



**HAL**  
open science

**Mesure des effets retest de l’“ Alouette ” et du “ L3 ”  
issu de la batterie “ Orlec ” chez des enfants  
normo-lecteurs de CM1 et CM2**

Ophélie Tavormina

► **To cite this version:**

Ophélie Tavormina. Mesure des effets retest de l’“ Alouette ” et du “ L3 ” issu de la batterie “ Orlec ” chez des enfants normo-lecteurs de CM1 et CM2. Médecine humaine et pathologie. 2021. dumas-03597170

**HAL Id: dumas-03597170**

**<https://dumas.ccsd.cnrs.fr/dumas-03597170>**

Submitted on 4 Mar 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MÉMOIRE présenté pour l'obtention du  
**CERTIFICAT DE CAPACITE D'ORTHOPHONISTE  
ET DU GRADE DE MASTER 2 D'ORTHOPHONIE**

Présenté et soutenu le 10 juin 2021 par TAVORMINA Ophélie

**Né(e) le 29/06/1996**

**MESURE DES EFFETS RETEST  
DE L'«ALOUETTE» ET DU « L3 » ISSU  
DE LA BATTERIE « ORLEC » CHEZ DES  
ENFANTS NORMO-LECTEURS DE CM1  
ET CM2**

Directeur de Mémoire : LELOUP Gilles

Co-directeur(s) de Mémoire : ESKINAZI Karine

Nice

**2021**

# Remerciements

« La chenille devient papillon » m'avait dit mon prof de maths de prépa quand j'ai eu le concours de Nice. Aujourd'hui, je finis ce mémoire, et j'aurai bientôt le tant attendu diplôme. Malgré les dures épreuves de ma vie, je me sens chanceuse car j'ai croisé votre route à tous. C'est par ce qui vous rend si spéciaux, que j'aimerais vous remercier vous qui avez rendu ce travail possible de près ou de loin.

**Savoir et guidance** : Merci Gilles Leloup d'avoir dirigé ce mémoire, et de nous avoir apporté vos connaissances précieuses et merci à vous Karine Eskinazi d'avoir codirigé nos travaux. Merci à Mme Manera pour son aide avec ma partie préférée ... les statistiques.

**Écoute et humanité** : Mme Auriane Gros, merci d'avoir toujours été à notre écoute. Dans les coups durs, vous avez su nous apaiser. Merci également à chaque membre du DON, de nous avoir accompagnées pendant ces 5 ans (à travers les vertes et les pas mûres), aujourd'hui vous avez solidement formé des professionnelles de santé fières de leur métier et de leur parcours.

**Amour et patience** : distance, covid, urgences, béquilles, multiples confinements, mental break-downs, mémoire nocturne, 140 km par jour, vie dans 26m<sup>2</sup>, déménagements, j'en passe... tu t'es toujours si bien occupé de moi, alors merci à toi qui partage ma vie, de m'avoir épaulée et soutenue cette année, de m'avoir motivée et remonté le moral quand c'était difficile, d'avoir toujours été attentionné, merci d'avoir quitté ton NORD pour mon sud et d'avoir été si présent dans cette dernière ligne droite, merci de faire autant rire, merci pour tout j'ai tant de chance.

**Amour et soutien** : merci à ma famille, à ma mère et à mon petit frère (plus si petit en 5ans...), pour votre amour et votre soutien sans faille, (et merci pour la mousse au chocolat maman). Merci à toi mon papa de me donner la force chaque jour, « nous savons que tu serais avec nous aujourd'hui si le ciel n'était pas si loin ».

**Bienveillance et confiance** : merci à tous les maîtres de stage bienveillants qui ont croisé ma route et qui m'ont donné confiance en moi, merci de m'avoir tant appris.

**Amitié et fous rires** : merci à ma famille niçoise, jamais ces études n'auraient été pareilles sans vous. Grandir à vos côtés a été magnifique, partager les joies et les peines avec vous a rendu ma vie meilleure. Merci pour ces fous rires qui donnent mal au ventre, ces soirées, ces journées BU (inter)minables, ces coups de soleil, ces faux anniversaires en terrasse. Merci de faire que je ne garde que le meilleur de tout ce que nous avons vécu en 5 ans, et ce n'est pas fini !



# Sommaire

<b>INTRODUCTION .....</b>	<b>1</b>
<b>PARTIE THEORIQUE .....</b>	<b>2</b>
<b>CHAPITRE 1 LA LECTURE .....</b>	<b>3</b>
<b>1. DEFINITION .....</b>	<b>3</b>
<b>2. MODELES COMPUTATIONNELS : LECTURE ET COMPETENCES COGNITIVES</b>	
<b>SOUS-JACENTES.....</b>	<b>4</b>
<b>3. COMPREHENSION : LES PROCESSUS MIS EN JEU EN LECTURE .....</b>	<b>6</b>
<b>CHAPITRE 2 L'EFFET RETEST .....</b>	<b>7</b>
<b>1. NOTION D'EFFET RETEST ET QUALITES PSYCHOMETRIQUES D'UN TEST .....</b>	<b>7</b>
1.1 NOTION D'EFFET RETEST.....	7
1.2 QUALITES PSYCHOMETRIQUES D'UN TEST : SENSIBILITE, SPECIFICITE, FIDELITE ET VALIDITE.	7
<b>2. ÉTAT DES LIEUX GLOBAL DE L'EFFET RETEST DANS LA LITTERATURE .....</b>	<b>8</b>
<b>3. TESTS ORTHOPHONIQUES : PERTINENCE DE L'EFFET RETEST DANS LA</b>	
<b>CLINIQUE</b>	<b>8</b>
<b>CHAPITRE 3 TESTS.....</b>	<b>10</b>
<b>1. L'«ALOUETTE» .....</b>	<b>10</b>
1.1 DESCRIPTION DU TEST.....	10
1.2 MODE DE PASSATION .....	11
1.3 COTATION .....	11
<b>2. L'ÉPREUVE « L3 » ISSU DE LA BATTERIE « ORLEC ».....</b>	<b>11</b>
2.1 DESCRIPTION DU TEST.....	11
2.2 MODE DE PASSATION .....	12
2.3 COTATIONS .....	12
<b>3. INTERET D'UN TEST SIGNIFIANT VS UN TEST NON SIGNIFIANT.....</b>	<b>12</b>
<b>PARTIE PRATIQUE .....</b>	<b>14</b>
<b>1. PROBLEMATIQUE ET OBJECTIFS.....</b>	<b>14</b>
<b>2. MATERIEL ET METHODE .....</b>	<b>14</b>
2.1 SUJETS .....	14
2.2 MATERIEL .....	15
2.3 METHODE .....	16

<b>3.</b>	<b>STRATEGIE D'ANALYSE DES DONNEES.....</b>	<b>17</b>
<b>4.</b>	<b>RESULTATS.....</b>	<b>18</b>
4.1	ANALYSE DES EFFETS RETEST A L'«ALOUETTE» ET AU « L3 » ISSU DE LA BATTERIE « ORLEC »	18
4.2	COMPARER L'EFFET RETEST DU TEST DE L'«ALOUETTE» A CELUI DU TEST « L3 ».....	19
	<b>DISCUSSION.....</b>	<b>21</b>
	<b>CONCLUSIONS ET PERSPECTIVES .....</b>	<b>24</b>
	<b>RÉFÉRENCES .....</b>	<b>25</b>

# INTRODUCTION

Le cadre méthodologique de la remédiation des troubles du langage écrit nécessite des évaluations répétées des compétences en lecture par la passation de tests en pré et post intervention. Afin de disposer de résultats comparables et étalonnés, les orthophonistes cliniciens utilisent alors les mêmes tests pour évaluer la progression des compétences travaillées et leur généralisation. Parfois espacées d'un intervalle de temps assez court, ces évaluations répétées peuvent provoquer un effet retest, qui n'est pas toujours mesuré ni décrit dans les propriétés métriques des tests utilisés. Si cet effet n'est pas pris en compte, cela induit des biais lors de l'analyse des résultats, qui pourraient être interprétés par l'examineur comme une normalisation ou un gain de performances, alors que les gains constatés sont en partie attribuables à un effet retest.

Partant de ce postulat, cette étude exploratoire a pour but de vérifier si un effet retest existe aux tests de « l'«Alouette» » et au test « L3 » issu de la batterie « Orlec », entre deux temps de passations à deux mois d'intervalle. Ces deux tests sont fréquemment utilisés en orthophonie pour poser un diagnostic de dyslexie, le premier étant un test de leximétrie non significatif, et le second un test mesurant la compréhension.

Ces deux tests sont donc conçus pour évaluer des compétences de lecture communes, comme l'identification des mots et la vitesse de lecture, mais possèdent un accès à la compréhension différent. La lecture repose sur l'association de ces compétences d'identification et de compréhension, qui seront décrites dans la première partie de notre revue de littérature. Puis, la description de modèles de lecture permettra d'illustrer les différents types de codages phonologiques, orthographiques, et sémantiques nécessaires à une lecture fluide. Ensuite, la notion d'effet retest, au travers des qualités psychométriques des tests et des études ayant validé des effets répétés, sera présentée afin de poser notre problématique. Les deux tests utilisés pour cette étude, « l'«Alouette» » et le « L3 » seront ensuite décrits en détaillant les particularités de chacun. Pour finir, la méthodologie adoptée pour cette étude sera présentée suivie des données recueillies, et d'une discussion des résultats.

