HAL
open science

Le numérique comme outil pédagogique pour l'apprentissage de la lecture et des mathématiques : du laboratoire à la salle de classe

Julie Lassault

To cite this version:

Julie Lassault. Le numérique comme outil pédagogique pour l'apprentissage de la lecture et des mathématiques : du laboratoire à la salle de classe. Education. Aix-Marseille Université (AMU), 2021. Français. NNT : . tel-03368809

HAL Id: tel-03368809

https://amu.hal.science/tel-03368809

Submitted on 7 Oct 2021

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.

NNT/NL : 0000AIXM0000/000ED000

THÈSE DE DOCTORAT

Soutenue à Aix-Marseille Université
le 25 Février 2021 par

Julie Lassault

Le numérique comme outil pédagogique pour l'apprentissage de la lecture et des mathématiques : du laboratoire à la salle de classe

Discipline
Psychologie

École doctorale
ED 356 – Cognition, Langage, Éducation

Laboratoire de Psychologie Cognitive
Aix – Marseille Université
Centre National de la Recherche Scientifique

Composition du jury

Bernard LÉTÉ — **Rapporteur**
Université Lumière Lyon 2
Laboratoire d'Étude des Mécanismes Cognitifs

Annie MAGNAN — **Rapporteuse**
Université Lumière Lyon 2
Laboratoire d'Étude des Mécanismes Cognitifs

Bruno DE CARA — **Examinateur**
Université Nice Sophia Antipolis
Laboratoire de Psychologie Cognitive et Sociale

Liliane SPRENGER-CHAROLLES — **Examinatrice**
Aix-Marseille Université – CNRS
Laboratoire de Psychologie Cognitive

Pascale COLÉ — **Présidente du jury**
Aix-Marseille Université – CNRS
Laboratoire de Psychologie Cognitive

Johannes ZIEGLER — **Directeur de thèse**
Aix-Marseille Université – CNRS
Laboratoire de Psychologie Cognitive

Affidavit

Je soussigné, Julie Lassault, déclare par la présente que le travail présenté dans ce manuscrit est mon propre travail, réalisé sous la direction scientifique de Johannes Ziegler, dans le respect des principes d'honnêteté, d'intégrité et de responsabilité inhérents à la mission de recherche. Les travaux de recherche et la rédaction de ce manuscrit ont été réalisés dans le respect à la fois de la charte nationale de déontologie des métiers de la recherche et de la charte d'Aix-Marseille Université relative à la lutte contre le plagiat.
Ce travail n'a pas été précédemment soumis en France ou à l'étranger dans une version identique ou similaire à un organisme examinateur.

Fait à Marseille, le 11 Décembre 2020

Cette œuvre est mise à disposition selon les termes de la Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International.

Résumé

La lecture est la colonne vertébrale de tous les apprentissages. Les études récentes de l'OCDE montrent que 40.5 % des collégiens ne maitrisent pas la lecture et que 21.5 % sont même en grande difficulté. Face à ce problème, les outils numériques peuvent offrir aujourd'hui des solutions « de première intention » au sein de l'école bien avant que les élèves soient en échec. Dans cette thèse, nous avons développé une version française du jeu sérieux finlandais GraphoGame que nous avons testé sur des élèves scolarisés dans des écoles de Réseaux d'Éducation Prioritaire. La première étude de validation randomisée et contrôlée, qui a été effectuée auprès d'enfants de CP et CE1 à risques de dyslexie, a permis de constater une progression en lecture de mots plus importante après l'entraînement GraphoGame qu'après un entraînement non-informatisé ou un entraînement mathématique informatisé. Une seconde expérimentation à grande échelle sur mille élèves de CP a montré que les élèves ayant suivi l'entraînement en lecture avec GraphoGame ont eu de meilleurs résultats aux tâches essentielles à l'apprentissage de la lecture (décodage, conscience phonologique et reconnaissance de mots) que les élèves ayant suivi un entrainement numérique en mathématiques. Cependant, nos résultats montrent aussi l'importance de la prise en compte du niveau initial de l'élève (les meilleurs progrès sont obtenus pour les élèves les plus faibles) et l'engagement dans le jeu, reflétant la persévérance et la motivation de l'élève. Nos résultats montrent donc des effets spécifiques du jeu qui se généralisent à des activités de lecture en dehors du jeu et qui semblent perdurer bien au-delà de l'expérimentation. Une version complète de GraphoGame Français est aujourd'hui gratuitement disponible sur toutes les plateformes en France.

Mots clés : lecture, décodage, école, numérique, jeu sérieux, GraphoGame

Abstract

Reading is the backbone of all learning. Recent OECD studies show that 40.5% of secondary school students are not proficient in reading and 21.5% are even in great difficulty. In response to this problem, digital tools can now offer "first-line" solutions within schools long before pupils fall behind. In this thesis, we have developed a French version of the Finnish serious game GraphoGame that we tested on pupils attending schools in Priority Education Networks. The first randomised controlled validation study, which was carried out on children at risk of dyslexia in first and second grade, found a greater progression in word reading after GraphoGame training than after non-computerised training or computerised mathematical training. A second large-scale experiment on one thousand first grade students showed that students who received reading training with GraphoGame performed better on tasks essential for learning to read (decoding, phonological awareness and word recognition) than students who received digital training in mathematics. However, our results also show the importance of taking into account the children's initial level (the greatest effects of intervention were obtained for the weakest children) and commitment to the game reflecting the pupil's perseverance and motivation. Our results therefore show specific effects of the game which generalised to reading activities outside of the game and which seem to last well beyond experimentation. A full version of French GraphoGame is now available for free on all platforms in France.

Keywords: reading, decoding, school, digital, serious game, GraphoGame

Remerciements

Je remercie d'abord la femme la plus extraordinaire qui existe sur cette Terre, ma Maman. Merci pour tout ce que tu as fait pour moi. Tous les mots de cette thèse ne sauraient exprimer la gratitude que j'ai pour toi et la fierté que j'ai d'être ta fille. Je remercie aussi mon Papa, pour m'avoir aimée et soutenue depuis toutes ces années, quel bonheur de t'avoir à mes côtés. Merci aussi à l'amour de toute ma vie pour ton soutien infaillible, tes conseils et ton écoute. Merci de m'avoir accompagnée dans la folle aventure de devenir les parents de la petite lentille, notre magnifique Margaux. Merci à Izy, le meilleur partenaire de télétravail.

Je remercie très chaleureusement tous les partenaires et acteurs de ce projet, les enseignants et les élèves, sans qui rien n'aurait été possible. Je veux remercier également tous les membres du LPC, les doctorants, post doctorants, ingénieurs et chercheurs que j'ai eu la chance de côtoyer quotidiennement. Un immense merci également à Colette et Fred, grâce à vous toutes les tâches administratives ont été tellement plus simples et les repas au labo tellement plus amusants.

Merci à mes collègues d'amour qui sont devenus mes amis. Merci Flower, Kam et Jon pour votre gentillesse et votre aide qui m'ont accompagnées tout au long de ce doctorat. J'attends maintenant vos remerciements pour vous avoir fait enfin découvrir de la vraie musique. Merci Ludivine, avec qui j'ai tissé des liens depuis le tout premier jour, pour m'avoir motivée et aidée à relativiser lorsque cela a été nécessaire. Bravo pour tout le travail que tu as accompli, j'ai hâte de clôturer ce merveilleux chapitre avec toi, et une coupe de champagne.

Je remercie de tout mon cœur ma team, Édouard et Jean-Patrice, la meilleure équipe avec laquelle on puisse espérer travailler. Quel plaisir de vous voir chaque matin, de boire du café dans nos tasses assorties, d'aller à Montréal ou à l'italien d'à côté. Merci pour ces heures à créer des fichiers Excel d'un milliard de lignes, à organiser des conférences et à tester tous les élèves de CP de Marseille. Merci Éd de m'avoir poussée à aller dans le détail de chaque chose, merci JP de m'avoir aidée à avoir un regard global sur toutes les situations. Aujourd'hui nous ne sommes plus officiellement collègues, mais on est toujours amis et j'ai hâte de vous retrouver.

J'adresse un immense merci à Liliane Sprenger-Charolles pour tout le savoir que tu m'as transmis. Je souhaite à tout le monde de travailler avec des personnes aussi passionnées et accessibles que toi. Je ne pouvais imaginer voir un jour mon nom à côté du tien sur une publication, j'en suis extrêmement fière.

Je remercie tout particulièrement les membres du jury, Bruno De Cara, Bernard Lété, Annie Magnan et la présidente Pascale Colé, pour leur temps consacré à la lecture et l'évaluation de cette thèse. Quelle fierté immense de vous présenter mes travaux.

Enfin j'adresse mes remerciements les plus sincères à mon directeur de thèse Johannes Ziegler. Depuis le début de mes études en psychologie j'ai été passionnée par l'apprentissage de la lecture. Je revois le temps où je lisais et j'étudiais tes travaux, sans penser un jour que j'aurais la chance immense de travailler avec toi. Merci de m'avoir fait confiance tout au long de ce magnifique projet. Merci de m'avoir guidée et accompagnée dans les multiples tâches qui façonnent une thèse et je te serai éternellement reconnaissante pour tout ce que tu m'as enseigné. Merci pour tout.

En mémoire à Jacques Ginestié,

qui a inlassablement œuvré pour un rapprochement entre terrain, recherche et formation.

Cette thèse n'aurait pu se faire sans lui, et la fin du projet Lemon marque le début d'une nouvelle aventure.

Table des matières

AFFIDAVIT 3

RESUME 5

ABSTRACT 7

REMERCIEMENTS 9

TABLE DES MATIERES 13

PREAMBULE 15

INTRODUCTION 17

1. APPRENTISSAGE DE LA LECTURE 19

1.1. MECANISMES 19
1.2. DIFFICULTES 27
1.3. INTERVENTIONS 30

2. GRAPHOGAME 33

2.1. PRESENTATION 33
2.2. ADAPTATION FRANÇAISE 36
2.3. OUTIL FINAL ET AMELIORATIONS 41

3. EXPERIMENTATIONS SUR LE TERRAIN 45

3.1. PREMIERE EXPERIMENTATION 45
3.2. EXPERIMENTATION LEMON 54

4. DISCUSSION GENERALE 73

BIBLIOGRAPHIE 81

ANNEXES 91

Préambule

Ce travail de thèse s'inscrit dans le cadre de l'appel à projet e-FRAN (Espaces de Formation, de Recherche et d'Animation Numérique) issu du Programme d'Investissement d'Avenir (PIA) 2. L'objectif était de soutenir des projets de transformation de l'École avec la volonté de créer des « territoires d'innovation numérique » en prenant appui sur la recherche. Dans ce contexte, l'action e-FRAN vise à mettre en évidence les conditions d'utilisations efficace du numérique au service de la réussite scolaire des élèves.

Parmi les lauréats, le projet LEMON (Lecture Mathématique Outils Numériques) propose la mise en place et l'évaluation à grande échelle d'un outil numérique d'aide à l'apprentissage de la lecture, au sein des classes de Cours Préparatoire (CP).

Introduction

Que ferions-nous dans un monde où nous serions incapables de comprendre ce que nous lisons ? Difficile de nous le représenter, car pour nous, lecteurs experts, la lecture est devenue un processus automatique. Nous ne nous rendons plus compte de la complexité de la machinerie cognitive qui se cache derrière cette capacité. Dans notre société actuelle, être incapable de lire est un véritable handicap car la lecture est la colonne vertébrale de tous les apprentissages, elle est nécessaire pour comprendre une simple notice, d'enrichir nos connaissances de manière autonome ou d'accéder facilement à un emploi. De nombreuses personnes éprouvent des difficultés au moment de l'apprentissage de la lecture qui peuvent persister jusqu'à l'âge adulte. Comment accompagner et aider les enfants à apprendre à lire, c'est la question à laquelle nous nous sommes intéressés.

Aujourd'hui en France, plus de 250 000 enfants souffrent de troubles d'apprentissage de la lecture et les conséquences de ces difficultés sont dramatiques : décrochage scolaire, détresse et avenir professionnel compromis. Plusieurs études nationales (Cycle des Évaluations Disciplinaires Réalisées sur des Échantillons, CEDRE, en 2015 ; Organisation de Coopération, et de Développement Économique, OCDE, en 2014) ont montré que la proportion des élèves en difficulté de lecture a augmenté significativement au cours des dix dernières années. Les résultats soulignent qu'à la fin de l'école primaire, un élève sur cinq a des difficultés pour comprendre un texte, proportion encore plus importante dans les établissements relevant de l'éducation prioritaire. Il a été montré qu'au sein des Réseaux d'Éducation Prioritaires (REP) la compréhension est altérée tout autant que les mécanismes de plus bas niveau, comme la reconnaissance des lettres (à la différence des écoles situées hors REP) (Daussin, Keskpaik & Rocher, 2011).

Face à l'ampleur du problème les enseignants peuvent se sentir démunis et les solutions extra-scolaires sont souvent insuffisantes. Les faibles liens qui existent entre le terrain, la formation et la recherche sont également vecteurs de difficultés. Dans une étude conduite en 2014, l'analyse du temps scolaire a permis de mettre en évidence que le volume de temps alloué à l'apprentissage de la lecture (en moyenne 7 minutes par jour) est largement insuffisant (Suchaut, Bougnères & Bouguen, 2014). Les auteurs encouragent à développer, dans les écoles, des pratiques pour optimiser au mieux les temps d'apprentissage. Il apparait ici comme primordial de proposer de nouveaux outils, notamment numériques, permettant une solution au sein même de la classe.

L'objectif de cette thèse était le développement et l'évaluation d'un outil informatisé d'aide à l'apprentissage de la lecture, notamment pour les enfants en

difficulté. De nombreux outils existent, mais ils sont rarement évalués scientifiquement et leur contenu est limité ou restreint. Il nous a donc semblé important de proposer aux enseignants un outil pertinent, complet, évalué et facilement accessible. Cet outil avait comme vocation de permettre une réponse individualisée, différenciée et adaptative de « première intention » au sein de l'école. Les expérimentations se sont déroulées au CP et au CE1, donc bien plus tôt que les interventions habituelles. Le second objectif était d'intervenir dans les établissements en REP et REP+, où le pourcentage d'élèves en grande difficulté de lecture atteint 30% (Billard *et al*, 2009 ; Fluss *et al.*, 2009). Nous avons fait l'hypothèse que l'utilisation intense de cet outil numérique devrait fortement réduire le pourcentage d'élèves en difficulté d'apprentissage de la lecture à la fin du CP, en améliorant leurs compétences en fluence, en décodage et en orthographe.

La thèse s'organise en quatre parties. La première partie est théorique et portera sur les mécanismes d'apprentissage de la lecture, de ses difficultés ainsi que des interventions qui existent pour prévenir ces difficultés. La deuxième partie présentera l'outil numérique finlandais GraphoGame et comment nous l'avons adapté pour la langue française. La troisième partie présentera les deux études de validation qui ont été conduites sur le terrain : la première auprès d'enfant « à risque » de dyslexie, la seconde au sein d'établissement situés en Réseaux d'Éducation Prioritaire. La dernière partie est consacrée à la discussion des résultats et les perspectives.

1. Apprentissage de la lecture

1.1. Mécanismes

La lecture représente l'une des plus formidables inventions de la civilisation humaine. Grâce à cette innovation, des symboles posés sur une feuille vont activer dans le cerveau du lecteur une série de processus qui aboutiront à la compréhension de ce qui vient d'être lu (Dehaene, 2007). Chez le lecteur expert, la reconnaissance des mots écrits est quasi reflexe : un quart de seconde suffit pour qu'un mot soit reconnu et choisi par l'ensemble des mots que le lecteur connait (Seidenberg & McClelland, 1989). Pour arriver à ce niveau d'automatisation, un apprentissage rigoureux est nécessaire (Ziegler, Perry & Zorzi, 2020). L'apprentissage de la lecture repose sur le langage oral pour lequel nous avons des prédispositions biologiques (Liberman, Cooper, Shankweiler & Studdert-Kennedy, 1967 ; Liberman & Whalen, 2000). En effet, le contact de l'enfant avec un milieu langagier même assez rudimentaire semble suffisant pour acquérir le langage oral (Pinker, 2003). En revanche, l'apprentissage du langage écrit doit impérativement être enseigné (Ferrand, Lété & Thevenot, 2018) et le seul contact avec l'écrit ne fera jamais d'un individu un lecteur expert.

<u>Code et décodage</u>

Le système d'écriture est une création culturelle qui permet de « fixer » le langage oral et de le rendre accessible par la modalité visuelle. Notre système d'écriture, qui est un système alphabétique, est un code dans lequel les lettres (ou ensembles de lettres, les graphèmes) correspondent à des sons de la parole (les phonèmes). Dans un système alphabétique parfait, chaque graphème correspond à un seul phonème et chaque phonème s'écrit avec un seul graphème (Ziegler, 2018). Un apprenti lecteur doit d'abord comprendre ce « principe alphabétique », puis l'appliquer pour associer à une lettre isolée ou à un groupe de lettres (graphème) le phonème correspondant, ce qu'on appelle le décodage (Share, 1995). C'est grâce à un enseignement explicite des liens qui existent entre les graphèmes et les phonèmes que l'enfant peut décoder des mots qu'il a peut-être déjà entendu mais jamais vu auparavant. Cependant, automatiser le décodage nécessite la répétition et de l'entrainement, d'autant plus que la majorité des systèmes alphabétiques, comme celui du français ou de l'anglais, ne sont pas parfaits et la même lettre peut avoir différentes prononciations et le même phonème différents orthographes (Ziegler, 2018, voir ci-dessous).

Il est aujourd'hui largement admis que « le rôle irremplaçable de la découverte et de la mise en œuvre du principe alphabétique » (Goigoux, 2000) est une étape initiale nécessaire dans l'apprentissage de la lecture (Morais, 1994 ; Sprenger-

Charolles & Casalis, 1996). Une enquête du National Reading Panel (2000) aux États-Unis a mis en évidence l'importance de l'enseignement systématique du décodage au début de l'apprentissage de la lecture pour devenir un lecteur expert. En parallèle, des études ont mis en évidence l'importance, au moment de l'apprentissage du décodage, de la capacité à identifier et à manipuler les unités phonologiques de la langue, notamment au niveau des phonèmes, ce qu'on appelle la « conscience phonémique » (pour une synthèse en français, voir Sprenger-Charolles & Colé, 2013). Une bonne conscience phonémique facilite l'apprentissage du décodage mais en même temps l'apprentissage du décodage développe la conscience phonémique, c'est une relation réciproque (Perfetti, Beck, Bell & Hughes, 1987). Quoi qu'il en soit, la conscience phonémique, même si elle est mesurée bien avant le début de la lecture, est un prédicteur puissant de l'apprentissage de la lecture (Holopainen, Ahonen & Lyytinen, 2001 ; Lyytinen, Erskine, Hamalainen, Torppa & Ronimus, 2015).

Différences inter-langues

En fonction du système orthographique de la langue dans laquelle on apprend à lire, le décodage sera mis en place plus ou moins rapidement. En effet, il existe des irrégularités au sein des systèmes d'écriture qui vont déterminer la vitesse avec laquelle la lecture sera apprise et maitrisée (Ziegler & Goswami, 2006). La consistance (ou régularité) des correspondances graphophonologiques varie d'une langue à l'autre en rendant certaines très régulières (transparentes) et d'autres très irrégulières (opaques) (voir tableau 1). Le degré d'irrégularité peut être quantifié différemment. Il est possible de compter le nombre de graphèmes et de phonèmes existant dans une langue, si le ratio n'est pas équivalent il est probable qu'à un phonème puisse correspondre plusieurs graphèmes et inversement. Le nombre de « violation » des règles de correspondance qui existent entre les graphèmes et les phonèmes est également un indicateur du niveau de régularité d'une langue (Ziegler, 2018).

		Profondeur Orthographique				
		Transparent				**Opaque**
Structure Syllabique	**Simple**	Finnois	Grec Italien Espagnol	Portugais	Français	
	Complexe		Allemand Norvégien Islandais	Néerlandais Suédois	Danois	Anglais

Tableau 1 : Classification des systèmes orthographiques de différentes langues selon leur complexité et leur profondeur/transparence (d'après Seymour, Aro & Erskine, 2003)

La consistance d'une langue peut être envisagée dans le sens de la lecture (nombre de fois qu'un graphème se prononce d'une façon) ou dans le sens de l'écriture (nombre de fois qu'un phonème s'écrit d'une façon) (Ziegler, 2018). Par exemple, dans le sens de la lecture la consistance du graphème « ch » qui se lit / ʃ / comme dans le mot « chemin », est de 97,58% car il n'existe que quelques mots où il se prononce /k/ comme dans le mot « chorale ». Dans le sens de l'écriture la consistance du phonème / ʃ / est de 99,3% car il s'écrit quasiment tout le temps avec le graphème « ch » mais on trouve des mots où c'est le graphème « sch » qui est utilisé comme dans le mot « schéma » (données issues de Manulex-Morpho (Peerman, Sprenger-Charolles & Messaoud-Galusi, 2013), base de données élaborée à partir de Manulex (Lété, Sprenger-Charolles & Colé, 2004)). La régularité qui existe entre les relations graphémiques et phonétique a donc un impact sur la vitesse avec laquelle le décodage sera mis en place (voir figure 1). Il sera maitrisé plus rapidement pour les systèmes d'écriture réguliers dans le sens de la lecture et de l'écriture, comme le finnois, l'italien ou l'espagnol.

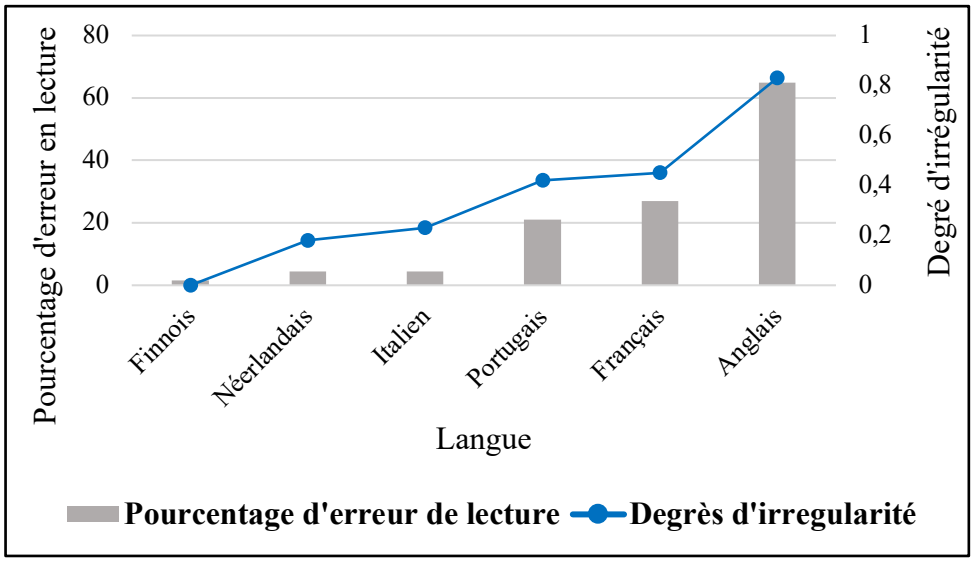


Figure 1 : Relation entre le niveau de lecture (pourcentage d'erreur) et le degré d'irrégularité d'une langue, adapté de Ziegler, 2018

Concernant les systèmes plus opaques comme le danois, le français ou l'allemand, le décodage est plus difficile à maitriser car si à chaque graphème correspond un phonème, à un phonème peut correspondre plusieurs graphèmes (par exemple, pour le français, le phonème / ã / peut s'écrire « an », « am », « ent », « em », « ean », « aen » etc.) (Seymour, Aro & Erskine, 2003). Enfin, certains systèmes d'écritures, comme l'anglais, sont irréguliers dans le sens de la lecture et de l'écriture ce qui rend son apprentissage encore plus difficile (Ziegler, 2018).

La structure syllabique des langues peut également être un critère d'efficacité dans la mise en œuvre du décodage. Elle est plus rapide lorsque la structure est simple que lorsqu'elle est complexe. Une langue aura un système d'écriture dit simple si, pour

la majorité des mots, peu de graphèmes sont nécessaires à la formation d'une syllabe (exemple en français : « chat », un mot, une syllabe, deux graphèmes). A l'inverse, un système est complexe si en majorité les mots de la langue sont composés de nombreux graphèmes (exemple en anglais : « sprint », un mot, une syllabe, cinq graphèmes) (Seymour, Aro & Erskine, 2003).

Pour la langue française, qui possède un système d'écriture relativement opaque, l'apprentissage des relations qui existent entre les graphèmes et les phonèmes est un des objectifs centraux de la première année d'école primaire (Cours Préparatoire, CP). Comme dans de nombreux pays, l'introduction d'une nouvelle correspondance graphophonologique par semaine est l'approche la plus courante mais peu de travaux existent sur la façon dont le rythme d'enseignement des correspondances contribue au développement des compétences en lecture. Une seule étude conduite en 2012 par Jones & Reutzel a montré que les enfants qui étaient exposés à une nouvelle CGP par jour, donc passaient plusieurs fois en revue chacune, réussissaient mieux à acquérir la connaissance des lettres que ceux suivant une approche plus traditionnelle. Cependant, les effets sur la lecture et d'autres compétences n'ont pas été étudiés. Une étude récente (Sunde, Furnes & Lundetræ, 2020) a donc examiné comment un rythme d'enseignement plus soutenu influence les compétences en lecture et en écriture au cours de la première année d'école. Les résultats généraux montrent qu'une introduction plus rapide est positivement corrélée aux capacités en lecture et écriture des élèves, en particulier pour ceux le plus en difficulté (connaissance lettres/sons ; vitesse de lecture ; reconnaissance de mots écrits et orthographe). Ces travaux ont été conduits en Norvège où la langue possède un système d'écriture sensiblement différent de celui du français. Vu le faible nombre d'études, il semble pertinent que ce domaine de recherche soit approfondi.

Modèles d'apprentissage de la lecture

Depuis plusieurs décennies, la recherche en psychologie cognitive et en psycholinguistique a fait émerger différents modèles permettant de rendre compte de l'apprentissage de la lecture (pour une revue, voir Ziegler, Perry & Zorzi, 2019). Le premier modèle d'apprentissage de la lecture était un modèle à stade (Frith, 1985). Ce modèle fait état de trois stades marqués par l'utilisation de stratégies différentes permettant d'aboutir à la maitrise de la lecture. Le premier stade est le stade logographique qui est caractérisé par l'utilisation de l'enfant d'indices fournis par le contexte et par les traits visuels caractéristiques des lettres (ou groupe de lettres). Toutefois, il a été largement questionné s'il s'agissait d'un vrai stade de lecture par qui tous les enfants devraient passer sur leur chemin de la lecture ou plutôt d'un phénomène de « pré-lecture » sans véritable importance pour l'apprentissage et la maitrise de la lecture par la suite (Sprenger-Charolles & Bonnet, 1996 ; Wimmer & Hummer, 1990). Le deuxième stade est le stade alphabétique (ou phonologique) qui est marqué par l'utilisation du décodage. Le troisième et dernier stade est le stade orthographique qui est caractérisé par l'analyse des mots en tant qu'unité

orthographique sans passer par la phonologie. Ce stade traduit le passage vers la lecture experte.

Ehri (1995) a également proposé un modèle développemental de l'apprentissage de la lecture comprenant quatre phases. La première est la phase pré-alphabétique. Dans cette phase, les enfants n'ont pas encore de connaissances alphabétiques et reconnaissent certains mots comme des images ou des ensembles. Ils utilisent des indices contextuels, des images et des stratégies de devinette pour identifier les mots. La deuxième phase est la phase alphabétique partielle qui est caractérisée par le début de l'utilisation du décodage. Mais dans cette phase, les enfants utilisent souvent le son porté la première lettre pour deviner des mots. La phase alphabétique partielle est plus fiable que l'utilisation de repères visuels, mais elle ne permet pas de lire des mots nouveaux. La troisième phase est la phase alphabétique complète qui se caractérise par l'utilisation du décodage. La quatrième et dernière phase est la phase alphabétique consolidée qui est considérée comme la phase finale dans le développement de la lecture des mots (Ehri & McCormick, 1998). La lecture des mots est rapide et sans effort et le lecteur est en mesure de se concentrer entièrement sur le sens de ce qui est lu.

Il existe également des modèles computationnels qui simulent l'apprentissage de la lecture de façon quantitative. Le plus connu de ces modèles est celui de Harm et Seidenberg (1999) qui va apprendre à « lire » en faisant correspondre l'orthographe des mots avec leur phonologie via un corpus d'entrainement de mots monosyllabiques. Les connexions du réseau sont modifiées lorsque cela est nécessaire pour diminuer les erreurs (différence entre les réponses fournies par le réseau et celles attendues). Les résultats montrent que le réseau est capable de lire correctement 98% des mots qui lui sont présentés et de généraliser ce qu'il a appris à des à la lecture de pseudo-mots (79% de réussite). Cependant, ce modèle a été critiqué par rapport à sa plausibilité développementale (Ziegler, Perry & Zorzi, 2014 ; 2019 ; 2020) puisque son régime d'apprentissage nécessite des millions d'essais supervisés (avec correction de l'erreur) ce qui est très différent de l'apprentissage des enfants pour qui souvent une ou deux rencontres avec un mot suffisent pour créer des représentations orthographiques (Share, 1995 ; 1999).

Le modèle de Share (1995) met en avant le mécanisme du décodage, que nous trouvons aussi dans les modèles de Frith (1985), Ehri (1995) et Ziegler et Goswami (2005), mais il rajoute un deuxième mécanisme, celui de l'auto-apprentissage. L'idée est simple. L'apprentissage de quelques règles graphophonologiques (plus au moins complexes en fonction de la langue) permet à l'enfant de décoder une grande majorité de mots, même ceux qu'il n'aura jamais rencontré, mais dont il connait déjà la forme phonologique et son sens grâce au langage oral. C'est ce que Share appelle « minimum number of rules, maximal generative power ». Puis, le mécanisme de décodage permet à l'enfant de lire en autonomie. Chaque fois que le lecteur parvient à décoder un mot qu'il retrouve dans son lexique phonologique et qui est correct dans

le contexte de ce qui est lu, il peut renforcer « les règles » qui lui ont permis ce décodage et créer une représentation orthographique (voir figure 2). On parle alors de phénomène « d'auto-apprentissage » car c'est la lecture qui renforce la lecture. En effet, un enfant qui lit seul améliore de plus en plus sa compréhension du code et va lire plus facilement et plus rapidement (Share, 1995).

