs dans la comparaison des performances des patients à la norme.

Le premier type d'écueil résulte du mode de construction des tranches. Deux possibilités s'offrent principalement aux concepteurs de tests. La première est d'opter pour des tranches d'âge dont les bornes correspondent aux bornes communément admises par le grand public pour les âges considérés. Par exemple, dans l'E.CO.S.SE (Lecocq, 1996), le groupe d'enfants intitulé « 7 ans » inclut des individus de 7 ; 0 à 7 ; 11 ans. Ce type d'intitulé peut donc être interprété intuitivement par les cliniciens. La seconde possibilité pour les créateurs de tests est de construire les tranches autour d'âges cibles. Par exemple, dans la N-EEL, les auteurs proposent la tranche suivante pour encadrer l'âge de 4 ans : 3 ans 7 mois et 1 jour à 4 ans 6 mois et 30 jours (Chevrie-Muller & Plaza, 2001). Ce mode de construction est moins intuitif et peut conduire les cliniciens à réaliser des erreurs dans le choix du groupe de référence si les bornes inférieures et supérieures exactes des tranches d'âge ne sont pas rappelées dans les supports privilégiés de cotation que sont les tableaux d'étalonnage ou les logiciels. Or, dans certains tests présentant ce type de normes d'âge, seule la valeur centrale de la tranche est mentionnée sur ces supports. C'est le cas du test ELDP (Macchi et al., 2012)³. Par exemple, il est uniquement indiqué « 6 ans » pour la tranche d'âge allant de 5 ; 6 ans et 0 jour à 6 ; 5 ans et 30 jours. Ce manque de détail peut amener certains praticiens à interpréter erronément qu'il s'agit des enfants de 6 ; 0 à 6 ; 11 ans.

Le second type d'écueil concerne les tests pour lesquels les bornes inférieures et supérieures des tranches d'âge se recouvrent partiellement, c'est-à-dire lorsque la borne supérieure d'une tranche est également présentée comme la borne inférieure de la tranche suivante. Dans ce cas, l'orthophoniste peut hésiter dans le choix du groupe d'étalonnage et comparer les performances de l'enfant à celles d'un groupe d'étalonnage inadéquat. Par exemple, pour un enfant de 4 ; 6 ans évalué à l'aide d'ELOLA (de Agostini et al., 1998), il n'est pas possible de savoir si l'on doit comparer ses performances au groupe de référence de 4 ; 0 à 4 ; 6 ans ou à celui de 4 ; 6 à 5 ; 0 ans car on ne sait pas à quelle tranche correspond l'âge « 4 ; 6 ans ».

Dans cet article, nous identifierons les tests qui présentent au moins une tranche d'étalonnage construite autour d'un âge cible et nous indiquerons si les bornes des tranches d'âge sont renseignées explicitement sur les supports de cotation de ces tests. Nous calculerons également la proportion de tests dont les bornes inférieures et supérieures se recouvrent partiellement.

De manière générale, notre étude vise à analyser méthodiquement des tests francophones de langage oral et/ou écrit destinés aux enfants. Notre examen porte sur les tests actuellement utilisés en clinique, et concerne 3 points de vigilance relatifs à la comparaison des scores des patients à la norme : l'adéquation entre les types de normes utilisés et la modalité langagière évaluée, l'adéquation entre l'étendue des tranches de développement et la vitesse de progression des compétences évaluées dans la population de référence, la clarté de la présentation des tranches d'âge dans les tableaux d'étalonnage ou les logiciels des tests.

³. L'ELDP a fait l'objet de modifications récentes (en juillet 2018) concernant certains écueils présentés dans cet article. Nous pensons que les cliniciens ne disposent vraisemblablement pour le moment que de l'ancienne version, c'est pourquoi nous avons analysé la version antérieure de l'ELDP et non la version révisée.

----- MÉTHODOLOGIE -----

Le but de cet article est de contribuer à améliorer l'usage des tests utilisés par les orthophonistes en leur fournissant des informations pratiques directement applicables au matériel dont ils disposent. Cette étude cible donc les tests utilisés sur le terrain.

Elle ne cible pas les tests les plus solides sur le plan psychométrique, pour 3 raisons. Premièrement, comme précédemment évoqué, notre objectif principal est d'aborder 3 points précis liés à la cotation, et non de traiter de la qualité psychométrique générale des tests. Deuxièmement, nous souhaitons traiter des tests utilisés par les orthophonistes et non d'un sous-échantillon de ces tests (qui seraient satisfaisants sur le plan psychométrique). Troisièmement, les qualités psychométriques générales des tests francophones de langage oral et écrit semblent assez peu satisfaisantes (Bouchard et al., 2009 ; Leclercq & Veys, 2014). De ce fait, si nous avons dû prendre en compte dans nos analyses uniquement les tests dont les qualités psychométriques des normes étaient correctes (ex. au niveau de la récence, la taille et la représentativité des échantillons, de l'adéquation des normes à l'usage local), il est probable que cela aurait considérablement réduit notre échantillon et limité la portée de notre étude.

Nous avons donc recherché une liste renseignant précisément l'identité des outils utilisés par les orthophonistes et leur fréquence d'utilisation. Toutefois, à notre connaissance, une telle liste n'existe pas, à ce jour. Par conséquent, nous avons sélectionné les tests sur la base de 2 listes officielles répertoriant les outils d'évaluation à disposition des orthophonistes / logopèdes de France et de Belgique francophone. La première est une liste des tests orthophoniques destinés à l'évaluation du langage oral et écrit, diffusée sur Internet par l'UNADREO : la société savante en orthophonie reconnue par le Ministère français de l'Enseignement Supérieur et de la Recherche (UNADREO, 2011). La seconde liste est diffusée par l'Institut National d'Assurance Maladie-Invalidité, l'organisme d'assurance maladie belge (INAMI, 2017b, 2017a, 2017c). En Belgique, l'emploi de ces tests conditionne le remboursement des interventions logopédiques par l'assurance de santé, en cas de dyslexie, dysorthographe, troubles du langage et dysphasie. Les tests qui y sont référencés sont utilisés sur le terrain par les logopèdes indépendants ayant signé une convention avec l'organisme d'assurance de santé. Ces 2 listes ciblent donc des tests utilisés par des professionnels susceptibles d'être influencés par les informations communiquées par l'UNADREO et l'INAMI, c'est-à-dire, globalement, des orthophonistes/logopèdes exerçant sur le territoire francophone de France et de Belgique.

À elles 2, ces listes comptent 75 tests. Nous avons uniquement étudié les tests commercialisés en 2018 (ou diffusés en accès gratuit sur Internet, ou dans des articles scientifiques), les tests destinés à un public francophone, et qui disposent de normes. Ces critères nous ont conduits à écarter 21 tests qui n'étaient plus diffusés en 2018, qui s'adressaient à un public non francophone ou ne disposaient pas de normes (annexe 1). Lorsqu'il existait plusieurs versions d'un même test (une originale et une révisée), nous avons examiné la version la plus récente. Au final, 54 tests ont été étudiés (tableau 1).

Environ la moitié de ces tests porte exclusivement sur le langage oral (51,85 %, soit 28 tests), un tiers porte uniquement sur le langage écrit (35,19 %, soit 19 tests). Une minorité (12,96 %, soit 7 tests) concerne les 2 modalités orale et écrite. Les normes de l'ensemble des 54 tests concernent en moyenne des enfants de 9 ; 0 ans (ET = 3 ; 10 ; Mdn = 9 ; 0 ; étendue : 0 - plus de 18 ans)⁴. En moyenne, ces tests ont été commercialisés ou diffusés en 2005, soit une

⁴. Pour les tests étalonnés en niveaux scolaires nous avons considéré l'âge moyen équivalent à chaque niveau

ancienneté de 13 ans (ET = 6 ; 3 ; Mdn = 2006 ; étendue : 1981-2014).

Nous avons consulté ces tests dans différentes testothèques. Les manuels ont été intégralement lus avec une attention particulière portée aux 3 thématiques abordées dans cet article. Puisqu'après leur publication il arrive que les tests fassent l'objet de recherches complémentaires (Hogan, 2017), les communications scientifiques publiées jusque fin 2017 à propos de ces tests ont également été consultées (ex. chapitres de livres, articles, communications affichées lors de colloques).

Tableau 1.

Tests inclus dans l'étude.