# **PARTIE THEORIQUE**



# CHAPITRE 1 LA LECTURE

## 1. Définition

La lecture est l'une des fonctions cognitives les plus étudiées, qu'il s'agisse de comprendre le fonctionnement de ce système chez le lecteur expert, les mécanismes en lien avec l'apprentissage de la lecture chez l'enfant ou l'origine des troubles de lecture (Phénix et al., 2017). La lecture est un mécanisme complexe qui repose sur l'intégration d'informations visuelles, orthographiques, phonologiques et sémantiques (Ziegler, 2019). Les descriptions de ces mécanismes font référence à divers type de codages phonologiques, orthographiques, sémantiques et morphologiques (Ferrand et Ayora, 2017) qui sont modélisés dans différents types de modèles de lecture computationnels.

Schématiquement, la lecture est décrite dans le modèle de Gough & Tunmer (1986) comme l'addition de l'identification des mots écrits et de la compréhension orale (Lecture = Identification x Compréhension :  $L = I \times C$ ). La compréhension écrite dépend de l'automatisation des procédures d'identification du mot écrit mais également du niveau de compréhension orale. L'identification des mots écrits (IME) dépend de différents traitements sous-jacents (phonologique, orthographique et sémantique) alors que la compréhension orale dépend des capacités de vocabulaire du lecteur. Cependant, le développement du vocabulaire est également influencé par les compétences d'IME, et particulièrement des codages phonologique (décodage des mots), orthographique (reconnaissance des mots) et sémantique. Cette importance du codage sémantique est particulièrement probante au début de l'apprentissage de la lecture, puis l'appui sur le contexte diminue avec l'âge et l'automatisation de l'IME (Stanovich et al., 1981).

Au début de l'apprentissage de la lecture, pour les langues ayant une écriture dite alphabétique, la reconnaissance des mots écrits se base sur le décodage : l'enfant apprend à associer les graphèmes aux phonèmes qui correspondent et chaque fois que le lecteur réussit à décoder un mot, les connexions utilisées pour décoder ce mot seront renforcées, comme un cercle vertueux d'apprentissage (Sprenger-Charolles & Ziegler, 2019). Ainsi, la lecture repose sur l'automatisation des procédures d'identification des mots écrits (Ziegler, 2019) et quand ces procédures sont suffisamment automatisées, le lecteur peut mobiliser ses ressources cognitives dans la compréhension de ce qui est lu (Castles et al., 2018).

## 2. Modèles computationnels : Lecture et compétences cognitives sous-jacentes

L'enjeu des recherches comportementales est d'une part, d'identifier les mécanismes cognitifs impliqués en lecture, mais aussi de comprendre comment ces mécanismes interagissent entre eux pour proposer une architecture des liens qu'ils entretiennent (Phénix et al., 2017).

Des modèles théoriques descriptifs, utilisés initialement en clinique, sont conçus à partir de l'analyse des erreurs de patients alexiques ayant subi des lésions cérébrales. Ces modèles représentent les deux voies de lecture : une voie phonologique qui permet la lecture de mots inconnus ou de pseudo-mots grâce à la conversion graphème-phonème, et une voie lexicale qui permet la lecture de mots connus ou irréguliers grâce à la reconnaissance globale des mots. Les modèles de Marshall & Newcombe (1973), et de Coltheart (1978) (qui évoluera ensuite vers une modélisation sur ordinateur en 2001), dits modèles à double voie, représentent la lecture et ses mécanismes sous-jacents. La modélisation se présente sous forme de schémas simples avec des flèches spécifiant les relations entre les différents mécanismes de lecture, alors représentés sous forme de boîtes reliées les unes aux autres. Le modèle théorique descriptif de Morton (1980) distingue également un principe de hiérarchie et un principe de dynamisme graduel. Le principe de hiérarchie sous-tend un niveau « perceptif » qui permet d'extraire les caractéristiques des lettres et un niveau « mot » ou « logogène » pour les connaissances lexicales. Le principe de dynamisme graduel sous-tend que l'information issue du niveau perceptif est progressivement accumulée au niveau « mot ». Cependant, ces types de modèles théoriques ne rendent pas compte des effets d'inconsistance pour la lecture de certains mots, c'est à dire qu'une même graphie peut se lire de différente manière (Phénix et al., 2017).

Plus récemment, les modèles computationnels rendent compte de la lecture à haute voix et tentent de formaliser les mécanismes cognitifs impliqués dans le traitement des mots isolés. Ces modèles de simulation sur ordinateur décrivent comment les mots sont encodés en mémoire, comment ils sont activés lors du processus de lecture et quel est le coût temporel de cette activation (Phénix et al., 2017). Les modèles computationnels reproduisent donc le comportement humain et les effets observables en lecture. Ces modèles computationnels simulent préférentiellement soit le traitement orthographique et l'analyse visuelle, soit le traitement phonologique et l'information auditive phonologique. Les modèles mixtes simulent à la fois les traitements orthographiques, phonologiques et sémantiques. Aucun modèle théorique ne parvient à simuler l'ensemble des mécanismes mis en jeu lors de l'IME.

Ces types de modèles représentent les différentes zones cérébrales et leurs connexions (Davis, 2010), initialement décrites dans les modèles dits connexionnistes comme celui de Rumelhart (1994). La majorité des modèles computationnels analyse l'identification perceptive et la décision lexicale,

avec ou sans amorçage, avec une mesure du temps de réaction et du taux d'erreur. Ces manipulations par simulations ont permis de mettre en évidence un effet de supériorité du mot, auparavant décrit par Reicher (1969), comme le fait qu'une lettre est davantage reconnue au sein d'un mot porteur de sens, plutôt que dans une chaîne de caractères alphabétiques sans signification. Ces manipulations confirment un effet de fréquence, explicable par l'identification plus rapide des mots plus souvent rencontrés, combiné à un effet de régularité.

La première version computationnelle d'un modèle à double voie est le DRC (*Dual Route Cascaded model*) de Coltheart et al. (2001). Pour la voie lexicale, ce modèle fait l'hypothèse de transmissions bidirectionnelles excitatrices et sérielles, entre les lexiques orthographique et phonologique. Une séquence écrite déclenche l'activation des traits qui composent les lettres et des unités lettres, puis l'information est propagée aux composantes des voies lexicales et sous lexicales, ensuite propagée aux unités « lettres » et « mots » du lexique orthographique. Les unités orthographiques activent celles du lexique phonologique qui elles-mêmes activent les phonèmes correspondants dans le système phonémique. Pour la voie phonologique, chaque lettre ou ensemble de lettres, est associé au phonème qui lui correspond le plus dans la langue grâce à la conversion graphème phonème. In fine, ces unités reçoivent donc des activations des deux voies de lectures. Ainsi, ce modèle met en avant des effets de lexicalité (mots lus plus vite que les pseudo-mots) grâce à l'utilisation des deux voies de lecture, mais aussi des effets de fréquence et des effets de régularité. Cet effet de régularité rend compte que les mots réguliers sont lus plus vite que les mots irréguliers, seulement si ces derniers sont de basse fréquence. Cependant, ces travaux ne rendent pas compte de l'effet de consistance étudié par Jared (2002), c'est à dire que certains mots réguliers peuvent contenir des graphèmes qui se prononcent de façon différente dans d'autres mots.

Parmi les différents types de modèles computationnels, le modèle d'activation interactive de McClelland & Rumelhart (1981), fait partie des tout premiers modèles connexionnistes. Ce modèle utilise des représentations locales, c'est-à-dire que chaque trait visuel et acoustique, chaque lettre et phonème ainsi que chaque mot, est représenté par une unité distincte. C'est un modèle hiérarchique à activation parallèle et interactive. L'architecture est triangulaire, représentant les codes orthographiques, phonologique et sémantique. Les unités orthographiques ou dites infralexicales (lettres ou groupes de lettres) sont connectées aux unités dites phonologiques (phonèmes ou groupes de phonèmes). Ces deux types d'unités sont eux-mêmes connectés aux unités dites lexicales (les mots stockés dans le lexique mental). Donc, des connexions directes existent entre ces différentes unités orthographiques et phonologiques. Le traitement du mot se fait grâce à des informations de bas niveau mais aussi grâce à des informations de haut niveau. Les unités composant ce modèle sont alors

interconnectées et bidirectionnelles. Deux types de connections sont recensées : excitatrices et inhibitrices. Les connexions excitatrices se retrouvent entre les différents niveaux tandis que les connexions inhibitrices sont directement à l'intérieur des niveaux. De ce fait, au niveau lexical, il peut y avoir un concept de compétition entre les mots, mais le mécanisme d'inhibition lexicale permet d'éliminer les autres candidats en identifiant le mot cible. Ce modèle a permis de simuler fidèlement l'effet de supériorité du mot mais également des pseudo-mots et rend également compte de l'effet de fréquence et de l'effet de densité du voisinage. Cependant un mécanisme permettant de simuler la décision lexicale n'est pas encore proposé, et le codage positionnel est strict et ne rend pas compte des effets de transposition des lettres en situation d'amorçage.