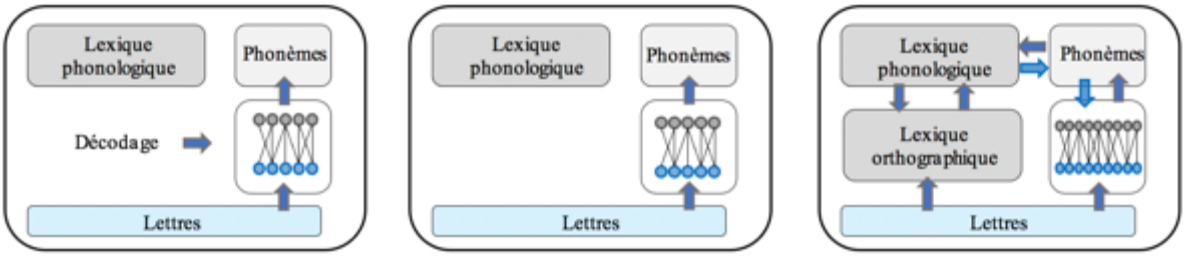


Figure 2 : Modèle d'apprentissage de la lecture basé sur le décodage et l'auto-apprentissage des représentations orthographiques, adapté de Ziegler, Perry et Zorzi (2014)

Pour démontrer que ce mécanisme peut fonctionner, Ziegler, Perry, et Zorzi (2014) ont implémenté un modèle d'apprentissage de la lecture qui apprend par décodage puis auto-apprentissage. Comme le montre la figure 2, le modèle part du principe que l'enfant possède un lexique phonologique (vocabulaire) bien avant la lecture, dans lequel sont stockés le son (la phonologie) et le sens (la sémantique) des mots. Puis l'enfant apprend les associations graphème-phonème : il apprend à décoder. Chaque décodage réussi, c'est-à-dire à chaque fois que l'enfant réussit à retrouver dans son lexique le bon mot en le lisant, cela permet de renforcer les connexions à l'origine de ce décodage, ce qui constitue un cercle vertueux d'apprentissage. L'apprentissage explicite (avec le soutien d'un maitre) devient alors un auto-apprentissage sans maitre. Ziegler et collaborateurs (2014) ont montré que l'implémentation de ce modèle sur ordinateur peut simuler l'apprentissage normal de la lecture. Ce passage du mode d'apprentissage explicite à l'auto-apprentissage est essentiel. Les enfants passent alors progressivement du statut de « novice » à « expert » en faisant correspondre automatiquement l'orthographe des mots à leur sens, sans avoir recours au décodage (Castles & Nation, 2006).

Enfin un dernier modèle qui met en avant le concept de la « qualité lexicale » a été proposé (Perfetti, 2007). Une bonne qualité lexicale se traduit par un stockage de la représentation d'un mot à la fois précis et flexible. La précision de la représentation - la connaissance de l'orthographe exacte - est importante car elle permet à un enfant de distinguer un mot écrit des autres mots d'apparence similaire, ce qui lui donne un accès direct à sa signification (différencier « chat » de « chut »). La flexibilité de la représentation est également importante car elle permet à l'enfant de s'adapter de manière dynamique aux différents sens portés par un même mot (différencier « faire les courses » de « faire la course »). La qualité lexicale joue un rôle majeur dans le passage de lecteur novice à expert car au fur et à mesure qu'elle se développe, les ressources cognitives allouées initialement au décodage se libèrent pour la

compréhension. La compréhension écrite est une tâche complexe qui mobilise énormément l'attention, la mémoire et les processus linguistiques de haut niveau (Castles, Rastle & Nation, 2018). Lorsque la qualité lexicale est élevée, les ressources cognitives du lecteur peuvent être largement orientées vers cette tâche difficile, car les mots individuels sont reconnus rapidement, automatiquement et avec un minimum d'effort conscient. En revanche, lorsque la qualité lexicale est faible, une partie des ressources cognitives limitées du lecteur doit être orientée vers la tâche plus fondamentale de la reconnaissance des mots, et la compréhension s'en trouve compromise.

Bases neurales de l'apprentissage de la lecture

L'imagerie cérébrale (notamment la Tomographie par Émission de Positons (TEP) et l'Imagerie par Résonnance Magnétique fonctionnelle (IRMf)), a permis d'étudier *in vivo* les réseaux neuronaux qui sont à la base des fonctions cognitives. Grâce à ces techniques non invasives, le vaste réseau dédié à la lecture a été découvert. Il s'étend du cortex visuel aux aires du langage dans le cortex frontal inferieur, le cortex temporal et le gyrus angulaire (voir figure 3).

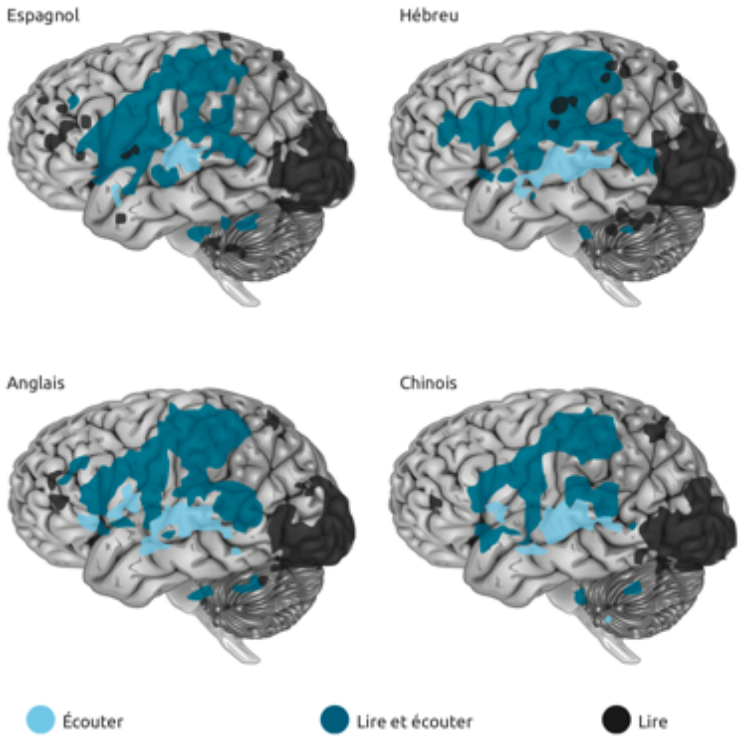


Figure 3 : Activation cérébrale de l'hémisphère gauche lors de la lecture et de la perception de la parole dans quatre langues (Rueckl et al., 2015), issu de Ziegler & Deleuze, 2018. (note : la zone noire est activée uniquement par la lecture. La zone bleue claire est active uniquement lors de l'écoute de la parole. La zone bleue foncée est activée conjointement lors de la lecture et l'écoute de la parole)

Le réseau cérébral de la lecture est très semblable à travers les différents systèmes d'écriture (lettres ou caractères qui codent pour des phonèmes, des syllabes

ou des graphèmes) (Rueckl *et al.*, 2015). Dans tous les systèmes d'écriture, alphabétique ou non (Bolger, Perfetti & Schneider, 2005) la reconnaissance visuelle des lettres (ou caractères) est réalisée dans une région spécifique de l'hémisphère gauche connue sous le nom d'aire de la forme visuelle des mots (Visual Word Form Area (VWFA)) (Cohen *et al.*, 2000). La fonctionnalité de cette zone n'est pas innée et se développe grâce à l'apprentissage de la lecture (Saygin *et al.*, 2016). Elle possède des propriétés fonctionnelles importantes pour la lecture comme l'invariance de position spatiale (Cohen *et al.*, 2002). La VWFA est initialement impliquée dans la reconnaissance des visages et des outils (Dehaene, 2007), mais sa réactivité aux visages décroit à mesure que la compétence en lecture s'améliore, suggérant une spécialisation dans la reconnaissance des lettres dans l'hémisphère gauche et la reconnaissance des visages dans l'hémisphère droit (mais voir une étude récente qui met en question cette hypothèse, Dehaene-Lambertz, Monzalvo & Dehaene, 2018).

Compréhension écrite

L'objectif de la lecture est de comprendre, c'est-à-dire saisir le sens de ce qui est lu. De manière schématique, la compréhension écrite d'un mot (puis d'un texte) peut être représentée sous la forme d'une multiplication simple (voir figure 4).

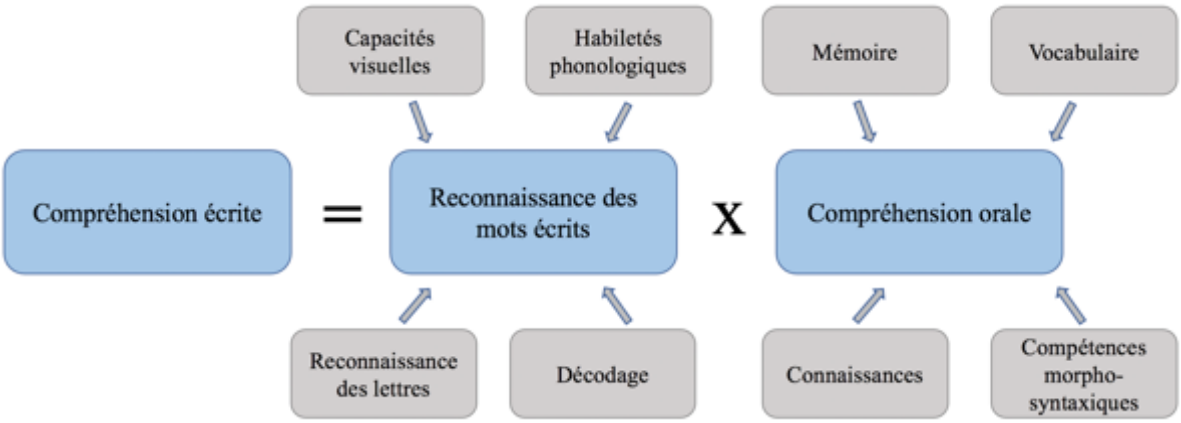


Figure 4 : Modèle simplifié de la compréhension écrite (Hoover & Gough, 1990) adapté de Sprenger-Charolles & Ziegler (2019)

Si l'une des deux variables est égale à zéro, le produit final sera lui aussi nul. Au commencement de l'apprentissage de la lecture, les facteurs intervenants dans la reconnaissance des mots écrits (comme la reconnaissance des lettres ou le décodage) expliquent la plus grande part de la variabilité observée en compréhension écrite (Catts *et al.*, 2015). Les troubles organiques de la vision et/ou de l'audition, la déficience intellectuelle et les carences environnementales telles qu'un pauvre environnement langagier, un enseignement de la lecture absent et un environnement socio-culturel défavorable expliquent également les difficultés d'apprentissage de la lecture (Gentaz, Sprenger-Charolles, Theurel & Colé, 2013). Lorsque le principe de décodage est

maitrisé et plus ou moins automatisé c'est la compréhension orale qui va prédire le mieux les compétences en compréhension écrite (Écalle *et al.*, 2013).

1.2. Difficultés

De nombreux paramètres influencent la reconnaissance des mots écrits et la compréhension orale. Comme le montre la figure 4, ils influencent également la compréhension écrite. La multiplicité des paramètres rend l'apprentissage de la lecture complexe et il n'est pas rare que les enfants rencontrent des difficultés. On utilise souvent le terme de « troubles dys » (Habib & Ziegler, 2016) pour parler de ces élèves, or la dyslexie développementale représenterait plutôt l'extrémité d'un continuum des difficultés de lecture. Selon différentes sources, sa prévalence en France varie entre 3,5 % et 8 % (Billard *et al.*, 2009 ; Habib, 2014 ; Sprenger-Charolles & Colé, 2003). Elle peut atteindre des niveaux plus élevés dans les milieux socio-économiques défavorisés (Fluss *et al.*, 2008).

Les grandes classifications internationales des maladies (CIM-10 et DSM-IV) définissent souvent la dyslexie comme « un trouble spécifique et durable de l'acquisition du langage écrit (incluant la lecture et l'orthographe), interférant de manière significative avec la réussite académique et/ou les activités de la vie quotidienne, mesurable sous la forme d'un écart par rapport aux performances attendues ou à l'âge et à l'intelligence du sujet, trouble qui ne peut être expliqué ni par un déficit sensoriel, ni par une affection neurologique ou psychiatrique, ni par un défaut d'intelligence, ni par un manque d'opportunité scolaire ». Comme le soulignent Habib et Joly-Pottuz (2008), cette définition possède des limites en ne faisant pas référence aux causes et aux mécanismes de la dyslexie qui sont pourtant bien identifiés. Ainsi, il est désormais souvent inclus dans la définition la nature génétique et l'origine neurobiologique (Norton *et al.*, 2015) mais surtout du lien étroit qui existe avec l'acquisition du langage oral et en particulier de la phonologie.

La théorie phonologique de la dyslexie repose sur l'idée que les difficultés sont dues à un double déficit (Sprenger- Charolles & Colé, 2013 ; Sprenger-Charolles, Siegel, Jimenez & Ziegler, 2011). On retrouve une représentation mentale dégradée des lettres et des phonèmes d'une part et des difficultés d'accès à ces représentations au sein du lexique d'autre part. Le traitement phonologique perturbé va empêcher d'établir correctement les correspondances avec les graphèmes et donc nuire à la mise en place du décodage. Le déficit phonologique est une cause directe des difficultés de lecture (pour une revue, voir Ziegler & Goswami, 2005).

Cependant, l'idée que la dyslexie est un trouble lié spécifiquement à la phonologie ne fait pas l'unanimité. D'autres théories prennent en compte un dysfonctionnement plus général qui affecte aussi la perception visuelle et les traitements visuo-attentionnels des enfants dyslexiques (Bosse, Tainturier & Valdois,

2007 ; Stein, 2014 ; Vidyasagar & Pammer, 2010). Une lecture fluide nécessite un traitement orthographique efficace. Si l'empan visuo-attentionnel est trop réduit, le traitement parallèle des lettres devient impossible tout comme l'exploitation des informations parafovéales (Bosse Tainturier & Valdois, 2007 ; Ziegler, Pech-Georgel, Dufau & Grainger, 2010). Si le codage de la position des lettres est déficitaire (Collis, Kohnen & Kinoshita, 2013) possiblement dû à des déficits de la voie magnocellulaire (Stein, 2014), le mécanisme de décodage est perturbé et l'apprentissage orthographique est fortement affecté. Si le traitement visuo-attentionnel est déficitaire, l'enfant peut souffrir d'un encombrement perceptif plus important (Zorzi *et al.*, 2012) et/ou d'un guidage oculomoteur perturbé.

Bien que la question de l'indépendance et de la prévalence de ces déficits fasse toujours débat (Saksida *et al.*, 2016), il semble aujourd'hui clair que les causes de la dyslexie sont multifactorielles (Menghini *et al.*, 2010 ; Pennington, 2006 ; Van Bergen, Van Der Leij & De Jong, 2014 ; Ziegler *et al.*, 2008). Avec une prévalence inférieure à 10%, la dyslexie n'explique pas les données recensées en difficultés de lecture par les études nationales. Il y a une population d'enfant pour qui la cause du faible niveau en lecture est ailleurs. La littérature suggère que les difficultés pourraient aussi résider dans la compréhension orale, le niveau de vocabulaire, ou encore le désavantage lié au milieu socio-économique (Billard *et al.*, 2008 ; Billard *et al.*, 2009 ; Fluss, *et al.*, 2009 ; Gentaz, Sprenger-Charolles, Theurel, & Colé, 2013).

Temps décodage et dédoublement classes

L'étude de Suchaut et collaborateurs (2014) a également mis en évidence un temps scolaire pour l'entrainement au décodage au CP insuffisant. Des travaux de références ont montré que l'apprentissage du code (et un traitement en phonologie indissociable) exige un temps d'entrainement individuel entre 36 heures et 53 heures (Vadasy, Jenkins, Antil, Wayne & O'Connor, 1997 ; Vadasy, Jenkins & Pool, 2000) alors que c'est environ une vingtaine d'heure qui sont en réalité consacré à cette tâche (l'autre moitié du temps est consacré à l'apprentissage du vocabulaire, de la grammaire *etc.*). *« Ces 20 heures annuelles éventuelles pendant lesquelles l'élève serait réellement engagé sur l'apprentissage du code apparaissent bien dérisoires face au défi que représente l'apprentissage de la lecture pour les élèves les plus fragiles. Ce temps d'engagement correspond à environ 7 minutes quotidiennes. »* (Suchaut, Bougnères & Bouguen, 2014).

La figure 5 offre une vision directe de la décomposition du temps scolaire et du faible volume horaire où l'élève s'entraine effectivement au décodage. Il est important de préciser toutefois, comme le font les auteurs, que ce graphique ne prend pas en compte les autres enseignements qui peuvent avoir un impact sur l'enseignement du code et son entrainement.

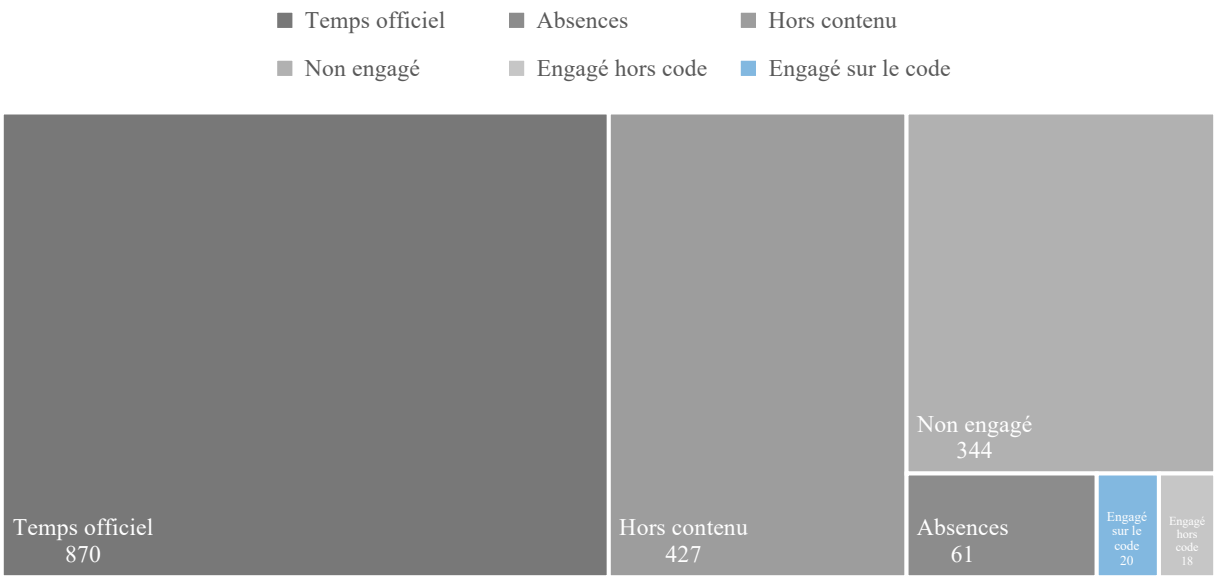


Figure 5 : Décomposition du temps scolaire au CP, adapté de Suchaut, Bougnères et Bouguen, 2014

Des solutions sont envisagées pour diminuer le nombre d'élève en difficultés, mais elles dépendent de la répartition du temps d'enseignement et à sa répartition parmi les élèves qui ont chacun des besoins spécifiques. À l'issue de leur étude, Suchaut et collaborateurs ont relevé une piste pertinente pour aider les apprentissages : la diminution des effectifs par classe, notamment dans les secteurs socialement défavorisés.

Dans le cadre donné par le Ministère de l'Éducation Nationale pour la priorité donnée à l'école primaire afin que 100% des élèves maitrisent les enseignements fondamentaux à l'entrée au collège (lire, écrire, compter et respecter autrui), le dédoublement des classes de CP en REP+ a commencé durant l'année scolaire 2017-2018. Elle s'est poursuivie pour le Cours Élémentaire 1 (CE1) dans ces REP + et pour les CP en REP durant l'année scolaire suivante. La réduction de la taille des classes a eu un fort impact positif sur les apprentissages des élèves. L'évaluation de la mesure a mis en évidence de meilleurs résultats en fin de CP pour les élèves issus de classes dédoublées par rapport à ceux d'élèves n'ayant pas étudié dans des classes réduites et la baisse de presque huit pourcents d'élève en très grande difficulté en français. En parallèle, les professeurs de CP dédoublées décrivent des classes « mieux disposées aux apprentissages scolaires [...] et plus favorables aux pratiques pédagogiques orientées vers la différenciation » et 98,5% d'entre eux reconnaissent une meilleure identification des besoins des élèves (Dédoublement des classes de CP en éducation prioritaire renforcée : première évaluation, 2020).

1.3. Interventions

Depuis des années 70, les outils numériques ont été développés pour faciliter les apprentissages (Ball, 1978 ; De Cara & Plaza, 2010 ; Dynarski *et al.*, 2007). Les propriétés mises en avant par ces outils sont, par exemple, la haute définition graphique, la qualité des stimulations auditives, le retour immédiat sur les performances, l'apprentissage par exploration et interaction active et l'aspect ludique qui permet de maintenir l'attention et la motivation des élèves (Mioduser, Tur-Kaspa & Leitner 2000). Ces outils permettent aussi de recueillir précisément les données de chaque utilisateur (le temps de réponse, les erreurs) afin de suivre de près leur évolution et adapter le contenu et la difficulté à la progression des utilisateurs.

Bien que les premières études sur les jeux numériques aient été conduites dans les années 70 aux États-Unis (Ball 1978), celles présentant une évaluation de l'efficacité de ces dispositifs sont apparues plus tardivement (De Aguilera & Mendiz Noguero 2003). En effet, peu d'outils ont été évalués systématiquement dans des études randomisées et contrôlées, sans doute parce que cette méthodologie est assez lourde à mettre en place (Girard, Écalle & Magnan 2013 ; Lassault & Ziegler, 2018).

En France, l'un des premiers logiciels d'entrainement à la lecture - *Play-On* - a été développé par Danon-Boileau et Barbier (2001) (voir figure 6). Ce logiciel porte essentiellement sur la discrimination auditive et visuelle de paires minimales du type /ba/ *vs* /pa/ (voisement). L'écran présente un terrain de basket avec deux paniers (un pour chaque son ; p. ex. /ba/ et /pa/). Un ballon est lancé au milieu de l'écran : l'enfant entend un son (p. ex. /ba/) et doit, à ce moment-là, envoyer le ballon dans le bon panier. Un second jeu reprend la même disposition, mais à chaque panier correspond soit «Oui» soit «Non» ; au jeté de ballon, l'enfant entend un mot suivi d'une syllabe, il doit envoyer le ballon vers le panier « Oui » lorsque la syllabe reprend la fin du mot (*bateau* - *to*) et vers le panier « Non » lorsque ce n'est pas le cas (*jardin* - *don*). Ce jeu a été testé auprès de 14 enfants dyslexiques âgés de 8 à 12 ans (Magnan, Écalle, Veuillet & Collet, 2004). L'entrainement a duré 10 heures réparties sur 5 semaines, à raison de 2 séances par jour de 15 minutes, 4 jours par semaine. Les enfants ont été partagés en 2 groupes de façon aléatoire, l'un bénéficiant de l'entrainement, l'autre non. À l'issue des 5 semaines, les groupes ont été inversés. Les résultats ont mis en évidence un bénéfice significatif de l'entrainement sur la tâche d'identification de mots « Timé 2 » (Écalle 2003).

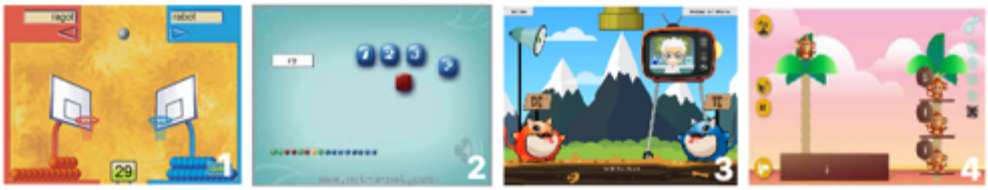


Figure 6 : Captures d'écran de logiciels d'entrainement à la lecture ; Play-On (1), Chassymo (2), RapDys® (3) et ELAN (4)

Un second logiciel d'entrainement de la conscience phonologique *Chassymo* a fait l'objet d'une étude sur 16 enfants faibles lecteurs en classe de CP répartis en 2 groupes (Écalle, Magnan & Jabouley, 2010). Dans ce jeu, l'enfant entend et voit une syllabe *(gi)*, puis, il entend un mot contenant cette syllabe *(gilet)*, il doit alors retrouver le bon emplacement de la syllabe dans le mot. L'entrainement a duré une dizaine d'heures à raison de 2 séances de 15 minutes par jour, 4 jours par semaine pendant 5 semaines. Les résultats ont permis de mettre en évidence un bénéfice de l'entrainement sur les performances en décodage et en reconnaissance de mots écrits des élèves du groupe ayant suivi l'intervention.

Un troisième outil d'entrainement de la perception des phonèmes a été évalué (Collet *et al.*, 2012). La méthode RapDys© vise à apprendre à l'enfant à discriminer et à identifier les différences acoustiques qui existe entre les phonèmes. Dans la première tâche, l'enfant doit dire si les deux phonèmes entendus sont identiques ou différents (tâche de discrimination). Dans la seconde tâche, l'enfant doit déterminer à quel graphème correspond le phonème qu'il vient d'entendre (/t/ ou /d/). La distance acoustique entre deux phonèmes est progressivement réduite en fonction des performances de l'enfant. Les résultats montrent que la méthode a des effets bénéfiques sur la lecture et tout particulièrement sur la compétence entrainée la conscience phonologique. Malgré des résultats prometteurs, ces études françaises ont été réalisées sur de petites populations.

Une étude récente a testé l'efficacité du jeu adaptatif ELAN (Watkins, Caporal, Merville, Kouider & Dehaene, 2020) sur 975 élèves français en classe de CP. Le logiciel est composé de plusieurs « leçons » dans lesquelles une correspondance graphème-phonème est introduite en trois étapes qui sont couramment utilisées par les enseignants. A la première étape, on trouve une vidéo d'un enfant prononçant le phonème étudié. A la deuxième étape, il est présenté une image dont le mot commence par le phonème travaillé. La troisième étape est une tâche haptique où l'enfant doit tracer ce phonème avec son doigt sur l'écran (en majuscule et en minuscule).

D'autres applications pour aider l'apprentissage de la lecture ont été développées. Certaines ciblent leur entrainement sur les habilités phonologiques, la connaissance des lettres ou le traitement syllabique (pour une revue, voir Écalle *et* al., 2016), d'autres ciblées sur l'attention visuelle (voir Valdois, 2017).

Les tablettes et les ordinateurs offrent des possibilités d'apprentissage, mais leur potentiel n'est pas plus grand que la qualité des logiciels qu'ils proposent (Watkins, Caporal, Merville, Kouider & Dehaene, 2020). L'une des premières études américaines qui a évalué systématiquement cinq logiciels de lecture en première année d'école primaire (2600 élèves) n'a constaté aucun effet significatif sur les résultats en lecture (Dynarski *et al.*, 2007). Une étude longitudinale (Campuzano *et al.*, 2009) a même montré que l'effet était négatif pour les élèves ayant de moins bons résultats. La méta-

analyse de 14 études (Slavin, Lake, Davis & Madden, 2011) n'a mis en évidence que des effets très faibles des programmes de technologie éducative pour les lecteurs du primaire en difficulté. L'une des raisons pourrait être liée à la motivation des élèves. En effet, dans une autre méta-analyse de 31 études, Wouters, van Nimwegen, van Oostendorp et van der Spek (2013) ont constaté que les jeux sérieux n'étaient pas plus motivants que les méthodes d'enseignement classiques.

2. GraphoGame

2.1. Présentation

Un logiciel d'aide à l'apprentissage de la lecture qui a été beaucoup développé, implémenté et testé ces dernières années est GraphoGame (Richardson & Lyytinen, 2014). GraphoGame est un programme informatisé créé par des chercheurs de l'Université de Jyväskylä en Finlande. Le logiciel a initialement été conçu pour les élèves présentant des troubles de l'apprentissage de la lecture ou étant à risque de développer une dyslexie, mais son potentiel en tant qu'outil d'aide à l'acquisition de la lecture et d'intervention précoce a été rapidement reconnu (Lyytinen, Erskine, Kujala, Ojanen, & Richardson, 2009). Au cours des dernières années, GraphoGame a été mis en œuvre et testé dans de nombreuses langues (voir tableau 2, pour une méta-analyse voir McTigue, Solheim, Zimmer, & Uppstad, 2020).

Version	Pays	Validation	Application
German	Autriche		
Portuguese	Brésil		
Canadian English	Canada		
Chile	Chili	+	
Pinyin	Chine	+	+
Mandarin Chinese	Chine		
Greek	Chypre		
Eesti	Estonie		
US English Rime	États-Unis	+	+
US Spanish	États-Unis	+	+
Ekapeli	Finlande	+	+
French	France	+	+
Greek	Grèce	+	
UK Phoneme	Hong-Kong		
Hungarian	Hongrie		
Els	Inde	+	
Bahasa	Indonésie	+	
Irish English	Irlande		
Arabic	Israël		
Kikuyu	Kenya	+	
Swahili	Kenya & Tanzanie	+	
Afrikaans	Namibie		
Dutch	Netherlands	+	+
Norwegian	Norvège		
Spanish Peru	Pérou		
Polish	Pologne	+	
Portuguese	Portugal	+	
UK Rime	Royaume-Uni	+	+
Russian	Russie		
Swedish	Suède	+	
German	Suisse	+	
Zhuyin	Taïwan		
Tonga	Zambie		
ChiNyanja	Zambie	+	

Tableau 2 : Versions de GraphoGame dans différents pays. (Les versions accompagnées du symbole « + » ont fait l'objet d'une validation scientifique / l'application numérique est disponible)

Le principe de GraphoGame (GG) est conforme aux théories de l'apprentissage de la lecture qui font consensus. La première étape consiste à apprendre les correspondances qui existent entre les graphèmes et leur unité équivalentes de la langue parlée, les phonèmes. C'est un moyen simple et parcimonieux pour retrouver la forme parlée, et donc le sens, de milliers de mots que les enfants ont stockés dans leur lexique phonologique avant la lecture (Share, 1995). En accord avec ces théories, l'idée clé de GG est d'introduire systématiquement ces correspondances, et, plus important encore, de trouver un moyen de les entrainer intensivement et de les automatiser. Le logiciel présente, de manière simultanée, des graphèmes et leurs phonèmes associés. Grâce au numérique les stimuli visuels et sonores sont d'excellente qualité. La tâche de l'enfant va être d'écouter le phonème proposé (ou mot, ou phrase) et de trouver le graphème (ou mot) correctement orthographié qui correspond. C'est cette présentation concomitante qui va permettre d'entrainer les capacités de décodage.