Langage oral	BEPL-A (Chevrie-Muller et al., 1997a), BEPL-B (Chevrie-Muller et al., 1997b), BILO Petits (Khomsî & Khomsî, 2009), BILO-2 (Khomsî, Khomsî et Pasquet, 2007), BILO-3C (Khomsî, Khomsî, Pasquet et Parbeau-Guêno, 2007), BILO-EC2 (Khomsî & Khomsî, 2007), BIMM (Gatignol et al., 2007), Boehm-3 (Boehm, 2009), DEDALE (Deltour & Hupkens, 2011), ECSP (Guidetti & Tourrette, 2009), ELDP (Macchi et al., 2012), ELO (Khomsî, 2001), ELOLA (de Agostini et al., 1998), Épreuve de décision lexicale (Maillart, 2003 ; Maillart & Schelstraete, 2004), Épreuve de discrimination phonologique (Maillart, 2003 ; Maillart & Schelstraete, 2004), Épreuves de dénomination et de désignation (Kremin & Dellatolas, 1995), EVALO 2-6 (Coquet et al., 2009), EVIP (Dunn et al., 1993), EXALang 3-6 (Helloin & Thibault, 2006), Hénin (Hénin, 1981), IFDC (Kern & Gayraud, 2010), ISADYLE (Piérart et al., 2010), Kikou 3-8 (Boutard & Bouchet, 2009a), N-EEL (Chevrie-Muller & Plaza, 2001), PEES 3-8 (Boutard & Bouchet, 2009b), Péléa (Boutard et al., 2011), T.L.O.C.C. (Maurin, 2006), THaPho (Ecalte, 2007).
Langage écrit	Alouette-R (Lefavrais, 2005), ANALEC (Inizan, 1998), BALE (Jacquier-Roux et al., 2010), BATELEM Révisée (Savigny et al., 2001), BELEC (Mousty & Leybaert, 1999 ; Wessang & Gariel, 2008), BELO (George & Pech-Georgel, 2008), BLI (Khomsî, 2003), CHRONOSDICTÉES (Baneath et al., 2006), ECHAS-C (Simonart, 1998a), ECL (Khomsî et al., 2005), La forme noire (Maeder, 2010), Le vol du P.C. (Boutard et al., 2006), LMC-R (Khomsî, 1999), ODÉDYS Version 2 (Jacquier-Roux et al., 2005), Orthographe au collège (Thibault, 2008), PEDAC (Simonart, 1998b), Phonolec collège (Plaza et al., 2011), Timé 2 (Ecalte, 2018), Timé 3 (Ecalte, 2006).
Langage oral et écrit	CLéA (Pasquet et al., 2014), E.CO.S.SE. (Lecocq, 1996), EXALang 5-8 (Thibault et al., 2010), EXALang 8-11 (Thibault et al., 2012), EXALang 11-15 (Lenfant et al., 2009), L2MA-2 (Chevrie-Muller et al., 2010), TCS (Gilles, 2006 ; Maeder, 2006 ; Maeder et al., 2009).

----- RÉSULTATS -----

Les caractéristiques détaillées de chacun des tests sont présentées dans les tableaux des annexes 2 et 3 pour le langage oral et le langage écrit, respectivement. Puisque les analyses varient en fonction de la modalité langagière, les tests concernant le langage à la fois oral et écrit apparaissent dans les 2 tableaux. La présentation des résultats qui suit est organisée comme les parties précédentes, c'est-à-dire par type d'écueil possible. Nous abordons tout d'abord la question de l'adéquation entre les types de normes et la modalité langagière évaluée. Puis nous présentons les résultats concernant l'adéquation entre l'étendue des tranches de développement et la vitesse de progression des compétences évaluées dans la population générale. Enfin, nous décrivons les résultats portant sur la clarté de la présentation des tranches d'âge dans les manuels, les tableaux d'étalonnage et les logiciels.

1. Adéquation entre les types de normes et la modalité langagière évaluée

Ces analyses ont pour but de déterminer la proportion de tests présentant des normes adaptées, compte tenu de l'objectif d'évaluation. Concernant le langage oral, les normes de développement les plus adaptées sont généralement les normes d'âge. Sur 35 tests permettant d'évaluer le langage oral, 20 proposent exclusivement des normes d'âge soit 57,14 %, 12 tests proposent exclusivement des normes scolaires soit 34,29 % et 3 proposent les 2 types de normes soit 8,57 %⁵. Il y a donc 65,71 % de tests de langage oral qui présentent des normes d'âge.

Pour les tests de langage oral ne proposant que des normes scolaires, nous avons suggéré qu'il était pertinent de vérifier si l'âge moyen des groupes d'étalonnage était disponible dans le manuel, afin de sélectionner le groupe de niveau scolaire le plus proche en âge de celui de l'enfant évalué. Sur les 12 tests de langage oral comprenant uniquement des normes scolaires, 10 présentent la moyenne d'âge des participants des différents groupes de référence, soit 83,33 %.

Concernant le langage écrit, nous avons indiqué que les normes de développement les plus adaptées semblent les normes scolaires. Parmi les 26 tests permettant d'évaluer le langage écrit, 19 proposent exclusivement des normes scolaires soit 73,08 %, 5 tests proposent exclusivement des normes d'âge soit 19,23 %, 2 proposent les 2 types de normes soit 7,69 %. Il y a donc 80,77 % des tests de langage écrit qui présentent des normes scolaires.

2. Adéquation entre l'étendue des tranches de développement et la vitesse de progression des compétences évaluées dans la population générale

Ces analyses ont pour but de déterminer la proportion de tests dont les tranches de développement présentent des étendues suffisamment étroites (précises) suivant l'âge ou le niveau scolaire des participants et l'objectif de l'évaluation. Les résultats sont développés ci-après à propos des tests de langage oral puis écrit.

⁵. La BIMM (Gatignol et al., 2007) présente des types de normes distincts suivant la population cible : des normes scolaires pour les adolescents et des normes d'âge pour les adultes. Compte tenu de notre sujet d'étude (les tests chez les enfants), nous avons choisi de considérer que ces normes sont scolaires.

a. Les tests de langage oral

Nous avons expliqué qu'il était souhaitable qu'un test de langage oral concernant les enfants de 5 ans ou moins, ou de grande section de maternelle ou moins, dispose de tranches de développement (âge ou niveau scolaire) inférieures à 1 an. Sur les 35 tests permettant d'évaluer le langage oral, 27 concernent les enfants de 5 ans ou moins, ou de grande section de maternelle ou moins, soit 77,14 %. Parmi ces 27 tests, 15 présentent des tranches de développement de moins de 1 an, soit 55,56 % et 12 tests présentent des tranches de développement de 1 an ou plus, soit 44,44 %.

b. Les tests de langage écrit

Il est souhaitable qu'un test de langage écrit adressé aux enfants de CP ou de CE1 (ou âge équivalent) présente des tranches de développement (âge ou niveau scolaire) inférieures à 1 an. Sur les 26 tests permettant d'évaluer le langage écrit, 16 concernent les enfants de CP et de CE1 (ou des âges équivalents) soit 61,54 %. Sur ces 16 tests, 2 présentent des tranches d'étalonnage inférieures à 1 an à la fois pour le CP et le CE1 soit 12,50 %, 3 présentent des tranches d'étalonnage inférieures à 1 an juste pour le CP soit 18,75 % et 11 présentent des tranches d'étalonnage de 1 an soit 68,75 %.

Par ailleurs, pour les tests de langage écrit concernant les enfants de CP ou de CE1 avec des tranches scolaires supérieures ou égales à 1 an, nous avons suggéré de tenir compte de la période de recueil des données, afin de sélectionner la tranche dont la période scolaire est la plus proche de celle de l'enfant. Douze tests de langage écrit avec des normes scolaires concernent le CP et le CE1 et possèdent au moins une tranche d'étalonnage d'une étendue de 1 an. Sur ces 12 tests, 8 indiquent les périodes de l'année au cours desquelles les données normatives ont été recueillies, soit 66,67 %.

Cette façon de sélectionner la tranche scolaire peut être appliquée à tout test de langage écrit dont une ou plusieurs tranches d'étalonnage sont égales à 1 an, afin d'améliorer la précision de la mesure. Dix-neuf tests de langage écrit, toutes classes confondues, présentant des normes scolaires, possèdent au moins une tranche d'étalonnage d'une étendue de 1 an. Sur ces 19 tests, 12 indiquent les périodes de l'année au cours desquelles les données normatives ont été recueillies, soit 63,16 %.

3. Les tranches d'âge : la clarté de leur présentation dans les manuels, les tableaux d'étalonnage et les logiciels

Nous avons indiqué que les tranches d'âge peuvent être exprimées soit d'une manière conventionnelle (ex. « 7 ans » : 7 ans et 0 mois à 7 ans et 11 mois), soit autour de l'âge cible (ex. « 4 ans » : 3 ans 7 mois et 1 jour à 4 ans 6 mois et 30 jours). Pour éviter toute erreur du clinicien, il semble important : (1) que les bornes des tranches d'âge soient explicitement présentées dans les manuels ou tableaux d'étalonnage, (2) que les bornes exactes des tranches d'âge soient rappelées dans les tableaux d'étalonnage. Parmi les 30 tests disposant de normes d'âge, 18 présentent des tranches dont les bornes correspondent à l'idée que l'on se fait usuellement des âges considérés soit 60,00 % ; 8 présentent des tranches construites autour d'un âge cible soit 26,67 % (ex. classe « 3 ans » allant de 2 ; 9 à 3 ; 3 ans dans EVALO 2-6) ; 4 ne présentent pas les bornes exactes au clinicien et ne renseignent qu'une valeur tant dans le manuel que dans les tableaux d'étalonnage soit 13,33 % (ex. classe « 7 ans » dans l'Alouette-

R). Parmi les 8 tests présentant des tranches d'âge construites autour d'un âge cible, 4 ne rappellent pas les bornes exactes des tranches dans les tableaux d'étalonnage ou les logiciels. Par ailleurs, il arrive que les normes d'âge présentent des intitulés équivoques, les tranches d'âge semblant se recouper. Sur l'ensemble des 30 tests disposant de normes d'âge, 8 présentent un intitulé équivoque, puisque les bornes supérieures des tranches recouvrent les bornes inférieures des suivantes, soit 26,67 %.