Si la capacité de ces modèles est de simuler le comportement expert en lecture, et les processus cognitifs sous-jacents, ils ne reflètent que partiellement la réalité de la lecture et du décodage car seule la lecture de mots isolés est simulée (Phénix et al., 2017).

### **3. Compréhension : les processus mis en jeu en lecture**

La compréhension est la finalité de la lecture. Comprendre un texte nécessite d'avoir à la fois un niveau de compréhension orale adéquat, et la maîtrise des mécanismes spécifiques à la lecture expliqués par les différents modèles théoriques et computationnels. La compréhension en lecture nécessite deux compétences indépendantes: la compréhension du langage oral et le décodage des mots écrits. Cette tâche cognitive, n'est pas aisée pour tous les enfants, car si comprendre le langage oral revient à comprendre le langage écrit, ce processus est déjà difficile pour les bons lecteurs (Bianco, 2016). La lecture repose sur l'automatisation des procédures d'identification des mots écrits (Ziegler, 2019), ainsi, lorsque ces procédures sont suffisamment automatisées, le lecteur peut mobiliser ses ressources cognitives dans la compréhension de ce qui est lu (Castles et al., 2018).

Dans ses travaux, Bianco (2016) expose plusieurs mécanismes en jeu dans la compréhension : l'identification des mots oraux et écrits, la mobilisation de diverses connaissances (langagières, structurelles et thématiques) d'un texte, la théorie de l'esprit, la mobilisation de l'attention et des fonctions exécutives, et enfin la construction d'habiletés spécifiques au traitement de texte. Cependant, des obstacles peuvent interférer au sein de ces mécanismes et troubler la compréhension, comme une mauvaise fluidité de lecture, de mauvaises stratégies de compréhension ou encore de pauvres connaissances lexicales. En effet d'après l'étude de Catts et al. (2006), les faibles compreneurs disposent de connaissances langagières moins étendues et approfondies.

# CHAPITRE 2 L'EFFET RETEST

## 1. Notion d'effet retest et qualités psychométriques d'un test

### 1.1 Notion d'effet retest

Lors d'une rééducation orthophonique, la mesure des progrès d'un patient s'effectue au moyen d'épreuves standardisées, notamment pour rendre compte de l'efficacité de la thérapie appliquée. En clinique, l'«effet retest» est consécutif au fait de faire passer le même un test à un individu à X temps d'intervalle et d'avoir une amélioration des scores, sans qu'il y ait eu d'entraînement (Bartels et al., 2010). Ce concept fut évoqué pour la première fois sous le nom « d'effet d'apprentissage », par Wechsler (2016) lors de la passation de la « Wechsler Intelligence Scale for Children » (WISC). Cet effet d'apprentissage des items dû au retest, augmentait les moyennes des individus et donc leurs notes brutes, ce qui faussait les résultats (Grégoire, 2009).

### 1.2 Qualités psychométriques d'un test : sensibilité, spécificité, fidélité et validité

Les notions de sensibilité, spécificité, fidélité et validité sont liées entre elles et intrinsèques à la notion d'effet retest. En effet, les scores obtenus au premier et au second test (donc au test et au retest) se doivent d'avoir une fidélité et une validité parfaite, or c'est quasiment impossible car lors du retest on ne peut pas recréer des conditions de passation exactement identiques à la première passation, ce qui crée des biais notables quant à l'interprétation des scores de test et retest (Lecerf et al., 2016).

Pour constater cette existence d'une variation entre le score du test et du retest et donc par conséquent une variation dans le fonctionnement cognitif de l'individu testé ; l'examineur doit donc prendre en compte les qualités psychométriques du test comme les effets retests, la sensibilité ou la fidélité des scores (Lecerf et al., 2016).

La notion de fidélité, « réfère au degré de constance ou d'exactitude avec lequel un(e) instrument ou technique mesure le concept ou le phénomène qu'il/elle est supposé(e) mesurer » (Fortin, 1994). La mesure de la stabilité des scores du test et du retest, est un indice de la qualité métrique d'un instrument nettement meilleur que tous les autres indicateurs de fidélité (McCrae et al., 2011).

La sensibilité mesure la probabilité que les résultats d'un test soient positifs alors que l'individu est malade. « Plus la sensibilité est proche de l'unité, moins il y a d'erreurs de détection des sujets malades (faux négatifs) » (Bertrand et al., 2010).

La spécificité mesure la probabilité que les résultats d'un test soient négatifs alors que l'individu est sain. « Plus la spécificité est proche de l'unité, moins il y a de faux positifs » (Bertrand et al., 2010).

## **2. État des lieux global de l'effet retest dans la littérature**

L'effet retest est peu présent dans la littérature, et les sources à citer sont peu nombreuses. Celles qui existent ne concernent que très peu les conduites de soin en orthophonie, et sont à propos des effets de tests répétés, et non pas du calcul fait à partir des scores de retests qui permet la validation d'épreuves. Les études existantes s'appliquent principalement au domaine de la psychologie clinique telles que McCrae et al. (2011), Scharfen et al. (2018a), Scharfen et al. (2018b) ; ou à celui de la neuropsychologie comme Hausknecht et al. (2007), Lecerf et al. (2016) ; ou encore au domaine de la médecine, comme l'étude de Midy (1996).

Parmi les quelques études menées dans le cadre du langage, pour tester l'effet retest de tests répétés, l'une des plus anciennes est celle de Flanagan (1953). Cette étude portait sur la vérification d'un effet retest lors de la lecture à voix haute pour des patients normo-lecteurs. Pour un texte d'une durée de 5 minutes lors de sa lecture à voix haute, la fiabilité de l'effet retest était de 80%, tandis que pour un texte plus long (8 minutes ou plus) la fiabilité de l'effet retest était de 90%. Grâce à ces données les chercheurs ont pu corrélérer que plus le texte est long, moins l'effet retest est fort. Une autre étude, celle-ci sur des enfants dans le domaine du langage écrit (Cirino et al., 2002), tend à démontrer un effet retest présent lors de tests de lecture répétés.

## **3. Tests orthophoniques : pertinence de l'effet retest dans la clinique**

Lors de la consultation initiale, l'orthophoniste fait passer un premier bilan au patient pour objectiver sa plainte en langage écrit et évaluer ses troubles, dans le but de poser un diagnostic. Cette évaluation se fera au moyen d'un test quantitatif normé qui situera le patient par rapport à une norme. L'intérêt d'utiliser des tests normés est qu'ils sont standardisés. La standardisation d'un test orthophonique se fait grâce à un recueil de données psychométriques et par la modélisation d'un protocole invariable de passations (Ianova & Hallowell, 2013).

Selon Seron (1979), lors d'une rééducation, des variables peuvent entraver les résultats de façon positive ou négative, d'où l'importance d'analyser tous les facteurs pouvant interférer dont un éventuel effet retest qui pourrait faire augmenter les scores. L'effet retest peut donc faire partie d'une remise en

question des résultats obtenus et être pris en compte dans le projet thérapeutique, qui peut lui aussi être remis en question, et tend à toujours évoluer selon la réaction du patient à la thérapie, et les progrès qu'il fait (Collette & Schelstraete, 2015).

# CHAPITRE 3 TESTS

Pour répondre aux besoins de dépistage et de diagnostic des troubles de la lecture, de nombreux outils d'évaluation du langage écrit ont été développés. Le plus connu d'entre eux en langue française est sans doute le test standardisé de lecture de l'«Alouette» (Lefavrais, 1967). D'autres tests ont vu le jour se focalisant sur différents aspects de l'acquisition du langage écrit comme le test « L3 » (Lobrot, 1987). Les tests orthophoniques en français pour la grande majorité, sont conçus pour le diagnostic et ne sont pas testés au niveau des qualités psychométriques évoquées précédemment (fidélité test/retest).

## 1. L'«Alouette»

Le test de leximétrie l'Alouette est destiné à évaluer la vitesse et la précision en lecture. Initialement créé par Lefavrais (1967) dans le but de différencier les mauvais lecteurs des dyslexiques. Ce test est le plus employé dans la recherche pour valider les inclusions des lecteurs dyslexiques versus les normo-lecteurs. Il est également fréquemment utilisé en pratique clinique pour évaluer les troubles de la lecture chez des individus de 6 à 16 ans (Olivier, 2007).

### 1.1 Description du test

Le texte est composé de 265 mots incluant des mots rares et des pièges orthographiques, des graphèmes ambigus, et des items proches phonologiquement (Pourcin & Colé, 2018). Le manuel de la seconde version (Lefavrais, 2005) reprend surtout le postulat du type d'erreurs et distingue les paralexies selon la classification des types de dyslexie (phonologique, surface et mixte) en référence au modèle à double voie (Coltheart et al., 2001). Cette classification repose sur le type d'erreurs selon quatre catégories : barbarisme (lecture d'un mot trop éloigné de la cible), conversion graphophonologique (erreurs dans la maîtrise de la combinatoire), paralexies verbales (remplacement du mot du texte par un mot visuellement proche), paralexies sémantiques (remplacement du mot du texte par un mot sémantiquement proche). Ce test permet d'évaluer les stratégies de lecture dans une situation de lecture de texte à voix haute, et de juger si ces troubles sont plutôt de l'ordre phonologique ou lexical (Bertrand et al., 2010).