La version finlandaise princeps était un jeu de lettres-sons qui a été testée auprès de lecteurs débutants à risque de développer des difficultés de lecture (Saine, Lerkkanen, Ahonen, Tolvanen, & Lyytinen, 2011). Les auteurs ont d'abord examiné 166 enfants et en ont ensuite sélectionné 50 (les 30 % les moins performants) pour l'étude d'intervention. Tous les enfants ont reçu une formation régulière de rattrapage basée sur la phonétique, étalée sur une période de 28 semaines, avec 4 séances hebdomadaires de 45 minutes. Chaque session comprenait des activités de pré lecture, de segmentation des mots, de décodage et d'orthographe et d'entraînement au vocabulaire. La moitié des enfants ont suivi cet entrainement classique (groupe « Remedial reading intervention », RRI). Pour l'autre moitié des enfants (sélectionnés aléatoirement), les activités de pré lecture ont été remplacées par une session de GraphoGame de 15 minutes (groupe « Computer-assisted remedial reading intervention », CARRI). Les résultats ont montré que les enfants du groupe CARRI ont fait des progrès pendant la première année et ont continué à progresser de la même manière lors des suivis effectués à 12 et 16 mois après la fin de l'intervention. Les enfants RRI ont également progressé, mais moins que ceux du groupe assisté par ordinateur. Les gains ont été observés en connaissance des lettres, en décodage, en précision, et également en fluidité et en orthographe.

Alors que la version originale finlandaise pourrait suffire dans un système d'écriture hautement transparent où littéralement tous les mots peuvent être correctement décodés sur la base de 20 correspondances lettres-sons (Landerl *et al.*, 2013 ; Ziegler *et al.*, 2010), la version française de GG tient compte des spécificités et des caractéristiques principales de l'orthographe du français. Elle introduit et entraîne ces correspondances à différentes tailles de grains, y compris les graphèmes, les syllabes, les rimes, les mots entiers et même les phrases (Kyle *et al.*, 2013 ; Lassault & Ziegler, 2018).

Les types d'animation sont identiques dans toutes les versions de GraphoGame, il en existe dix : le sous-marin, la nuit étoilée, le jeu de télévision, la

grenouille, la balle, le pirate, le fantôme, le chercheur de diamant, les ballons et un fond simple pour les jeux de mots à trous, formation de mots et formation de phrases. Le type d'exercice proposé peut varier d'une version à une autre mais la construction de base reste la même. En ce qui concerne la version française il existe cinq types de jeux (voir figure 7) :

Classique : il y a un son (phonème ou mot) cible et des distracteurs. Les distracteurs peuvent être définis et choisi en fonction de la compétence à travailler, par exemple « peau » et « beau » ; ou proposés aléatoirement comme les mots « cirque », « crique » et « risque » qui sont soit cibles soit distracteurs les uns des autres.

Multi instances : le son cible est à trouver trois fois parmi six ou neuf propositions. Ce type de jeu est souvent utilisé pour les mots courts ou courants, comme le pronom personnel « le », qui peuvent être facilement reconnus. La composante visuo-spatiale augmente la difficulté.

Mots à trous : les graphèmes qui composent le mot qui ne sont pas travaillés spécifiquement lors d'une séquence sont fixés (l'enfant n'a pas à les retrouver) pour favoriser l'attention sur la CGP travaillée. Dans ce type d'exercice il est possible d'ajouter un ou plusieurs distracteurs et augmenter ainsi la difficulté. Par exemple le mot « sapin », seul le 's' est à placer, 'apin' est déjà fixé, avec 'z' comme distracteur, si on travaille le son /S/ porté par la lettre 's' en début de mot.

Formation de mots : à la manière des mots à trous, l'enfant doit placer les graphèmes dans le bon ordre pour former le mot entendu. En revanche ici tous les graphèmes composant le mot sont à placer, rien n'est fixé. L'ajout de distracteurs augmente ici aussi la difficulté de l'exercice.

Formation de phrases : la tâche ici est d'ordonner les mots pour former la phrase entendue. Dans les premiers niveaux, les phrases étant composées de deux mots il n'y a pas de majuscules pour le premier mot et de points à la fin. Ils sont néanmoins présents quand la phrase à construire contient plus de mots. Encore une fois il est ici également possible d'ajouter des distracteurs qui augmentent la difficulté.

Figure 7 : Les cinq types d'exercice de la version française de GraphoGame

2.2. Adaptation française

Le matériel français a été conçu grâce à la base de données Manulex-MorphO (Peerman, Sprenger-Charolles & Messaoud-Galusi, 2013), issue de Manulex (Lété, Sprenger-Charolles & Colé, 2004), corpus qui recense environ deux millions de formes non lemmatisées contenues dans cinquante-quatre manuels scolaires du primaire. La nouveauté de Manulex-MorphO a été de prendre en compte la morphologie (marques de flexions et de dérivations) et de retenir les formes lexicales les plus fréquentes de Manulex, ce qui représente vingt pourcents des entrées lexicales mais quatre-vingt-dix-huit pourcents des deux millions d'occurrences contenues dans les manuels examinés. Manulex-MorphO recense de nombreux indices dont deux particulièrement intéressants pour nous ici : la fréquence des mots et la consistance des correspondances graphophonologiques (CGP) (l'indice de consistance des correspondances phonographologiques est également disponible). La consistance des CGP a été calculée via toutes les formes possibles présentes dans le corpus (par exemple « parle », et également « parler » et « parles »). La fréquence textuelle (rapportée pour un million de mot) est donnée par deux indices « F » et « U », le dernier prend en compte la fréquence de mots dans les manuels scolaires mais aussi le nombre de manuels différents dans lesquels ils sont rencontrés. C'est ce dernier indice qui a été utilisé pour la création du contenu Français. La progression choisie repose sur les recommandations proposées par Sprenger-Charolles (2017) (voir tableau 3). La progression alors « optimale » au sein de GraphoGame se découpe en plusieurs modules et séquences de difficulté croissante. Les CGP les plus fréquentes et les plus régulières sont présentées dès le début, les CGP contextuelles, les graphèmes rarement rencontrés, les marques de flexion et les supports de dérivation portés par les lettres muettes en fin de mots sont introduits au milieu et à la fin du jeu. Toutefois, il est important d'introduire relativement tôt les mots constitués de graphèmes fréquents (mots fonction, noms, verbes) qui sont composés de CGP peu consistantes (« avec », « cet », « dix », « il est » *etc.*) dans la mesure où ils sont nécessaires à la construction de phrases simples (Sprenger-Charolles, 2017).

	Voyelles			
	CGP	Exemple	Fréquence	Consistance
Voyelles orales	a: /A/	Bal	143.280	99.9
	ou : /u/	Fou	50.04	96.8
	o : /O/	Bol	49.707	99.9
	é : /E/	Été	38.875	100.0
	eu : /Œ/	Deux	23.315	97.6
Voyelles nasales	on : /õ/	Oncle	33.125	92.2
	an : /ã/	Tante	29.147	99.9
	un : /ẽ/	Un	19.074	100.0
	in : /ɛ̃/	Fin	8.310	100.0
Allographes des voyelles	au : /O/	Auto	17.003	100.0
	eau : /O/	Eau	6.102	100.0
	ai : /E/	Faire	22.816	98.0
	et : /E/	Effet	20.290	98.0
	è : /E/	Lève	11.773	100.0
	ê : /E/	Arrêt	5.932	100.0
Semi voyelle	oi : /wa/	Loi	20.725	100.0
Cas particuliers	e : /ə/	Le (fin de mot)	218.004	96.4
	e : /#/	Vie (fin de mot)	8.257	3.6
	i : /i/	Lire	97.911	79.2
	u : /y/	Lune	51.248	81.9

	Consonnes			
	CGP	Exemple	Fréquence	Consistance
Consonnes	r : /R/	Rire	171.890	99.7
	l : /l/	Lilas	160.533	99.7
	p : /p/	Pipe	81.615	94.6
	m : /m/	Maman	58.937	99.9
	v : /v/	Vivre	52.447	100.0
	n : /n/	Nominal	45.517	100.0
	qu : /K/	Quelque	29.668	99.9
	b : /b/	Barbe	28.564	98.3
	f : /f/	Faufiler	27.396	99.2
	ch : /ʃ/	Cherche	18.614	99.2
	j : /ʒ/	Jour	16.507	99.9
	d : /d/	Dinde	100.406	93.0
Cas particuliers	s : /s/	Sel (début de mot)	47.867	100.0
	ss : /ss/	Passé	11.826	100.0
	s : /s/	Toute position	67.552	32.4
	s : /z/	Rose	14.408	6.9
	s : /#/	Amis (fin de mot)	126.443	60.7
	t : /t/	Tarte	99.257	71.0
	t : /#/	Droit (fin de mot)	37.784	27.0
	t : /s/	Addition	2.783	2.0

Tableau 3 : Exemples de correspondances graphophonologiques fréquentes et consistantes (adapté de Sprenger-Charolles, 2017)

GraphoGame est composé de soixante-sept séquences réparties en deux modules (modules non visibles par le joueur, voir tableau 4). Le module 1 a fait l'objet d'une expérimentation au sein de classes avec des élèves en difficultés d'apprentissage de la lecture (voir Ruiz *et al.,* 2017), il s'étend des séquences 1.1 à 5.2 (voir tableau 4, séquences grisées). Les premières séquences sont composées des voyelles 'a', 'é', 'o', 'eu', 'ou', 'e' (de « je ») 'i' et 'u' qui sont toutes fréquentes avec une prononciation régulière (sauf pour 'i' et 'u' lorsqu'elles sont suivies par une voyelle, cas de figure non présenté au début). Sont introduites aussi les consonnes fricatives au début 'j', 'f', 'l', 'r' car elles sont très consistantes et leur reconnaissance est plus faciles que les consonnes occlusives car elles peuvent se prononcer par un son continu (Sprenger-Charolles, 2017).

Séquences	Compétences entrainées		
1.1	Voyelle (a-i-o-u-é-eu-ou-e) Consonne (j-f-l-r)	11	Voyelle (es-ez-er-et)
1.2	Répétition de la séquence 1	12.1	Consonne contextuelle (g - son : /g/)
1.3	Ajout de 's' en début de mot	12.2	Consonne contextuelle (g - son : /Z/)
1.4	Répétition de la séquence 1.3	13	Consonne (gn)
1.5	Mots voyelles-Consonne (il) et C-V-C (sur, seul)	14.1	Voyelle nasale (ant-and-ent)
1.6	Lettre finale muette (lit) et mots fréquents (et, elle)	14.2	Voyelles nasale (om-im-am-em)
2.1	Consonnes (p-m-d-v)	14.3	Voyelle nasale (ain-ein-aim)
2.2	Consonnes (b-n-t-ch)	15.1	Voyelle (e) devant double consonne (ell-ett-err-eff-ess)
2.3	Répétition des séquences 2.1 et 2.2	15.2	Voyelle (e) en milieu de mot (ec-es-er)
2.4	Mots Consonnes-Consonnes-Voyelles (tri, trou …)	16	Consonne (ph)
2.5	Répétion de la séquence 2.4	17.1	Voyelle (i) devant une Voyelle (son : /j/)
2.6	Bilan de la séquence 2	17.2	Voyelle (i) devant un ou deux "l" (il - ill - son : /j/ - exemple : paille)
3.1	Discriminations visuelles (u-n, ...)	18.1	Voyelle (y) entre deux Consonnes (son : /i/)
3.2	Discriminations phonémiques (t-d, …)	18.2	Voyelle (y) devant une Voyelle
3.3	Discriminations mixtes (p-b, …)	19	Exceptions. (ill - son /il/ - exemple : ville)
4.1	Voyelles nasales (an-on)	20.1	Lettre muette support de dérivation (exemple : grand-grande)
4.2	Voyelles nasales (in-un)	20.2	Consonne (s) finale, marque du pluriel
5.1	Bilan du module 1 - (partie 1)	21	Consonne (x)
5.2	Bilan du module 1 - (partie 2)	22.1	Consonne (ç, sc)
6.1	Voyelle (oi) et Voyelle nasale (oin)	22.2	Consonne (t - son : /s/)
6.2	Voyelle (ui)	23.1	Exceptions. Consonne (t) prononcée en fin de mot
6.3	Voyelle (ai)	23.2	Exceptions. Consonnes (s et c) prononcées en fin de mot
6.4	Voyelle (ei)	24	Consonne (ch - son /k/)
6.5	Voyelle (au-eau)	25	Voyelle (oeu)
6.6	Consonne muette (h)	26	Exception (puzzle, femme …)
6.7	Voyelles (é-è-ê)	27.1	Présent des verbes du premier groupe (je chante, tu joues,...)
7.1	Discriminations visuelles (u-n, ...)	27.2	Présent des verbes du premier groupe (phrases)
7.2	Discriminations phonémiques (t-d, …)	27.3	Présent du verbe être
7.3	Discriminations mixtes (p-b, …)	27.4	Pronom personnel sujet et verbe être
8	Consonne (qu-k)	27.5	Présent du verbe avoir
9.1	Consonne contextuelle (c - son : /k/)	27.6	Pronom personnel sujet et verbe avoir
9.2	Consonne contextuelle (c - son : /s/)	28	Bilan : genre, nombre et verbe
10.1	Consonne (z)		
10.2	Consonne contextuelle (s - son : /z/)		
10.3	Consonne contextuelle (s, ss - son : /s/)		

Tableau 4 : progression au sein de GraphoGame (grisé : module 1- blanc : module 2)

L'introduction simultanée de voyelles et de consonnes est essentielle pour pouvoir présenter des petits mots ou non mots (de type Consonne-Voyelle (CV) ou Voyelle-Consonne (VC)) et non des graphèmes isolés le plus tôt possible (voir tableau 5 et figure 8). Les premières séquences sont répétées rapidement afin de ne pas présenter de nouvelles CGP, même si pour la plupart elles ont été enseignées en classe, et de renforcer celles déjà présentées.
Viennent ensuite progressivement des mots fréquents, de nouvelles consonnes et de nouvelles structures syllabiques plus complexes (de type CCV) et les voyelles nasales. Dans l'étude de Ruiz et collaborateurs (2017) seul le module 1 était présenté c'est pourquoi il se termine par des séquences de « bilan ».

Séquence 1.1	Infos	Voyelles (a-i-o-u-é-eu-ou-e) / Consonnes (j-f-l-r)				Type de jeu
Niveau 1	V	a	i	é		Classique
Niveau 2	V	u	o	ou		Classique
Niveau 3	V	a	o	eu		Classique
Niveau 4	CV	fa	fi	fu	fo	Classique
Niveau 5	CV	ja	ji	ju	jo	Classique
Niveau 6	CV	la	li	lu	lo	Classique
Niveau 7	CV	ra	ri	ro	ré	Classique
Niveau 8	CV	jeu	reu	feu	leu	Classique
Niveau 9	CV	jou	fou	lou	rou	Classique
Niveau 10	CV	le	la			Multi Instance
Niveau 11	CV	le	la	fou	feu	Formation de mots
Niveau 12	Phrase	le fou	le feu	le jeu		Formation de phrases

Tableau 5 : Création du contenu de la séquence 1.1 de GraphoGame

Figure 8 : Exercices de GG tirés des niveaux 9 à 12, de la séquence 1.1

L'un des objectifs de ce travail de doctorat a été de créer du nouveau contenu francophone au logiciel GraphoGame à travers de séquences qui présenteraient l'ensemble des CGP non entrainées dans le module 1. Du nouveau contenu était également nécessaire dans la mesure où le temps d'entrainement proposé dans l'étude à grande échelle serait plus long que celui de l'expérimentation précédente (16 semaines *vs.* 5 semaines).
Au sein du second module, qui pourrait correspondre à la seconde phase de l'apprentissage de la lecture sont abordées 5 grandes spécificités de la langue française. D'abord, les graphèmes contextuels, qui se lisent différemment selon le contexte où ils se trouvent comme le 'c' de « centre » et celui de « conte » (séquence 9.1 et 9.2, voir tableau 6) ou le 's' de « sur » (déjà vu dans le module 1) ou de « festin » et celui de «ruse » (séquence 10.2 et 10.3).

Séquence 9.1	Infos	La consonne 'c' se lit /k/ devant 'a', 'o', 'u' ou devant une consonne						Type de jeu
Pré		co, ce, ci	car, cire, certe	cure, cire, cerf	contre, cintre, centre	écurie, cent, ciao	écume, censé, nocif	Classique
		cane, cime, fonce	cause, sauce, épicé	caisse, cygne, cyan	coin, cèpe, acier	coiffe, celui, ciel	colle, ceci, cil	
Niveau 0	Consigne	cou (tou)	ca (ga)	cau (sau)	so (co)	su (cu)	soin (coin)	Classique
Niveau 1	/k/	s k ieur (c)	c as qu e	k ébab (c)	re qu in (c)	c al qu e	c o qu in	Formation de mots
Niveau 2	ca	case	casse	cassé	café			Classique
Niveau 3	co / cu	cousin	coussin	cousine	cuisine			Classique
Niveau 4	c + consonne	clé	clou	cri	cru			Multi instances
Niveau 5	c+consonne	c ro* qu er*	c la* qu e*	c ra* qu e*	c riti* qu e*	c lini* qu e*	c lo* qu e*	Mots à trous
Niveau 6	c / qu	ch a* qu e* (c)	ch è* qu e* (c)	ch a* c un* (s)	ch a* c une* (s)	ch imi* qu e* (c)	c o* qu eluche* (s)	Mots à trous
Niveau 7	c / qu	claque	calque	casque	craque	croque		Classique
Niveau 8	c / qu	c l a qu e	c a l qu e	c a s qu e	c r a qu e	c r o qu e		Formation de mots
Niveau 9	c /qu	quatre	cadre	quart	carte			Classique
Niveau 10	c / qu	ris* qu e* (c)	c ro* qu is*	c omi* qu e*	li* qu ide* (c)	c o* qu eli* c ot*	c ubi* qu e*	Mots à trous
Niveau 11	c / qu	c oni* qu e*	mé* c ani* qu e*	c o* qu ine*	c li* qu er*			Mots à trous
Niveau 12	Phrase	Qui a cassé le micro ?	Carole va à l'école.	Pourquoi le placard est-il fermé à clé ?	Il porte un masque de carnaval.	Il écrit une carte à sa cousine Corinne.		Formation de phrases
Post		co, ce, ci	car, cire, certe	cure, cire, cerf	contre, cintre, centre	écurie, cent, ciao	écume, censé, nocif	Classique
		cane, cime, fonce	cause, sauce, épicé	caisse, cygne, cyan	coin, cèpe, acier	coiffe, celui, ciel	colle, ceci, cil	

Séquence 9.2	Infos	La consonne 'c' se lit /s/ devant 'e', 'i', 'y'						Type de jeu
Pre		cap, cou, ce	car, cire, cri	cure, cerf, crête	contre, cloitre, cintre	écurie, cent, canal	écume, nocif, coing	Classique
		cane, cime, côte	cause, ronce, corne	cyprès, clair, Chine	coin, cèpe, croute	coiffe, celui, croire	hasard, charité, caprice	
Niveau 0	Consigne	ce (che)	ci (zi)	cy (fy)	cé (fé)	ki (ci)	kê (cê)	Classique
Niveau 1	c / s	gla* c e* (s)	pou* c e* (s)	o* c éan* (s)	voi* c i* (s)	a* c ide* (s)	ly* c ée* (s)	Mots à trous
Niveau 2	c / s	s oi* (c)	s ouk* (c)	s on* (c)	s oin* (c)	i* c i* (s)	pu* c e* (s)	Mots à trous
Niveau 3		crique	cirque	signe	risque			Multi instances
Niveau 4	s ou c	s anté* (c)	s oldat* (c)	s apin* (c)	s umo* (c)	ré* c ent* (s)	ri* c in* (s)	Mots à trous
Niveau 5	s / c / ss	c a s e	s au c e	c e r i s e	s e c ou ss e	c au s e	s ou c i	Formation de mots
Niveau 6	multi	cause	sauce	case	casse			Classique
Niveau 7		p r in c e	r e c e tt e	o c é an	c e r i s e	s au c i ss e		Formation de mots
Niveau 8		prin* c e* (s, ss)	re* c ette* (s,ss)	o* c éan* (ss, s)	c eri* s e* (ss)	s au* c isse* (s)	c ui* s ine* (ss)	Mots à trous
Niveau 9	cy / ci	li* qu ide* (c,	c ygne* (k, ss)	c ylindre* (k, ss)	c ycle* (k, ss)	k yrielle* (c, h)	k yste* (c, h)	Mots à trous
Niveau S		Il fait un signe de la main .	Il y a un cygne sur le lac .	Elle raconte un conte .	Il compte sur ses doigts .	Il suce son pouce .	La cerise pousse sur le cerisier .	Formation de phrases
Niveau S2	Distracteurs homophones	Il* fait* un* signe de* la* main* .*	Il* y* a* un* cygne sur* le* lac* .* (signe)	Elle* raconte* un* conte .* (compte)	Il* compte sur* ses* doigts* .* (conte)	Il* suce* son* pouce .* (pousse)	La* cerise* pousse sur* le* cerisier* .* (pouce)	Formation de phrases
Post		cap, cou, ce	car, cire, cri	cure, cerf, crête	contre, cloitre, cintre	écurie, cent, canal	écume, nocif, coing	Classique
		cane, cime, côte	cause, ronce, corne	cyprès, clair, Chine	coin, cèpe, croute	coiffe, celui, croire	hasard, charité, caprice	

Tableau 6 : Création du contenu des séquences 9.1 et 9.2 de GraphoGame
(note : les graphèmes entre parenthèses constituent les distracteurs définis ; les graphèmes suivis d'une astérisque sont fixés dans les exercices mots à trous)

On aborde également les graphèmes qui ont des allographes, où plusieurs orthographes produisent le même son comme 'au' et 'eau' (séquence 6.5, voir tableau 7) ou 'ant', 'ent' et 'and' (séquence 14.1).

Séquence 6.5	Infos	Allographes 'au' et 'eau' du son /o/							Type de jeu
Pre		b* eau coup* (edu, an, ou)	anim* au x* (av, ou, an)	ch* au de* (an, av, ou)	ois* eau (edu, an, ou)	f* au x* (av, an, ou)	vél* o (a, on, ou)		Classique
		j* au ne* (ou, av, an)	tabl* eau (ean, an, eu)	g* au che* (ou, av, an)	mot* o (on, ou, a)	b* eau (ean, an, eu)	cad* eau (edu, ou, an)		
Niveau 1	au / eau	au (ai)	eau (eu)	au (ai)	eau (eu)				Classique
Niveau 2	au	autre	autour	aube	autruche				Classique
Niveau 3	au	f au x	m au v ai s	s au t e	s au f				Formation de mots
Niveau 4	au	s* au terelle* (av, an, ou)	crap* au d* (av, an, ou)	s* au mon* (av, an, ou)	dinos* au re* (av, an, ou)	ch* au ssure* (av, an, ou)	d* au phin* (av, an, ou)	ch* au ssette* (av, an, ou)	Mots à trous
Niveau 5	au	j au n e	g au ch e	ch au d	ch au d e	v é l o	m o t o		Formation de mots
Niveau 6	au	j* au ne* (av, an, ou)	g* au che* (an, av, ou)	ch* au d* (ou, an, av)	ch* au de* (an, av, ou)	vél* o (a, ou,on)	mot* o (a, on, ou)		Mots à trous
Niveau 7	eau	veau	seau	beau	peau				Classique
Niveau 8	eau	p in c eau	n ou v eau	t a b l eau	b a t eau				Formation de mots
Niveau 9	eau	radeau (rateau, rideau)	rateau (radeau, rideau)	rideau (rateau, radeau)	réseau (roseau, ruisseau)	ruisseau (réseau, roseau)	roseau (réseau, ruisseau)		Classique
Niveau 10	eau	chapeau (chapiteau,	chapiteau (château,	château (chapeau, chapiteau)	bouleau (bureau, bateau)	bureau (bateau, bouleau)	bateau (bouleau, bureau)		Classique
Niveau 11	eau	ch a m eau	t r ou p eau	c o r b eau	m oi n eau	t au r eau	a gn eau	oi s eau	Formation de mots
Niveau 12	eau	cham* eau (ean, edu, eu)	troup* eau (ean, edu, eu)	corb* eau (ean, edu, eu)	moin* eau (ean, edu, eu)	taur* eau (ean, edu, eu)	agn* eau (ean, edu, eu)	ois* eau (ean, edu, eu)	Mots à trous
Niveau 13	au / eau	ch au m i è r e	ch au d r on	c a d eau	g â t eau	m o r c eau	c ou t eau		Formation de mots
Niveau 14	au / eau	ch* au mière* (av, an, ou)	ch* au dron* (av, an, ou)	cad* eau (eu, ean, edu)	gat* eau (eu, ean, edu)	morc* eau (eu, ean, edu)	cout* eau (eu, ean, edu)		Mots à trous
Niveau 15	au / eau	cis* eau x*	rat* eau	trav* au x*	mart* eau	au tomne*	b* eau coup*	cerv* eau	Mots à trous
Niveau 16	au / eau	cis* eau x* (ean, eu, edu)	rat* eau (ean, eu, edu)	trav* au x* (av, an, ou)	mart* eau (ean, eu, edu)	au tomne* (av, an, ou)	b* eau coup* (ean, eu, edu)	cerv* eau (ean, eu, edu)	Mots à trous
Niveau 17	au / eau	v au t ou r	ch e v r eau	a n i m au x*	f au c on	ch e v au x*			Formation de mots
Niveau S		Il est beau ton manteau.	Il a un nouveau chapeau.	Julie porte un joli chapeau jaune.	Elle va au bureau en moto.	Il a vu un troupeau de taureaux.	Ils sont partis sur l'eau dans un beau bateau.	Le mot 'six' se termine comme le mot 'dix'.	Formation de phrases
Post		b* eau coup* (edu, an, ou)	anim* au x* (av, ou, an)	ch* au de* (an, av, ou)	ois* eau (edu, an, ou)	f* au x* (av, an, ou)	vél* o (a, on, ou)		Classique
		j* au ne* (ou, av, an)	tabl* eau (ean, an, eu)	g* au che* (ou, av, an)	mot* o (on, ou, a)	b* eau (ean, an, eu)	cad* eau (edu, ou, an)		

Tableau 7 : Création du contenu des séquences 6.5 de GraphoGame
(note : les graphèmes entre parenthèses constituent les distracteurs définis ; les graphèmes suivis d'une astérisque sont fixés dans les exercices mots à trous)

Les lettres doubles (ou lettres géminées) sont traitées dans une séquence précise, pour la majorité ce sont des consonnes ou le 'e' qui précède se prononcent /E/ comme dans « effort » (séquence 15.1).
Les autres irrégularités de l'orthographe du français comme les mots « femme » ou « sept » sont vues plus tard dans le jeu (séquences 23.1, 26 ...).
Enfin, les marques phonologiques nominales, de genre et de nombre comme le 'e' et le 's' de « amies » (séquences 20.1 et 20.2) ; les marques phonologiques verbales comme le 's' de « tu danses » (séquences 27.1 et 27.2) sont également travaillés à la fin du jeu.
On retrouve également au sein du second module des séquences spécifiques portant sur les confusions qui existent en français : les confusions visuelles : 'u'/'n' , les confusions phonémiques : 't'/'p' et les confusions mixtes : 'b'/'d' (séquences 3.1, 3.2 et 3.3 / 7.1, 7.2 et 7.3).

Certaines caractéristiques spécifiques du français sont complexes et doivent être expliquées et explicitées par un enseignant pour faciliter leur apprentissage (c'est le cas des graphèmes contextuels par exemple). Au sein de GraphoGame de nombreuses consignes explicites et la nature variée des exercices peut être une aide pour assimiler ces notions, mais le logiciel reste un outil d'entrainement. En revanche, comme le souligne Sprenger-Charolles (2017) d'autres spécificités de la langue doivent passer par la mémorisation de ces instances (comme les lettres géminées).

Très tôt, dès la séquence 1.2, apparaissent des niveaux de « pré test » avant la séquence et de « post test » à la fin. Ces niveaux sont identiques et vont permettre d'apprécier l'évolution et le progrès des élèves à très court terme, au sein de la séquence travaillée. Dans la plupart des cas, ces niveaux contiennent uniquement des mots travaillés dans la séquence.

2.3. Outil final et améliorations

L'outil GraphoGame tel qu'il existe aujourd'hui est disponible gratuitement pour tablettes et smartphones (sur les plateformes de téléchargement des applications Google Play et Apple Store). Il est constitué des soixante-sept séquences fixes qui suivent une progression « optimale » pour aider les élèves en difficulté d'apprentissage de la lecture (Sprenger-Charolles, 2017). Chaque séquence cible une compétence particulière et est composée d'une dizaine d'exercices qui durent deux à trois minutes chacun. Chaque séquence, entourée par les niveaux de pré et post test dure alors environ trente minutes. La version qui a été testée à grande échelle n'était pas personnalisable (voir chapitre suivant) dans la mesure où il était nécessaire que chaque élève s'entraine en passant strictement par les mêmes étapes. En effet, chaque niveau de jeu devait être réussi à quatre-vingts pourcents pour passer au suivant. Sans ce pourcentage de réussite les niveaux supérieurs étaient inaccessibles. Cependant, pour ne pas que les élèves restent bloqués sur le même niveau pendant des semaines, à l'issue de cinq échecs consécutifs (c'est-à-dire moins de quatre-vingts pourcents de réussite), le niveau suivant était automatiquement débloqué. Les niveaux de pré et post tests ne sont pas répétés même si l'enfant est en dessous du seuil de réussite.