----- DISCUSSION-----

Les résultats sont discutés par type d'écueil, dans le même ordre que pour les sections précédentes.

1. Adéquation entre les types de normes et la modalité langagière évaluée

Dans cette partie, nous abordons l'adéquation entre les types de normes et la modalité langagière évaluée pour les tests de langage oral puis écrit. Par la suite, nous présentons certaines limites de notre travail sur cette thématique.

a. Les tests de langage oral

En introduction, nous avons souligné que les normes établies par tranche d'âge sont généralement préférables aux normes scolaires pour les tests de langage oral. Parmi les tests analysés, environ 66 % présentent effectivement des normes d'âge. Il paraît regrettable qu'environ 34 % proposent exclusivement des normes scolaires, comme l'ELO (Khomsî, 2001) ou les EXALang 5-8, 8-11 et 11-15 (Lenfant et al., 2009 ; Thibault et al., 2010, 2012). Dans ce cas, nous suggérons de comparer les performances du patient au groupe de référence dont la moyenne d'âge est la plus proche de l'âge du patient. Pour cela, le clinicien peut prendre connaissance de l'âge moyen des participants de chaque groupe de référence dans environ 83 % des tests sans normes d'âge. Nous avons identifié ces tests et relevé les pages des manuels où se trouve l'information afin d'en faciliter la recherche par les orthophonistes (annexes 2 et 3). Seules la L2MA-2 (Chevrie-Muller et al., 2010) et la BIMM (Gatignol et al., 2007) ne mettent pas ces informations à disposition des cliniciens. Cela implique d'interpréter avec prudence les mesures d'écart à la norme des patients, pour ces 2 tests.

b. Les tests de langage écrit

S'agissant des tests de langage écrit, nous avons justifié en introduction qu'il paraissait plus adéquat de disposer de normes scolaires. Parmi les tests analysés, environ 81 % présentent des normes scolaires et environ 19 % sont exclusivement étalonnés en âge, comme l'ANALEC (Inizan, 1998). Pour ces derniers, il semble préférable de comparer les performances du patient avec celles du groupe de référence dont l'âge correspond le mieux à l'âge moyen observé dans la classe du patient. Cela est particulièrement important lorsque l'enfant a redoublé.

L'adéquation entre les types de normes des tests et les compétences évaluées paraît un peu meilleure pour le langage écrit que pour le langage oral. En effet, le pourcentage de tests de langage écrit proposant des normes scolaires (81 %) est plus important que le pourcentage de tests de langage oral proposant des normes d'âge (66 %). Pour les tests de langage oral, le recrutement des participants pour l'établissement de normes s'effectue généralement en écoles ou collèges, par commodité. Il est possible que ce mode de recrutement incite les auteurs des tests de langage oral à agencer leurs normes en fonction de l'organisation du recueil des données,

c'est-à-dire par niveaux scolaires.

c. Limites de notre analyse

Notre analyse présente plusieurs limites. Elle est fondée sur une distinction entre les tests de langage oral qui devraient idéalement présenter des normes par tranches d'âge (car l'apprentissage du langage oral se fait par l'exposition à une langue maternelle, dans tous les milieux de vie de l'enfant), et des tests de langage écrit qui devraient présenter des normes par niveaux scolaires (car son apprentissage dépend essentiellement des programmes scolaires). Ces principes se vérifient pour la plupart des tests. Par exemple, la compréhension orale des mots tels que « cochon » ou « courgette » extraits du test de langage oral ELO, dépend vraisemblablement davantage de l'âge chronologique que des programmes scolaires. Inversement, la capacité à écrire des mots ou des phrases sous dictée dans le test BALE, dépend certainement plus du niveau scolaire que de l'âge. Toutefois, ces généralités ne se vérifient pas systématiquement. En particulier, certains auteurs sélectionnent volontairement, pour leurs tests de langage oral, des items fréquemment utilisés dans le cadre scolaire, afin d'évaluer si la maîtrise orale des jeunes est suffisante pour réussir leur scolarité. Par exemple, l'auteur du T.L.O.C.C. écrit « Nous avons (...) effectué une sélection de mots (...) en étudiant plus spécifiquement le vocabulaire employé à partir de la 4^e et jusqu'en seconde » (Maurin, 2006, p. 38). Parmi les items de ce test, l'on trouve des mots assez liés aux programmes scolaires, tels que « pétrolier » ou « kamikaze ». En somme, l'utilisation de normes scolaires pour certains tests de langage oral paraît justifiée par la sous-compétence évaluée et le type d'items utilisé.

Par ailleurs, considérer que les normes des tests de langage oral doivent obligatoirement être exprimées en âges et celles des tests de langage écrit en niveaux scolaires présente une autre limite. En effet, ce principe ne paraît pas adapté lorsqu'un clinicien souhaite comparer les performances d'un patient dans un même domaine, sur 2 modalités différentes (ex. compréhension orale vs écrite). Dans l'idéal, cela nécessite de comparer les performances du patient à celles d'une seule et même population de référence, à l'aide d'une même batterie de langage oral et écrit. Dans le cas contraire, les mesures d'écart à la norme ne seraient pas tout à fait comparables, puisqu'elles correspondraient à 2 populations de référence différentes. Lorsqu'un clinicien souhaite comparer les performances d'un même patient à l'oral et à l'écrit, il est donc justifié qu'il utilise, pour les 2 épreuves d'une même batterie, soit des normes d'âge, soit des normes scolaires. S'il choisit les normes d'âge, inévitablement, la comparaison pour la modalité écrite n'est pas idéale. Inversement, s'il choisit des normes scolaires, c'est la comparaison à la norme pour la modalité orale qui ne l'est pas. Mais dans les 2 cas, la comparaison des performances entre une modalité et l'autre paraît correcte, sur le plan méthodologique.

2. Adéquation entre l'étendue des tranches de développement et la vitesse de progression des compétences évaluées dans la population générale

Pour obtenir une précision optimale dans l'évaluation de la distance à la norme des performances des patients, nous avons indiqué qu'il serait idéal de disposer de normes avec des tranches d'étendue inférieure à 1 an, pour les enfants de moins de 5 ans en langage oral, et pour le CP et le CE1 en langage écrit.

a. Les tests de langage oral

Près de 56 % des tests de langage oral pour des enfants de 5 ans ou moins ou de grande section

ou moins présentent des tranches d'étalonnage d'une étendue de moins de 1 an. Dans 44 % des cas, les tranches ont une étendue de 1 an ou plus. De ce fait, l'on peut regretter que pour environ la moitié des tests de cet âge, l'appréciation de la distance à la norme des performances du patient manque de précision.

Huit tests de langage oral avec un étalonnage par niveau scolaire, concernent les enfants scolarisés en maternelle. Parmi ces 8 tests, seuls 2 proposent des tranches d'étalonnage inférieures à 1 an : le THaPho (Ecalte, 2007) et la Boehm-3 (Boehm, 2009). Les 6 autres tests cumulent donc 2 limites : (1) présenter des normes scolaires alors que les normes d'âge semblent plus adaptées aux tests de langage oral, (2) présenter des tranches d'étalonnage larges alors que l'évolution des compétences testées est rapide à cet âge. Il paraît donc d'autant plus important pour ces tests de vérifier la moyenne d'âge de chaque groupe de niveau afin de ne pas amplifier la marge d'erreur lors de la cotation. Les pages des manuels où apparaissent ces informations sont renseignées dans l'annexe 2.

b. Les tests de langage écrit

Parmi les tests concernant les enfants de CP et de CE1, il est étonnant de constater qu'une majorité, environ 69 %, ne présente que des tranches d'étalonnage de 1 an. Uniquement 31 % des tests pour les enfants de CP et de CE1 comportent des tranches d'étalonnage inférieures à 1 an, pour 1 ou 2 de ces niveaux scolaires. Les tests BELO (George & Pech-Georgel, 2008), EXALang 5-8 (Thibault et al., 2010) et PEDAI C (Simonart, 1998b) présentent des étalonnages par trimestre ou par semestre uniquement pour le CP. Seuls le Timé 2 (Ecalte, 2018) et la BATELEM Révisée (Savigny et al., 2001) proposent des tranches d'étalonnage d'étendue inférieure à 1 an, à la fois pour le CP et le CE1. Ces outils permettent donc une évaluation précise des compétences des patients en langage écrit, puisqu'ils tiennent compte de l'importance de la progression des performances en reconnaissance des mots écrits et en transcription, au moins au cours du CP voire en CE1. Pour les tests étalonnés par année scolaire complète, pour ces 2 niveaux, il est tout de même possible dans environ 67 % des cas de choisir le groupe de comparaison le plus adapté en prenant connaissance des périodes de recueil des données normatives. Si ce procédé paraît crucial lorsque le clinicien utilise des tests ciblant des enfants de niveau CP et CE1, il reste pertinent pour l'évaluation des performances d'enfants de niveaux scolaires supérieurs, afin de connaître précisément la mesure de l'écart à la norme. En ce sens, il est possible de consulter la période de recueil des données normatives dans 63 % des tests avec des normes scolaires, dont au moins une tranche d'étalonnage possède une étendue de 1 an. Le tableau de l'annexe 3 présente, pour les tests de langage écrit avec des normes scolaires, ces périodes précises de recueil lorsqu'elles sont disponibles dans les manuels. Nous espérons que cet outil pratique facilitera la recherche de cette information par les cliniciens.