Le texte comprend des pièges pour un lecteur qui utiliserait des anticipations contextuelles (« poison » au lieu de « poisson » après « lac ») ou des mauvaises conduites d'identification (items proches



visuellement ou phonologiquement, comme « amie/Annie » ou « gai/geai »). Les mots du début du texte sont considérés comme facilement compréhensibles (le toit, la mousse...) par rapport à un contenu sémantique riche et peu usuel, ésotérique pour certains (les bardeaux, la lume...). Le texte est entouré de dessins favorisant les erreurs contextuelles (dessin d'un écureuil près du mot « écueil »).

## **1.2 Mode de passation**

L'enfant est invité à lire à voix haute le texte le plus vite possible et le plus correctement possible. La lecture est limitée à 3 minutes. Les mots qui étaient rares pour les enfants il y a 40 ans le sont encore actuellement.

## **1.3 Cotation**

La cotation la plus récente, « Alouette-R », (Lefavrais, 2005) prend en compte différents indicateurs : le temps de lecture (TL), le nombre de mots lus (M), d'erreurs (E), de mots corrects (C), l'indice de précision (CM) et l'indice de vitesse (CTL). L'« Alouette » permet également une analyse qualitative des erreurs. L'efficacité en lecture peut ensuite être calculée en écart type ou en centiles et analysée selon l'âge ou le niveau scolaire de l'enfant. Les résultats en temps et précision permettent d'analyser les troubles de la lecture, selon les types d'erreurs commises. Les indices de vitesse et de précision permettent d'obtenir un pourcentage de réussite du nombre de mots correctement lus, par rapport au nombre de mots lus au total et également par rapport à la contrainte temporelle de 3 minutes.

Ce test a également été validé pour dépister les adultes dyslexiques avec un niveau universitaire (Cavalli et al., 2018).

## **2. L'épreuve « L3 » issu de la batterie « ORLEC »**

Test pédagogique issu de la batterie « Orlec », l'épreuve « L3 » (Lobrot, 1987) est l'un des 4 sous test de cette batterie qui évalue les compétences de lecture.

### **2.1 Description du test**

Validé par Pierart & Gregoire en 2004, le subtest « L3 » (ou « L3-R ») est une tâche de lecture de phrases lacunaires avec décisions lexicales mettant en jeu des processus de recherche de sens et de mécanismes inhérents à la compréhension. Chez le lecteur dyslexique, les compétences en

compréhension écrite sont moins altérées que les compétences en identification de mots. Le « L3 » (ou « L3-R ») contribue donc au diagnostic différentiel de dyslexie (Casalis et al., 2019).

Ce subtest est composé de 36 phrases lacunaires, à compléter parmi un choix lexical de 5 mots. Sur les 36 items, en fonction des mots proposés pour répondre, 25 explorent des confusions de sons (par exemple pour une phrase dont le contexte de lecture est une promenade dans la forêt, le mot cible attendu est « châtaignes » mais le mot « champions » est parmi les distracteurs), et 13 proposent des confusions entre des éléments lexicaux appartenant au même champ sémantique (par exemple pour une phrase dont le contexte de lecture est le printemps, le mot cible attendu est « jonquille » mais le mot « feuilles » est parmi les distracteurs).

## **2.2 Mode de passation**

L'enfant est invité à lire en lecture silencieuse les 36 phrases et doit souligner avec un stylo le mot qui complète chaque phrase parmi 5 mots proposés.

## **2.3 Cotations**

Un score brut est établi, par rapport au nombre de phrases lues et au nombre d'erreurs, qui sera à situer dans une norme en fonction du sexe et de l'âge de l'enfant.

## **3. Intérêt d'un test signifiant vs un test non signifiant**

Comme énoncé précédemment, l'«Alouette» est un test de leximétrie non signifiant et le test « L3 » est un test mesurant la compréhension écrite de phrases signifiantes. En accord avec la littérature, l'intérêt de comparer les performances de lecture à ces deux tests devrait permettre de valider si un test signifiant et mesurant la compréhension serait plus sensible à un effet retest qu'un test de lecture sans signification.

En effet, dans l'étude de Stanovich et al. (1981), des enfants de primaire, normo-lecteurs, sont testés en lecture de mots au sein de trois types de phrases. Des phrases porteuses de sens, par exemple "Le pilote pilotait l'avion.", qui aident à la lecture par rapport au contexte (comme le test « L3 »). Des phrases dénuées de sens, comme « le pilote a caché l'os », où l'on ne peut pas s'appuyer sur le contexte pour lire le mot (comme dans l'« Alouette»). Des phrases neutres au niveau du sens comme « l'avion » où le contexte n'influence pas le lecteur. Cette étude démontre que le contexte joue un rôle

prépondérant pour l'identification du mot, car les mots sont mieux lus au sein des phrases avec contexte. Avec de l'entraînement, il est aussi démontré que le lecteur a moins besoin de s'appuyer sur le contexte pour lire le mot cible et obtient de bons scores quel que soit le type phrase. De ce fait, pour l'« Alouette» le lecteur ne peut pas s'appuyer sur le contexte comme aide pour lire, tandis que pour le test « L3 » le contexte soutient la décision lexicale qui est demandée à la fin de la lecture de chaque phrase lacunaire.

La lecture nécessite de nombreuses compétences sous-jacentes et fait appel à divers mécanismes. Pour les évaluer, différents tests existent. Cependant, la passation trop rapprochée d'un même test peut fausser les résultats obtenus. Ainsi, les effets retest possibles sont à prendre en compte lors de l'analyse des scores, comme le démontrent les différentes sources de la littérature. Pour cela, ont été choisis dans cette étude, le test de l'« Alouette » et le « L3 » issu de la batterie « Orlec » présentés précédemment. La démarche d'analyse des données et de la méthodologie employée ainsi que les résultats obtenus feront l'objet de la prochaine partie.

# PARTIE PRATIQUE

## 1. Problématique et objectifs

Lors d'évaluations pré et post entraînement ou ligne de base, une partie des gains de scores des patients pourraient être attribués à un effet retest. Cette étude exploratoire consiste à vérifier si un éventuel effet retest au test de l'«Alouette» et au test «L3» issu de la batterie «Orlec» existe, et lequel a le plus grand effet retest.

L'objectif primaire est de vérifier si un effet retest, au test de l'«Alouette» et au test «L3» existe chez des enfants normo-lecteurs scolarisés en CM1 et CM2. Pour l'«Alouette» les critères de mesures sont : les moyennes des scores bruts mesurées en T1 comparées à celles mesurées T2 en temps et en précision (soit le nombre d'erreurs). Pour le test «L3» les critères de mesures sont les moyennes des scores bruts mesurées en T1 comparées à celles mesurées T2 en nombre de phrases lues et en nombre de phrases correctement lues.

Le second objectif est de vérifier si l'effet retest est plus prononcé dans une épreuve non signifiante (l'«Alouette») en comparaison avec une épreuve qui teste la compréhension (le test «L3»). Les critères de mesure pour ces objectifs sont la comparaison de la progression entre T1 et T2 des scores de l'«Alouette» en précision (soit le nombre d'erreurs) aux scores de «L3» en précision (soit le nombre d'erreurs).

## 2. Matériel et méthode

### 2.1 Sujets

Dans le cadre de l'étude, 90 enfants ont été testés, 45 enfants de CM1 et 45 enfants de CM2 ont été inclus. En T1, l'âge moyen des CM1 était de 9 ans 6 mois, l'âge moyen des CM2 était de 10 ans 6 mois. T2 a eu lieu à deux mois de distance. Ont été inclus 72 enfants, soit 30 CM1 (19 filles et 11 garçons) et 42 CM2 (20 filles et 22 garçons).

Finalement, 18 enfants ont été exclus, notamment certains qui étaient inclus lors du T1 et qui n'ont pas pu être inclus en T2 du fait d'une classe fermée dans le contexte de la pandémie de COVID 19.

Les critères d'inclusions étaient les suivants : les enfants devaient être âgés entre 8 ans 9 mois et 11 ans 9 mois et scolarisés en classe de CM1 ou CM2. Le consentement éclairé devait être signé par l'un des deux parents ou le représentant de l'autorité parentale.

Les critères de non-inclusion étaient les suivants : retrait du consentement de participation à l'étude, absence prolongée tombant pendant les périodes de passation, présence d'un trouble spécifique des apprentissages selon le DSM V (Manuel diagnostique et statistique des troubles mentaux, 2013), déterminé par au moins un score pathologique à une des deux épreuves qui mesuraient la leximétrie (« Alouette » ou « Pingouin ») et/ou à l'épreuve de compréhension écrite (le test « L3 »), et au moins un score pathologique à une des épreuves évaluant les compétences sous-jacentes à la lecture.

## **2.2 Matériel**

Lors du premier temps de passation (T1), les enfants ont passé les prétests d'inclusion de lecture et de compétences reliées à la lecture afin d'écartier un trouble spécifique du langage écrit.