GraphoGame étant un « jeu sérieux », il a été développé dans une intention pédagogique qui est l'entrainement au décodage. Toutefois, afin de maintenir l'attention et la motivation des enfants il est important que le logiciel garde une part ludique (Mioduser, Tur-Kaspa & Leitner, 2000). Cela commence par la présence d'un environnement coloré et attractif et la création de son avatar. De plus, à chaque item réussi au sein des exercices, il est possible de récolter des pièces qui vont permettre d'aller personnaliser son avatar (coiffure, vêtements et accessoires), et de collectionner des stickers à coller dans un livre d'image virtuel. Ces « magasins » ne sont accessibles qu'au bout de quinze minutes de temps de travail effectif et ne restent ouvert que deux minutes, pour que l'objectif principal de l'outil reste l'entrainement au décodage.

La version désormais disponible est plus personnalisable dans la mesure où il est possible de choisir au début, le mode de jeu que l'on souhaite. Il y a d'une part le mode « linéaire » où les séquences se suivent comme dans la version testée dans les écoles. Ce mode est compatible avec un entrainement en classe, ou en séance de suivi avec un orthophoniste ou un neuropsychologue par exemple. D'autre part, on

trouve le mode « aventure » où l'avatar de l'enfant se déplace sur une carte et peut faire les exercices proposés qu'il souhaite sans forcément suivre une ligne directrice (voir figure 9). Notons toutefois que le contenu de ces deux modes est identique.

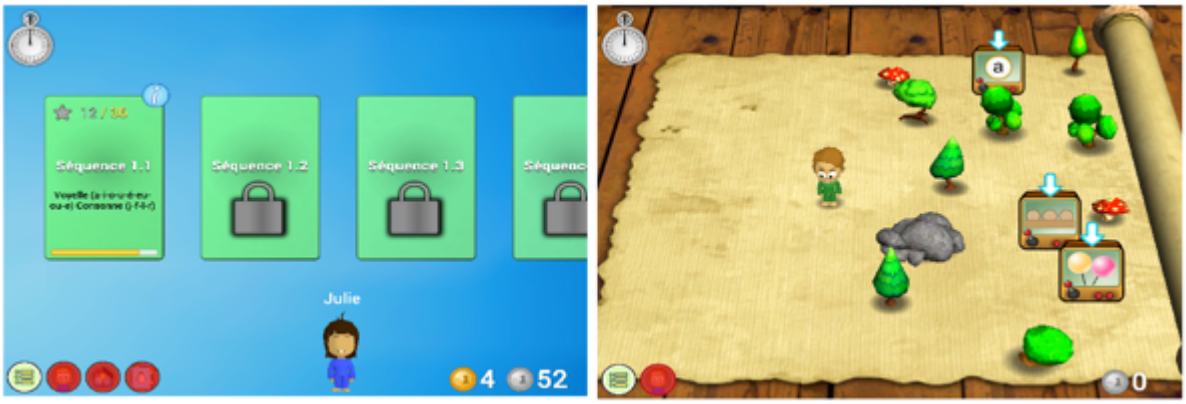


Figure 9 : Deux modes de jeu sur GraphoGame. À gauche le mode linéaire, à droite le mode aventure

L'autre nouveauté est de pouvoir choisir le point de départ en fonction de son niveau. Il y a trois points de départ possibles qui vont correspondre environ à 3 niveaux de lecture différents. Le niveau « débutant » commence à la séquence 1.1, le niveau « intermédiaire » à la séquence 1.5 et le niveau « confirmé » à la séquence 2.5.
De nouvelles améliorations sont possibles, comme celles qui ont été proposées par les enseignants qui ont travaillé dans le projet, comme le fait d'avoir une « tablette enseignant » qui pourrait permettre de suivre les progrès des élèves et de proposer du contenu adapté à chacun. Il pourrait être intéressant également de pouvoir choisir la séquence pour la présenter à ses élèves juste après l'introduction d'une nouvelle CGP et l'entrainer via un autre support. Dans sa dernière mise à jour, l'option « déverrouiller le contenu » a été activée pour justement accéder directement à la séquence que l'on souhaite faire travailler.

L'application, dans sa version actuelle, permet déjà de suivre l'évolution des enfants via une fiche récapitulative (voir figure 10) qui recense le temps de jeu total, le temps d'entrainement effectif, la date de dernière connexion, le pourcentage de réussite et le nombre de niveaux réalisés. Une « alerte » est également disponible qui propose notamment d'augmenter le temps de jeu lorsque celui-ci est trop faible.

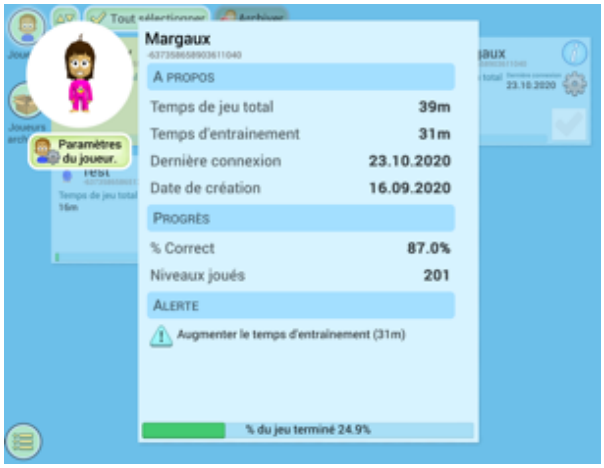


Figure 10 : Fiche de synthèse d'un joueur dans GraphoGame

Des recommandations sur l'ergonomie et les critères d'utilisabilité de l'application (position des éléments, taille, brillance *etc.* voir figure 11) ont été maquettée et testée (test utilisateur) sur vingt utilisateurs de CP dans le Travail d'Étude et de Recherche d'une étudiante en Master d'ergonomie (Master Facteurs humains et ingénierie des systèmes d'information, Aix-Marseille Université). Certaines ont été mis en place dans la version actuelle de GraphoGame.

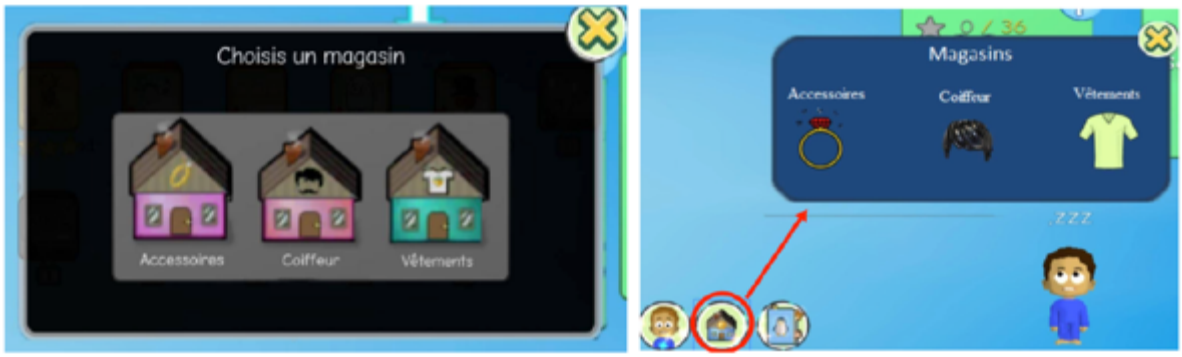


Figure 11 : Proposition d'amélioration de l'ergonomie de GraphoGame. Les petits symboles peu visibles sur les magasins sont remplacés par un « menu magasin » où ils sont seuls, plus gros et alors plus clairement identifiables

3. Expérimentations sur le terrain

3.1. Première expérimentation

Pour évaluer l'efficacité de GraphoGame Français, nous avons conduit une première étude sur des élèves de Cours Préparatoire (CP) et de Cours Élémentaire 1 (CE1) en difficulté d'apprentissage de la lecture. Cette étude a fait l'objet d'une publication [Ruiz, J.P., Lassault, J., Sprenger-Charolles, L., Richardson, U., Lyytinen, H., & Ziegler, J. C. (2017). GraphoGame : un outil numérique pour enfants en difficultés d'apprentissage de la lecture. Approche Neuropsychologique des Apprentissages chez l'Enfant (A.N.A.E.), 148, 333-343].

3.1.1. Cours Préparatoire (CP)

Au CP, 34 enfants issus de deux écoles élémentaires du département des Alpes de Haute-Provence ont participé à l'étude. Les enfants ont été sélectionnés à partir de signalements des enseignants. Les élèves ont été assignés à un groupe de façon aléatoire : le groupe 1 a commencé l'entrainement avec GraphoGame (GG) tandis que l'autre suivait le travail classique réalisé lors des Activités Pédagogiques Complémentaires (APC). Le tableau 8 résume la répartition des élèves par groupe et l'appariement des groupes en termes de nombre d'enfant, d'âge, de sexe ainsi que les résultats moyens au tests proposés en début d'expérimentation.

	Groupe 1	Groupe 2	Total
Nombre d'élèves	19	15	34
Âge (mois)	77.3	74.9	76.0
Sexe (G/F)	10/9	8/7	18/16
TVAP	39.5	43.0	41.0
PM47	22.1	22.4	22.2
Lecture (M/min)	12.4	13.9	13.0

Tableau 8 : Composition des groupes de CP et résultats au Test de Vocabulaire Actif Passif (TVAP), aux matrices progressives de Raven (PM47) et au test de lecture en une minute

La durée des entrainements était d'une heure par semaine (2 fois 30 minutes ou 3 fois 20 minutes) pour une durée de 5 semaines. À l'issu de ces 5 semaines les groupes étaient inversés. Les deux groupes d'entrainements randomisés ont été comparés afin d'éviter un biais de sélection. L'utilisation d'un groupe contrôle permet de neutraliser d'autres facteurs pouvant intervenir et influencer les performances des élèves. Le design de cette étude permet à chaque enfant d'être son « propre contrôle » (voir figure 12).

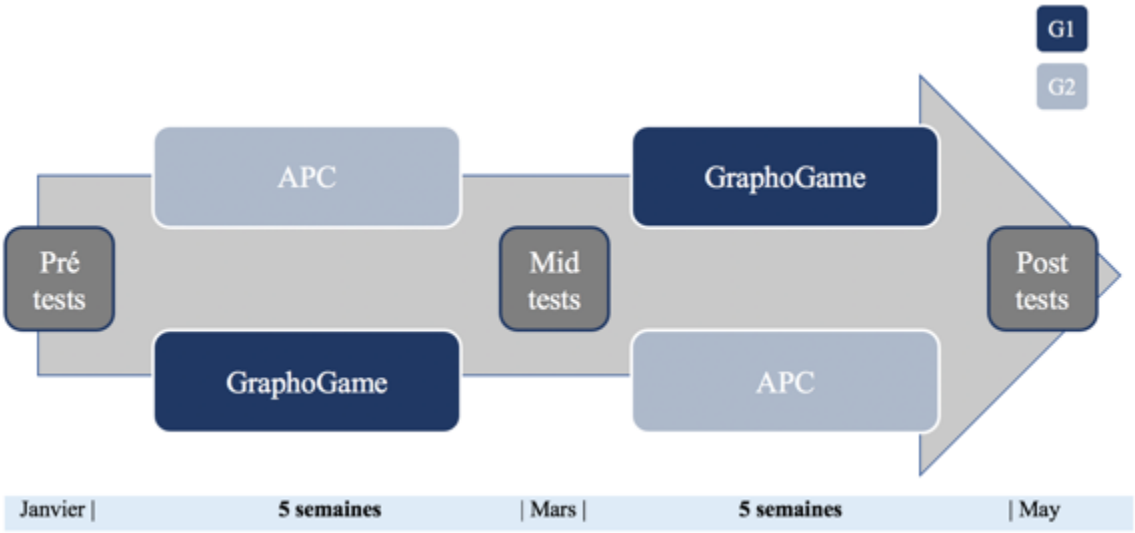


Figure 12 : Design expérimental des classes de CP

Tests d'évaluation

Les matrices progressives de Raven (PM 47, Raven, 1976). Pour évaluer le niveau cognitif non verbal des enfants, les matrices progressives de Raven couleur (qui testent le raisonnement analogique) ont été choisies. Ce test, qui est l'un des plus sensibles pour évaluer le facteur général d'intelligence (Prabhakaran, Smith, Desmond, Glover & Gabrieli, 1997), se compose de 3 séries de 12 items proposées aux enfants sans limite de temps. Cette version, destinée aux enfants de 4 à 11 ans, est basé sur le fait que le sujet doit être capable de donner un sens aux éléments de la figure afin de la compléter correctement, grâce à l'une des 6 propositions présentées (voir figure 13). Les normes ont été établies sur une population française de 1 064 enfants âgés de 4 à 11 ans et demi. Aucun indice de fiabilité (alpha de Cronbach) n'est donné pour ce test en français. Ce test est présenté uniquement au début de l'expérimentation.

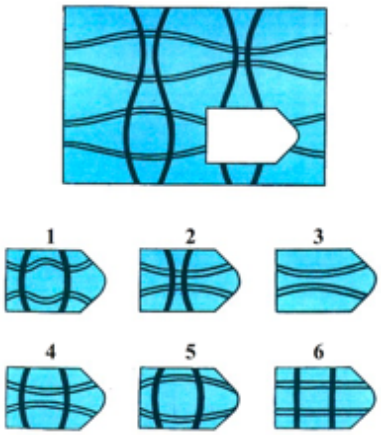


Figure 13 : Exemple de planche des PM47 (Raven, 1976)

Test de vocabulaire actif et passif (TVAP 5-8 ans, Deltour & Hupkens, 1984). Pour évaluer le niveau d'intelligence verbale (et écarter un trouble spécifique du langage oral avec retentissement sur la lecture), nous avons choisi un test de vocabulaire (TVAP 5-8 ans). Nous avons évalué uniquement le niveau de vocabulaire en réception, c'est à dire la capacité à associer un mot prononcé par l'expérimentateur avec la bonne image parmi 6 représentées (voir figure 14) : l'image correcte, une proche et quatre distracteurs. Par exemple, pour « château », l'image d'un château (choix correct, 2 points), d'une grande maison (choix proche, 1 point), d'un gâteau, d'un chat, d'une tour de contrôle et d'un roi. Aucun indice de fiabilité (alpha de Cronbach) n'est donné pour ce test. Ce test est présenté uniquement au début de l'expérimentation.

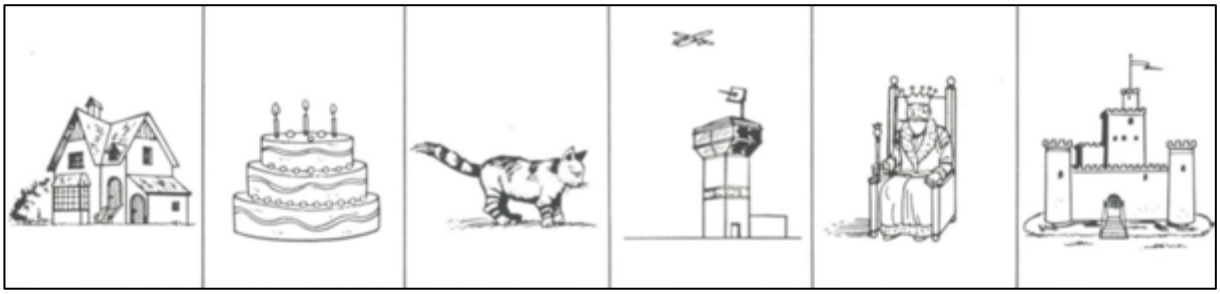


Figure 14 : Exemple de planche du TVAP (Deltour & Hupkens, 1984)

Test de lecture en 1 min (Gentaz, Sprenger-Charolles, Theurel & Colé, 2013). Le test utilisé au CP pour évaluer le niveau de lecture au cours des 10 semaines d'entraînement est un test de lecture de mots familiers et de mots inventés pour le début du CP. Il consiste à lire à haute voix le plus de mots possibles en 1 minute pour chaque catégorie (mots familiers et mots inventés, voir figure 15). Les mots familiers (35mots , 5 par ligne) sont fréquents et figurent parmi les 1 000 premiers mots des manuels de CP (Lété, Sprenger-Charolles & Colé, 2004). Ils sont courts, mono- ou bi-syllabiques, réguliers sur le plan des correspondances graphème-phonème (mais 5 d'entre eux ont un graphème dont la prononciation dépend d'une règle contextuelle, '*c*' ou '*g*') et ils comportent peu de groupes consonantiques. Les mots inventés, ou pseudo-mots, (30 mots, 5 par ligne) sont appariés aux mots fréquents en longueur, en structure syllabique et en difficulté orthographique. Ils sont tous réguliers sur le plan des correspondances graphème-phonème, mais comme pour les mots, 5 d'entre eux ont un graphème dont la prononciation dépend d'une règle contextuelle ('*c*' ou '*g*'). Le coefficient de fidélité pour ce test (*Cronbach alpha*) est de .92. Lorsque l'enfant a lu moins de quatre mots correctement sur les 10 premiers mots d'une catégorie, l'épreuve est arrêtée. Ce test a été passé à trois reprises lors des phases de pré test, à 5 semaines (mid-test) et à la fin des 10 semaines d'entraînement (post test).

Mots Familiers

à	où	la	au	tu
un	il	été	on	mur
ni	sur	qui	vélo	par
feu	ce	peur	ami	moto
peau	lune	gare	lire	bon
mardi	col	avril	roi	faire
facile	cheval	vrai	ligne	porte

Mots Inventés

o	i	bi	ul	til
ja	ol	ata	dik	nar
vaf	zou	dul	lèvo	tur
veur	co	noil	opa	jal
neau	timé	cal	tibe	lupe
pinde	gor	ablir	sar	vaipe

Figure 15 : Planches du Test de lecture en une minute (Gentaz, Sprenger-Charolles, Theurel & Colé, 2013)

Résultats

Progression au sein de GG. La progression des élèves au sein du jeu a été évaluée en comparant le pourcentage de bonnes réponses du premier niveau de chaque séquence (pré test) au pourcentage de bonnes réponses du dernier niveau (post test) de chaque séquence. Comme le montre la figure 16, les élèves ont progressé en moyenne de 4% entre le pré test et le post test de chaque séquence ($t(33) = 4.53$; $p < .0001$).

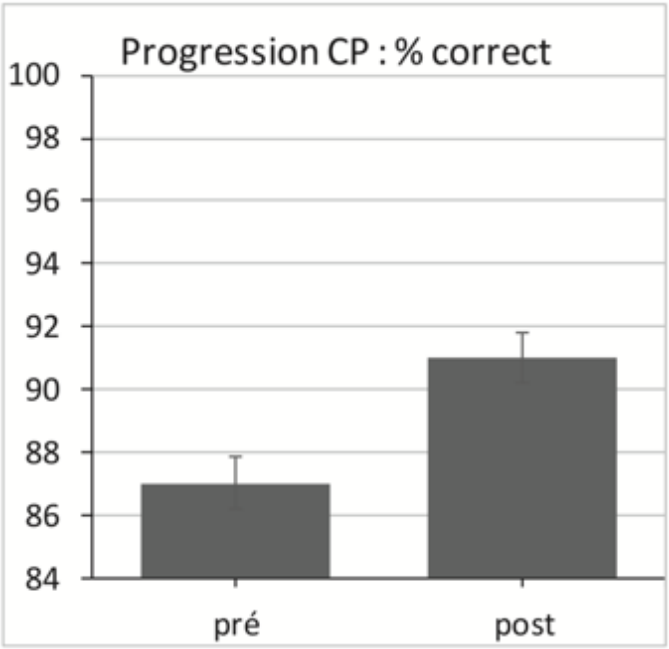


Figure 16 : Progression au sein de GG entre le premier et dernier niveau de chaque séquence. Les barres d'erreurs correspondent à des erreurs standards intra-sujets

Progression au test de lecture. Les résultats du test de lecture en 1 minute réalisé avant et après l'entraînement sont présentés dans la figure 17, pour la fluence

(nombre de mots correctement lus en 1 minute) et la précision (pourcentage de mots correctement lus). Pour la fluence, l'ANOVA montre un effet du facteur Progression ($F(1,33) = 94.4$, $p < .0001$) reflétant le fait que l'ensemble des enfants ont progressé entre le pré test et le post test. Plus important, l'interaction entre le facteur Progression et le facteur Type d'intervention est significative ($F(1,33) = 12.8$, $p < .001$) indiquant que les enfants ont davantage progressé après les séances GG qu'après un entraînement APC classique. Pour la précision, l'ANOVA montre également un effet significatif du facteur Progression ($F(1,33) = 32.4$, $p < .001$) et une interaction significative entre le facteur Progression et le facteur Type d'intervention ($F(1,33) = 8.2$, $p < .01$). Cette interaction reflète le fait que la précision de lecture était meilleure après les séances de GG (+ 13.01 %) qu'après l'APC classique (+ 2.23 %). En ce qui concerne la lecture de pseudo mots, l'ANOVA montre un effet significatif du facteur Progression pour la vitesse ($F(1,33) = 75.8$, $p < .0001$) et la précision ($F(1,33) = 24,2$, $p < .0001$) mais l'interaction entre le facteur Progression et le facteur Type d'Intervention n'est pas significative ni pour la fluence ni pour la précision ($F<1$).

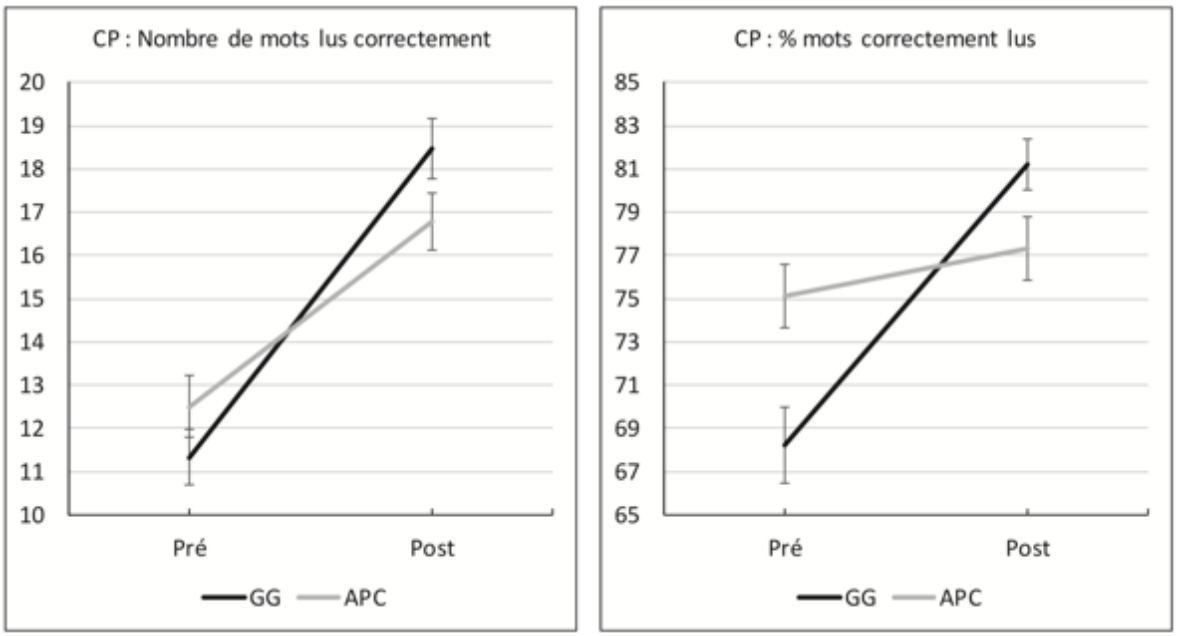


Figure 17 : Nombre de mots familiers correctement lus en 1 minute (gauche) et pourcentage de mots correctement lus (droite) aux pré-tests et post-tests du test de lecture en 1 minute. Les barres d'erreurs correspondent à des erreurs standards intra-sujets

3.1.2. Cours Élémentaire 1 (CE1)

Au CE1, 35 enfants issus de six écoles élémentaires du département des Alpes de Haute-Provence ont participé à l'étude. Les enfants ont été sélectionnés à la fin du CP. Les élèves ont été assignés à un groupe de façon aléatoire : le groupe 1 a commencé l'entrainement avec GraphoGame (GG) tandis que l'autre suivait un entrainement en mathématiques avec le logiciel Toké Maths. Le tableau 9 résume la

répartition des élèves par groupe et l'appariement des groupes en termes de nombre d'enfant, d'âge, de sexe ainsi que les résultats moyens au tests proposés en début d'expérimentation.

	Groupe 1	Groupe 2	Total
Nombre d'élèves	18	17	35
Âge (mois)	87.9	89.1	88.5
Sexe (G/F)	11/7	7/10	18/17
TVAP	46.5	44.7	45.6
PM47	24.9	25.6	25.2
Lecture (M/min)	12.1	13.4	12.8

Tableau 9 : Composition des groupes de CE1 et résultats au Test de Vocabulaire Actif Passif (TVAP), aux matrices progressives de Raven (PM47) et au test de lecture l'Alouette

La durée des entrainements était d'une heure par semaine (2 fois 30 minutes ou 3 fois 20 minutes) pour une durée de 5 semaines. Comme pour les élèves de CP, à l'issu de ces 5 semaines les groupes étaient inversés permettant également à chaque enfant d'être son « propre contrôle » (voir figure 18).

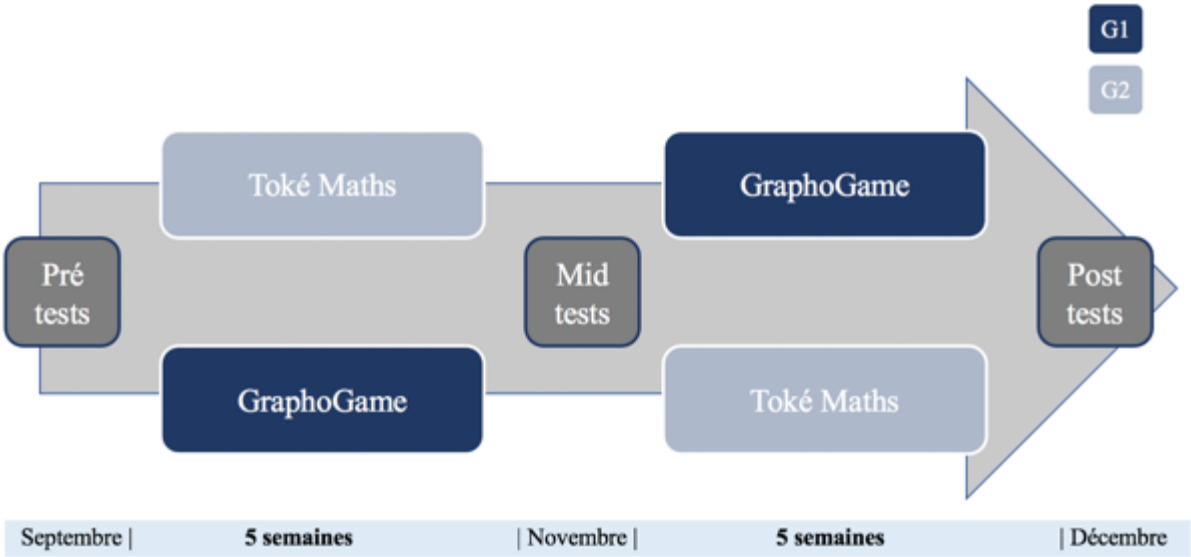
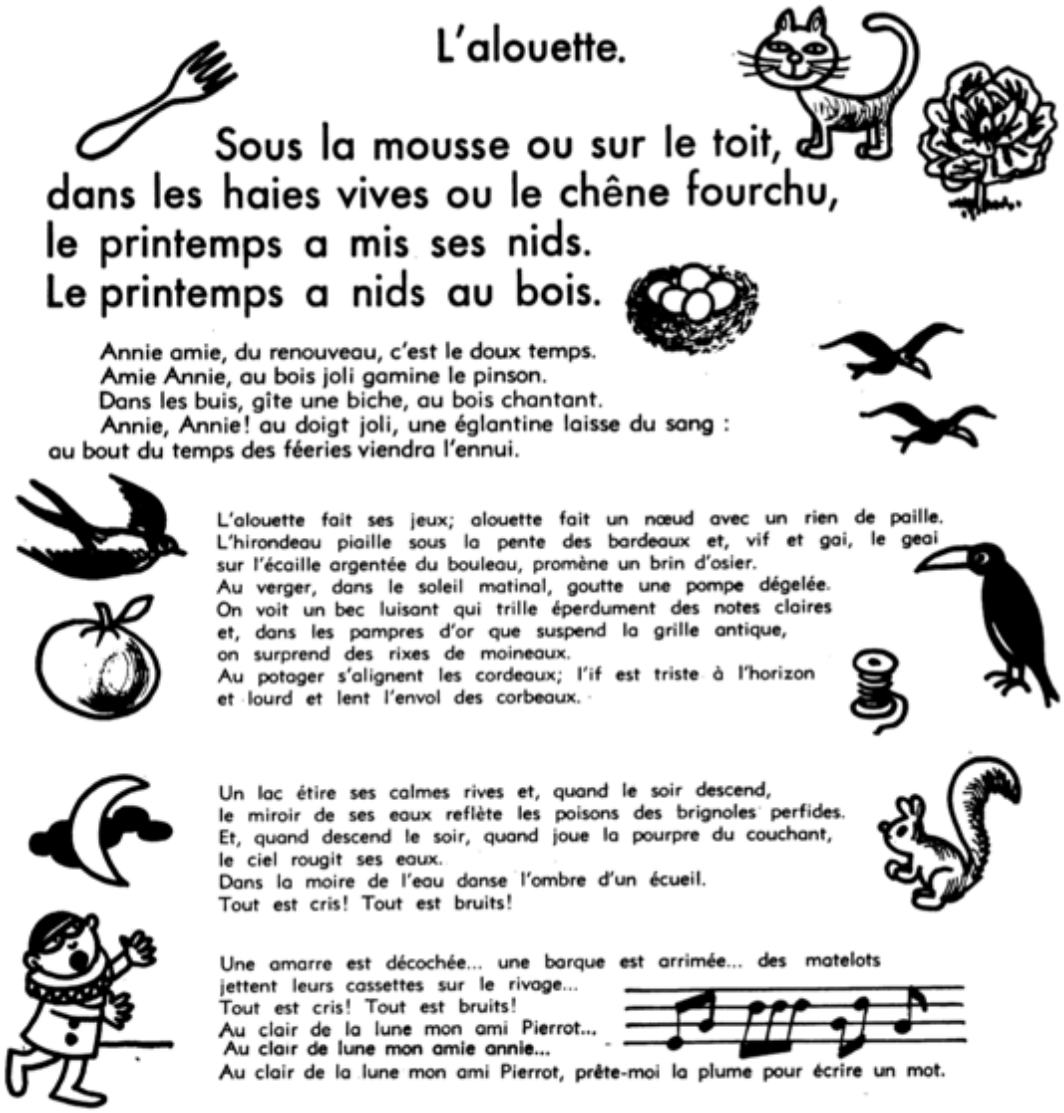


Figure 18 : Design expérimental des classes de CE1

Tests d'évaluation

Les matrices progressives de Raven (PM47) et le test de vocabulaire actif passif (TVAP) ont été proposés dans les mêmes conditions de passation que pour les élèves de CP, en pré test. Seul le test de lecture proposé à trois reprises (pré, mid et post) est différent.