c. Limites de notre analyse

Il nous paraît idéal de disposer de normes avec des tranches d'étendue inférieure à 1 an, pour les enfants de moins de 5 ans en langage oral, et pour le CP et le CE1 en langage écrit. Toutefois, cette idée très générale porte sur le langage (oral ou écrit) dans son ensemble. Il est probable, en réalité, que chacune des sous-compétences langagières (orales ou écrites) ne se développe pas exactement à la même vitesse que les autres, durant ces périodes. Par exemple pour le langage oral, l'essentiel de l'articulation est mis en place vers 5 ans (McLeod & Crowe, 2018). À partir de 5 ans, on peut donc s'attendre à ce que les enfants progressent peu au niveau articulatoire alors que des progrès seraient attendus aux niveaux lexical et morphosyntaxique. En effet, le vocabulaire se développe tout au long de la vie et toutes les structures syntaxiques

ne sont pas acquises à la fin de l'école élémentaire (Dehaene, 2019). Pour le langage écrit, la reconnaissance des mots écrits et la transcription de mots réguliers et de pseudomots se développent considérablement en CP et en CE1. En revanche, les compétences en transcription de mots irréguliers évoluent de façon plus progressive tout au long de l'école élémentaire (Sprenger-Charolles et al., 2003). Il serait donc moins important de disposer de tranches étroites pour la transcription de mots irréguliers que pour la transcription de mots réguliers et de pseudomots. En définitive, disposer de normes avec des tranches d'étendue inférieure à 1 an, pour les enfants de moins de 5 ans en langage oral, et pour le CP et le CE1 en langage écrit est une règle à ajuster en fonction du type de sous-compétence évaluée.

3. Les tranches d'âge : la clarté de leur présentation dans les manuels, les tableaux d'étalonnage et les logiciels

Deux types d'écueil liés à la construction des tranches d'âge et leur présentation dans les manuels, les tableaux d'étalonnage et les logiciels peuvent être décrits.

Le premier type d'écueil résulte du mode de construction des tranches d'âge. En effet, les tranches d'âge peuvent être construites de 2 manières : soit en utilisant les bornes communément admises pour les âges considérés, soit en les construisant autour d'un âge cible. Ce second mode de construction est moins intuitif et nécessite que les concepteurs de tests rappellent clairement les bornes des tranches d'âge dans les tableaux d'étalonnage ou les logiciels en complément des informations présentées dans les manuels. Pourtant, dans 4 tests, seule une valeur centrale de la tranche est indiquée dans au moins 1 des supports courants de cotation (ex. les tableaux d'étalonnage). Nous invitons donc les cliniciens à consulter attentivement tous les supports de cotation, lorsqu'ils utilisent ces tests. Il s'agit de la BEPL-B (Chevrie-Muller et al., 1997b), d'EVALO 2-6 (Coquet et al., 2009), de l'ELDP (Macchi et al., 2012) et d'EXALang 3-6 (Helloin & Thibault, 2006). Dans 4 autres tests, les auteurs ne renseignent pas non plus dans les manuels les bornes exactes des tranches d'âge et ne donnent qu'une valeur d'âge générale. En conséquence, il n'est pas possible pour le clinicien de déterminer si la tranche d'âge est construite de façon intuitive (ex. « 12 ans » : 12 ; 0 à 12 ; 11 ans) ou autour d'un âge cible (ex. « 12 ans » : 11 ; 6 ans à 12 ; 6 ans). Il est donc difficile de choisir le groupe de référence adapté pour comparer les scores du patient à la norme. Par exemple, dans *Le vol du P.C.* (Boutard et al., 2006), il n'est pas possible, à notre connaissance, de savoir avec certitude si le groupe de référence dont l'intitulé d'âge est « 12 ans » correspond à des enfants de 11 ; 6 ans à 12 ; 6 ans ou de 12 ; 0 à 12 ; 11 ans. Concernant l'*Alouette-R* (Lefavrais, 2005), cet écueil ne devrait pas impacter la cotation étant donné que les normes scolaires sont disponibles et peuvent être utilisées préférentiellement aux normes d'âge, puisqu'il s'agit d'un test de langage écrit.

Le second type d'écueil possible concerne les tests pour lesquels la borne supérieure d'une tranche est également la borne inférieure de la tranche suivante (ex. de 3 ; 3 à 3 ; 9 ans, de 3 ; 9 à 4 ; 3 ans). Dans ce cas, il semble préférable de considérer la borne supérieure présentée comme exclue de la tranche, car cela correspond à l'usage le plus fréquent des intitulés d'âge. Les tests concernés sont : la BEPL-A (Chevrie-Muller et al., 1997a), la BEPL-B (Chevrie-Muller et al., 1997b), l'ELOLA (de Agostini et al., 1998), l'EVALO 2-6 (Coquet et al., 2009), l'EXALang 3-6 (Helloin & Thibault, 2006), le test de Hénin (Hénin, 1981), le Péléa (Boutard et al., 2011), les épreuves de dénomination et de désignation (Kremin & Dellatolas, 1995). La superposition des bornes peut être une source d'erreur car elle peut conduire le clinicien à comparer les performances du patient à un groupe de référence erroné, dans le cas (malgré tout relativement rare) où l'âge du patient correspond précisément à l'une des bornes.

Au final, si l'orthophoniste considère qu'il dispose d'informations insuffisantes ou peu claires dans le manuel d'un test pour pouvoir coter et interpréter correctement les réponses d'un patient, le guide d'utilisation des tests édité par la Commission Internationale des Tests recommande de ne pas utiliser cet outil (Vrignaud et al., 2003). Par ailleurs, une présentation non-équivoque que les futurs concepteurs de tests pourraient s'approprier apparaît dans le Kikou 3-8 (Boutard & Bouchet, 2009a) et le PEES 3-8 (Boutard & Bouchet, 2009b). Dans ces tests, les tranches sont présentées clairement dans les tableaux d'étalonnage et leurs bornes apparaissent précisément comme incluses ou exclues. Par exemple, les performances d'enfants de 3 ans et 4 ans sont à comparer aux classes respectives suivantes : [3 ; 4[ans et [4 ; 5[ans. Cette présentation permet d'éviter les confusions décrites précédemment.

4. Vers des standards de qualité pour les tests de langage

Sur le plan international, il semble qu'une réflexion de fond soit en cours, de la part des chercheurs et des cliniciens, à propos de la qualité des tests. En effet, de nombreux guides de bonne pratique ont été publiés par la Commission Internationale des Tests, en particulier le guide d'utilisation des tests traduit en 17 langues (Vrignaud et al., 2003). Il existe également un ouvrage à l'attention des concepteurs et des utilisateurs des tests (American Educational Research Association et al., 2014). Il décrit les standards de qualité des outils d'évaluation. La création de grilles d'évaluation des tests témoigne également de cette préoccupation pour la qualité des outils (Evers et al., 2013 ; Hilaire-Debove, 2017), de même que les études concernant les qualités psychométriques des tests anglophones de langage oral et écrit pour l'enfant (McCauley & Swisher, 1984 ; Friberg, 2010 ; Kirk & Vigeland, 2014 ; Flipsen & Ogiela, 2015). Néanmoins, dans ce champ de recherche, une étude à la fois récente et étendue sur les qualités psychométriques des tests francophones de langage oral et écrit destinés aux enfants paraît nécessaire.

5. Limites de notre étude et remarques complémentaires

Le but de cet article est de traiter de tests utilisés par les professionnels du langage, afin de mettre à leur disposition des informations pratiques sur l'usage des normes de ces tests et leur cotation. En l'absence de document renseignant l'identité et la fréquence d'utilisation des tests par les orthophonistes, nous avons sélectionné les outils d'évaluation étudiés grâce à 2 listes pré-établies. L'une des 2 est française, l'autre concerne des tests dont l'utilisation conditionne le remboursement des interventions logopédiques par l'assurance de santé belge. Les tests qui y sont référencés sont utilisés par les logopèdes indépendants ayant signé une convention avec l'assurance de santé. De plus, chacun des 54 tests était commercialisé ou diffusé en accès gratuit sur Internet en 2018. Pour ces 2 raisons, il est probable que ces outils soient actuellement utilisés sur le terrain. Toutefois, notre sélection ne garantit pas que les tests répertoriés soient les plus fréquemment utilisés par l'ensemble des orthophonistes ou logopèdes exerçant sur les territoires francophones français et belge. Il se peut que nous ayons analysé plusieurs tests peu employés et qu'à l'inverse quelques tests fréquemment administrés sur le terrain aient été exclus.

Par ailleurs, cette étude s'intéresse à un champ très spécifique de la conception des tests. Elle a pour objectif l'analyse de thématiques précises afin d'optimiser l'utilisation des tests par les praticiens et ne porte pas d'appréciation générale sur chacun des tests susmentionnés. Elle invite les cliniciens à rester vigilants lors du choix du groupe de comparaison et lors de l'interprétation des scores.

Enfin, une évaluation ne se réduit pas à des informations quantitatives sur l'écart entre les

compétences du patient et celles d'une population de référence. Un examen fin et complet intègre les diverses informations quantitatives mais également qualitatives délivrées par les tests, à d'autres composantes moins formelles de l'évaluation, par exemple l'analyse des informations anamnestiques ou le comportement de l'enfant lors du jeu libre.