La passation des épreuves de lecture débutent par des tâches de lecture de mots réguliers, irréguliers et pseudo-mots courts et longs, issues de la batterie informatisée « EVALEC-primaire » (Sprenger-Charolles et al., 2010) (items à lire le plus vite et précisément possible présentés via le logiciel). L'épreuve d'identification (via une liste sur papier) de mots courts et longs, fréquents et non fréquents « EVAL2M », issue de la batterie informatisée « EVALEO 6-15 » (Launay et al., 2018) était ensuite proposée, puis l'épreuve de l'« Alouette-R » (Lefavrais, 2005) (lecture à voix haute d'un texte non signifiant présenté sur papier). L'épreuve du « Pingouin » issue de la batterie informatisée « EVALEO 6-15 » (lecture à voix haute d'un texte signifiant présenté sur feuille, le plus vite et le plus correctement possible) est par la suite proposée. Enfin, la passation des tests de lecture se clôturait avec l'épreuve « L3 » (Lobrot, 1987) (lecture silencieuse de phrases lacunaires présentées sur papier, avec décision lexicale à entourer au stylo).

Les épreuves de compétences reliées à la lecture sont issues de la batterie informatisée « EVALEC-primaire » (Sprenger-Charolles et al., 2010), et prennent en compte la précision et la rapidité de réalisation des tâches de segmentation syllabique et phonémique (consonne/voyelle/consonne et consonne/consonne/voyelle), et la répétition de pseudo-mots de 3 à 6 syllabes (évaluation de la mémoire à court terme phonologique). Les autres prétests sont issus de la batterie informatisée d'évaluation du langage oral et du langage écrit chez les sujets de 6 à 15 ans, « EVALEO 6-15 » (Launay et al., 2018) : la répétition de chiffres endroits puis envers (écoute de séquences de chiffres à

répéter dans l'ordre entendu puis dans l'ordre inverse), et l'évaluation de l'empan visuo attentionnel (séquence de 5 chiffres présentés visuellement à l'écran à restituer oralement).

Lors du 2<sup>e</sup> temps (T2), seules les épreuves de lectures ont été à nouveau proposées.

## 2.3 Méthode

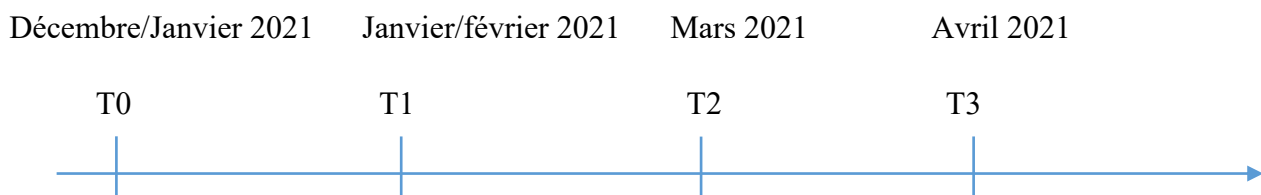
### 2.3.1 Déroulement général de l'étude

Afin de constituer une large base de données, les passations des enfants ont été réparties équitablement entre les 4 différentes examinatrices (entre 20 et 25 enfants chacune). L'examinatrice restait la même pour chaque enfant en T1 et T2, et les passations se sont déroulées dans des salles individuelles.

Les deux temps de passation se sont déroulés dans deux écoles primaires sous-contrats avec l'accord de leurs directrices respectives, et le consentement parental des enfants participant à l'étude via un formulaire de consentement éclairé (voir annexe I). L'étude était décrite oralement aux participants et par écrit à leurs responsables légaux. Les épreuves décrites étaient utilisées lors de l'inclusion des enfants (1 heure de passation à T1), puis lors de la phase retest à deux mois (30 minutes de passation à T2).

### 2.3.2 Temporalité de l'étude

Ce projet s'est déroulé en plusieurs temps :



**Figure 1** : ligne temporelle des différents temps de l'étude

Le T0 correspond au temps de la préparation de l'étude. La familiarisation avec les tests par les examinatrices a été nécessaire, étant donné que l'éventail de tests était assez large et que ces derniers pouvaient provenir de plusieurs batteries distinctes, et nécessiter plusieurs supports différents (ordinateur ou papier). Des cahiers de passation ont été créés regroupant l'ensemble des épreuves, ainsi qu'un tableau Excel pour reporter les données. Ce tableau comprenait les scores bruts et les déviations standards de chaque épreuve. Enfin, l'anonymisation de chaque élève a été faite, pour créer les profils de chacun sur les batteries informatisées. Les données personnelles (sexe, âge, classe) des participants ont été intégrées de façon anonyme dans le tableau de report des données.

En T1 les premières passations ont eu lieu : les prétests de segmentation syllabique/phonémique et de répétitions de pseudos mots issus de la batterie informatisée « EVALEC-primaire » (Sprenger-Charolles et al., 2010) ont d'abord été passés, puis ce furent les épreuves de répétition de chiffres endroits/envers et l'empan visuo attentionnel, issues de la batterie informatisée « EVALEO 6-15 » (Launay et al., 2018). Ensuite les épreuves de lecture ont été passées dans l'ordre suivant : lecture de mots issus de la batterie informatisée « EVALEC-primaire », lecture de la liste de mots « EVAL2M » et du texte « Le Pingouin » issus de la batterie informatisée « EVALEO 6-15 », puis lecture du texte de l'« Alouette » et enfin test du « L3 ».

Puis, en T2 à 2 mois d'intervalle +/- 4 jours, les secondes passations se sont déroulées, avec seulement les tests de lecture en ordre inverse du T1 : le test « L3 », l'«Alouette», « Le pingouin », « EVAL2M », et la lecture de mots issus de la batterie informatisée « EVALEC ». Les enfants vus en derniers lors du T1 ont été vus en premier lors du T2 pour éviter un effet de protocole.

Pour finir en T3 les données ont été reportées dans le tableau Excel destiné à cet effet, puis analysées.

### **3. Stratégie d'analyse des données**

Pour les deux épreuves le test de normalité « Kolmogorov-Smirnov » a été utilisé pour analyser les données et vérifier si les scores suivaient une loi normale.

Pour répondre à l'objectif primaire qui était de vérifier s'il existe un effet retest au test de l'«Alouette» et au test « L3 » chez des enfants normo-lecteurs de CM1 et CM2, les scores obtenus à l'«Alouette» mesurés en T1 ont été comparés avec les scores mesurés T2, et les scores obtenus au test « L3 » mesurés en T1 ont été comparés avec les scores mesurés en T2. Ici, les scores à l'«Alouette» se mesuraient en précision (nombre d'erreurs totales faites au cours de la lecture) et en temps de lecture en secondes. Les scores au test « L3 » se mesuraient en nombre de phrases lues (au total), ainsi qu'en nombre de phrases correctement lues.

Pour l'«Alouette» : le T test student pour échantillons appariés a été utilisé pour comparer les scores de temps de lecture et de précision en T1 et T2. Enfin, pour comparer la taille de l'effet en T1 par rapport à T2, des moyennes de temps et des moyennes de précision, le d de Cohen a été utilisé.

Pour le test « L3 » : le test de rangs signés de Wilcoxon a été utilisé pour comparer les scores du nombre de phrases lues en T1 et en T2. Le T test student pour échantillons appariés a été utilisé pour comparer les scores du nombre de phrases correctement lues en T1 et en T2. Enfin, pour comparer la

taille de l'effet en T1 par rapport à T2 des moyennes du nombre de phrases lues, et des moyennes du nombre de phrases correctement lues, le d de Cohen a été utilisé.

Pour répondre aux objectifs secondaires qui étaient de comparer l'effet retest de l'«Alouette» par rapport à l'effet retest du test « L3 », et donc comparer l'effet retest d'un test non significatif à un test mesurant la compréhension, le taux de progression a été calculé en pourcentage en fonction de la différence entre les moyennes des scores en précision (nombre d'erreurs) de T1 par rapport à T2 pour chacune des deux épreuves.

#### 4. Résultats

N	T1	T2			
71	M (SD)	M (SD)	p value	d de Cohen	Progression entre T1 et T2 en %
<b>Alouette</b>					
Précision	12.46 (6.14)	10.04 (6.31)	p < 0.001	d= 0.39	19.42%
Temps	139.33 (30.79)	132.14 (30.18)	p < 0.001	d= 0.23	5.16%
<b>L3 (ORLEC)</b>					
Nombre de phrases lues	30.61 (5.30)	32.01 (4.42)	p < 0.001	d= 0.28	4.74%
Nombre de phrases correctement lues	28.29 (4.95)	29.95 (4.26)	p > 0.05	d= 0.30	5.98%
Précision	2.32 (1.70)	2.05 (1.49)	p=0.377	d=0.14	10.39%

**Tableau 1** : Résultats obtenus au test de l'«Alouette» et au test « L3 » à T1 et T2

#### 4.1 Analyse des effets retest à l'«Alouette» et au « L3 » issu de la batterie « Orlec »

Pour la première analyse, les moyennes et les écarts types des scores obtenus en temps et en précision ont été calculés pour le test de l'«Alouette», puis les moyennes et les écarts types des scores des phrases lues et des phrases lues correctement ont été calculés pour le test « L3 ».

##### 4.1.1 L'«Alouette»

Les scores de temps, faisant l'objet d'une distribution normale, ont été analysés avec le T test student. Pour le temps de lecture, la moyenne obtenue lors du T1 ( $m_{t1} = 139.33$ ;  $EC = 30.79$ ) est significativement supérieure à celle obtenue lors du T2 ( $m_{t2} = 132.14$ ;  $EC = 30.18$ ) ( $t_{70} = 3.77$ ;  $p < 0.001$ ).