L'Alouette (Lefavrais, 1965). Ce test a été utilisé pour évaluer le niveau de lecture. Il propose un texte asémantique de 265 mots avec une présence accrue de mots peu fréquents obligeant l'enfant à mettre en œuvre la procédure de décodage graphophonologique plutôt que ses connaissances lexicales (voir figure 19). Ce test est très sensible aux difficultés de lecture (Bertrand, Fluss, Billard & Ziegler, 2010 ; Cavalli *et al.*, 2017).



L'alouette.

Sous la mousse ou sur le toit,
dans les haies vives ou le chêne fourchu,
le printemps a mis ses nids.
Le printemps a nids au bois.

Annie amie, du renouveau, c'est le doux temps.
Amie Annie, au bois joli gamine le pinson.
Dans les buis, gîte une biche, au bois chantant.
Annie, Annie! au doigt joli, une églantine laisse du sang :
au bout du temps des féeries viendra l'ennui.

L'alouette fait ses jeux; alouette fait un nœud avec un rien de paille.
L'hirondeau piaille sous la pente des bardeaux et, vif et gai, le geai
sur l'écaille argentée du bouleau, promène un brin d'osier.
Au verger, dans le soleil matinal, goutte une pompe dégelée.
On voit un bec luisant qui trille éperdument des notes claires
et, dans les pampres d'or que suspend la grille antique,
on surprend des rixes de moineaux.
Au potager s'alignent les cordeaux; l'if est triste à l'horizon
et lourd et lent l'envol des corbeaux.

Un lac étire ses calmes rives et, quand le soir descend,
le miroir de ses eaux reflète les poisons des brignoles perfides.
Et, quand descend le soir, quand joue la pourpre du couchant,
le ciel rougit ses eaux.
Dans la moire de l'eau danse l'ombre d'un écueil.
Tout est cris! Tout est bruits!

Une amarre est décochée... une barque est arrimée... des matelots
jettent leurs cassettes sur le rivage...
Tout est cris! Tout est bruits!
Au clair de la lune mon ami Pierrot...
Au clair de lune mon amie annie...
Au clair de la lune mon ami Pierrot, prête-moi la plume pour écrire un mot.

Figure 19 : Planche de passation de l'Alouette (Lefavrais, 1965)

Résultats

Progression au sein de GG. La progression des élèves au sein du jeu a été évaluée en comparant le pourcentage de bonnes réponses du premier niveau de chaque séquence (pré test) au pourcentage de bonnes réponses du dernier niveau (post test) de chaque séquence. Comme le montre la figure 20, les élèves ont

progressé en moyenne de 2,75% entre le pré test et le post test de chaque séquence ($t(35) = 3.8$; $p < .001$).

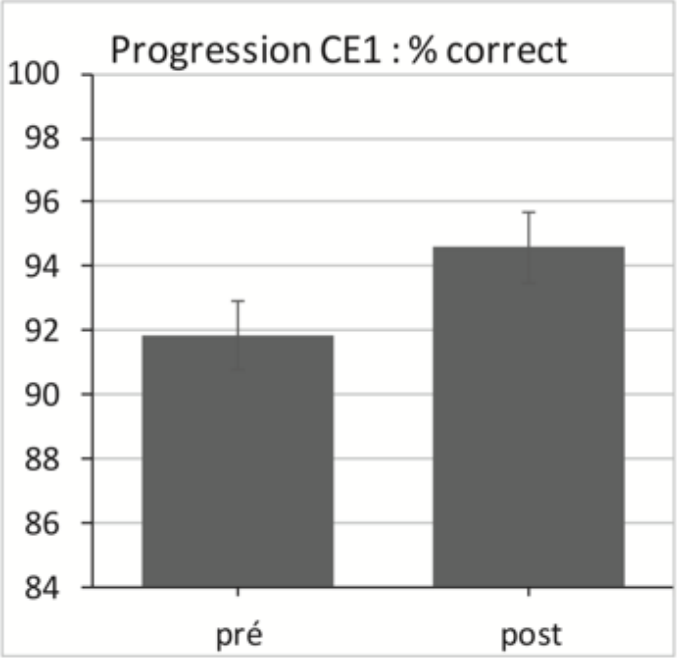


Figure 20 : progression au sein de GG entre le premier et dernier niveau de chaque séquence. Les barres d'erreurs correspondent à des erreurs standards intra-sujets

Progression au test de lecture. Les résultats du test de l'Alouette passés avant et après l'entraînement sont présentés dans la figure 21, pour la fluence (nombre de mots correctement lus en 3 minutes) et la précision (% de mots correctement lus). L'analyse de la fluence montre que l'ensemble des enfants ont progressé entre le pré test et le post test avec un effet significatif du facteur Progression ($F(1,34) = 116.2$, $p < .0001$). L'interaction entre le facteur Progression et le facteur Type d'intervention est également significative ($F(1,34) = 7.6$, $p < .01$) indiquant que les élèves ont davantage progressé après l'entraînement GG qu'après l'entraînement mathématique. Pour la précision, l'ANOVA montre un effet significatif du facteur Progression ($F(1,34) = 13.4$, $p < .001$) mais l'interaction entre le facteur Progression et le facteur Type d'intervention n'était pas significative ($F < 1$).

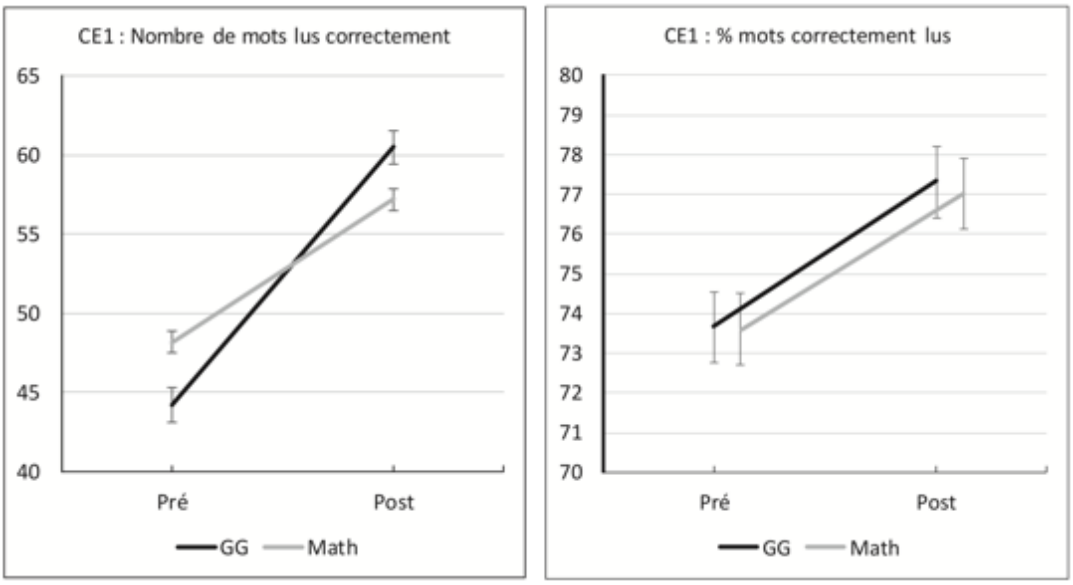


Figure 21 : Nombre de mots correctement lus en 3 minutes (gauche) et pourcentage de mots correctement lus (droite) aux pré-tests et post-tests du test de l'Alouette. Les barres d'erreurs correspondent à des erreurs standards intra-sujets

Discussion

L'objectif de cette étude consistait à tester et valider par un protocole contrôlé et randomisé l'efficacité de l'adaptation française de GG auprès d'enfants «à risques» de dyslexie scolarisés au CP et au CE1 et participant à une prise en charge complémentaire au sein de l'école. Les résultats obtenus permettent d'affirmer que GG a apporté un bénéfice spécifique pour ces élèves. En effet, pour les enfants du CP, les résultats du test de lecture en 1 minute, proposé avant et après chaque phase d'entraînement, montrent que la vitesse et la précision de lecture de mots familiers ont progressé plus après la phase GG qu'après la phase APC classique. Ces bénéfices ont été obtenus après seulement 5 semaines d'entraînement. En revanche, aucun effet spécifique de l'entraînement n'a été obtenu pour la lecture de mots inventés (pseudo-mots). Un entraînement avec GG génère donc un effet de transfert à la lecture à voix haute de mots familiers mais pas de pseudo-mots. Il est possible que le transfert à la lecture de pseudo-mots nécessite un entraînement plus intense.

Les résultats du CE1 montrent également des effets d'entraînement spécifique à GG pour la vitesse de la lecture dans le test de l'Alouette mais pas sur la précision de la lecture : les enfants ont lu significativement plus de mots après les phases d'entraînement GG qu'après les phases d'entraînement *TokéMath*. Ces effets ont été obtenus après seulement 5 semaines d'entraînement. Il convient de noter que les deux tests de lecture impliquent la lecture à voix haute tandis que GG n'entraîne pas explicitement la lecture à voix haute. Il est donc possible que les effets de transfert de l'entraînement aient été plus importants dans un test de lecture silencieuse basé sur reconnaissance visuelle des mots, tel que *Timé2* (Écalle, 2003), ou dans un test de choix orthographique (Siegel, Share & Geva, 1995).

3.2. Expérimentation LEMON

La seconde étude a été conduite sur près de 1000 élèves de CP de neuf écoles de Marseille. Cette étude a fait l'objet d'un article soumis pour publication en décembre 2020 [Lassault, J., Sprenger-Charolles, L., Albrand, J.P., Alavoine, E., Richardson, U., Lyytinen, H. & Ziegler, J.C. (2020). Positive effects of GraphoGame reading intervention in French: the importance of initial reading level and engagement. Article soumis pour publication.]. Cette étude a été précédée d'une étude pilote conduite dans trois classes de CP.

Étude Pilote

L'étude pilote a été conduite dans trois classes de CP de l'école des Abeilles (Marseille centre). Après une session de pré test évaluant le décodage, le vocabulaire, la compréhension orale, les compétences visuo-attentionnelles et des notions de calcul, les élèves de deux classes ont suivi un entrainement avec GraphoGame, alors que les élèves de la troisième classe ont suivi un entrainement en mathématique (deux applications : *Toké math* et Attrape nombre). Après 6 semaines d'entrainement (par session de 15 minutes, 4 jours par semaines), nous avons proposé une séance de post test (tests de décodage, vitesse de lecture) afin d'apprécier l'évolution de chaque élève.

L'objectif de l'étude pilote était de tester notre matériel (tests, tablettes, casques) et la logistique, avant d'étendre le dispositif à grande échelle. Cela a permis de modifier le choix de nos tests (fort effet plafond pour les tests de calcul) et surtout d'ajuster notre design expérimental. En effet, les résultats ont montré une augmentation significative des habiletés en décodage (test de lecture en 1 minute et Alouette) des élèves ayant suivi l'entrainement en mathématique. Ces résultats s'expliquent par une augmentation du temps de classe alloué au décodage et aux ateliers de lecture dans le groupe « math ». Ce design expérimental a placé l'enseignant dans une position d'évaluation de ses compétences et l'a conduit à presque doubler le temps d'enseignement de lecture dans sa classe. Afin de contourner ce biais et de ne pas instaurer un climat de concurrence, nous avons adapté le design pour l'étude à grande échelle.

Étude principale

Population

Neuf cent vingt et un enfants ont participé à cette étude à grande échelle, de deux ans. Quatre cent cinquante et un participants de 9 écoles (1 en Réseau d'éducation prioritaire (REP) ; 1 hors REP ; 7 REP+) répartis en 37 classes se sont entrainés avec le logiciel de lecture GraphoGame durant la première année. Quatre cent soixante-dix enfants se sont entrainés avec un logiciel en mathématique, Fiete

Math, au cours de la deuxième année. Ils ont été scolarisés dans les mêmes établissements que l'année précédente, répartis en 36 classes (voir tableau 10).

		Année 1	Année 2	Total
Hors REP	Abeilles	39	44	83
REP	Airbel	55	77	132
REP +	Aygalades	46	63	109
	Cadenat	47	56	103
	HLM Perrin	35	36	71
	Malpassé	56	48	104
	Maurelette	75	74	149
	Révolution	41	28	69
	St Just Corot	57	44	101
		451	470	**921**

Tableau 10 : Effectif des élèves de chaque école pour les deux années d'expérimentation

Pour chaque groupe, la durée des entrainements était d'une heure par semaine (15 minutes par jour, 4 séances par semaine), pendant 16 semaines, en demi-groupe. Une moitié de la classe s'entraînait avec le logiciel alors que l'autre travaillait avec l'enseignant, puis, après 15 minutes, les groupes s'intervertissaient. La première année tous les élèves de l'expérimentation LEMON se sont entrainés avec le logiciel GraphoGame, tandis que l'année suivante la nouvelle cohorte d'élève s'est entrainé dans le même conditions avec le logiciel de mathématique Fiete Math (voir figure 22).

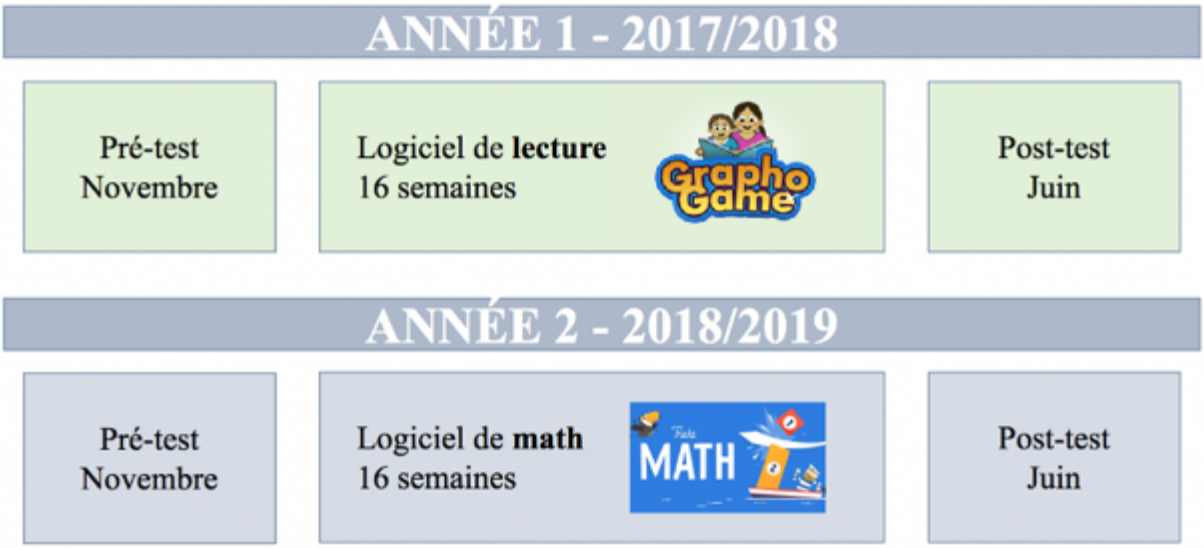


Figure 22 : Design expérimental de l'expérimentation LEMON

Le logiciel d'entraînement aux compétences en mathématiques choisi est Fiete Math (Ahoiii Entertainment). Dans ce jeu, les enfants utilisent la cargaison du marin Fiete pour calculer, grâce à des boîtes de nombres qui peuvent être assemblées ou séparées. Les tâches sont simples : les navires qui arrivent doivent être chargés de

manière appropriée (voir figure 23). Comme des blocs de construction, les enfants peuvent assembler les caisses dans n'importe quelle forme et les diviser afin d'atteindre la quantité demandée sur le chargement (qui change à chaque essai). Afin de maintenir la motivation, les enfants peuvent débloquer de nouvelles terres et de nouveaux ports, gagner des médailles et des navires plus grands. Chaque port a une compétence arithmétique à travailler : compter (comptage, reconnaitre des groupes), fractionner (fractionner les chiffres différemment, apprendre les nombres jusqu'à 20), valeur du 5 (calculer plus facilement les chiffres en base 5, passage à la dizaine, élaboration du système décimal), techniques opératoires (multiplier et diviser par 2, opérations commutatives, opérations proches, opérations additives et soustractives à termes différents), addition (suites numériques logiques, opérations simples à difficiles, inversions), soustraction (idem qu'addition) et calcul libre (s'exercer aux opérations additives et soustractives). Le jeu contient plus de 1000 exercices avec les nombres allant jusqu'à 30. Lors du positionnement de la collection sur le bateau, si on ne la lâche pas, on peut voir la zone derrière le bateau devenir rouge en cas d'erreur (avant que la collection soit rejetée) ou vert en cas de réussite donnant un feedback immédiat à l'élève.

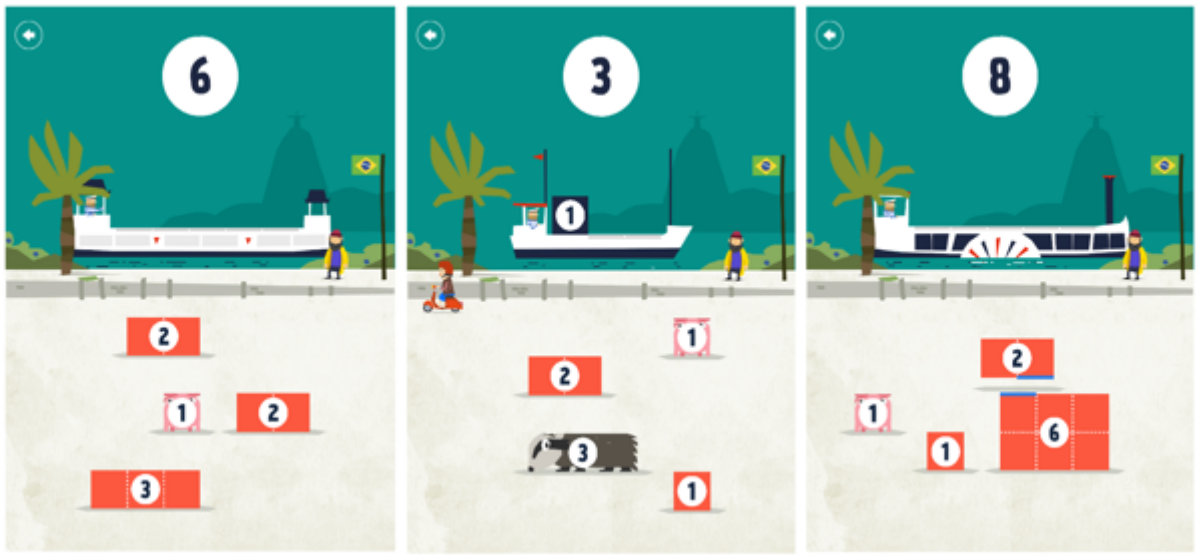


Figure 23 : Exemples d'exercices issus de Fiete Math

Procédure et évaluations

La procédure suivie a été la même pendant les deux années de l'expérience. En novembre (période 1) et en juin (période 2), chaque enfant a été testé pour connaitre ses capacités en décodage, en conscience phonologique, en compréhension orale et écrite, en identification des mots, en raisonnement non verbal et en mathématiques. Les sessions se sont déroulées en 2 étapes et ont duré environ 1 heure. Elles ont commencé par les 2 tests de décodage, puis par la conscience phonologique et le vocabulaire. Elles se sont poursuivies avec des tâches de compréhension orale, de raisonnement non verbal et de mathématiques. Le raisonnement non verbal et le vocabulaire n'ont été évalués qu'en période 1 et la

compréhension écrite qu'en période 2. L'ordre choisi ici était arbitraire, mais il a été le même pour tous. Toutes les tâches ont été administrées par des étudiants en sciences de l'éducation ou en psychologie qui ont suivi des sessions de formation et ont été accompagnés sur place. Les enfants ont été testés individuellement, sauf pour l'évaluation de l'identification des mots, qui a été faite en demie-classe.

Tests

Décodage : Deux tâches ont été évaluées : la maîtrise des mots isolés et des pseudo-mots et la maîtrise des mots en contexte. La fluidité est une mesure composite basée sur la précision et le temps (nombre d'éléments correctement lus dans un temps donné). Pour la tâche de mots isolés, le test de lecture de mots en une minute a été utilisé (Gentaz, Sprenger-Charolles, Theurel & Colé, 2013), pour la lecture de mots « en contexte », le test présenté était "L'Alouette" (Lefavrais, 1965). Pour une description de ces tests, voir la partie précédente.

Conscience phonologique : Elle a été évaluée avec 3 tâches de la batterie EVALEC (Sprenger-Charolles, Colé, Bechennec & Kipffer-Piquard, 2005) (voir figure 24). Le principe est le même dans chaque exercice, les enfants doivent supprimer le premier élément des pseudo mots dits par l'examinateur. Ils doivent prononcer le pseudo mot sans la première syllabe dans la tâche 1 (10 pseudo mots trisyllabiques structure consonne-voyelle (CV)). Ils doivent prononcer le pseudo mot sans le premier phonème dans la tâche 2 (12 pseudo mots monosyllabiques, structure CVC) et dans la tâche 3 (12 pseudo mots monosyllabiques, structure CCV).

Sup_Syllabique			*Sup_Phonémique_CVC*			*Sup_Phonémique_CCV*		
ITEM	*REPONSE*	*SCORE (0/1)*	*ITEM*	*REPONSE*	*SCORE (0/1)*	*ITEM*	*REPONSE*	*SCORE (0/1)*
ESSAI			ESSAI			ESSAI		
tonira	nira	-	zak	ak	-	fla	la	-
varéla	réla	-	chor	or	-	spo	po	-
TEST			TEST			TEST		
povidu	vidu		puf	uf		klo	lo	
tokali	kali		jor	or		pra	ra	
tipango	pango		zil	il		sri	ri	
banidé	nidé		kip	ip		tsé	sé	
zofitu	fitu		dour	our		blo	lo	
kossila	sila		bir	ir		sti	ti	
buliva	liva		tal	al		pso	so	
rétouda	touda		gof	of		flin	lin	
valoté	loté		fëk	ék		sla	la	
soguté	guté		sat	at		vri	ri	
-	-	-	chol	ol		spa	pa	
-	-	-	vaf	af		grou	rou	
Total /10			**Total /12**			**Total /12**		
Temps (s)			**Temps (s)**			**Temps (s)**		

Figure 24 : Trois tâches de conscience phonologique de la batterie EVALEC (Spenger-Charolles, Colé, Bechennec & Kipffer-Piquard, 2005)

Compréhension : la compréhension orale a été évaluée en période 1 et la compréhension écrite en période 2. La tâche de compréhension orale contenait 25 phrases de l'"ECoSSe" (Lecocq, 1996) (voir figure 25) : l'examinateur prononçait la phrase à haute voix puis montrait 4 images à l'enfant qui devait choisir celle correspondant à ce qu'il venait d'entendre. Six structures de phrases ont été utilisées (active, pronom, double négation, termes spatiaux, relative et passive). En ce qui concerne la compréhension écrite, la procédure est similaire, sauf qu'au lieu d'entendre les phrases, les élèves doivent les lire. Les mêmes structures de 6 phrases ont été utilisées pour les 14 items du test.

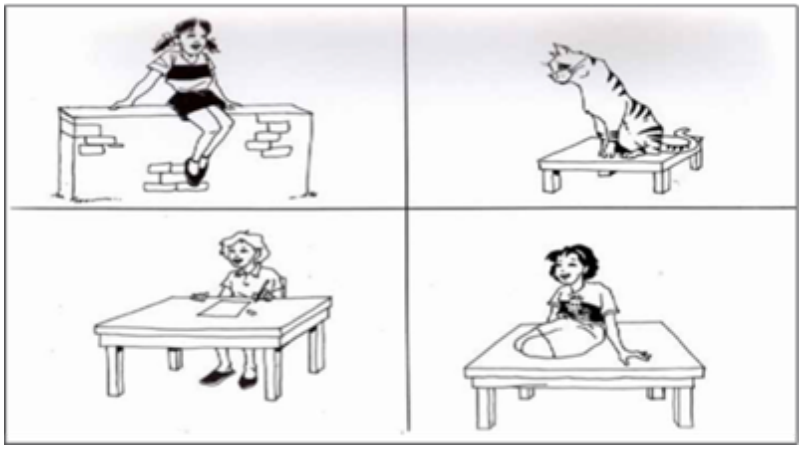


Figure 25 : Exemple de planche de compréhension orale (l'enfant entend) ou de compréhension écrite (il lit) : « La fille est assise sur la table. »

Identification des mots : Le Timé 2 (Écalle, 2003) (voir figure 26) teste l'identification de mots écrits dont le développement et le degré d'automaticité constituent une des clés de compréhension des textes. Il est divisé en 3 tâches (12 items chacune). Dans la première, l'élève doit trouver le mot cible, c'est-à-dire celui prononcé par l'examinateur parmi 5 propositions. Chaque distracteur est soit un homophone, soit un pseudo mot visuellement proche, soit un voisin orthographe, soit un pseudo mot non conventionnel. Dans la deuxième tâche, il est demandé à l'enfant d'identifier le mot correspondant à l'image présentée. Enfin, dans la dernière tâche, il doit associer deux mots sémantiquement liés (par exemple, trouver le mot "mer" associé à "bateau", au lieu de "mre", "merle", "mère" ou "men"). La passation de ce test était semi-collective.

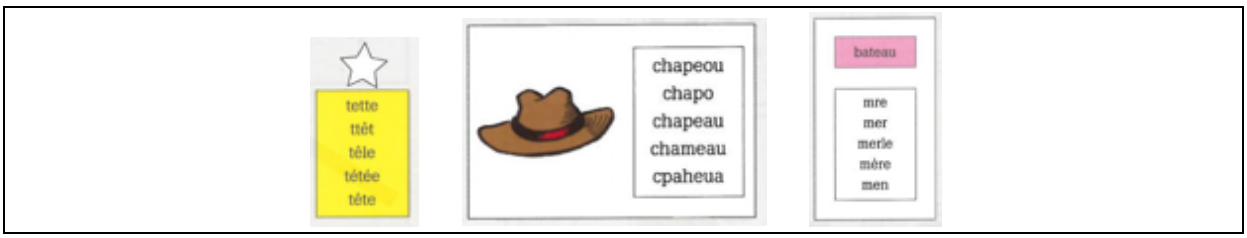


Figure 26 : Planches d'exemples des tâches du Timé 2 (à gauche, tâche 1, au centre tâche 2 et à droite tâche 3)

Vocabulaire : le test de vocabulaire actif et passif TVAP (Deltour & Hupkens, 1984) a été utilisé. Voir sa description dans la partie précédente.

Raisonnement non verbal : Les matrices progressives de Raven (Raven, 1976) ont été utilisées. Pour une description voire la partie précédente.

Capacités mathématiques : Trois tâches ont été tirées de la batterie d'évaluation mathématique Tedi-Math (Van Nieuwenhoven, Grégoire & Noël, 2001) (voir figure 27): 6 opérations imagées où l'examinateur décrit une image et pose une question à l'enfant (par exemple "il y a deux ballons rouges et trois ballons bleus. Combien y a-t-il de ballons en tout ?") ; 8 additions simples et 4 additions à trous.

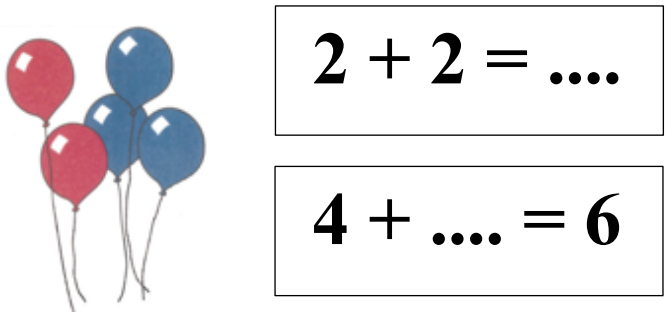


Figure 27 : Exemple d'items des tâches du Tedi-Math (Van Nieuwenhoven, Grégoire & Noël, 2001)

Une autre tâche consistant à comparer des paires de nombres a été évaluée. Sur une tablette, 36 paires de nombres ont été présentées une par une, les enfants devaient appuyer le plus rapidement possible sur le plus grand nombre de la paire. La liste contenait 4 paires par différence de chiffres (de 1 à 9), avec 2 réponses correctes à droite et 2 à gauche. Le support numérique a permis de recueillir la précision et le temps de la réponse (en ms). Le test a été réalisé grâce à OpenSesame (Mathôt, Schreij & Theeuwes, 2012).

Résultats

Les résultats sur l'ensemble des tests pour les deux groupes (GG et Fiete) et les deux temps de mesure (pré test et post test) sont présentés dans le tableau 11.

La stratégie d'analyse générale est la suivante : étant donné que chaque individu i est emboîté dans une classe j, nous avons utilisé des modèles mixtes linéaires hiérarchiques pour prendre en compte l'élève et la classe comme variables aléatoires en utilisant la procédure IBM SPSS MIXED (voir équation 1) (voir Heck, Thomas & Tabata, 2009). Pour chaque variable critique (par exemple, les scores de suppression de phonèmes, la fluence en lecture), nous avons prédit les scores en post

test par les scores obtenus en pré test. Cela nous permet d'exprimer le gain global d'un enfant sur un test donné en fonction de son niveau pré test.