----- CONCLUSION -----

Dans cet article, nous avons évoqué plusieurs écueils potentiels liés à la conception et à l'utilisation des normes des tests de langage oral et écrit chez l'enfant. Dans un nombre non négligeable de cas, ces normes ne sont pas adaptées à la modalité langagière évaluée, à la vitesse d'évolution des compétences testées, et les intitulés des tranches d'âge peuvent être équivoques dans les tableaux d'étalonnage, les logiciels ou les manuels.

Nous avons décrit ces aspects et formulé quelques conseils pour contourner ces écueils ou s'en prémunir, afin de tendre vers une évaluation langagière optimale. Nous proposons aux cliniciens 2 tableaux en annexes, un pour le langage oral et un pour le langage écrit. Ces tableaux leur permettront de retrouver les résultats présentés ci-dessus pour un grand nombre des tests utilisés. Lorsque le clinicien choisit le groupe de référence permettant de comparer les performances de son patient à la norme, il nous semble important qu'il reste vigilant et examine attentivement les manuels pour un usage adapté et critique des tests : « le mauvais usage des mesures n'est pas intrinsèque à l'outil lui-même, mais est de la responsabilité de l'utilisateur de cet outil » (Grégoire, 2014, p. 462). Par ailleurs, nous espérons que ces réflexions pourront être utiles lors de la conception de nouveaux tests. Enfin, les aspects étudiés dans cette étude ne suffisent assurément pas à eux seuls pour apprécier la qualité des tests. Les propriétés psychométriques sont par exemple des caractéristiques très importantes à prendre en compte (Leclercq & Veys, 2014; Perdrix, 2018) de même que des qualités non-psychométriques, comme l'aspect ludique et attrayant du matériel.

----- REMERCIEMENTS -----

Les auteurs remercient les professionnels de l'équipe pluridisciplinaire du Centre Régional de Diagnostic des Troubles d'Apprentissage de Lille, pour les échanges particulièrement riches et animés qui s'y développent et qui représentent une aide précieuse dans l'élaboration de la pensée. Ils remercient également S. Ravez de leur avoir facilité la consultation de nombreux tests à la testothèque du département d'orthophonie de Lille, et M. Chetaneau pour sa réflexion sur les propriétés psychométriques des tests de langage oral et écrit.

----- BIBLIOGRAPHIE -----

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.

American Psychiatric Association. (2013). *DSM-5. Diagnostic and Statistical Manual of Mental Disorders (5th ed.)*. American Psychiatric Association Publishing.

American Psychiatric Association. (2015). *DSM-5. Manuel diagnostique et statistique des troubles mentaux (5e édition)*. Elsevier Masson.

- Aubert, J. et Blanchard, S. (1991). *L'évaluation des compétences d'un lecteur*. EAP.
- Autesserre, D., Deltour, J.-J. et Lacert, P. (1988). *EDP 4-8. Epreuve de Discrimination phonémique pour enfants de 4 à 8 ans*. EAP.
- Baneath, B., Boutard, C. et Alberti, C. (2006). *Chronodictées. Outil d'évaluation des performances orthographiques avec et sans contrainte temporelle, du CE1 à la 3e*. OrthoÉdition.
- Bassano, D. (2008). Présentation du DLPF : Inventaire du Développement du Langage de Production en Français. *Cahiers de la Société Belge des Logopèdes Universitaires*, 30(3), 3-12.
- Bassano, D. (2007). Émergence et développement du langage : enjeux et apports des nouvelles approches fonctionnalistes. Dans E. Demont et M.-N. Metz-Lutz (dir.), *L'acquisition du langage et ses troubles* (p. 13-46). Solal. <https://hal.archives-ouvertes.fr/hal-00601129/document>
- Bassano, D., Labrell, F., Champaud, C., Lemétayer, F. et Bonnet, P. (2005). Le DLPF : un nouvel outil pour l'évaluation du développement du langage de production en français. *Enfance*, 57(2), 171-208. <https://www.cairn.info/revue-enfance1-2005-2-page-171.htm>
- Beitchman, J. H., Wilson, B. A., Brownlie, E., Walters, H. et Lancee, W. J. (1996). Long-term consistency in speech/language profiles : I. Developmental and academic outcomes. *Journal of the American Academy of Child & Adolescent Psychiatry*, 35(6), 804-814. <https://doi.org/10.1097/00004583-199606000-00021>
- Bernicot, J., Laval, V., Bateau, B. et Lacroix, A. (2005). L'évaluation des capacités pragmatiques chez l'enfant : présentation de nouveaux outils. Dans B. Piérart (dir.), *Le langage de l'enfant, comment l'évaluer ?* (p. 161-177). De Boeck. <https://www.cairn.info/le-langage-de-l-enfant--9782804145620-page-161.htm>
- Bishop, D. V. M. (1998). Development of the Children's Communication Checklist (CCC): A method for assessing qualitative aspects of communicative impairment in children. *Journal of Child Psychology and Psychiatry*, 39(6), 879-891. <https://doi.org/10.1017/S0021963098002832>
- Bishop, D. V. M., Snowling, M. J., Thompson, P. A. et Greenhalgh, T. (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>
- Boehm, A. E. (2009). *Test des concepts de base de Boehm-3* (3e édition). ECPA.
- Bouchard, M.-E. G., Fitzpatrick, E. M. et Olds, J. (2009). Psychometric analysis of assessment tools used with francophone children. *Canadian Journal of Speech-Language Pathology and Audiology*, 33(3), 129-139. https://cjslpa.ca/files/2009_CJSLPA_Vol_33/No_03_113-160/Bouchard_Fitzpatrick_Olds_CJSLPA_2009.pdf
- Boutard, C. et Bouchet, M. (2009a). *Kikou 3-8. Protocole d'évaluation de la compréhension*

syntaxique et narrative chez l'enfant de 3 à 8 ans 11 mois. OrthoÉdition.

Boutard, C. et Bouchet, M. (2009b). *PEES 3-8. Protocole d'Évaluation de l'Expression Syntaxique de 3 à 8 ans 11 mois.* OrthoÉdition.

Boutard, C., Claire, I. et Gretchanovsky, L. (2006). *Le vol du P.C. Évaluation fonctionnelle de la lecture chez les sujets de 11 à 18 ans.* OrthoÉdition.

Boutard, C., Guillon, A. et Charlois, A.-L. (2011). *Pélée. Protocole d'Évaluation du Langage Élaboré de l'Adolescent.* OrthoÉdition.

Chevrie-Muller, C. et Plaza, M. (2001). *N-EEL. Nouvelles Épreuves pour l'Examen du Langage de 3 ans 7 mois à 8 ans 7 mois.* ECPA.

Chevrie-Muller, C., Maillart, C., Simon, A.-M. et Fournier, S. (2010). *L2MA-2. Batterie Langage oral, Langage écrit, Mémoire, Attention, du CMI à la 6e* (2e édition). ECPA.

Chevrie-Muller, C., Simon, A.-M., Le Normand, M.-T. et Fournier, S. (1997a). *Batterie d'Évaluation PsychoLinguistique. BEPL-A et BEPL-A « forme courte » de 2 ans 9 mois à 4 ans 3 mois.* Pearson Canada Assessment.

Chevrie-Muller, C., Simon, A.-M., Le Normand, M.-T. et Fournier, S. (1997b). *Batterie d'Évaluation Psycholinguistique. BEPL-B de 2 ans 9 mois à 4 ans 3 mois.* Pearson Canada Assessment.

Coquet, F., Ferrand, P. et Roustit, J. (2009). *EVALO 2-6. ÉVALuation du développement du Langage Oral chez l'enfant de 2 ans 3 mois à 6 ans 3 mois.* OrthoÉdition.

De Agostini, M., Metz-Lutz, M.-N., Van Hout, A., Chavance, M., Deloche, G., Pavao-Martins, I. et Dellatolas, G. (1998). Batterie d'évaluation du langage oral de l'enfant aphasique (ELOLA): standardisation française (4-12 ans). *Revue de Neuropsychologie*, 8(3), 319-367.

De Renzi, E. et Ferrari, C. (1978). The Reporter's Test: A sensitive test to detect expressive disturbances in aphasics. *Cortex*, 14(2), 279-293.
[https://doi.org/10.1016/S0010-9452\(78\)80054-9](https://doi.org/10.1016/S0010-9452(78)80054-9)

Dehaene, S. (dir.) (2019). *La science au service de l'école.* Odile Jacob.

Dehaene, S. (2007). *Les neurones de la lecture.* Odile Jacob.

Deltour, J.-J. (2002). *TCG-R. Test de closure grammaticale. Version révisée.* ATM.

Deltour, J.-J. et Hupkens, D. (2011). *DEDALE. Batterie du Dépistage Expérimental des Difficultés d'Apprentissage de la Lecture et de l'Écriture.* Eurotests éditions.