La moyenne du temps de lecture en secondes a diminué de 5.16% de T1 à T2. Selon le calcul du d de Cohen, la taille de l'effet est faible ( $< 0.50$ ) entre T1 et T2 avec  $d=0.23$ .

Les scores de précision, faisant l'objet d'une distribution normale, ont été analysés avec le T test student. Pour les scores en précision, la moyenne obtenue lors du T1 ( $m_{t1}= 12.46$  ;  $EC = 6.14$ ) est significativement supérieure à celle obtenue lors du T2 ( $m_{t2}= 10.07$  ;  $EC= 6.32$ ) ( $t_{70}=3.87$  ;  $p<0.001$ ). La moyenne des scores de précision en nombre d'erreurs a diminué de 19.42% de T1 à T2. Selon le calcul du d de Cohen, la taille de l'effet est faible ( $< 0.50$ ) entre T1 et T2 avec  $d=0.39$ .

#### *4.1.2 Le test « L3 »*

Les scores du nombre de phrases lues, faisant l'objet d'une distribution non normale, ont été analysés avec le test des rangs signés de Wilcoxon.

Pour le score du nombre de phrases lues la moyenne du T1 ( $m_{t1}= 30.59$  ;  $EC = 5.34$ ) a une différence statistiquement significative par rapport à celle du T2 ( $m_{t2}= 32.04$  ;  $EC = 4.45$ ) ( $Q_{ob70}=206$  et  $p<0,001$ ). La moyenne du nombre de phrases lues a augmenté de 4.74% entre T1 et T2. Selon le calcul du d de Cohen, la taille de l'effet est faible ( $< 0.50$ ) entre T1 et T2 avec  $d =0.29$ .

Les scores du nombre de phrases correctement lues, faisant l'objet d'une distribution normale, ont été analysés avec le T test student.

Pour le nombre de phrases correctement lues, la moyenne du T1 ( $m_{t1}= 28.28$  ;  $EC = 4.99$ ) n'a pas une différence statistiquement significative par rapport à la moyenne en T2 ( $m_{t2}= 29.97$  ;  $EC =4.29$ ) ( $t_{70} = -4.95$  et  $p>4.91$ ). La moyenne du nombre de phrases correctement lues a augmenté de 5.98 % entre T1 et T2. Selon le calcul du d de Cohen, la taille de l'effet est faible ( $< 0.50$ ) entre T1 et T2 avec  $d=0.3$ .

## **4.2 Comparer l'effet retest du test de l'«Alouette» à celui du test « L3 »**

Les scores de précision de l'«Alouette», faisant l'objet d'une distribution normale, ont été analysés avec le T test student.

Pour les scores en précision à l'«Alouette», la moyenne obtenue lors du T1 ( $m_{t1}= 12.46$  ;  $EC = 6.14$ ) est significativement supérieure à celle obtenue lors du T2 ( $m_{t2}= 10.07$  ;  $EC= 6,32$ ) ( $t_{70}=3.87$  ;  $p<0.001$ ). La moyenne des scores de précision en nombre d'erreurs a diminué de 19.42% de T1 à T2. Selon le calcul du d de Cohen, la taille de l'effet est faible ( $< 0.50$ ) entre T1 et T2 avec  $d=0.39$ .

Les scores de précision au test « L3 » faisant l'objet d'une distribution non normale ont été analysés avec le test des rangs signés de Wilcoxon.

Pour les scores en précision au test « L3 », la moyenne du T1 ( $m_{t1} = 2.31$  ;  $EC = 1.71$ ) a une différence statistiquement significative par rapport à la moyenne en T2 ( $m_{t2} = 2.07$  ;  $EC = 1.5$ ) ( $Q_{qob70} = 606.5$  et  $p = 0.46$ ). Le taux d'erreur a diminué de 10.39 % de T1 à T2. Selon le calcul du  $d$  de Cohen, la taille de l'effet est faible ( $< 0.50$ ) entre T1 et T2 avec  $d = 0.14$ .

## DISCUSSION

Le premier objectif de cette étude était de rechercher un effet retest aux tests de l'«Alouette» et au test « L3 » issu de la batterie « Orlec », chez des enfants normo-lecteurs de CM2 et CM2. Les résultats ont montré un effet retest existant pour l'«Alouette» de T1 à T2 : le temps de lecture a diminué (-5.16%), et la précision a augmenté (+19.42%), donc une amélioration des performances est notable de T1 par rapport à T2. Pour le test « L3 », de T1 à T2, la moyenne du nombre de phrases lues a augmenté (+4.74%) et la moyenne du nombre de phrases correctement lues a augmenté également (+5.98 %), là encore on remarque une amélioration des performances. Ensuite l'effet retest au test de l'« Alouette », test de leximétrie non significatif, a été comparé à celui du test « L3 », test mesurant la compréhension. Les résultats ont montré que l'effet retest est plus important à l'«Alouette» qu'au test « L3 ». Ces gains sont caractérisés par une meilleure précision entre T1 et T2 au test de l'«Alouette» qu'au test « L3 » avec une augmentation de 20.39% pour le premier test contre une augmentation de 9.09% pour le second.

Des facteurs purement qualitatifs peuvent en partie expliquer les effets retest. Globalement, les enfants ont énoncé avoir trouvé les épreuves plus faciles en T2 du fait de les avoir déjà réalisées en T1 et se souvenaient des tests. Lors de la deuxième passation en T2, les enfants avaient déjà vu leur examinatrice, connaissaient déjà les consignes et sous quelle forme allaient se dérouler les tests, par conséquent les conditions étaient plus rassurantes et moins stressantes, cela a pu se ressentir au niveau des performances. Les enfants des deux écoles où les tests ont été performés avaient pour l'ensemble, un très bon niveau de lecture.

Dans la littérature, peu de données permettent d'expliquer les effets retest des deux épreuves utilisées. Pourquoi une épreuve montrerait plus d'effets retest que l'autre ? Comme décrite dans le modèle de Gough & Tunmer (1986), la lecture fait appel à la compréhension or pour chacun de ces tests, l'accès à la compréhension est différent. En effet, le test de l'« Alouette » est non significatif et limite l'accès à la compréhension. Son auteur (Lefavrais, 1967) l'a conçu pour dépister les dyslexies et tester la fluence de lecture. Considéré comme un « Gold Standard », ce test n'est toutefois pas fondé sur un modèle théorique descriptif ou computationnel validé. Néanmoins, le modèle de Coltheart et al. (2001) rend tout de même compte des processus mis en jeu lors de la lecture d'un texte tel que l'« Alouette », avec l'effet de régularité. En outre, les deux études de Stanovich et al. (1981,1988) à propos du modèle de compensation interactif du développement de la fluence, peuvent également expliquer le fonctionnement du test de l'« Alouette ». L'effet contextuel en lecture augmente quand

un mot est moins fréquent. Le parallèle peut ici être fait avec ce test composé de mots rares, de pièges orthographiques, et de graphèmes ambigus (Pourcin & Colé, 2018).

Le test « L3 » (Lobrot, 1987) est quant à lui, un test signifiant évaluant la compréhension de phrases. Le parallèle peut être fait avec le modèle connexionniste de Rumelhart (1994), qui décrit un codage phonologique, orthographique et sémantique simultané. Le lecteur a besoin de ce triple processus pour la lecture des phrases lacunaires et pour la décision lexicale parmi les cinq propositions par phrase. Cette épreuve de closure de phrase fait aussi appel aux fonctions exécutives telles que l'inhibition, la planification, et la flexibilité pour faire un choix de réponse inférentielle, ce qui représente un coût cognitif supplémentaire. Le test « L3 » en faisant appel à la compréhension devrait obtenir un effet de sémantisation notable selon les travaux de McArthur (2007), et donc avoir un effet retest plus important que celui du test de l'« Alouette ». Or, les résultats obtenus dans cette étude démontrent le contraire.

Les travaux de Scharfen et al., (2018a) pourraient expliquer cet effet retest plus marqué au test de l'« Alouette » versus le test « L3 », en démontrant que l'automatisation des stratégies de résolution spécifiques à chaque test est plus longue pour une tâche complexe que pour une tâche simple. Ainsi, la complexité de la tâche appellerait davantage à la mémorisation du retest.