	Pré tests - Moyenne (ET)			Post tests - Moyenne (ET)		
	GG	Fiete	Total	GG	Fiete	Total
Décodage						
Fluence - Mots	7.94 (6.89)	11.18 (9.69)	9.59 (8.59)	35.38 (21.64)	38.30 (22.93)	36.87 (22.34)
Fluence - pesudo mots	7.92 (5.56)	11.01 (8.00)	9.50 (7.08)	26.93 (16.23)	29.72 (16.06)	26.36 (16.20)
Fluence - Alouette	11.21 (12.67)	17.69 (20.16)	14.51 (17.21)	66.31 (40.85)	73.02 (44.64)	69.73 (42.93)
Conscience phonologique						
Sup. syllabique	6.34 (3.43)	5.91 (3.28)	6.12 (3.36)	7.65 (3.02)	6.54 (3.31)	7.08 (3.22)
Sup. phonemique CVC	4.80 (4.63)	5.68 (4.09)	5.25 (4.39)	8.55 (3.86)	8.2((3.94)	8.40 (3.90)
Sup. phonemique CCV	4.66 (4.03)	5.57 (3.41)	5.13 (3.75)	7.25 (3.81)	7.01 (3.46)	7.13 (3.63)
Ecosse						
Oral	18.36 (4.61)	18.45 (4.43)	18.41 (4.52)	-	-	-
Écrit	-	-	-	7.98 (3.36)	8.63 (3.18)	8.31 (3.28)
Timé 2						
Oral	3.40 (2.05)	4.17 (2.26)	3.80 (2.19)	7.90 (2.60)	7.59 (2.54)	7.75 (2.57)
Visuel	2.59 (1.67)	2.78 (1.89)	2.69 (1.79)	5.42 (2.41)	5.14 (2.28)	5.28 (2.35)
Semantique	2.43 (1.53)	2.56 (1.50)	2.50 (1.51)	4.40 (2.42)	4.14 (2.31)	4.27 (2.37)
TVAP	34.67 (10.04)	34.51 (7.70)	34.59 (8.92)	-	-	-
PM 47	18.58 (5.15)	17.97 (4.70)	18.27 (4.93)	-	-	-
Mathématiques						
Opération images	5.14 (1.24)	5.07 (1.22)	5.11 (1.23)	5.45 (0.95)	5.45 (0.83)	5.45 (0.89)
Additions simples	4.48 (2.42)	5.07 (2.21)	4.78 (2.33)	6.09 (1.94)	6.60 (1.74)	6.35 (1.85)
Additions à trous	1.63 (1.68)	1.93 (1.63)	1.79 (1.66)	2.62 (1.54)	3.20 (1.28)	2.92 (1.44)
Comparaison paire nombre	31.64 (5.82)	32.48 (5.08)	32.07 (5.46)	33.62 (3.69)	33.58 (4.15)	33.60 (3.93)

Tableau 11 : résultats aux tests (pré et post) des deux groupes d'élèves

$$Yij\ (post) = \beta 0j + \beta 1jxij(pre) + \varepsilon ij \qquad (1)$$

Ensuite, nous avons ajouté l'intervention comme un effet fixe parce que nous voulions savoir si les scores au post test étaient différemment affectés par les deux interventions (GG contre Math) lorsque le niveau initial de l'enfant (son score au pré test) était pris en compte (voir l'équation 2). Nous avons également ajouté l'interaction entre les scores de l'intervention et ceux du pré test car on peut s'attendre à ce que l'efficacité de l'entrainement dépende du niveau initial de l'enfant. Les termes d'interaction n'ont été ajoutés que s'ils augmentaient de manière significative l'ajustement du modèle (rapport de vraisemblance logarithmique) dans une procédure de sélection de modèle à terme.

$$Yij\ (post) = \beta 0j + \beta 1jxij(pre) + \beta 2j\ interventionij + \varepsilon ij \qquad (2)$$

Ensuite, nous avons ajouté l'engagement dans l'entraînement (plus haut niveau atteint) comme un autre effet fixe parce que nous nous attendions à ce qu'elle affecte l'efficacité de l'intervention (équation 3).

$$Y_{ij}\,(post) = \beta_{0j} + \beta_{1j}x_{ij}(pre) + \beta_{2j}\,intervention_{ij} + + \beta_{3j}x_{ij}(engagement) + \varepsilon_{ij} \qquad (3)$$

Nous avons également ajouté l'interaction entre intervention * engagement et la triple interaction entre intervention * pré test * engagement, mais seulement si elles augmentent significativement les ajustements du modèle.
Dans les sections suivantes, nous présentons les résultats ANOVA III des effets fixes et des interactions, les estimations paramétriques des modèles hiérarchiques à effets mixtes finaux et les données ajustées pour les quatre tests critiques qui ont évalué les effets de l'intervention en dehors du jeu : suppression des phonèmes, reconnaissance de mots (test Timé 2), fluence de lecture de texte (Alouette) et lecture de mot (test de lecture en une minute). Pour avoir une idée de la taille d'effet (ES) de l'intervention GG, nous avons calculé les coefficients normalisés, obtenus en normalisant (c'est-à-dire moyenne = 0 ; écart type = 1) chaque variable avant l'analyse.

Suppression phonémique (Evalec). Le modèle final comprend les effets fixes de l'intervention (GG contre Math), le niveau initial au pré-test et l'engagement de l'étudiant dans le jeu ainsi que tous les termes d'interaction. Les analyses ANOVA III ont montré des effets fixes significatifs de l'intervention (F (1, 290.1) = 23.27, p < .0001), du niveau pré-test (F (1, 897.8) = 191.2, p < .0001) et de l'engagement (F(1, 891.9) = 70.8, p < .0001). Il y a eu des doubles interactions significatives entre les effets de l'intervention et du pré-test (F(1, 897.8) = 13.4, p < .0001) et l'intervention et l'engagement (F(1, 891.9) = 4.04, p < .05), ainsi qu'une triple interaction entre l'intervention, le niveau en pré test et l'engagement (F(1, 865.3) = 9.35, p < .0001). Pour apprécier la direction des effets et des interactions, il est important d'examiner les estimations des paramètres, qui sont présentées dans le tableau 12.

Parameters	Betas	t	Sig.
Random Effect			
Class	3.602	3.510†	0.000
Intercept	10.026	17.101	0.000
Fixed Effects			
[Intervention=GG]	3.755	4.825	0.000
[Intervention=MATH]	0.000	.	.
Pre-test	0.484	11.712	0.000
Engagement	2.021	4.166	0.000
[Intervention=GG] * Pre-test	-0.202	-3.659	0.000
[Intervention=Math] * Pre-test	0.000	.	.
[Intervention=GG] * Engagement	1.451	2.224	0.026
[Intervention=Math] * Engagement	0.000	.	.
[Intervention=GG] * Engagement * Pre-test	-0.118	-3.781	0.000
[Intervention=Math] * Engagement * Pre-test	-0.079	-2.097	0.036

† z-value (Wald)

Tableau 12 : Estimations paramétriques non standardisées du modèle final prédisant les scores de suppression des phonèmes après le test sur la fonction de l'intervention, le niveau pré-test et l'engagement dans l'entraînement

Comme on peut le voir dans le tableau 12, l'intervention GG a un effet positif important (+3,75 points) au-dessus de l'intervention Math. L'interaction entre l'intervention et le niveau pré test suggère que les enfants ayant initialement des compétences plus faibles en matière de suppression des phonèmes bénéficient davantage de l'intervention GG (-0,20 point, ES =0,23 ; voir également la figure 28). L'interaction entre l'intervention et l'engagement montre que les bénéfices de l'intervention GG augmentent avec l'engagement dans l'intervention (+1,45 points, ES =0,16). La triple interaction suggère que les plus grands bénéfices de l'intervention GG ont été obtenus lorsque le niveau initial des enfants était faible et l'engagement élevé. La figure 28 montre les résultats du post test ajusté au modèle en fonction du niveau initial du pré-test pour trois niveaux d'engagement : faible engagement (1 ET en dessous de l'engagement moyen), engagement moyen et engagement élevé (1 ET au-dessus de l'engagement moyen).

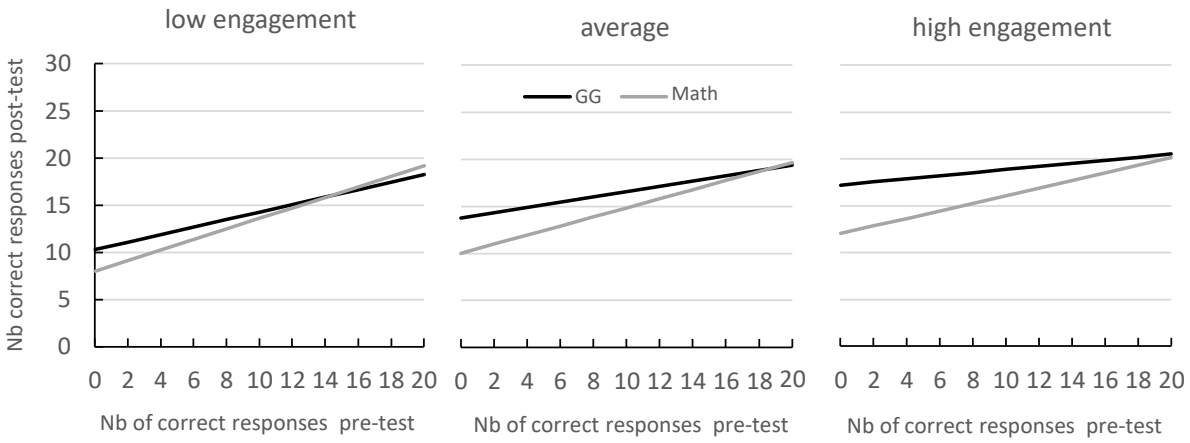


Figure 28 : Résultats du post test de suppression des phonèmes en fonction de l'intervention, du niveau pré-test et de l'engagement dans l'entraînement

<u>Reconnaissance de mots (Timé 2).</u> Le modèle final comprend les effets fixes de l'intervention, le niveau en pré-test et l'engagement, ainsi que les deux doubles interactions. Les analyses ont montré des effets fixes significatifs de l'intervention (F1, 533,9) = 43,86, $p < 0,0001$), du pré-test (F (1, 897,04) = 123,44, $p < 0,0001$) et de l'engagement (F(1, 891,60) = 319,44, $p < 0,0001$). Il y a eu des interactions significatives entre les effets de l'intervention et du pré-test (F(1, 897.04) = 29.00, $p <$ .0001) et les effets de l'intervention et de l'engagement (F(1, 891.80) = 57.46, $p <$ .0001). Les estimations des paramètres sont présentées dans le tableau 13.

Parameters	Betas	t	Sig.
Random Effect			
Class	1.251	2.600†	0.009
Intercept	10.415	18.087	0.000
Fixed Effects			
[Intervention=GG]	5.465	6.623	0.000
[Intervention=MATH]	0.000		
Pre-test	0.670	12.586	0.000
Engagement	1.698	7.518	0.000
[Intervention=GG] * Pre-test	-0.438	-5.386	0.000
[Intervention=MATH] * Pre-test	0.000		
[Intervention=GG] * Engagement	2.501	7.581	0.000
[Intervention=MATH] * Engagement	0.000		

† *z*-value (Wald)

Tableau 13 : Estimations des paramètres du modèle final prédisant les résultats du post test de reconnaissance de mots en fonction de l'intervention, du niveau pré test et l'engagement dans l'entrainement

Les résultats montrent un effet fixe clair de l'intervention, GG produisant des scores de reconnaissance de mots plus élevés (+5,4 points, ES =0,25) que l'intervention en mathématiques, toutes choses égales par ailleurs (c'est-à-dire niveau de pré-lecture et engagement). Il est important de noter qu'il y a eu une interaction entre l'intervention et le niveau initial de reconnaissance des mots, ce qui reflète le fait que l'intervention GG a été plus bénéfique pour les lecteurs ayant de moins bonnes compétences de reconnaissance des mots avant le test (β = -.0438, ES =0,30 ; voir figure 29). De même, il y a eu une interaction significative entre les effets de l'intervention et l'engagement, ce qui reflète le fait que l'intervention GG était plus efficace (β = +2,5 points, ES =0,40) lorsque l'engagement dans l'intervention était élevé. La figure 29 montre les résultats du post test adapté au modèle en fonction du niveau initial de reconnaissance des mots et de trois niveaux d'engagement dans l'entraînement.

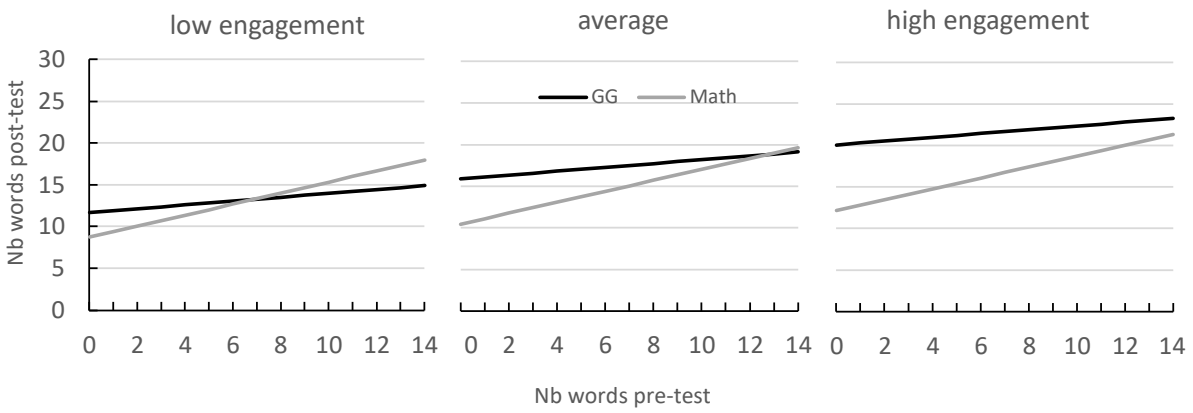


Figure 29 : Résultats du post test de reconnaissance de mots adapté au modèle en fonction de l'intervention, du niveau pré-test et de l'intensité de l'entrainement

Lecture de texte à voix haute (Alouette). Le modèle final comprend les effets fixes de l'intervention, le niveau en pré test et l'engagement, la double interaction entre l'intervention et l'engagement et la triple interaction entre les effets de l'intervention, l'engagement et le niveau du pré-test. L'effet fixe de l'intervention n'a pas atteint le seuil de significativité (F1 (1, 81,91) = 2,21, p = 0,141). Cependant, il y a eu une interaction significative entre les effets de l'intervention et de l'engagement (F(1, 897.5 = 20.8, p < .0001) et une triple interaction significative entre les effets de l'intervention, de l'engagement et du niveau en pré test (F(1, 897.5) = 23.8, p < .0001). Les effets fixes du niveau en pré test et de l'engagement étaient également significatifs (F(1, 896.7) = 300.8, p < .0001 et F(1, 897.1) = 228.03, p < .0001, respectivement). Les estimations des paramètres sont présentées dans le tableau 14.

Parameters	Betas	t	Sig.
Random Effect			
Class	63.460	2.708†	0.007
Intercept	49.895	21.374	0.000
Fixed Effects			
[Intervention=GG]	4.268	1.488	0.141
[Intervention=MATH]	0.000 .	.	
Pre-test	1.468	17.346	0.000
Engagement	14.478	7.777	0.000
[Intervention=GG] * Engagement	12.549	4.562	0.000
[Intervention=Math] * Engagement	0.000 .	.	
[Intervention=GG] * Engagement * Pre-test	-0.564	-6.023	0.000
[Intervention=Math] * Engagement * Pre-test	-0.355	-4.408	0.000

† z-value (Wald)

Tableau 14 : Estimations des paramètres du modèle final prédisant les résultats du post test de lecture de texte sur la fonction de l'intervention, le niveau pré-test et l'engagement dans l'entrainement

Comme on peut le voir dans le tableau 14, le principal résultat de l'amélioration de la lecture de texte est une interaction significative entre les effets de l'intervention et l'engagement, ce qui suggère que l'intervention de GG a obtenu des scores de lecture plus élevés en post test uniquement lorsque l'engagement était moyen ou élevé (+12,5 scores, ES =0,22 ; voir également la figure 30). La triple interaction avec le niveau de lecture en pré test suggère que les gains de GG par rapport aux mathématiques étaient les plus élevés pour les lecteurs ayant initialement de faibles scores en pré test (faibles lecteurs) et diminuaient pour les lecteurs ayant des scores élevés en pré test (bons lecteurs).

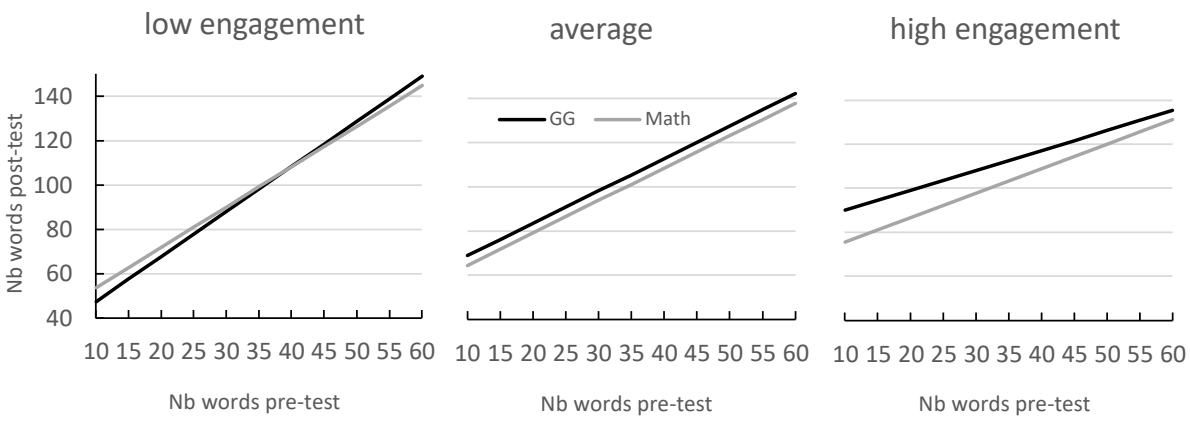


Figure 30 : Résultats du post test de lecture de texte adapté au modèle (test Alouette) en fonction de l'intervention, des résultats du pré-test et de l'engagement dans l'entraînement

Lecture de mots à voix haute (test de lecture en une minute). Le modèle final comprend les effets fixes de l'intervention, le niveau en pré test et l'engagement, la double interaction entre l'intervention et l'engagement et la triple interaction entre l'intervention, l'engagement et le niveau en pré test. L'effet de l'intervention n'était pas significatif (F (1, 93,5) = 2,85, p = 0,095). Cependant, il y a eu une interaction significative entre les effets de l'intervention et de l'engagement (F(1, 887.5) = 6.95, $p < .01$) et une triple interaction significative entre les effets de l'intervention, de l'engagement et du niveau en pré test (F(1, 870.2) = 12.31, $p < .0001$). Les effets fixes du niveau en pré test et de l'engagement étaient également significatifs (F(1, 899.8) = 290.6, $p < .0001$ et F(1, 886.4) = 149.6, $p < .0001$, respectivement). Les estimations des paramètres sont présentées dans le tableau 15.

Parameters	Betas	t	Sig.
Random Effect			
Class	17.201	2.820†	0.005
Intercept	23.624	17.746	0.000
Fixed Effects			
[Intervention=GG]	2.475	1.643	0.104
[Intervention=MATH]	0.000 .	.	
Pre-test	1.475	17.008	0.000
Engagement	9.926	8.951	0.000
[Intervention=GG] * Engagement	4.481	2.784	0.005
[Intervention=Math] * Engagement	0.000 .	.	
[Intervention=GG] * Engagement * Pre-test	-0.526	-6.050	0.000
[Intervention=Math] * Engagement * Pre-test	-0.471	-5.582	0.000

† z-value (Wald)

Tableau 15 : Estimations des paramètres du modèle final prédisant les résultats du post test de lecture de texte en fonction de l'intervention, du niveau et de l'intensité de l'entrainement

Les principaux résultats de la lecture de mots isolés peuvent être résumés ainsi : bien que le principal effet de l'intervention n'ait pas atteint le seuil de significativité (ES =0,11), il y a eu une interaction significative entre les effets de

l'intervention et l'engagement, montrant qu'un fort engagement a produit des gains significatifs de GG par rapport à l'intervention en mathématiques (β = +4,48 points, ES =0,18). La triple interaction entre les effets de l'intervention, du niveau en pré test et de l'engagement était également significative, ce qui suggère que le fort engagement dans l'entraînement pour les moins bons lecteurs a produit les plus grands gains de GG par rapport à l'intervention en mathématiques (voir figure 31).

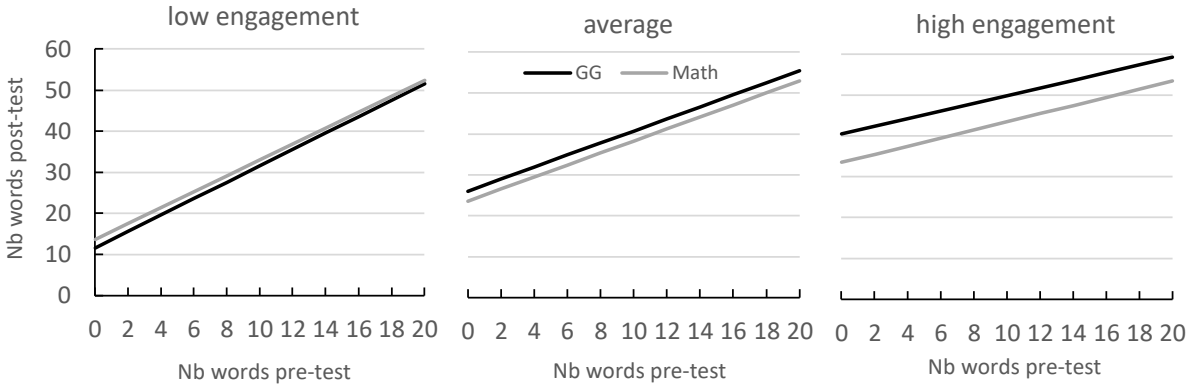


Figure 31 : Résultats du post test de lecture de mots (test de lecture en une minute) modélisés en fonction de l'intervention, du niveau pré-test et de l'engagement dans l'entraînement

Résultats en mathématiques. Les capacités mathématiques des élèves ont notamment été évaluées via trois tâches de résolution d'opération issues de la batterie Tedi-Math. Nous avons comparé les résultats obtenus aux trois tâches (somme des scores obtenus) et nous remarquons que les élèves ayant suivi l'entrainement en mathématiques ont significativement de meilleurs résultats par rapport aux élèves ayant suivi l'entrainement avec GG ($p < .004$) (voir figure 32).

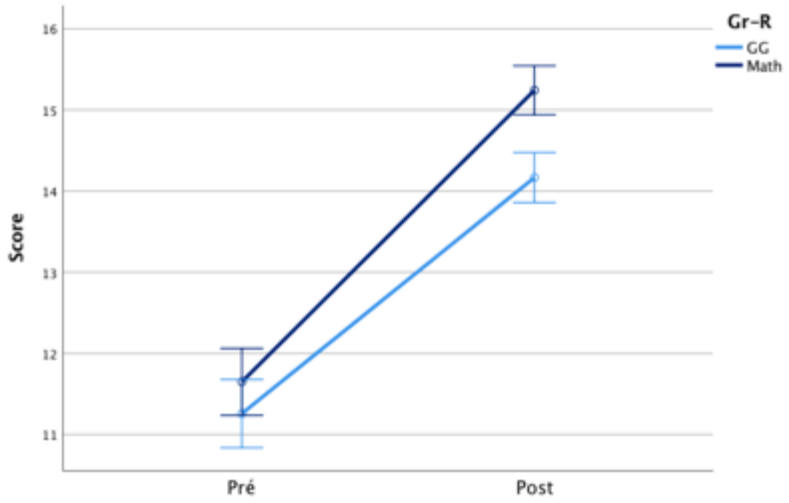


Figure 32 : Analyse des performances des élèves des deux groupes (GG et Math) aux trois tâches de mathématiques en pré test et en post test

Corrélations. Pour mieux comprendre l'influence importante de l'engagement dans les résultats obtenus ci-dessus, nous avons regardé les corrélations de cette variable avec les gains (différence pré-post) pour le nombre de mots dans deux des tests. On trouve une corrélation positive entre l'engagement (le plus haut niveau atteint) dans GraphoGame et le gain observé en nombre de mots lus en 3 minutes (test de l'Alouette, r = 0,57) et avec le gain observé en nombre de mots lus en 1 minute (test de lecture de mots en 1 minute, r = 0,56) (voir figure 33). Cette relation de corrélation pourrait s'expliquer par le fait que les élèves qui sont allés le plus loin dans le jeu sont ceux qui avaient déjà de bonnes compétences en lecture. Or, on ne retrouve pas cette corrélation positive dans les tâches de mathématiques (somme des 3 tâches, r = -0,02).

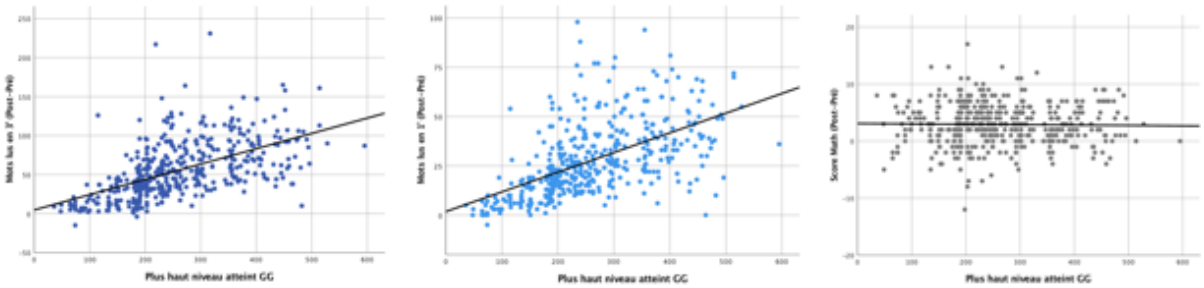


Figure 33 : Corrélation entre le plus haut niveau atteint dans GraphoGame et le gain en nombre de mots correctement lus en 3 minutes (test de l'Alouette à gauche) ; et le gain en nombre de mots correctement lus en 1 minute (test de lecture de mots en 1 minute au centre) ; et le gain dans les tâches de mathématiques (à droite)

On trouve également une corrélation positive entre le pourcentage d'essais réussis (essais corrects) dans GraphoGame et le gain observé en nombre de mots lus en 3 minutes (test de l'Alouette, r = 0,56) et avec le gain observé en nombre de mots lus en 1 minute (test de lecture de mots en 1 minute, r = 0,54) (voir figure 34). De même, cette corrélation positive n'est pas observée avec les tâches de mathématiques.

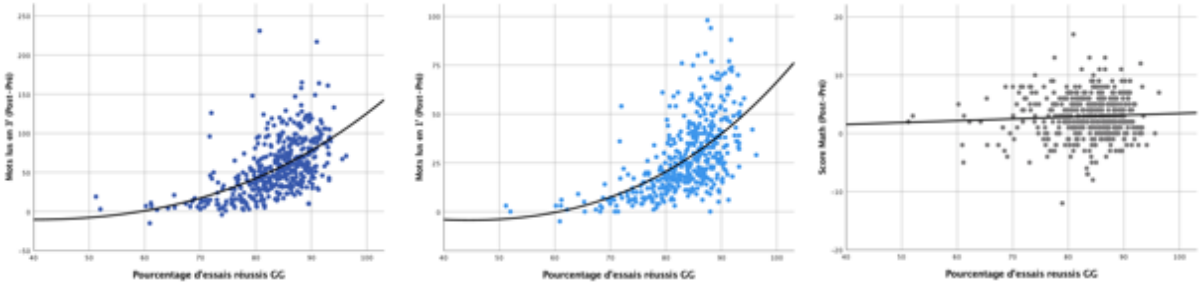


Figure 34 : Corrélation entre le pourcentage d'essais réussis dans GraphoGame et le gain en nombre de mots correctement lus en 3 minutes (test de l'Alouette à gauche) ; et le gain en nombre de mots en 1 minute (test de lecture de mots en 1 minute, au centre) ; et le gain dans les tâches de mathématiques (à droite)

En résumé, ces corrélations mettent clairement en évidence le lien qui existe entre le temps passé et l'engagement au sein du jeu (plus haut niveau atteint), le taux de réussite dans le jeu (pourcentage d'essais réussis) et l'amélioration des performances aux tests d'évaluation en dehors du jeu. Ces résultats soulignent que les gains observés sont dus aux activités de jeu et que l'entrainement avec GraphoGame a eu un effet très spécifique car aucun lien ne ressort pour l'entrainement en mathématique.

Progression au sein de GG**.** Enfin, nous pouvons regarder le progrès des élèves au sein du jeu. La progression des élèves au sein du jeu a été évaluée en comparant le pourcentage de bonnes réponses du premier niveau de chaque séquence (pré test) au pourcentage de bonnes réponses du dernier niveau (post test) de chaque séquence. Comme le montre la figure 35, les élèves ont progressé en moyenne de 13% entre le pré-test et le post test de chaque séquence ($t(15955) = 13.46$; $p < .0001$). Ces résultats reflètent les progrès à court terme des élèves et mettent en évidence un bénéfice immédiat d'une séquence d'entrainement sur leurs performances.

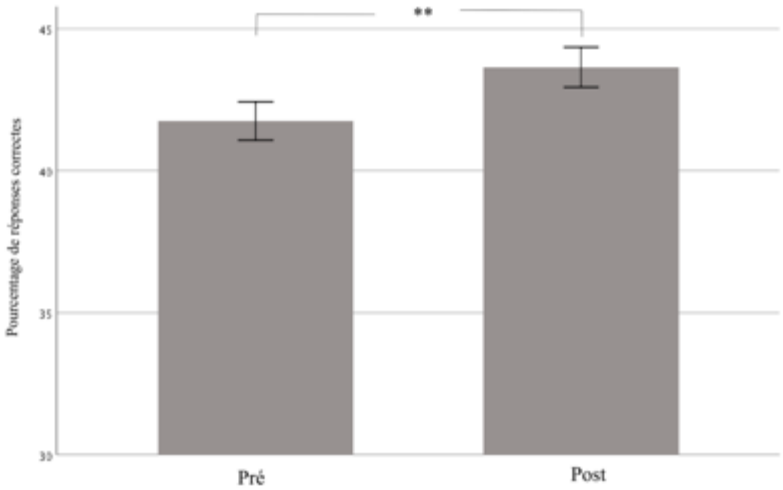


Figure 35 : Progression au sein de GG entre le premier et dernier niveau de chaque séquence. Les barres d'erreurs correspondent à des erreurs standards intra-sujets

Résultats en CE1

Âge de Lecture. Les élèves ayant suivi l'entrainement avec GG au CP, ont passé une nouvelle séance de « post test » à la fin de leur année de CE1 (suivi longitudinal). Les tests proposés étaient identiques à ceux de l'année précédente : lecture de mots en une minute, alouette, Timé 2 et épreuves du Tedi math. Le test de l'Alouette nous a permis de quantifier le nombre de mots que les élèves lisaient correctement en trois minutes, et il nous fournit également un âge de lecture par enfant. La comparaison de l'âge chronologique (âge réel) des élèves avec l'âge de lecture obtenu grâce à l'Alouette est représenté dans la figure 36.