Dewart, H. et Summers, S. (1997). *The pragmatics profile of everyday communication skills in childhood.* NFER-Nelson.
http://complexneeds.org.uk/modules/Module-2.4-Assessment-monitoring-and-evaluation/All/downloads/m08p080c/the_pragmatics_profile.pdf

- Doutriaux, F. et Lepez, R. (1980). *TNO. Test de Niveau en Orthographe*. ECPA.
- Dunn, L. M., Thériault-Whalen, C. et Dunn, L. M. (1993). *Échelle de vocabulaire en images Peabody. Adaptation française du Peabody Picture Vocabulary test-revised*. Psycan.
- Ecalte, J. (2006). *Timé 3. Test d'identification de mots écrits pour enfants de 7 à 15 ans*. Mot à Mot.
- Ecalte, J. (2007). *THaPho, Test d'habiletés phonologiques pour enfants de 5 à 8 ans*. Mot à Mot.
- Ecalte, J. (2018). *Timé 2. Test d'identification de mots écrits de 6 à 8 ans*. Adeprio.
- Ecalte, J. et Magnan, A. (2002). *L'apprentissage de la lecture : Fonctionnement et développement cognitifs*. Armand Colin.
- Evers, A., Hagemester, C., Høstm, A., Lindley, P., Muñiz, J., et Sjöberg, A. (2013). EFPA review model for the description and evaluation of psychological and educational tests. EFPA.
- Flipsen, P. et Ogiela, D. A. (2015). Psychometric characteristics of single-word tests of children's speech sound production. *Language, Speech, and Hearing Services in Schools*, 46(2), 166-178. https://doi.org/10.1044/2015_LSHSS-14-0055
- Friberg, J. C. (2010). Considerations for test selection: How do validity and reliability impact diagnostic decisions? *Child Language Teaching and Therapy*, 26(1), 77-92. <https://doi.org/10.1177/0265659009349972>
- Garcia, L. J., Paradis, J., Sénécal, I. et Laroche, C. (2006). Use of and satisfaction with assessment tools for evaluating communication disorders in French. *Canadian Journal of Speech-Language Pathology and Audiology*, 30(4), 239-249. https://cjslpa.ca/files/2006_JSLPA_Vol_30/No_04_209-263/Garcia_Paradis_Senecal_Laroche_JSLPA_2006.pdf
- Gatignol, P., Marin-Curtoud, S. et ERU 16. (2007). *BIMM. Batterie Informatisée du Manque du Mot*. ECPA.
- George, F. et Pech-Georgel, C. (2008). *BELO. Batterie d'Evaluation de Lecture et d'Orthographe CP-CE1*. De Boeck Solal.
- Gilles, P.-Y. (2006). *Test de compréhension syntaxique (ERU 10). Résultats CMI à 3ème (2004 & 2005)*. OrthoEdition. [https://www.orthoedition.com/upload/291120061506RESULTATSTCS2\(tous\).pdf](https://www.orthoedition.com/upload/291120061506RESULTATSTCS2(tous).pdf)
- Grégoire, J. (2014). L'examen diagnostique est-il normatif ? *A.N.A.E.*, 26(132-133), 459-465.
- Guidetti, M. et Tourrette, M. (2009). *ECSP. Echelle d'Evaluation de la Communication Sociale Précoce*. Eurotests éditions.
- Haute Autorité de Santé. (2017). Comment améliorer le parcours de santé d'un enfant avec troubles spécifiques du langage et des apprentissages ?

https://www.has-sante.fr/portail/upload/docs/application/pdf/2018-01/guide_tsla_vf.pdf

Helloin, M.-C. et Thibault, M.-P. (2006). *EXALang 3-6. Batterie d'examen des fonctions langagières chez l'enfant de 3 à 6 ans*. Happyneuron.

Hilaire-Debove, G. (2017). *Pourquoi et comment évaluer les outils d'évaluation en orthophonie*. Dans P. Gatignol et T. Rousseau (dir.), *Actes des XVIIes rencontres internationales d'orthophonie: Efficacité des thérapies*. OrthoEdition.

https://www.unadreo.org/wp-content/uploads/2020/03/04_HilaireDebove.pdf

Hénin, N. (1981). Étude de la motricité et des praxies oro-faciales chez l'enfant de 2 ans et demi à 12 ans et demi. *Bulletin d'Audiophonologie*, 12, 17-42.

Hilton, L. M. (1989). Identification et évaluation des différences pragmatiques du langage à partir de méthodes non-psychométriques. *Glossa*, 18, 14-20.

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, 26(1), 55-88. <https://doi.org/10.1016/j.dr.2005.11.002>

Hogan, T. P. (2017). *Introduction à la psychométrie*. (N. Parent & R. Stephenson, Trad.) (2e édition). Chenelière Éducation.

Hutchinson, T. A. (1996). What to look for in the technical manual : Twenty questions for users. *Language, Speech, and Hearing Services in Schools*, 27(2).

<https://doi.org/10.1044/0161-1461.2702.109>

INAMI. (2017a). Liste limitative des tests pour l'évaluation de la lecture. <http://www.riziv.fgov.be/SiteCollectionDocuments/liste-logopedes-tests-dyslexie.pdf>

INAMI. (2017b). Liste limitative des tests pour l'évaluation de l'orthographe. <http://www.riziv.fgov.be/SiteCollectionDocuments/liste-logopedes-tests-dysorthographe.pdf>

INAMI. (2017c). Liste limitative des tests pour l'évaluation du langage oral et de la dysphasie. http://www.riziv.fgov.be/SiteCollectionDocuments/liste_logopedes_tests_langage_dysphasie_2021.pdf

Inizan, A. (1998). ANALEC. *Analyse du savoir lire de 8 ans à l'âge adulte*. ECPA.

Inserm. (2007). *Expertise collective. Dyslexie. Dysorthographe. Dyscalculie. Bilan des données scientifiques*. Les éditions Inserm.

Jacquier-Roux, M., Lequette, C., Pouget, G., Valdois, S. et Zorman, M. (2010). *BALE. Batterie Analytique du Langage Ecrit*. Cogni-sciences.

Jacquier-Roux, M., Valdois, S. et Zorman, M. (2005). *ODÉDYS Version 2. Outil de DÉpistage des DYSlexies*. Cogni-sciences.

Kerbel, D. et Grunwell, P. (1998a). A study of idiom comprehension in children with semantic-pragmatic difficulties. Part I: Task effects on the assessment of idiom comprehension in children. *International Journal of Language & Communication Disorders*, 33(1), 1-22.

<https://doi.org/10.1080/136828298247901>

Kerbel, D. et Grunwell, P. (1998b). A study of idiom comprehension in children with semantic-pragmatic difficulties. Part II: Between-groups results and discussion. *International Journal of Language & Communication Disorders*, 33(1), 23-44.
<https://doi.org/10.1080/136828298247901>

Kern, S. et Gayraud, F. (2010). *IFDC. Inventaire Français du Développement Communicatif. 8/30 mois*. Editions de la Cigale.

Kerr, A., Guildford, S. et Bird, E. K.-R. (2003). Standardized Language Test Use : A Canadian Survey. *Canadian Journal of Speech-Language Pathology and Audiology*, 27(1), 10-28.
https://cjslpa.ca/files/2003_JSLPA_Vol_27/No_01_1-92/Kerr_Guildford_Kay-Raining_Bird_JSLPA_2003.pdf

Khomsi, A. (2003). *BLI. Bilan de Lecture Informatisé*. ECPA.

Khomsi, A. (2001). *ELO. Evaluation du Langage Oral de la PSM au CM2*. ECPA.

Khomsi, A. (1999). *LMC-R. Epreuve d'évaluation de la compétence en lecture*. ECPA.

Khomsi, A. (1987). *O-52. Évaluation des stratégies de compréhension en situation orale*. ECPA.

Khomsi, A. et Khomsi, J. (2009). *BILO Petits. Bilan Informatisé de Langage Oral pour le Cycle 1*. ECPA.

Khomsi, A. et Khomsi, J. (2007). *BILO-EC2. Évaluation des Contraintes pour le cycle 2*. ECPA.

Khomsi, A., Khomsi, J. et Pasquet, F. (2007). *BILO-2. Bilan Informatisé de Langage Oral pour le cycle 2*. ECPA.

Khomsi, A., Khomsi, J., Pasquet, F. et Parbeau-Guéno, A. (2007). *BILO-3C. Bilan Informatisé de Langage Oral pour le cycle 3 et le collège*. ECPA.

Khomsi, A., Nanty, I., Parbeau-Guéno, A. et Pasquet, F. (2005). *ECL. Évaluation des Compétences Linguistiques au collège*. ECPA.

Kirk, C. et Vigeland, L. (2014). A psychometric review of norm-referenced tests used to assess phonological error patterns. *Language, Speech, and Hearing Services in Schools*, 45(4), 365-377. https://doi.org/10.1044/2014_LSHSS-13-0053

Kremin, H. et Dellatolas, G. (1995). L'accès au lexique : une étude de standardisation chez l'enfant d'âge pré-scolaire. *Revue de Neuropsychologie*, 5(3), 309-338.

Launay, L. (2018). Du DSM-5 au diagnostic orthophonique : élaboration d'un arbre décisionnel. *Rééducation Orthophonique*, 273, 71-92.
<https://www.fno.fr/wp-content/uploads/2019/01/RO-2018-Du-DSM5-au-diagnostic-orthoq-LL-1.pdf>

Laveault, D. et Grégoire, J. (2014). *Introduction aux théories des tests en psychologie et en*

sciences de l'éducation. De Boeck.