Certaines limites ayant pu fausser les résultats sont à relever. Des nuisances sonores ont pu gêner les enfants lors des tests: passage des personnels de l'école dans les salles de passations, les salles de classe adjacentes qui pouvaient générer du bruit, et enfin des travaux dans les deux écoles (lors du T2). En outre, le temps de passation était assez long en T1 (1heure environ) ce qui a pu majorer l'effet de fatigue des enfants. Avec la situation sanitaire actuelle (pandémie de COVID 19) nous avons dû porter des masques (examineurs et participants à l'étude) ce qui a pu gêner les enfants pour articuler. La limite de l'atteinte du seuil plafond se pose également, car au test de l'«Alouette» et au test « L3 », certains enfants avaient atteint les scores maximum en précision dès la première passation, aucune progression ne pouvait être relevée en T2. Une limite importante à cette étude était la situation sanitaire avec la pandémie de COVID 19. A cause des différents confinements ayant provoqué la fermeture des écoles, le laps de temps pour faire les passations a été réduit, et cela a beaucoup rapproché la fin des deuxièmes passations de la date de rendu de ce mémoire, et le temps a manqué pour faire des analyses supplémentaires : vérifier par item pour les deux tests, lesquels avaient été échoués/réussis et tenter de comprendre pourquoi, ou vérifier par enfant si les items réussis/échoués en T1 l'étaient aussi en T2. Néanmoins, cette étude exploratoire rend compte de l'effet retest de deux épreuves orthophoniques très utilisées, et ici testées de façon rapprochée ce qui reproduit les situations

que l'on pourrait rencontrer en exercice en faisant des évaluations pré et post entraînement. Cela permettra dorénavant, de mesurer au plus juste les gains des performances des patients pour ces tests-là. De plus, l'échantillon de participants retenu (72) permet d'obtenir des données représentatives pour cette étude.

## CONCLUSIONS ET PERSPECTIVES

Dans la littérature, peu de données concernent l'effet retest des tests utilisés par les orthophonistes pour évaluer le langage écrit. Dans cette présente étude, nous avons mesuré à deux mois d'intervalle, chez des enfants normo-lecteurs scolarisés en CM1 et CM2, les éventuels effets retest de deux tests orthophoniques très utilisés pour dépister les dyslexies : le test de l'«Alouette» et le test « L3 » issu de la batterie « Orlec ».

Le premier est un test de leximétrie non signifiant et le second un test signifiant de lecture mesurant la compréhension. Nous avons tout d'abord comparé les performances à ces deux tests entre les deux passations séparées par deux mois. Les résultats vont dans le sens de l'hypothèse d'un effet retest à ces deux tests de lecture, caractérisé par des gains aux différentes mesures de ces tests. Ensuite, l'effet retest du test de l'«Alouette» a été comparé à celui du test « L3 ». Les résultats démontrent un effet retest beaucoup plus prononcé pour le test non signifiant l'« Alouette », que pour le test signifiant « L3 ».

Cette étude doit se poursuivre avec une nouvelle passation à deux mois de la deuxième passation qui permettra de confirmer ou infirmer les données actuelles. Il est également prévu une nouvelle passation à 6 mois. Plus généralement, il semble pertinent de mesurer les effets retest aux outils utilisés dans l'évaluation du langage écrit et du langage oral.

## RÉFÉRENCES

- Bartels, C., Wegrzyn, M., Wiedl, A., Ackermann, V., & Ehrenreich, H. (2010). Practice effects in healthy adults: A longitudinal study on frequent repetitive cognitive testing. *BMC Neuroscience*, *11*(1). <https://doi.org/10.1186/1471-2202-11-118>
- Bertrand, D., Fluss, J., Billard, C., & Ziegler, J. C. (2010). Efficacité, sensibilité, spécificité : Comparaison de différents tests de lecture. *L'Année psychologique*, *110*(02), 299. <https://doi.org/10.4074/S000350331000206X>
- Bianco, M. (2016). *Lire et comprendre : Développement et enseignement*. 9.
- Casalis, Leloup, & Bois Parriaud. (2019). *Prise en charge des troubles du langage écrit chez l'enfant* (Elsevier Masson). <https://www.elsevier-masson.fr/prise-en-charge-des-troubles-du-langage-ecrit-chez-lenfant-9782294754203.html>
- Castles, A., Rastle, K., & Nation, K. (2018). Ending the Reading Wars : Reading Acquisition From Novice to Expert. *Psychological Science in the Public Interest*, *19*(1), 5-51. <https://doi.org/10.1177/1529100618772271>
- Catts, H. W., Adlof, S. M., & Ellis Weismer, S. (2006). Language deficits in poor comprehenders : A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research: JSLHR*, *49*(2), 278-293. [https://doi.org/10.1044/1092-4388\(2006/023\)](https://doi.org/10.1044/1092-4388(2006/023))
- Cavalli, E., Leloup, G., El AHMADI, A., Poracchia, F., Sprenger-Charolles, L., & Colé, P. (2018). *Le test de l'Alouette validé pour les jeunes adultes dyslexiques*. *139*, 12-15.
- Cirino, P. T., Rashid, F. L., Sevcik, R. A., Lovett, M. W., Frijters, J. C., Wolf, M., & Morris, R. D. (2002). Psychometric Stability of Nationally Normed and Experimental Decoding and Related Measures in Children with Reading Disability. *Journal of Learning Disabilities*, *35*(6), 526-539. <https://doi.org/10.1177/00222194020350060401>
- Collette, E., & Schelstraete, M.-A. (2015). Rééducation de l'orthographe dans le cas d'une dyslexie développementale : Etude de cas clinique chez un étudiant. *Rééducation Orthophonique*, *91*.
- Coltheart. (1978). *Lexical access in simple reading tasks*. (Strategies of information processing). London ; New York: Academic Press.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC : A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204-256.

<https://doi.org/10.1037/0033-295x.108.1.204>

Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review*, 117(3), 713-758. <https://doi.org/10.1037/a0019738>

*DSM V, Manuel diagnostique et statistique des troubles mentaux*. (2013).

Ferrand et Ayora. (2017). *Psychologie cognitive de la lecture : Reconnaissance des mots écrits chez l'adulte*. <https://www.deboecksuperieur.com/ouvrage/9782804193348-psychologie-cognitive-de-la-lecture>

Flanagan. (1953). *Durational characteristics of oral reading effects of sample on test retest reliability*.

Fortin, F. (1994). *PROPRIÉTÉS MÉTROLOGIQUES DES INSTRUMENTS DE MESURE (FIDÉLITÉ - VALIDITÉ)*. 5.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading and reading disability. *Remedial and Special Education*, 6-10.

Grégoire, J. (2009). *L'examen clinique de l'intelligence de l'enfant*. Mardaga. <https://www.cairn.info/l-examen-clinique-de-l-intelligence-de-l-enfant--9782804700164.htm>

Hausknecht, J. P., Halpert, J. A., Di Paolo, N. T., & Moriarty Gerrard, M. O. (2007). Retesting in selection : A meta-analysis of coaching and practice effects for tests of cognitive ability. *Journal of Applied Psychology*, 92(2), 373-385. <https://doi.org/10.1037/0021-9010.92.2.373>

Ianova, Hallowell. (2013). *A tutorial on aphasia test development in any language : Key substantive and psychometric considerations*. [https://www.researchgate.net/publication/256102423\\_A\\_tutorial\\_on\\_aphasia\\_test\\_development\\_in\\_any\\_language\\_Key\\_substantive\\_and\\_psychometric\\_considerations](https://www.researchgate.net/publication/256102423_A_tutorial_on_aphasia_test_development_in_any_language_Key_substantive_and_psychometric_considerations)

Jared, D. (2002). Spelling-Sound Consistency and Regularity Effects in Word Naming. *Journal of Memory and Language*, 46(4), 723-750. <https://doi.org/10.1006/jmla.2001.2827>

Launay, L., Maeder, C., Roustit, J., & Touzin, M. (2018). *EVALEO 6-15 – Batterie d'évaluation du langage oral et du langage écrit chez les sujets de 6 à 15 ans*. 43.

Lecerf, T., Kieng, S., & Geistlich, S. (2016). WISC-IV : Valeurs seuils pour des changements significatifs des scores de différence test-retest. *Pratiques Psychologiques*, 23. <https://doi.org/10.1016/j.prps.2016.07.003>



Lefavrais. (1967). *Manuel du test de l'Alouette*.

Lefavrais. (2005). *ALOUETTE-R - Test d'analyse de la lecture et de la dyslexie* (Paris: ECPA).  
<https://www.pearsonclinical.fr/test-danalyse-de-la-lecture-et-de-la-dyslexie>

Lobrot, L. (1987). *Lire, avec épreuves pour évaluer la capacité de lecture (D-OR-LEC)* (ESF).  
<https://www.librairiedialogues.fr/livre/1412925-lire-avec-epreuves-pour-evaluer-la-capacite-de--michel-lobrot-esf>

Marshall, J. C., & Newcombe, F. (1973). Patterns of paralexia : A psycholinguistic approach. *Journal of Psycholinguistic Research*, 2(3), 175-199. <https://doi.org/10.1007/BF01067101>

McArthur, G. (2007). Test–retest effects in treatment studies of reading disability : The devil is in the detail. *Dyslexia*, 13(4), 240-252. <https://doi.org/10.1002/dys.355>

McClelland, J., & Rumelhart, D. (1981). *An interactive activation model of context effects in letter perception : I. An account of basic findings*.

McCrae, R. R., Kurtz, J. E., Yamagata, S., & Terracciano, A. (2011). Internal Consistency, Retest Reliability, and Their Implications for Personality Scale Validity. *Personality and Social Psychology Review*, 15(1), 28-50. <https://doi.org/10.1177/1088868310366253>

Midy, F. (1996). *Validité et fiabilité des questionnaires d'évaluation de la qualité de vie : Une étude appliquée aux accidents vasculaires cérébraux*. 38 p., Table, ref. bib. : 3 p. 1/4.