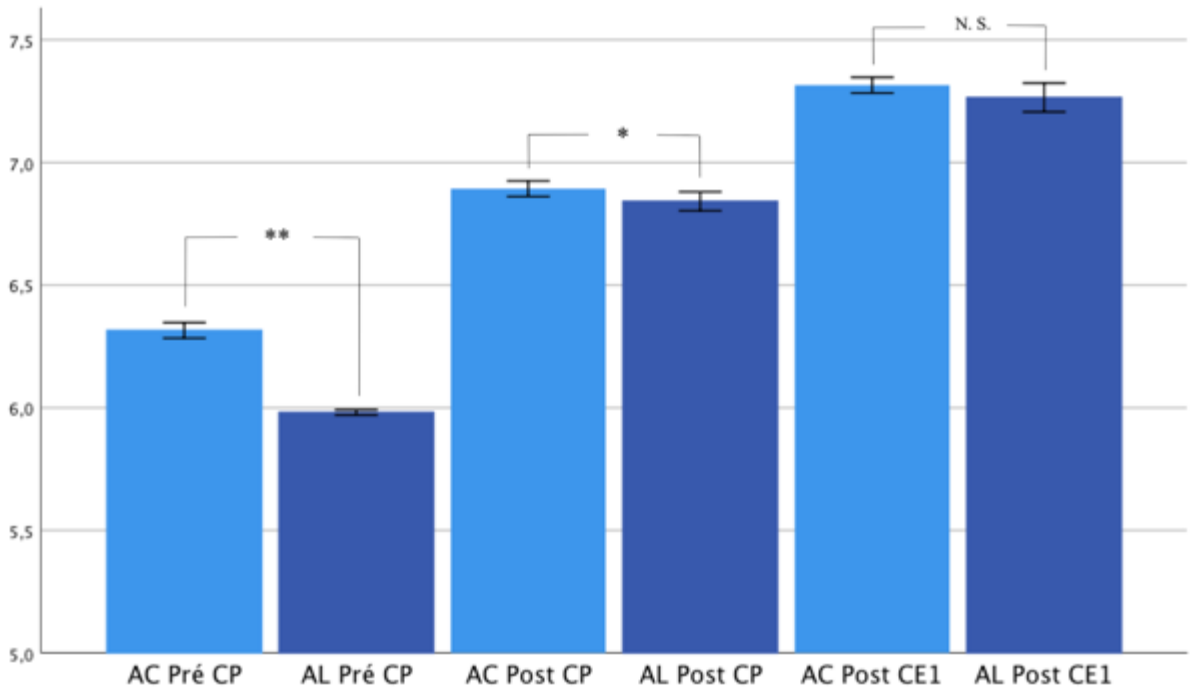


Figure 36 : *Comparaison des âges chronologiques (AC) aux âges de lecture (AL) obtenus grâce au test de l'Alouette (Lefavrais, 1965) en pré test à gauche, en post test 1 au centre et en post test 2 à droite des élèves ayant suivi l'entrainement avec GG*

Les analyses montrent qu'au début de l'apprentissage de la lecture (pré test) l'âge de lecture des enfants est significativement inférieur à leur âge chronologique (AC) (p <.0001) puisqu'ils sont apprentis lecteurs. À la fin de CP, la différence entre l'AC et l'AL est encore significative (p <.043) mais l'écart entre les deux diminue. À la fin du CE1, on observe que cette différence n'est plus significative (p <.137), ce qui montre que les élèves qui ont de meilleures compétences en décodage rattrapent leur « retard », et que cet effet se maintient un an après.

ÉvalAide. Le dispositif ÉvalAide mis en place permet d'évaluer les compétences cognitives des élèves en début et milieu de CP ainsi qu'en début de CE1. Il a pour but d'identifier les besoins spécifiques de chaque élève afin de mieux adapter l'intervention pédagogique des enseignants. Les évaluations proposées fournissent des repères précis sur les compétences acquises et les progrès dans différents domaines (langage et mathématiques). L'objectif des épreuves est de détecter ou d'anticiper d'éventuelles difficultés scolaires et réduire le délai entre un diagnostic et une intervention pédagogique (ÉvalAide, 2019).
Concernant les épreuves de français en début de CE1, on retrouve des tâches de comparaison de suites de lettres, de dictée de syllabes et de mots, de lecture à voix haute de mots et de textes, de compréhensions de mots et de phrases lues par l'enseignant ou lues par l'élève (voir figure 37).

Figure 37 : Exemples d'items issus des évaluations nationales en début de CE1. Tâche de compréhension de mots lus par l'enseignant à gauche ; tâche de compréhension de phrases lues par l'élève au centre ; tâche de lecture à voix haute par l'élève à droite.

Nous avons pu comparer les résultats[1] à des épreuves d'élèves de CE1 ayant suivi l'entrainement avec GG au cours de leur année de CP aux résultats obtenus à ces mêmes tâches par des sujets « contrôles », issus de la même classe, mais n'ayant pas suivi l'entrainement avec GG (voir figure 38). Les élèves contrôles viennent d'autres classes de CP, non incluses dans le projet, et ne se sont donc pas entrainés avec GG.

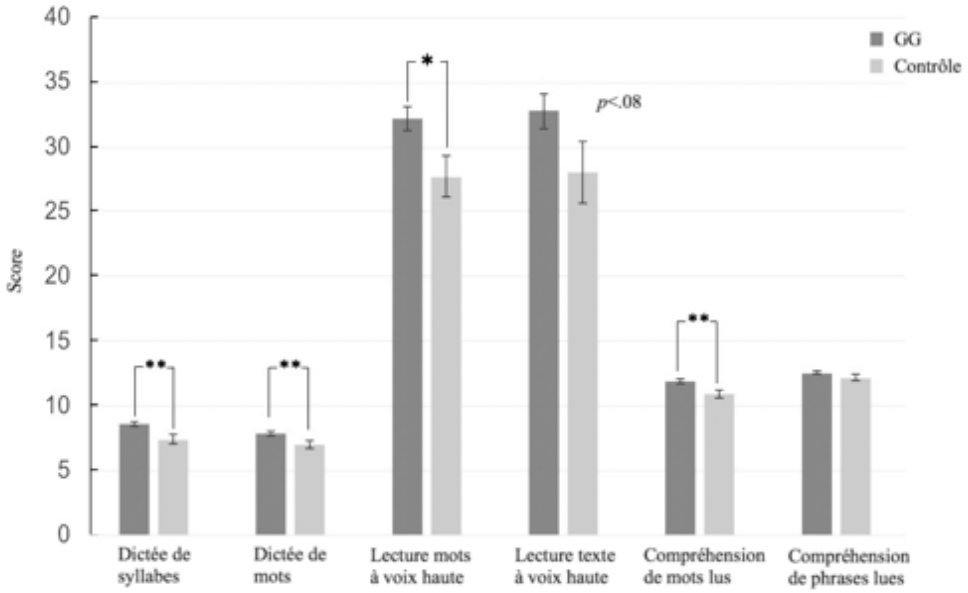


Figure 38 : résultats aux épreuves de « lecture » des évaluations nationales d'élèves de CE1 ayant suivi l'entrainement GG au cours de leur année de CP (en gris foncé) par rapport à des élèves « contrôles » n'ayant pas suivi l'entrainement (en gris clair)

[1] Les données des évaluations pour les élèves de nos classes LEMON ont été mis à disposition de façon anonymisée par la DEPP et le rectorat d'Aix-Marseille

Les tâches de dictées de syllabes et de mots, de lecture de mots à voix haute et de compréhension orale, ont été significativement mieux réussies par les élèves ayant suivi l'entrainement avec GG (la différence pour la tâche de lecture de texte est partiellement significative). Il semble qu'un entrainement intensif des correspondances graphophonologiques permet aux élèves, un an après, de mieux lire, écrire et comprendre des mots et des phrases.

4. Discussion générale

La lecture est un prérequis indispensable pour accéder à l'ensemble des domaines de l'apprentissage. Bon nombre d'élève éprouve des difficultés au moment de cet apprentissage qui peuvent avoir un impact négatif sur leur scolarité et leur vie personnelle et professionnelle. La prise en charge de ces difficultés reste encore insuffisante et non systématique malgré la mise en place de dispositifs prometteurs (dispositif ÉvalAide, évaluation de méthode d'enseignement de la lecture, outils numériques).

Au sein de ces dispositifs d'aide, les logiciels éducatifs proposent des entrainements ciblés destinés aux enfants en difficultés et les aider à atteindre les objectifs fixés. Les résultats des recherches sur le numérique comme outils d'aide à l'apprentissage de la lecture sont mitigés (Campuzano, *et al.*, 2009 ; Dynarski *et al.*, 2007 ; Slavin, Lake, Davis & Madden, 2011). Toutefois, certains logiciels orientés sur les compétences alphabétiques, l'enseignement explicite de la conscience phonologique et le séquençage des phonèmes figurent parmi les programmes les plus efficaces pour améliorer les résultats en lecture (Cheung & Slavin, 2011 ; Torgesen *et al.,* 2010). Parmi ces programmes qui se sont révélés prometteurs, quelques-uns se sont distingués par leur engagement à fusionner la recherche et la pratique comme le logiciel GraphoGame (Donnelly *et al.,* 2020). Graphogame a été mise en œuvre et testé dans de nombreuses langues (allemand, anglais, espagnol, portugais...) et se base sur la correspondance « son-lettres » pour entrainer les compétences phonologiques d'abord, puis la morphologie et l'orthographe (Richardson & Lyytinen, 2014). Elle utilise un environnement gamifié qui permet de produire des niveaux élevés de motivation et d'engagement, tant à l'école qu'à la maison (McTigue *et al.*, 2020 ; Ronimus & Lyytinen, 2015).

L'objectif de cette thèse était d'adapter au français le contenu de GraphoGame de le tester et de le valider. Des résultats probants ont été obtenus dans notre première étude auprès d'enfants « à risques » de développer une dyslexie, bénéficiant d'une prise en charge au sein de l'école (Activité Pédagogiques Complémentaires, APC). En effet, nous avons observé un bénéfice spécifique de l'entrainement avec GraphoGame avec de meilleurs résultats des élèves du groupe GG en décodage et en vitesse de lecture par rapport aux élèves du groupe APC, et ce après seulement 5 semaines d'entrainement.

La seconde étude, que nous avons réalisée avec près de 1000 élèves de CP, avait quatre objectifs. Le premier objectif était de créer le contenu francophone de GraphoGame, qui couvre l'ensemble des correspondances graphophonologiques étudiées au CP, qui suit une progression optimale (Sprenger-Charolles, 2017) en tenant compte des irrégularités et des spécificités de notre langue (Peereman & Sprenger-Charolles, 2018). Nous avons élaboré le contenu de sorte que les correspondances graphème-phonème les plus consistantes et les plus fréquentes

soient introduites progressivement et qu'elles soient ensuite formées dans le contexte de syllabes, de mots et de phrases alors constitués de GPC préalablement apprises (Lassault & Ziegler, 2018). Ainsi, la version française de GraphoGame va au-delà de l'enseignement des correspondances lettres-sons classique. Le deuxième objectif était d'évaluer l'efficacité de ce logiciel à grande échelle, à cause de la forte corrélation négative qui existe en la taille de l'échantillon et la taille d'effet des interventions, où les études avec un nombre restreint de sujets ont tendance à surestimer les effets (Cheung & Slavin, 2016). Nous voulions aussi tester l'efficacité au sein d'écoles de Réseaux d'Éducation Prioritaire (REP et REP+) car le besoin d'une intervention précoce au sein de l'école est particulièrement élevé, notamment en France (Fluss *et al.,* 2008 ; Fluss *et al.*, 2009). Le troisième objectif était d'évaluer l'intervention GraphoGame en lecture par rapport à une intervention en mathématique tout aussi engageante, en conservant les effets sur les enseignants et l'école constants. Il établit que les effets sur les enseignants sont le facteur dominant qui affecte le gain scolaire des élèves (Sanders, Wright, & Horn, 1997). Le quatrième objectif de cette étude était d'analyser les résultats par des modèles hiérarchiques à effets mixtes, qui prennent en compte les élèves et les classes comme facteurs aléatoires, ainsi que certaines variables modératrices, telles que le niveau de lecture initial de l'enfant et l'engagement. Étonnamment, peu d'études antérieures comportaient des variables permettant d'estimer l'engagement et la motivation des élèves (McTigue *et al.*, 2020).

Les résultats que nous avons obtenus montrent que les élèves qui se sont entrainés avec GraphoGame ont significativement de meilleurs résultats aux tâches de conscience phonologique, d'identification de mots (oral et écrits) et de décodage de mots isolés ou en contexte que les élèves ayant suivi l'entrainement en mathématiques. L'interaction significative qui existe entre le type d'intervention et les résultats en pré test a mis en évidence l'impact du niveau initial des élèves sur leur progression, avec une amélioration des performances encore plus importante chez les faibles décodeurs. Ce sont ces derniers qui bénéficient le plus d'un entrainement au décodage précoce et intensif. Les résultats ont également montré l'effet indiscutable de l'engagement au sein du logiciel car les élèves qui se sont le plus investit dans l'outil ont amélioré leurs compétences en conscience phonologique, décodage et identification de mots. L'analyse des corrélations entre les données recueillies au sein du jeu et les capacités de décodage a mis en évidence que ce sont probablement les activités réalisées par les élèves au sein même du jeu (réussite, temps de jeu, engagement) qui font progresser les élèves, car l'engagement dans le GG corrèle avec la progression en lecture ce qui n'est pas le cas avec l'engagement dans le jeu Math qui ne corrèle pas avec les bénéfices en lecture.

L'analyse en fin de CE1 de la différence entre l'âge chronologique des élèves et l'âge de lecture obtenu par le test de l'Alouette très significative en début de CP (âge de lecture faible) moins significative en fin de CP et plus significative en fin de CE1 met un évidence une amélioration des compétences en décodage qui perdure l'année suivante. Nous avons également eu la chance de pouvoir analyser les données de nos

élèves qui ont passé 5 mois après la fin de l'expérimentation des évaluations nationales de CE1. L'analyse montre que les résultats en dictée, en lecture de mots et en compréhension étaient supérieurs pour les élèves ayant suivi l'entrainement avec GG au CP par rapport aux élèves de la même classe de CE1 n'ayant pas suivi l'entrainement au CP. Ces résultats, qui corroborent ceux que nous avons obtenus, suggèrent que le bénéfice de l'intervention avec GG se maintient un an après et s'observe en dehors de notre protocole expérimental (environnement de condition de passation des épreuves différentes, examinateur différent *etc.*).

Bien que l'évaluation de l'efficacité de l'intervention en mathématiques ne fût pas l'objectif principal de notre étude, nous avons néanmoins cherché à savoir si l'intervention en mathématiques a produit une augmentation significative des compétences mathématiques après l'entrainement par rapport aux effets de l'intervention en lecture. En post test des compétences mathématiques, nous avons obtenu des effets significatifs de l'intervention en mathématiques et la taille de l'effet était comparable à celle de l'intervention GG (c'est-à-dire ES = 0,29). Il est intéressant de noter que l'intervention en mathématiques a montré exactement la même triple interaction que l'intervention GG, dans laquelle les gains de l'intervention en matière d'aptitudes mathématiques étaient plus importants pour les enfants ayant initialement des aptitudes mathématiques plus faibles, et surtout lorsque l'engagement au sein du jeu était élevé.

L'importance de la prise en compte du niveau initial et de l'engagement

Comment interpréter nos résultats dans le contexte de la méta-analyse très critique de McTigue *et al.* (2020) qui a fortement mis en question l'efficacité de GraphoGame ? En effet, par rapport aux résultats négatifs rapportés dans cette méta-analyse, les résultats obtenus au cours de ces deux expérimentations sont rassurants même si la taille d'effet reste modeste (ES ≃ 0,20 - 0,30). Cela dit, la taille d'effet de la mesure phare du dédoublent du CP étaient inférieure (ES ≃0,16) ((Dédoublement des classes de CP en éducation prioritaire renforcée : première évaluation, 2020). Selon cette méta-analyse, la majorité des études, à part celles de Saine *et al.* (2010, 2011, 2012), n'observent pas de gains spécifiques pour la lecture par rapport à un entrainement contrôle actif (maths) ou un groupe de contrôle actif. Selon les auteurs, les résultats positifs de Saine *et al.*, qui ont été obtenus en Finlande, peuvent s'expliquer par le fait que la Finlande est une « exception » dans le domaine de l'éducation car elle possède un système éducatif très efficace, un statut socio-économique général élevé et aussi une orthographe transparente (voir Zielger *et al.*, 2018). Il est possible qu'une intervention qui a fonctionné dans cet environnement puisse ne pas être facilement transposée à d'autres langues, même si l'enseignement est adapté. McTigue et collaborateurs avancent également la piste de l'absence d'utilisation dans l'analyse des résultats de modèles mixtes qui peuvent aider à

explorer les processus d'apprentissage contextualisés et socialement situés, ainsi qu'à comprendre les différences entre les groupes.

En effet, il est souvent supposé implicitement que les interventions efficaces doivent bénéficier à tous les enfants dans une mesure similaire (« all size fits all »). Nos résultats mettent en avant que tel n'est pas les cas et la prise en compte des différences interindividuelles semblent nécessaires pour obtenir des résultats claires et robustes (voir aussi Perry, Zorzi et Ziegler, 2019). En effet, nous avons trouvé que l'intervention a profité le plus aux enfants dont les compétences étaient initialement faibles. Ce résultat est en accord avec McTigue et collaborateurs qui avancent également que les études antérieures ont mis l'accent sur la technologie elle-même au détriment d'autres paramètres pertinents, comme le niveau de l'élève mais également l'engagement et la motivation au sein du jeu. En effet, concernant ce dernier point, nos résultats montrent bien que seuls les enfants ayant un engagement élevé dans le jeu présentent des avantages importants en termes d'intervention, notamment en ce qui concerne la fluidité en lecture. Ces résultats montre donc qu'on ne peut pas supposer que tous les enfants seront engagés et motivés de la même façon dans l'utilisation de Graphogame tout simplement car c'est un outil numérique (Wouters, van Nimwegen, van Oostendorp & van der Spek,2013).

Cette interaction entre le gain de lecture et l'engagement de l'élève dans le jeu est rassurant pour une autre raison. En effet, elle met en évidence que les bénéfices de notre intervention GG sont directement liés au comportement des utilisateurs dans le jeu et ne sont donc pas un artefact d'une autre variable externe qui n'est pas imputable au jeu. Sur la base de ces résultats, il nous semble primordiale que l'engagement et la réussite dans le jeu tout comme le niveau initial de l'élève doivent être prises en compte pour une évaluation de l'efficacité d'un outil d'intervention en classe.

Enfin, les auteurs de la méta-analyse mettent en avant l'importance fondamentale de la participation des adultes lorsque les élèves s'entraînent avec un jeu sérieux. Ils montrent que l'absence d'effets d'apprentissage était plus marquée lorsque le jeu était utilisé dans le strict respect de ses prescriptions (travail individuel) que quand il était intégré aux pratiques pédagogiques de classe. En effet, l'apprentissage de la lecture est tellement complexe qu'il est possible que l'entrainement informatique doive être utilisé en plus des activités de lecture « plus authentiques et sociales » (Huemer, Landerl, Aro & Lyytinen, 2008). De plus, les interactions avec les enseignants peuvent aider les élèves à activer des connaissances en les verbalisant (et non seulement en les appliquant dans le jeu). McTigue et collaborateurs mettent également en avant le rôle de l'implication des adultes dans la motivation des élèves, avec des effets positifs largement renforcés lorsque l'enseignant encourage ses élèves. Ils précisent également que le soutien pédagogique des enseignants pendant l'utilisation outils numériques aide les apprenants à sélectionner et à organiser plus efficacement les nouvelles informations

apprises (Wouters & van Oostendorp, 2013). C'est pourquoi, tout au long de notre expérimentation nous avons proposé des sessions de formation sur l'utilisation de GraphoGame en classe (sa conception, son contenu, sa fonction) destinées spécifiquement aux enseignants du projet. Nous avons notamment expliqué pourquoi nous préconisions l'utilisation de GG en demi-groupe, ce qui a permis l'augmentation des interactions (un enseignant pour 6 élèves dans les classes dédoublées) et le calme au sein de la classe.

Valorisation

Dans le but de diffuser largement nos travaux, nous avons réalisé un travail de vulgarisation important. Les formations dispensées aux enseignants (apprentissage de la lecture, ses difficultés, la conception de GraphoGame, mais aussi l'apprentissage des mathématiques ou l'objectif des évaluations nationales) sont toutes disponibles en ligne sur la plateforme de l'université AMUpod (https://amupod.univ-amu.fr/search/?q=projet+LEMON). Nous nous sommes rendus dans les classes à chaque période scolaire pour rencontrer les enseignants afin d'avoir fréquemment leurs retours sur l'utilisation des outils. Nous avons aussi créé et mis en ligne un site web (www.grapholearn.fr) qui recense les grands principes de la conception du jeu, l'accès facile aux publications en français relatives au projet et un lien de téléchargement de l'application. En effet, depuis la validation du dispositif, GraphoGame est disponible gratuitement sur toutes les plateformes de téléchargement en France, avec trois points de départ possibles du jeu, en fonction du niveau initial de l'enfant. Pour accompagner cette mise en ligne, nous proposons un manuel d'utilisation détaillé qui présente le jeu, ses principes et guide les utilisateurs tout au long de leur utilisation (création d'un joueur, paramètres, bonus.).

Dans le contexte de confinement que nous traversons et avec l'enseignement à distance de nombreuses familles et enseignants ont dû de plus en plus s'orienter vers les ressources numériques. Le Comité Scientifique de l'Éducation Nationale, sur la plateforme du Réseau-Canopé, a recommandé des ressources pédagogiques qui pouvaient être utilisés en compléments des temps de classe, notamment GraphoGame (https://www.reseau-canope.fr/conseil-scientifique-de-leducation-nationale/outils-pedagogiques/recommandations-de-ressources-pedagogiques.html).

Limites

Nos études présentent certaines limitations qu'il convient de souligner. En ce qui concerne la première étude, le taux d'entraînement reste assez faible par rapport aux études qui ont rapportées des effets d'entraînement impressionnants tant au niveau comportemental que cérébral. Par exemple, dans l'étude de Eden et collaborateurs (2004), les adultes dyslexiques ont bénéficié d'un entraînement de 112

heures (3 heures par jour pendant 8 semaines). Dans l'étude de Simos et collaborateurs (2002), les enfants dyslexiques ont eu un entraînement de 80 heures (1-2 heures par jour pendant 8 semaines). Dans l'étude de Franceschini et collaborateurs (2013), les enfants dyslexiques ont été entrainés pendant 12 heures. Notre taux d'entraînement a été largement en dessous avec seulement 5 heures.

Puis, nous n'avons pas pu suivre ces enfants pour savoir si les effets d'entraînement persistent dans la durée et lesquels de ces enfants « à risque » ont été diagnostiqués dyslexiques par la suite. En effet, la persistance des effets d'entraînements dans la durée est un enjeu majeur et peu d'études d'entraînement ont étudié les effets d'entraînement à long terme (Elbro & Petersen, 2004).

Une autre limite est, comme discuté plus haut, le fait que nous ayons uniquement utilisé des tests de lecture à voix haute. Or, nous pouvons nous attendre à des effets plus importants sur la lecture silencieuse (Ziegler, Bertrand, Lété & Grainger, 2014). Enfin, nous n'avons pas pu vérifier si l'entraînement a un effet bénéfique sur la compréhension de textes lors de la lecture (Gough & Tunmer, 1986). Cette première validation de GraphoGame pour les enfants à risque en contexte scolaire a été assez encourageante.

Dans la seconde étude, la première limite est que le design expérimental n'est pas randomisé et contrôlé. Initialement, le design expérimental devait être le même que dans la première étude de validation, avec au sein des neuf écoles deux tiers des classes de CP s'entrainant avec GraphoGame et un tiers avec Fiete Math (design randomisé contrôlé). Toutefois, il n'était pas prévu de changement au milieu de l'expérimentation pour que tous les élèves bénéficient d'un temps d'entrainement important. Nous avons procédé ainsi pour l'étude pilote réalisée avant de déployer le dispositif à grande échelle, mais à l'issue de celle-ci nous avons observés de meilleurs résultats en décodage pour les élèves de la classe avec entrainement Math. Ces observations s'expliquent par une augmentation des temps et des ateliers de lecture mis en place dans cette classe. Le protocole expérimental mis en place a induit un changement de comportement de l'enseignant qui s'est senti évalué par rapport aux deux autres enseignant qui utilisaient GraphoGame dans leur classe. Pour éviter ce biais et ne pas instaurer de climat d'évaluation il a été choisi de procéder comme mentionné précédemment, avec alors un appariement plus faible du groupe contrôle.
Nous pouvons relever une seconde limite concernant l'utilisation successive des mêmes tests pour les différentes phases de phases de test. En effet, les tests de décodage (lecture de mots familiers et inventés en une minute) ont été présentés aux élèves du projet deux fois (pré test CP, post test CP). Il est apparu que ces tests de lecture de mots utilisés au sein de notre expérimentation étaient les mêmes que ceux utilisés au sein des évaluations nationales en CE1 du dispositif ÉvalAide. Il est alors possible que les résultats à ces tests obtenus avec les évaluations nationales, soient significativement supérieurs pour les élèves du projet que pour les élèves contrôles à cause d'un effet test retest car ils lisaient ces mots pour la troisième fois.

Une autre limite ou danger que nous avons essayé d'éviter est la tendance à penser que le logiciel est un substitut à toutes les autres activités de lecture. Au cours de la mise en place de l'étude nous avons fait le choix de réduire le temps d'utilisation des logiciels et de commencer l'entrainement quatre mois après la rentrée des classes, afin que la dynamique « classique » d'apprentissage de la lecture soit mise en place par chaque enseignant au sein de sa classe. Il était primordial que GraphoGame soit utilisé comme un outil d'aide à l'apprentissage de la lecture, en plus du travail sur l'enseignement explicite par le professeur des correspondances graphophonologiques et sur la compréhension. En effet, le logiciel ne permet pas un travail de lecture à voix haute et ne possède pas d'exercices ciblés sur la compréhension. Le travail de Suchaut et collaborateurs (2014) a mis en évidence un temps d'engagement dans le code trop restreint au CP et le logiciel permet de doubler de ce temps de décodage mais il ne remplace pas toutes les autres activités essentielles et nécessaires (reconnaissance de lettres, vocabulaire, compréhension orale *etc.*) pour comprendre correctement un texte.

Perspectives

GraphoGame français est actuellement utilisé dans d'autres projets comme le projet Fluence (Laboratoire de Psychologie et Neurocognition, Université Grenoble Alpes) dont l'objectif est le développement et la validation de dispositifs numériques innovants, utilisés comme outils pédagogiques au sein de la classe. Il est aussi utilisé dans une étude randomisée contrôlée auprès d'enfants dyslexiques (Laboratoire Anthropologie, Psychologie Cliniques Cognitives et Sociales, Université Nice Sophia Antipolis). Tous les projets autour de Graphogame visent à tester son efficacité en compléments des activités réalisées au sein de la classe ou chez un orthophoniste. En effet, GraphoGame a montré de réels effets positifs, mais les bénéfices de son utilisation dépendront de sa capacité à compléter ce que les élèves apprennent en classe. La littérature sur les technologies éducatives souligne l'importance de l'intégration du contenu des logiciels aux programmes scolaires existants (Tare, Shell & Jackson, 2020 ; Taylor *et al.*, 2020) et de la dynamique entre les deux. Il serait alors intéressant de créer d'un manuel scolaire d'apprentissage de la lecture tenant compte de la progression au sein de GraphoGame et/ou d'un livret d'exercices qui pourraient favoriser l'apprentissage de la lecture. L'utilisation concomitante de ces trois supports, comme méthode de lecture, dès la rentrée du CP, permettrait d'identifier, d'aider et ainsi diminuer le nombre d'élève en grande difficultés de lecture. Comme énoncé précédemment, à l'heure actuelle GraphoGame permet uniquement le travail en lecture silencieuse. Il pourrait être intégré un logiciel de reconnaissance de la parole avec la création de séquences spécifiques pour entrainer la lecture à voix haute qui est fondamentale. Enfin il serait aussi pertinent d'inclure, plus tard dans le jeu, des séquences de compréhension orale et écrite où l'enfant entend ou lit une phrase et doit décider si elle est vrai ou fausse. Le logiciel pourrait être aussi encore plus

adaptatif avec un accès direct à une séquence par exemple, ce qui permettrait aux enseignants d'individualiser le travail de chacun et d'adapter au mieux leurs pratiques aux besoin spécifiques de chaque élève.

Bibliographie

AMERICAN PSYCHIATRIC ASSOCIATION. (2000). *Diagnostic and Statistical Manual of Mental Disorders, 4th Edition text revised. (DSM IV-TR)*. American Psychiatric Association. Washington DC.

BALL H.G. (1978). Telegames teach more than you think. *Audiovisual Instruction,* 23 (5), 24-26.

BERTRAND, D., FLUSS, J., BILLARD, C. & ZIEGLER, J.C. (2010). Efficacité, sensibilité, spécificité: comparaison de différents tests de lecture [Efficiency, sensitivity, specificity: Comparison of different reading tests]. *Année psychologique, 110*, 299-320.

BILLARD, C., FLUSS, J., DUCOT, B., BRICOUT, L., RICHARD, G., ECALLE, J., . . . ZIEGLER, J.C. (2009). Deficits in reading acquisition in primary school: Cognitive, social and behavioral factors studied in a sample of 1062 children. *Revue D Epidemiologie Et De Sante Publique, 57*(3), 191-203. doi:10.1016/j.respe.2009.02.205

BILLARD, C., FLUSS, J., DUCOT, B., WARSZAWSKI, J., ÉCALLE, J., MAGNAN, A., . . . ZIEGLER, J.C. (2008). Study of causal factors of reading impairment in a sample of 1062 7 to 8-years-old children. *Archives de Pédiatrie, 15*(6), 1058-1067. doi:10.1016/j.arcped.2008.02.020

BOLGER, D. J., PERFETTI, C. A., & SCHNEIDER, W. (2005). Cross-cultural effect on the brain revisited: Universal structures plus writing system variation. *Human brain mapping*, *25*(1), 92-104.