Leclercq, A.-L. et Veys, E. (2014). Réflexions sur le choix de tests standardisés lors du diagnostic de dysphasie. *A.N.A.E.*, 131, 374-382. <http://hdl.handle.net/2268/180427>

Lecocq, P. (1996). *L'E.CO.S.SE. Une Épreuve de COmpréhension Syntaxico-SEmantique*. Presses Universitaires du Septentrion.

Lefavrais, P. (2005). *Alouette-R. Test d'analyse de la lecture et de la dyslexie*. ECPA.

Lenfant, M., Thibault, M.-P. et Helloin, M.-C. (2009). *EXALang 11-15. Batterie informatisée du langage oral, langage écrit, compétences transversales. Collégiens - Adolescents*. Happyneuron.

Lieberman, A. M. (1992). The relation of speech to reading and writing. Dans R. Frost et L. Katz (dir.), *Orthography, phonology, morphology, and meaning*. Elsevier.

Lobrot, M. (1967). *Batterie d'épreuves pour mesurer la lecture et l'orthographe (ORLEC)*. EAP.

Macchi, L., Descours, C., Girard, E., Guitton, E., Morel, C., Timmermans, N. et Boidein, F. (2012). *ELDP. Épreuve Lilloise de Discrimination Phonologique destinée aux enfants de 5 ans à 11;6 ans*. <https://medecine.univ-lille.fr/orthophonie/reseau-professionnel>

Maeder, C. (2010). *La forme noire*. OrthoÉdition.

Maeder, C. (2006). *TCS. Test de Compréhension Syntaxique*. OrthoÉdition.

Maeder, C., Gilles, P.-Y. et Médina, F. (2009). Test de Compréhension Syntaxique (TCS): résultats de la validation en version orale, analyse et intérêt dans la pratique clinique. *Glossa*, 107, 1-16.

Maillart, C. (2003). *Origine des troubles morphosyntaxiques chez des enfants dysphasiques*. Thèse de doctorat en logopédie, non publiée. Université Catholique de Louvain. <http://orbi.ulg.ac.be/bitstream/2268/73809/1/thesemaillart2.pdf>

Maillart, C. et Schelstraete, M.-A. (2004). L'évaluation des troubles phonologiques : illustration de la démarche diagnostique. Dans M.-A. Schelstraete et M.-P. Noël (dir.), *Les troubles du langage et du calcul chez l'enfant. Une approche psycholinguistique et neuropsychologique* (p. 113-147). E.M.E. & InterCommunications. <https://orbi.uliege.be/handle/2268/8004>

Maurin, N. (2006). *T.L.O.C.C. Test de Langage Oral Complexe pour Collégiens*. OrthoÉdition.

McCauley, R. J. et Swisher, L. (1984). Psychometric review of language and articulation tests for preschool children. *Journal of Speech and Hearing Disorders*, 49(1), 34-42. <https://doi.org/10.1044/jshd.4901.34>

McLeod, S. et Crowe, K. (2018). Children's consonant acquisition in 27 languages : A cross-linguistic review. *American Journal of Speech-Language Pathology*, 27(4), 1546-1571.

https://doi.org/10.1044/2018_AJSLP-17-0100

Monfort, M., Juarez, A. et Juarez Sanchez, I. (2005). *Les troubles de la pragmatique chez l'enfant*. OrthoÉdition.

Mousty, P. et Leybaert, J. (1999). Évaluation des habiletés de lecture et d'orthographe au moyen de BELEC : données longitudinales auprès d'enfants francophones testés en 2e et 4e années. *Revue Européenne de Psychologie Appliquée*, 49(4), 325-342.

Parisse, C. et Maillart, C. (2004). Le développement morphosyntaxique des enfants présentant des troubles de développement du langage : données francophones. *Enfance*, 56(1), 20-35. <https://doi.org/10.3917/enf.561.0020>

Pasquet, F., Parbeau-Guéno, A. et Bourg, E. (2014). *CLéA. Communiquer, Lire et Écrire pour Apprendre*. ECPA.

Perdrix, R. (2018). Pour une contribution éclairée et raisonnée de l'évaluation standardisée et normalisée aux diagnostics de troubles développementaux du langage. Quelques éléments pour choisir, comprendre et exploiter les tests en orthophonie. *Rééducation Orthophonique*, 273, 47-70.

Phelps-Terasaki, D. et Phelps-Gunn, T. (2007). *TOPL-2. Test of pragmatic language (2nd edition)*. Pro-Ed.

Piérart, B., Comblain, A., Grégoire, J. et Mousty, P. (2010). *Batterie ISADYLE*. De Boeck Solal.

Plaza, M., Oudry, M., Gatignol, P. et Robert, A.-M. (2011). *Phonolec Collège*. Adeprio.

Puyuelo Sanclemente, M., Solanas Pérez, A., Wiig, E.H. et Renom Pinsachs J. (2007). *BLOC-SR. Bateria de lenguaje objetiva y criterial screening revisada*. Masson.

Ravard, F. et Ravard, J.-C. (1992). *NBNO-C. Test scolaire d'évaluation du niveau en orthographe (usage et grammaire) et de dépistage des dysfonctionnements de l'écrit*. Editions scientifiques et psychologiques.

Ravard, J.-C., Ravard, F. et Anglade, J.-C. (1993). *Test des compétences verbales et métalinguistiques. NBTL*. EAP.

Rousseau, T. (2016). Labellisation des outils d'évaluation utilisés en orthophonie. *L'orthophoniste*, 361, 34-35.

Savigny, M., Barbier, C., Coupey Le Roy, R., Girard, J. et Roussel, G. (2001). *BATELEM Révisée. Batterie d'épreuves pour l'école élémentaire. Cycle II et première année du cycle III*. ECPA.

Shulman, B. B. (1986). *Test of pragmatic skills (Revised)*. Communication Skill Builders.

Simon, C. S. (1986). *Evaluating communicative competence: A functional pragmatic procedure*. Communication Skill Builders.

- Simonart, G. (1998a). *ECHAS-C: Echelle d'apprentissages scolaires primaires*. Eurotests éditions.
- Simonart, G. (1998b). *PEDAIC. Tests pédagogiques de premier cycle primaire*. Eurotests éditions.
- Sprenger-Charolles, L., Siegel, L. S., Béchennec, D. et Serniclaes, W. (2003). Development of phonological and orthographic processing in reading aloud, in silent reading, and in spelling: A four-year longitudinal study. *Journal of Experimental Child Psychology*, 84(3), 194-217. [https://doi.org/10.1016/S0022-0965\(03\)00024-9](https://doi.org/10.1016/S0022-0965(03)00024-9)
- Thibault, M.-P. (2008). *Orthographe au collège*. Orthomotus.
- Thibault, M.-P., Helloin, M.-C. et Croteau, B. (2010). *EXALang 5-8. Batterie informatisée pour l'examen du langage oral et écrit chez l'enfant de 5 à 8 ans*. Happyneuron.
- Thibault, M.-P., Lenfant, M. et Helloin, M.-C. (2012). *EXALang 8-11. Batterie informatisée d'examen du langage oral, langage écrit, mémoire, attention, compétences transversales*. Happyneuron.
- UNADREÓ. (2011). Liste des tests orthophoniques utilisés pour l'évaluation du langage oral et langage écrit. <https://www.unadreo.org/la-recherche/tests-en-orthophonie/>
- Van Rompaey, M., Schelstraete, M.-A. et Mariol, M. (2004). Apprentissage explicite de règles implicites en orthographe lexicale : Une étude de cas. Dans M.-A. Schelstraete et M.-P. Noël (dir), *Les troubles du langage et du calcul chez l'enfant. Une approche psycholinguistique et neuropsychologique* (p. 185-205). E.M.E. & Intercommunications.
- Vrignaud, P., Castro, D. et Mogenet, J.-L. (2003). Recommandations internationales sur l'utilisation des tests. Version 2000. *Pratiques Psychologiques*, Numéro spécial hors-série.
- Wessang, L. et Gariel, P. (2008). Lecture et écriture de mots isolés en 6^e : étalonnage de la BELEC. *Glossa*, 106, 52-68.
- Wetherby, A. M. (1986). Ontogeny of communicative functions in autism. *Journal of Autism and Developmental Disorders*, 16(3), 295-316. <https://doi.org/10.1007/BF01531661>
- Wetherby, A. M. et Prutting, C. A. (1984). Profiles of communicative and cognitive-social abilities in autistic children. *Journal of Speech & Hearing Research*, 27(3), 364-377. <https://doi.org/10.1044/jshr.2703.364>

----- ANNEXES -----

Annexe 1. Tests exclus de l'étude

Langage oral	BLOC-SR (Puyuelo Sanclemente et al., 2007), CCC (Bishop, 1998), Compréhension d'expressions imagées (Kerbel & Grunwell, 1998a, 1998b), CSBS (Wetherby, 1986 ; Wetherby & Prutting, 1984), DLPF (Bassano, 2008 ; Bassano et al., 2005), EDP 4-8 (Autesserre et al., 1988), EEC (Simon, 1986), ESO (Dewart & Summers, 1997), L.E.C.P.C. (Bernicot et al., 2005), Liste de contrôle du langage pragmatique (Hilton, 1989), NBTL (Ravard et al., 1993), O-52 (Khomsî, 1987), Profil des troubles pragmatiques (Monfort et al., 2005), Reporter's Test (De Renzi & Ferrai, 1978), TCG-R (Deltour, 2002), Test of pragmatic skills (Shulman, 1986), TOPL-2 (Phelps-Terasaki & Phelps-Gunn, 2007)
Langage écrit	Évaluation des compétences d'un lecteur (Aubert & Blanchard, 1991), NBNO (Ravard & Ravard, 1992), ORLEC (Lobrot, 1967), TNO (Doutriaux & Lepez, 1980)