Morton, J. (1980). The logogen model and orthographic structure. En U. Frith (Ed.) : Cognitive processes in Spelling. *The Logogen Model and Orthographic Structure in Cognitive Processes in Spelling*, 117-133.

Olivier, D. (2007). Etude comparative de deux épreuves de lecture : L'Alouette et Lecture en une minute. *Etude comparative de deux épreuves de lecture : l'Alouette et Lecture en une minute*.

Phénix, T., Diard, J., & Valdois, S. (2017). *LES MODÈLES COMPUTATIONNELS DE LECTURE*. 36.

Pierart, B., & Gregoire, J. (2004). *Déchiffrer et comprendre : Le test de closure en lecture revisité. Etalonnage belge du L3 de Lobrot*.

Pourcin, L., & Colé, P. (2018). L'évaluation cognitive de la lecture au collège : Synthèse des principaux outils de dépistage et de diagnostic des troubles spécifiques de la lecture et cognitifs associés. *Psychologie Française*, 63(2), 105-128. <https://doi.org/10.1016/j.psfr.2016.07.001>

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, 81(2), 275-280. <https://doi.org/10.1037/h0027768>

Rumelhart, D. E. (1994). *An Interactive Activation Model of Context Effects in Letter Perception : Part 2. The Contextual Enhancement Effect and Some Tests and Extensions of the Model*. 18.

Scharfen, J., Blum, D., & Holling, H. (2018). Response Time Reduction Due to Retesting in Mental Speed Tests : A Meta-Analysis. *Journal of Intelligence*, 6(1), 6. <https://doi.org/10.3390/jintelligence6010006>

Scharfen, J., Peters, J. M., & Holling, H. (2018). Retest effects in cognitive ability tests : A meta-analysis. *Intelligence*, 67, 44-66. <https://doi.org/10.1016/j.intell.2018.01.003>

Seron, X. (1979). *Aphasie et neuropsychologie : Approches thérapeutiques*. Editions Mardaga.

Sprenger-Charolles, L., Colé, P., Piquard-Kipffer, A., Pourcin, L., & Leloup, G. (2010). *Manuel Evalec Primaire & Collège : Logiciel d'évaluation des troubles spécifiques d'apprentissage de la lecture du CP à la 3ème EDITEUR : HappyNeuron*. 32.

Sprenger-Charolles, L., & Ziegler, J. (2019). *Apprendre à lire : Contrôle, automatismes et auto-apprentissage*.

Stanovich, K. E. (1988). *Explaining the differences between the dyslexic and the garden-variety poor reader : The phonological-core variable-difference model*.

Stanovich, K. E., West, R. F., & Feeman, D. J. (1981). A longitudinal study of sentence context effects in second-grade children : Tests of an interactive-compensatory model. *Journal of Experimental Child Psychology*, 32(2), 185-199. [https://doi.org/10.1016/0022-0965\(81\)90076-X](https://doi.org/10.1016/0022-0965(81)90076-X)

Wechsler. (2016). *WISC-V - Échelle d'intelligence de Wechsler pour enfants et adolescents—5ème édition*. <https://www.pearsonclinical.fr/wisc-v>

Ziegler, J. C. (2019). *L'art de lire et d'enseigner la lecture*. 20.

# Annexes

## Annexe I : formulaire de consentement éclairé



Département d'Orthophonie  
28, av. Valombrose  
06107 Nice Cedex



Recueil du consentement des parents et de l'enfant à l'étude d'effets retest de la lecture pour les enfants scolarisés en CM1 et CM2.

### Promoteur

Laboratoire Cobtek (EA 7276)

Institut Claude Pompidou/CMRR

10 rue Molière

06100 Nice

Secrétariat et Contact

04 93 37 76 30,

[Francoise.pasturaud@unice.fr](mailto:Francoise.pasturaud@unice.fr)

### Investigateur coordinateur

Gilles Leloup

Cobtek (EA 7276) - CHU Nice

(C.E.R.T.A)

Contacts

[gilles.leloup@univ-cotedazur.fr](mailto:gilles.leloup@univ-cotedazur.fr)

### **Chère Madame, Cher Monsieur,**

Bonjour et merci de lire ces quelques informations avant de donner votre accord pour que votre enfant participe à l'étude sur les effets retests des tests de lecture. A la suite de l'annonce que vous avez reçue, la loi n°2012-300 du 5 mars 2012 relative aux recherches impliquant la personne humaine (dite loi Jardé) nécessite que vous donniez votre consentement pour que votre enfant participe à cette étude sur les effets retests de lecture.

Nous devons vous rappeler l'objectif, la durée, le lieu de passation, les tests de lecture, les mentions de récoltes des résultats, de l'archivage des données, les bénéfices et les risques de cette étude.

*Objectif:* en pratique clinique ou lors des études expérimentale sur les troubles de la lecture (dyslexie), les cliniciens et les chercheurs doivent prendre des mesures répétées de vitesse et de précision de lecture. Or, la répétition de ces mesures peut invalider les résultats si nous ne connaissons pas le gain moyen qu'obtiennent les enfants à ces passations répétées. C'est pour cette raison que

sollicitons votre enfant, son consentement et le votre pour obtenir ces informations qui nous permettrons d'avoir des barèmes de mesures de lecture plus fiables.

*Durée* : au total d'1h30 répartie sur 5 passations. La première passation est de 30 minutes, les quatre suivantes de 20 minutes chacune réparties sur un an.

*Lieu de passation* : son établissement scolaire

*Tests* : 5 tests de lecture et 5 tests pour évaluer ses habiletés langagières et visuelles, et des questions sur ce que votre enfant pensera des tests qu'il va lire. Plus un questionnaire de renseignements destiné aux parents et un questionnaire destiné aux enseignants.

*Récoltes des résultats* : anonyme (nom de votre enfant codé en 4 lettres), stockage des données au CHU-Lenval (CERTA). Possibilité sur simple mail ou courrier de détruire toutes les données concernant les données sur votre enfant. Une copie des questionnaires et des résultats des passations sera conservée par l'investigateur pendant le temps de l'étude et de la rédaction d'une publication scientifique. A la suite, toutes les données seront archivées.

*Participation volontaire* : à tout moment votre enfant lors d'une passation ou vous-mêmes pouvez faire part par mail, courrier, le souhait de suspendre la participation à l'étude.

*Bénéfices et Risques* : les possibles bénéfices sont de vous informer et d'informer votre enfant sur son niveau de lecture. Dans le cas d'une suspicion de difficultés, nous vous informerons directement. A notre connaissance pas de risque, et si ces passations inquiètent trop votre enfant, elles seront de suite suspendues à sa demande et à la vôtre.

Dans le cadre de la validation des effets retests chez les enfants scolarisés en CM1 et CM2, j'autorise mon enfant (nom et prénoms).....,né(e) ....., scolarisé(e) en classe de ..... ,à l'école....., à participer à la passation du questionnaire.

Le

A

Tavormina Ophélie

Mesure des effets retest de l'«Alouette» et du « L3 » issu de la batterie « ORLEC » chez des enfants normo-lecteurs de CM1 et CM2.

**Résumé :**

La notion d'effet retest implique de faire passer le même un test à un individu à X temps d'intervalle et d'avoir une amélioration des scores. La passation rapprochée de tests pré et post entraînement, peut induire des biais d'interprétation des résultats, (gain ou normalisation des scores) attribuables à un effet retest. Les données concernant les effets retest sont peu nombreuses dans la littérature, notamment en orthophonie, c'est pourquoi ces effets ont été analysés chez des enfants normo-lecteurs de CM1 et CM2, avec deux tests très utilisés pour dépister les dyslexies: l'« Alouette » et le « L3 » issu de la batterie « Orlec ». Le premier est un test de leximétrie non signifiant et le second mesure la compréhension. Nous avons d'abord vérifié si chacun des tests possédait un effet retest, grâce à deux temps de passations espacés de deux mois, puis nous avons ensuite comparé les effets retest de chacun. Cette étude exploratoire sur une large cohorte, révèle qu'il existe effectivement un effet retest pour ces deux tests, et que celui de l'« Alouette » est supérieur à celui du « L3 ». Une nouvelle passation à deux mois de la deuxième est prévue pour confirmer ou infirmer les données recueillies.

**Mots-clés :** effets retest, Alouette, Orlec, orthophonie, lecture

**Abstract :**

The notion of retest effect implies that an individual takes the same test at X time interval and that there is an improvement of the scores. The close timing of pre and post-training tests can lead to biases in the interpretation of results (gain or normalization of scores) that can be due to a retest effect. In literature retest effects are scarce, especially in speech therapy, and that is why we have analyzed these effects among normal-readers children in their fourth and fifth year of primary school, with two tests widely spread to detect dyslexia: the "Alouette" and the "L3" from the "Orlec" battery. The first one is a fluency test with little significance and the second one measures comprehension. We first checked whether each of the tests had a retest effect, thanks to two time periods with an interval of two months, and then we compared the retest effects of each. This exploratory study on a large cohort has revealed that there is indeed a retest effect for both tests, and that the retest effect of the "Alouette" is superior to that of the "L3". A new test has been scheduled after an interval of two months from the second one to confirm or refute the data collected.

**Keywords :** retest effects, Alouette, Orlec, speech therapy, reading