Annexe 2. Caractéristiques des tests permettant d'évaluer le langage oral

Test	Norme A = âge S = scolaire	Âge min. d'applica tion (année ; mois)	Âge max. d'applica tion (année ; mois)	Étendue des tranches d'âge pour les enfants ≤ 5 ans	Tranche d'âge construite autour d'un âge cible	Recouvre ment des bornes	Niveau scolaire min. d'applica tion	Niveau scolaire max. d'applica tion	Étendue des tranches scolaires ≤ GSM	Moyenne d'âge par niveau scolaire (tests S) : si présentée dans le manuel, page indiquée ci-dessous
BEPL-A	A	2 ; 9	4 ; 1	6 mois en 6 mois	Oui	Oui	–	–	–	–
BEPL-B	A	2 ; 9	4 ; 1	6 mois en 6 mois	Oui	Oui	–	–	–	–
BILO Petits	A	3 ; 0	5 ; 5	6 mois en 6 mois	Non	Non	–	–	–	–
BILO-2	S	5 ; 4	7 ; 5	–	–	–	GSM	CE1	Par an	p. 25
BILO-3C	S	7 ; 8	15 ; 6	–	–	–	CE1	3e	–	p. 26
BILO-EC2	S	5 ; 4	7 ; 5	–	–	–	GSM	CE1	Par an	p. 21
BIMM	S	11-12	14-15	–	–	–	6e	3e	–	Non renseignée
Boehm-3	S	5 ; 3	7 ; 11	–	–	–	GSM	CE1	2 niveaux : début et fin d'année	p. 47
CléA (modalité orale)	A	2 ; 6	14 ; 11	6 mois en 6 mois	Non	Non	–	–	–	–
DEDALE	S	4 ; 6	6 ; 6	–	–	–	MSM	CP	Par an	p. 37
E.CO.S.SE. (modalité orale)	A	4 ; 0	12 ; 11	6 mois en 6 mois	Non	Non	–	–	–	–
ECSP	A	0 ; 3	2 ; 6	6 mois en 6 mois	Non	Non	–	–	–	–
ELDP	A	4 ; 6	11 ; 5	Par an	Oui	Non	–	–	–	–
ELO	S	3 ; 4	10 ; 4	–	–	–	PSM	CM2	Par an	p. 3
ELOLA	A	4 ; 0	12 ; 0	6 mois en 6 mois	Non	Oui	–	–	–	–
Épreuve de décision lexicale (Maillart & Schelstraete, 2004)	A	4	7	1,5 ans d'écart	Non renseignée	Non renseigné	–	–	–	–
Épreuve de discrimination phonologique (Maillart & Schelstraete, 2004)	A	4	7	1,5 ans d'écart	Non renseignée	Non renseigné	–	–	–	–
Épreuves de dénomination et de désignation (Kremin & Dellatolas, 1995)	A	3 ; 0	6 ; 0	6 mois en 6 mois	Non	Oui	–	–	–	–
EVALO 2-6	A	2 ; 3	6 ; 3	6 mois en 6 mois	Oui	Oui	–	–	–	–

EVIP	A	2 ; 6	18 ; 0 et plus	6 mois en 6 mois	Non	Non	–	–	–	–
EXALang 3-6	A	2 ; 8	5 ; 10	6 mois en 6 mois	Oui	Oui	–	–	–	–
EXALang 5-8 (modalité orale)	S	6 ; 1	8 ; 2	–	–	–	GSM	CE1	Par an	p. 9
EXALang 8-11 (modalité orale)	S	8 ; 9	10 ; 10	–	–	–	CE2	CM2	–	p. 9
EXALang 11-15 (modalité orale)	S	11 ; 6	15 ; 0	–	–	–	6e	3e	–	p. 12
Hénin	A	2 ; 6	12 ; 6	Par an	Oui	Oui	–	–	–	–
IFDC	A	0 ; 8	2 ; 5	1 mois en 1 mois	Non	Non	–	–	–	–
ISADYLE	A	3 ; 0	11 ; 11	6 mois en 6 mois	Non	Non	–	–	–	–
Kikou 3-8	A	3 ; 0	8 ; 11	Par an	Non	Non	–	–	–	–
L2MA-2 (modalité orale)	S	7-8	11-12	–	–	–	CE1	6e	–	Non renseignée
N-EEL	A	3 ; 7	8 ; 7	Par an	Non	Non	–	–	–	–
PEES 3-8	AS	3 ; 0	8 ; 11	Par an	Non	Non	PSM	CE2	Par an	–
Péléa	A	11 ; 0	18 ; 0 et plus	–	Non	Oui	–	–	–	–
T.L.O.C.C.	AS	9 ; 6	16 ; 9	–	Non	Non	6e	3e	–	–
TCS (modalité orale)	AS	8 ; 6	15 ; 11	–	Oui	Non	CM1	3e	–	–
THaPho	S	5 ; 3	7 ; 6	–	–	–	GSM	CE1	2 niveaux : début et milieu d'année	p. 12

Note. Les tirets correspondent à des tests non concernés par l'analyse.

Annexe 3. Caractéristiques des tests permettant d'évaluer le langage écrit

Test	Norme A = âge S = scolaire	Âge min. d'applica tion (année ; mois)	Âge max. d'applica tion (année ; mois)	Étendue des tranches d'âge CP/CE1	Tranche d'âge construite autour d'un âge cible	Recouv rement des bornes	Niveau scolaire min. d'applica tion	Niveau scolaire max. d'applica tion	Étendue des tranches scolaires	Période de recueil des normes pour les tests avec des normes scolaires
Alouette-R	AS	6 ; 0	16 ; 0	Par an	Non renseignée	Non renseigné	CP	Term	Par an	Non renseignée
ANALEC	A	7 ; 7	15 ; 6	–	Oui	Non	–	–	–	–
BALE	S	–	–	–	–	–	CE1	CM2	Par an	Informations apparemment contradictoires : de décembre à avril (p. 8), de novembre à février (p. 17)
BATELEM Révisée	S	–	–	–	–	–	GSM	CE2	Début et fin d'année scolaire	Du 15 au 30 octobre et du 15 au 30 mai
BELEC	S	–	–	–	–	–	CE1	6 ^e	Par an	Non renseigné
BELO	S	–	–	–	–	–	CP	CE1	1 norme par trimestre pour le CP	CP - Trimestre 1 : fin novembre à fin janvier CP CP - Trimestre 2 : février à avril CP
BLI	S	–	–	–	–	–	CE1	CM2	Par an	CP - Trimestre 3 : mai CP à décembre CE1 CE1 : février CE1 à novembre CE2
CHRONOSDICTÉES	S	–	–	–	–	–	CE1	3 ^e	Par an	De septembre à décembre
CLéA (modalité écrite)	A	2 ; 6	14 ; 11	Par an	Non	Non	–	–	–	–
E.CO.S.SE. (modalité écrite)	A	7 ; 0	14 ; 5	Par an	Non	Non	–	–	–	–
ECHAS-C	S	–	–	–	–	–	CE2	6 ^e	Par an	Du 15 mai au 30 juin
ECL	S	–	–	–	–	–	6 ^e	3 ^e	Par an	Non renseignée
EXALang 5-8 (modalité écrite)	S	–	–	–	–	–	GSM	CE1	Mi et fin CP 1 seule norme pour tout le CE1	En juin et en septembre (fin d'année scolaire) et mars pour "mi CP"

EXALang 8-11 (modalité écrite)	S	-	-	-	-	-	CE2	CM2	Par an	De février à avril
EXALang 11-15 (modalité écrite)	S	-	-	-	-	-	6e	3e	Par an	De janvier à mai
L2MA-2 (modalité écrite)	S	-	-	-	-	-	CE1	6e	Par an	Non renseignée
LA FORME NOIRE	A	9 ; 0	11 ; 11	-	-	Non	-	-	-	-
Le vol du P.C.	A	11	18 et plus	-	-	Non renseignée	-	-	-	-
LMC-R	S	-	-	-	-	-	CE1	5e	Par an	De septembre à décembre
ODÉDYS Version 2	S	-	-	-	-	-	CE1	5e	Par an (/!\ pas de groupe contrôle en classe de 6 ^e)	De novembre à février
Orthographe au collège	S	-	-	-	-	-	6e	3e	Par an	Milieu d'année scolaire
PEDA1C	S	-	-	-	-	-	CP	CE1	Mi et fin CP 1 seule norme pour tout le CE1	En juin
Phonolec collègue	S	-	-	-	-	-	6e	3e	Par an	D'octobre à avril
TCS (modalité écrite)	AS	8 ; 6	15 ; 5	-	-	Non	CM1	3e	Par an	Non renseignée
Timé-2	S	-	-	-	-	-	CP	CE2	Début et milieu d'année	De septembre à octobre et de janvier à février
Timé-3	S	-	-	-	-	-	CE1	3e	Par an	De septembre à décembre

Note. Les tirets correspondent à des tests non concernés par l'analyse.