BOSSE, M.L, TAINTURIER, M.J. & VALDOIS, S. (2007). Developmental dyslexia: the visual attention span deficit hypothesis. *Cognition, 104* (2), 198-230.

CAMPUZANO, L., DYNARSKI, M., AGODINI, R., & RALL, K. (2009). Effectiveness of reading and mathematics software products: Findings from two student cohorts (NCEE 2009-4041). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

CASTLES, A. & NATION, K. (2006). How does orthographic learning happen? In ANDREWS S. (Ed.), *From inkmarks to ideas: Current issues in lexical processing* (pp. 151-179). Hove: Psychology Press.

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.

CATTS, H.W., HERRERA, S., NIELSEN, D.C., & BRIDGES, M.S. (2015). Early prediction of reading comprehension within the simple view framework. *Reading and Writing*, *28*(9), 1407-1425.

CAVALLI, E., COLÉ, P., LELOUP, G., PORACCHIA-GEORGE, F., SPRENGER-CHAROLLES, L. & EL AHMADI, A. (2017). Screening for Dyslexia in French-Speaking University Students: An Evaluation of the Detection Accuracy of the Alouette Test. *Journal of Learning Disabilities*, 22219417704637.

CEDRE. (2015). « Nouvelle évaluation en fin de collège : compétences langagières et littératie. Évaluation des acquis des élèves », *Note d'information* 21, juillet 2016, Direction de l'Évaluation, de la Prospective et de la Performance (DEPP).

CHEUNG, A.C., & SLAVIN, R.E. (2011). The Effectiveness of Education Technology for Enhancing Reading Achievement: A Meta-Analysis. *Center for Research and reform in Education*.

CHEUNG A.C. & SLAVIN R.E. (2016). How methodological features affect effect sizes in Education. *Educational Researcher* 45(5), 283-292.

CLASSIFICATION STATISTIQUE INTERNATIONALE DES MALADIES ET DES PROBLEMES DE SANTE CONNEXES. DIXIEME REVISION (CIM 10) (1993). Organisation Mondiale de la Santé. Genève.

COHEN, L., DEHAENE, S., NACCACHE, L., LEHERICY, S., DEHAENE-LAMBERTZ, G., HENAFF, M.A., & MICHEL, F. (2000). The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain*, *123*(2), 291-307.

COHEN, L., LEHÉRICY, S., CHOCHON, F., LEMER, C., RIVAUD, S., & DEHAENE, S. (2002). Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain*, *125*(5), 1054-1069.

COLLET, G., COLIN, C., SERNICLAES, W., HOONHORST, I., MARKESSIS, E., DELTENRE, P., & LEYBAERT, J. (2012). Effect of phonological training in French children with SLI: perspectives on voicing identification, discrimination and categorical perception. *Research in developmental disabilities*, *33*(6), 1805-1818

COLLIS, N.L., KOHNEN, S. & KINOSHITA, S. (2013). The role of visual spatial attention in adult developmental dyslexia. *Quarterly Jour- nal of Experimental Psychology (Hove), 66* (2), 245-260.

DANON-BOILEAU L. & BARBIER D. (2001). *Play-On : un logiciel d'entraînement à la lecture*. Paris : Audivimédia.

DAUSSIN, J.M., KESKPAIK, S., & ROCHER, T. (2011). L'évolution du nombre d'élèves en difficulté face à l'écrit depuis une dizaine d'années. *France, portrait social*, 137-152.

DAVIES, P. (1999). What is Evidence-based Education? *British Journal of Educational Studies, 47* (2), 108-121.

DE AGUILERA M. & MENDIZ NOGUERO A. (2003). Video games and education: (Education in the face of a 'parallel school'). *Computers in Entertainment* 1 (1), 10-14.

DE CARA, B., & PLAZA, M. (2010). Les outils informatisés d'aide à la lecture : un bilan des recherches. *ANAE, 107*, *108*, 184-190.

DEDOUBLEMENT DES CLASSES DE CP EN EDUCATION PRIORITAIRE RENFORCEE : PREMIERE EVALUATION. (2020). Consulté à l'adresse : https://www.education.gouv.fr/dedoublement-des-classes-de-cp-en-education-prioritaire-renforcee-premiere-evaluation-11879

DEHAENE S. (2007). *Les neurones de la lecture*. Paris: Odile Jacob.

DEHAENE-LAMBERTZ, G., MONZALVO, K., & DEHAENE, S. (2018). The emergence of the visual word form: Longitudinal evolution of category-specific ventral visual areas during reading acquisition. *PLoS biology*, *16*(3), e2004103.

DELTOUR, J.J. & HUPKENS, D. (1984*). Test de vocabulaire actif et passif pour enfants (5 à 8 ans) (TVAP).* Braine-le-Château : éditions de l'Application des techniques modernes.

DONNELLY, P.M., GIJBELS, L., LARSON, K., MATSKEWICH, T., LINNERUD, P., KUHL, P.K., & YEATMAN, J.D. (2020). A symbolic annotation of vowel sounds for emerging readers. https://doi.org/10.31234/osf.io/akjdr

DYNARSKI, M., AGODINI, R., HEAVISIDE, S., NOVAK, T., CAREY, N., CAMPUZANO, L., . . . SUSSEX, W. (2007). *Effectiveness of Reading and Mathematics Software Products: Findings from the First Student Cohort*. Research report - Report number NCEE 2007-4005. Retrieved from https://telearn.archives-ouvertes.fr/hal-00190019

ÉCALLE J. (2003). Développement des processus d'identification de mots écrits : une étude transversale entre 6 et 8 ans. *Rééducation Orthophonique* 213, 77-96.

ÉCALLE J., MAGNAN A. & JABOULEY D. (2010). *Chassymo : un logiciel d'aide au traitement syllabique.* Châteauroux : Adeprio Diffusion.

ÉCALLE, J. (2003). *Timé-2 : test d'identification de mots écrits pour enfants de 6 à 8 ans.* Paris : Centre de psychologie appliquée.

ÉCALLE, J., BOUCHAFA, H., POTOCKI, A., & MAGNAN, A. (2013). Comprehension of written sentences as a core component of children's reading comprehension. *Journal of Research in Reading*, *36*(2), 117-131.

ÉCALLE, J., KLEINSZ, N. & MAGNAN, A. (2013). Computer-assisted learning in young poor readers: The effect of grapho-syllabic training on the development of word reading and reading comprehension. *Computers in Human Behavior, 29* (4), 1368-1376.

ÉCALLE, J., NAVARRO, M., SUAREZ-LABAT, H., GOMES, C., CROS, L., & MAGNAN, A. (2016). Concevoir des applications sur tablettes tactiles pour stimuler l'apprentissage de la lecture : avec quelles hypothèses scientifiques?. *Sciences et Technologies de l'Information et de la Communication pour l'Éducation et la Formation*, *23*(2), 33-56.

EDEN, G.F., JONES, K.M., CAPPELL, K., GAREAU, L., WOOD, F.B., ZEFFIRO, T.A. & FLOWERS, D.L. (2004). Neural Changes following Remediation in Adult Developmental Dyslexia. *Neuron, 44* (3), 411-422.

EHRI, L.C. (1995). Phases of development in learning to read words by sight. *Journal of Research in Reading, 18*(2), 116-125.

EHRI, L.C., & MCCORMICK, S. (1998). Phases of word learning: Implications for instruction with delayed and disabled readers. *Reading & Writing Quarterly, 14*(2), 135-163.

ELBRO, C. & PETERSEN, D.K. (2004). Long-Term Effects of Phone- me Awareness and Letter Sound Training: An Intervention Study With Children at Risk for Dyslexia. *Journal of Educational Psychology, 96* (4), 660-670.

EVALAIDE (2019). Evaluer pour mieux aider : un dispositif scientifique de prévention des difficultés en lecture et en mathématiques au CP et au CE1. Rédaction réalisée sous la direction de L.Sprenger-Charolles à partir (pour la partie lecture) de contributions d'elle-même ainsi que de S.Andreu, M.Bianco, S.Dehaene, M.Fayol, F.Ramus et J.Ziegler.

FERRAND, L., LÉTÉ, B. & THEVENOT, C. (2018). *Psychologie cognitive des apprentissages scolaires*. Paris : Dunod.

FLUSS, J., ZIEGLER, J.C., WARSZAWSKI, J., DUCOT, B., RICHARD, G., & BILLARD, C. (2009). Poor Reading in French Elementary School: The Interplay of Cognitive, Behavioral, and Socioeconomic Factors. *Journal of Developmental and Behavioral Pediatrics, 30*(3), 206-216.

FLUSS, J., ZIEGLER, J.C., ECALLE, J., MAGNAN, A., WARSZAWSKI, J., DUCOT, B., ... & BILLARD, C. (2008). Prévalence des troubles d'apprentissages du langage écrit en début de scolarité : l'impact du milieu socioéconomique dans 3 zones d'éducations distinctes. *Archives de pédiatrie*, *15*(6), 1049-1057.

FRANCESCHINI, S., GORI, S., RUFFINO, M., VIOLA, S., MOLTENI, M. & FACOETTI, A. (2013). Action Video Games Make Dyslexic Child- ren Read Better. *Current Biology, 23* (6), 462-466.

FRITH, U. (1985). The usefulness of the concept of unexpected reading failure: Comments on « Reading retardation revisited. ». *British Journal of Developmental Psychology, 3*, 15-17.

GENTAZ, E., SPRENGER-CHAROLLES, L., THEUREL, A., & COLÉ, P. (2013). Reading comprehension in a large cohort of French first graders from low socio-economic status families: A 7-month longitudinal study. *PLoS One*, *8*(11), e78608

GIRARD, C., ÉCALLE, J. & MAGNAN, A. (2013). Serious games as new educational tools: how effective are they? A meta-analysis of recent studies. *Journal of Computer Assisted Learning, 29* (3), 207-219.

GOIGOUX, R. (2000). Apprendre à lire à l'école : les limites d'une approche idéovisuelle. *Psychologie française*, *45*(3), 233-244.

GOUGH, P.B. & TUNMER, W.E. (1986). Decoding, Reading, and Reading Disability. *Remedial and Special Education, 7* (1), 6-10.

HABIB, M. (2018). *La constellation des dys: bases neurologiques de l'apprentissage et de ses troubles*. Paris-Bruxelles : De Boeck Supérieur.

HABIB, M. & ZIEGLER, J.C. (2016). Dyslexie et troubles apparentés : une revue critique de 15 ans de recherche scientifique Perspectives Thérapeutiques. In PINTO, S. & SATO, L. (Eds.), *Traité de Neurolinguistique* (pp 329-336). Paris : De Boeck Supérieur.

HABIB, M., & JOLY-POTTUZ, B. (2008). Dyslexie, du diagnostic à la thérapeutique : un état des lieux. *Revue de Neuropsychologie*, *18*(4), 247-325.

HARM, M.W., & SEIDENBERG, M.S. (1999). Phonology, reading acquisition, and dyslexia: insights from connectionist models. *Psychological review*, *106*(3), 491.

HECK, R.H, THOMAS, S.L., & TABATA, L.N. (2009). An Introduction to Multilevel Modeling Techniques (Quantitative Methodology Series). Mahwah, NJ: Erlbaum.

HOLOPAINEN, L., AHONEN, T., & LYYTINEN, H. (2001). Predicting delay in reading achievement in a highly transparent language. *Journal of Learning Disabilities, 34*(5), 401-413.

HOOVER, W.A., & GOUGH, P.B. (1990). The simple view of reading. *Reading and writing*, *2*(2), 127-160.

HUEMER, S., LANDERL, K., ARO, M., & LYYTINEN, H. (2008). Training reading fluency among poor readers of German: Many ways to the goal. *Annals of Dyslexia*, *58*(2), 115–137. https://doi.org/10.1007/s11881-008-0017-2

JONES, C. & REUTZEL, R. (2012). Enhanced alphabet knowledge instruction: Exploring a change of frequency, focus, and distributed cycles of review. *Reading Psychology*, *33*(5), 448–464.

KYLE, F., KUJALA, J., RICHARDSON, U., LYYTINEN, H., & GOSWAMI, U. (2013). Assessing the Effectiveness of Two Theoretically Motivated Computer-Assisted Reading Interventions in the United Kingdom: GG Rime and GG Phoneme. *Reading Research Quarterly, 49*, 59–74.

LANDERL, K., RAMUS, F., MOLL, K., LYYTINEN, H., LEPPÄNEN, P.H., LOHVANSUU, K., ... & KUNZE, S. (2013). Predictors of developmental dyslexia in European orthographies with varying complexity. *Journal of Child Psychology and Psychiatry*, *54*(6), 686-694.

LASSAULT, J., & ZIEGLER, J.C. (2018). Les outils numériques d'aide à l'apprentissage de la lecture. *Langue française*, (3), 111-121.

LECOCQ, P. (1996). *L'ECOSSE* Villeneuve d'Ascq*: Presses Universitaires du Septentrion*.

LEFAVRAIS, J. (1965). *Test de l'Alouette (révision 2005)*. Paris : ECPA.

LETE B., SPRENGER-CHAROLLES, L., & COLÉ, P. (2004). Manulex: A lexical database from French readers. Behavioral Research Methods. *Instruments and Computers, 36*(1), 156-166.

LIBERMAN A.M., COOPER F.S., SHANKWEILER D.P., & STUDDERT-KENNEDY M. (1967). Perception of the speech code. *Psychol. Rev.* 74:431-61

LIBERMAN A.M., & WHALEN D.H. (2000). On the relation of speech to language. *Trends Cogn. Neurosci.* 4:187-96

LYYTINEN, H., ERSKINE, J., HAMALAINEN, J., TORPPA, M., & RONIMUS, M. (2015). Dyslexia-Early Identification and Prevention: Highlights from the Jyvaskyla Longitudinal Study of Dyslexia. *Curr Dev Disord Rep, 2*(4), 330-338. doi:10.1007/s40474-015-0067-1

LYYTINEN, H., ERSKINE, J., KUJALA, J., OJANEN, E., & RICHARDSON, U. (2009). In search of a science-based application: A learning tool for reading acquisition. *Scandinavian Journal of Psychology, 50*(6), 668-675. doi:10.1111/j.1467-9450.2009.00791.x

MAGNAN A., ÉCALLE, J., VEUILLET, E., & COLLET, L. (2004). The effects of an audio-visual training program in dyslexic children. *Dyslexia* 10 (2), 131-140.

MATHÔT, S., SCHREIJ, D., & THEEUWES, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior research methods*, *44*(2), 314-324.

MCTIGUE, E. M., SOLHEIM, O.J., ZIMMER, W.K., & UPPSTAD, P.H. (2020). Critically reviewing GraphoGame across the world: Recommendations and cautions for research and implementation of computer-assisted instruction for word-reading acquisition. *Reading Research Quarterly*, *55*(1), 45-73.

MENGHINI, D., FINZI, A., BENASSI, M., BOLZANI, R., FACOET- TI, A., GIOVAGNOLI, S., & VICARI, S. (2010). Different underlying neurocognitive deficits in developmental dyslexia: a comparative study. *Neuropsychologia, 48* (4), 863-872.

MIODUSER D., TUR-KASPA H., & LEITNER I. (2000), "The learning value of computer-based instruction of early reading skills", *Journal of Computer Assisted Learning* 16 (1), 54-63.

MORAIS, J. (1994). *L'art de lire*. Paris: Odile Jacob.

NATIONAL READING PANEL. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups* Washington, DC: National Institute of Child Health and Human Development.

NORTON, E.S., BEACH, S.D., & GABRIELI, J.D. (2015). Neurobiology of dyslexia. *Current opinion in neurobiology*, *30*, 73-78.

OCDE. (2014). *PISA 2012 Results: What Students Know and Can Do - Student Performance in Mathematics, Reading and Science*, vol. I, PISA, OECD Publishing.

PEERMAN R., SPRENGER-CHAROLLES, L., & MESSAOUD-GALUSI, S. (2013). The contribution of morphology to the consistency of spelling-to-sound relations: A quantitative analysis based on French elementary school readers. *L'Année psychologique, 113*(1), 3-33.

PENNINGTON, B.F. (2006). From single to multiple deficit models of developmental disorders. *Cognition, 101* (2), 385-413.

PERFETTI, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific studies of reading*, *11*(4), 357-383.

PERFETTI, C.A., BECK, I., BELL, L.C., & HUGHES, C. (1987). Phonemic knowledge and learning to read are reciprocal: A longitudinal study of first grade children. *Merrill-Palmer Quarterly (1982-)*, 283-319.

PINKER, S. (2003). *The language instinct: How the mind creates language*. Penguin UK.

PRABHAKARAN, V., SMITH, J.A., DESMOND, J.E., GLOVER, G.H., & GABRIELI, J.D. (1997). Neural substrates of fluid reasoning: an fMRI study of neocortical activation during performance of the Raven's Progressive Matrices Test. *Cognitive Psychology, 33* (1), 43-63.

RAVEN, J.C. (1976). *Standard progressive matrices: Sets A, B, C, D & E*. Oxford: Psychologists Press.

RICHARDSON, U., & LYYTINEN, H. (2014). The GraphoGame method: The theoretical and methodological background of the technology-enhanced learning environment for learning to read. *Human Technology*, *10*(1).

ROMERO M. (2016), *Jeux numériques et apprentissages*, Paris, Éditions JFD.

RONIMUS, M., & LYYTINEN, H. (2015). Is school a better environment than home for digital game-based learning? The case of GraphoGame. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments*.

RUECKL, J.G., PAZ-ALONSO, P.M., MOLFESE, P.J., KUO, W.J., BICK, A., FROST, S.J., ... & LEE, J.R. (2015). Universal brain signature of proficient reading: Evidence from four contrasting languages. *Proceedings of the National Academy of Sciences*, *112*(50), 15510-15515.

RUIZ, J.P., LASSAULT, J., SPRENGER-CHAROLLES, L., RICHARDSON, U., LYYTINEN, H., & ZIEGLER, J.C. (2017). GraphoGame: un outil numérique pour enfants en difficultés d'apprentissage de la lecture. *Approche Neuropsychologique des Apprentissages chez l'Enfant (A.N.A.E.),* 148, 333-343.

SAINE, N.L., LERKKANEN, M.K., AHONEN, T., TOLVANEN, A., & LYYTINEN, H. (2011). Computer-Assisted Remedial Reading Intervention for School Beginners at Risk for Reading Disability. *82*(3), 1013-1028. doi:10.1111/j.1467-8624.2011.01580.x

SAKSIDA, A., IANNUZZI, S., BOGLIOTTI, C., CHAIX, Y., DEMO- NET, J.F., BRICOUT, L., & RAMUS, F. (2016). Phonological skills, visual attention span, and visual stress in developmental dyslexia. *Developmental Psychology, 52* (10), 1503-1516.

SANDERS, W.L., WRIGHT, S.P., & HORN, S.P. (1997). Teacher and Classroom Context Effects on Student Achievement: Implications for Teacher Evaluation. *Journal of Personnel Evaluation in Education, 11*(1), 57-67.

SAYGIN, Z.M., OSHER, D.E., NORTON, E.S., YOUSSOUFIAN, D.A., BEACH, S.D., FEATHER, J., ... & KANWISHER, N. (2016). Connectivity precedes function in the development of the visual word form area. *Nature neuroscience*, *19*(9), 1250-1255.

SEIDENBERG, M.S., & MCCLELLAND, J.L. (1989). A distributed, developmental model of word recognition and naming. *Psychological review*, *96*(4), 523.

SEYMOUR, P.H., ARO, M., ERSKINE, J.M., & COLLABORATION WITH COST ACTION A8 NETWORK. (2003). Foundation literacy acquisition in European orthographies. *British Journal of psychology*, *94*(2), 143-174.

SHARE, D.L. (1995). Phonological recoding and self-teaching: Sine qua none of reading acquisition. *Cognition*, *55*(2), 151-218.

SHARE, D.L. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology, 72*(2), 95-129. doi:10.1006/jecp.1998.2481

SIEGEL, L.S., SHARE, D., & GEVA, E. (1995). Evidence for superior orthographic skills in dyslexics. *Psychological Science, 6* (4), 250-254.

SIMOS, P.G., FLETCHER, J.M., BERGMAN, E., BREIER, J.I., FOORMAN, B.R., CASTILLO, E.M., & PAPANICOLAOU, A.C. (2002). Dyslexia-specific brain activation profile becomes normal following successful remedial training. *Neurology*, *58*(8), 1203-1213.

SLAVIN, R.E., LAKE, C., DAVIS, S., & MADDEN, N.A. (2011). Effective programs for struggling readers: A best-evidence synthesis. *Educational Research Review, 6*(1), 1-26.

SPRENGER-CHAROLLES, L. (2017). Une progression pédagogique construite à partir de statistiques sur l'orthographe du français (d'après Manulex-Morpho) : pour les lecteurs débutants et

atypiques. *Approche Neuropsychologique des Apprentissages chez l'Enfant (A.N.A.E)*, *148*, 247-256.

SPRENGER-CHAROLLES, L., & COLÉ, P. (2013). Lecture et dyslexie (2ème édition, revue et actualisée). Paris: Dunod.

SPRENGER-CHAROLLES, L., & CASALIS, S. (1996). *Lire. Lecture et écriture : acquisition et troubles du développement*. Paris: Presses universitaires de France.

SPRENGER-CHAROLLES, L., & BONNET, P. (1996). New doubts on the importance of the logographic stage. *Current Psychology Of Cognition*, *15*, 173-208.

SPRENGER-CHAROLLES, L., & ZIEGLER, J.C. (2019). Apprendre à lire : contrôle, automatismes et auto-apprentissage. In A. Bentollila & B. Germain, editor, *L'apprentissage de la lecture*. Nathan, Septembre.

SPRENGER-CHAROLLES, L., COLÉ, P., BÉCHENNEC, D., & KIPFFER-PIQUARD, A. (2005). French normative data on reading and related skills from EVALEC, a new computerized battery of tests (end Grade 1, Grade 2, Grade 3, and Grade 4). *European Review of Applied Psychology*, *55*(3), 157-186.

SPRENGER-CHAROLLES, L., SIEGEL, L.S., JIMENEZ, J.E., & ZIEGLER, J.C. (2011). Prevalence and Reliability of Phonological, Surface, and Mixed Profiles in Dyslexia: A Review of Studies Conducted in Languages Varying in Orthographic Depth. *Scientific Studies of Reading, 15* (6), 498-521.

STEIN, J. (2014). Dyslexia: The Role of Vision and Visual Attention. *Current Developmental Disorders Reports, 1* (4), 267-280.

SUCHAUT B., BOUGNERES A., & BOUGUEN A. (2014). Sept minutes pour apprendre à lire : à la recherche du temps perdu. *Document de travail,* École d'économie de Paris, Institut des Politiques Publiques.

SUNDE, K., FURNES, B., & LUNDETRÆ, K. (2020). Does Introducing the Letters Faster Boost the Development of Children's Letter Knowledge, Word Reading and Spelling in the First Year of School?. *Scientific Studies of Reading*, *24*(2), 141-158

TARE, M., SHELL, A.R., & JACKSON, S.R. (2020). Student engagement with evidence-based supports for literacy on a digital platform. *Journal of Research on Technology in Education*, 1-11.

TAYLOR, D.B., HANDLER, L.K., FITZPATRICK, E., & WHITTINGHAM, C.E. (2020). The device in the room: Technology's role in third grade literacy instruction. *Journal of Research on Technology in Education*, 1-19.

TORGESEN, J.K., WAGNER, R.K., RASHOTTE, C.A., HERRON, J., & LINDAMOOD, P. (2010). Computer-assisted instruction to prevent early reading difficulties in students at risk for dyslexia: Outcomes from two instructional approaches. *Annals of dyslexia*, *60*(1), 40-56.

VADASY, P.F., JENKINS, J.R., ANTIL, L.R., WAYNE, S.K., & O'CONNOR, R.E. (1997). The effectiveness of one-to-one tutoring by community tutors for at-risk beginning readers. Learning Disability Quarterly, 20 (1).

VADASY, P.F., JENKINS, J.R., & POOL, K. (2000). Effects of tutoring in phonological and early reading skills on students at risk for reading disabilities. Journal of Learning Disabilities, 33, 579-590

VALDOIS, S. (2017). Entraîner l'attention visuelle pour remédier aux troubles de la lecture. *ANAE*, *148*, 1-11.

VAN BERGEN, E., VAN DER LEIJ, A., & DE JONG, P.F. (2014). The intergenerational multiple deficit model and the case of dyslexia. *Frontiers in human neuroscience, 8* (346).

VAN NIEUWENHOVEN, C., NOËL, M.P., & GREGOIRE, J. (2001). *Tedi-math : test diagnostique des compétences de base en mathématiques*. Paris : ECPA.

VIDYASAGAR, T.R., & PAMMER, K. (2010). Dyslexia: a deficit in visuo-spatial attention, not in phonological processing. *Trends in Cognitive Sciences, 14* (2), 57-63.

WATKINS, C. P., CAPORAL, J., MERVILLE, C., KOUIDER, S., & DEHAENE, S. (2020). Accelerating reading acquisition and boosting comprehension with a cognitive science-based tablet training. *Journal of Computers in Education*, 1-30

WIMMER, H., & HUMMER, P. (1990). How German-speaking first graders read and spell: Doubts on the importance of the logographic stage. *Applied Psycholinguistics, 11*(4), 349-368. doi:10.1017/S0142716400009620

WOUTERS, P., VAN NIMWEGEN, C., VAN OOSTENDORP, H., & VAN DER SPEK, E.D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology, 105*(2), 249-265. doi:10.1037/a0031311

WOUTERS, P., & VAN OOSTENDORP, H. (2013). A meta-analytic review of the role of instructional support in game-based learning. *Computers & Education*, *60*(1), 412–425. https://doi.org/10.1016/j.compedu.2012.07.018

ZIEGLER, J.C. (2018). Différences inter-linguistiques dans l'apprentissage de la lecture. *Langue française*, *199*(3), 35-49.

ZIEGLER, J.C., BERTRAND, D., LÉTÉ, B., & GRAINGER, J. (2014). Orthographic and phonological contributions to reading development: Tracking developmental trajectories using masked priming. *Developmental Psychology, 50* (4), 1026-1036.

ZIEGLER, J.C., BERTRAND, D., TÓTH, D., CSÉPE, V., REIS, A., FAÍSCA, L., ... & BLOMERT, L. (2010). Orthographic depth and its impact on universal predictors of reading: A cross-language investigation. *Psychological science*, *21*(4), 551-559.

ZIEGLER, J.C., CASTEL, C., PECH-GEORGEL, C., GEORGE, F., ALARIO, F.X., & PERRY, C. (2008). Developmental dyslexia and the dual route model of reading: Simulating individual differences and subtypes. *Cognition, 107*, 151-178. doi:10.1016/j.cognition.2007.09.004

ZIEGLER, J.C., & DELEUZE, A. (2018). Lire. In A. HOUDE, O. & BORST, G. (Eds), *Le cerveau et les apprentissages* (pp.47-68). Paris : Nathan.

ZIEGLER, J.C., & GOSWAMI, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychological bulletin*, *131*(1), 3.

ZIEGLER J.C. & GOSWAMI U. (2006). Becoming literate in different languages: similar problems, different solutions. *Developmental Science, 9*(5), 429-436.

ZIEGLER, J.C., PECH-GEORGEL, C., DUFAU, S., & GRAINGER, J. (2010). Rapid processing of letters, digits, and symbols: What purely visual-attentional deficit in developmental dyslexia? *Developmental Science, 13*, F8-F14.

ZIEGLER, J.C., PERRY, C., & ZORZI, M. (2014). Modelling reading development through phonological decoding and self-teaching: implications for dyslexia. *Philos Trans R Soc Lond B Biol Sci, 369*(1634), 20120397. doi:10.1098/rstb.2012.0397

ZIEGLER, J.C., PERRY, C., & ZORZI, M. (2019). Modeling the Variability of Developmental Dyslexia. In C. Perfetti, K. Pugh, & L. Verhoeven (Eds.), *Developmental Dyslexia across Languages and Writing Systems* (pp. 350-371). Cambridge: Cambridge University Press

ZIEGLER, J.C., PERRY, C., & ZORZI, M. (2020). Learning to Read and Dyslexia: From Theory to Intervention Through Personalized Computational Models. *Current Directions in Psychological Science, 29*(3), 293-300. doi:10.1177/0963721420915873

ZORZI, M., BARBIERO, C., FACOETTI, A., LONCIARI, I., CARROZZI, M., MONTICO, M. & ZIEGLER, J.C. (2012). Extra-large letter spacing improves reading in dyslexia. *Proceedings of the National Academy of Sciences, 109* (28), 11455-11459.

Annexes

Annexe 1

Lassault, J., & Ziegler, J. C. (2018). Les outils numériques d'aide à l'apprentissage de la lecture. *Langue française*, (3), 111-121.

Annexe 2

Ruiz, J. P., Lassault, J., Sprenger-Charolles, L., Richardson, U., Lyytinen, H., & Ziegler, J. C. (2017). GraphoGame : un outil numérique pour enfants en difficultés d'apprentissage de la lecture. *Approche Neuropsychologique des Apprentissages chez l'Enfant (A.N.A.E.),* 148, 333-343.

Annexe 3

Lassault, J., Sprenger-Charolles, L., Albrand, J.P., Alavoine, E., Richardson, U., Lyytinen, H. & Ziegler, J.C. (2020). Positive effects of GraphoGame reading intervention in French: the importance of initial reading level and engagement. Article soumis pour publication.

Julie Lassault
Aix-Marseille Université & CNRS, Laboratoire de Psychologie Cognitive (UMR 7290) & Fédération de recherche 3C (Marseille)

Johannes C. Ziegler
Aix-Marseille Université & CNRS, Laboratoire de Psychologie Cognitive (UMR 7290) & Fédération de recherche 3C (Marseille)

Les outils numériques d'aide à l'apprentissage de la lecture

1. INTRODUCTION [1]

La lecture est le pré-requis indispensable pour accéder à l'ensemble des domaines de l'apprentissage. Le rapport de l'Organisation de Coopération et de Développement Économique (OCDE) montre qu'au niveau international, la France se positionne au 21e rang (OCDE 2014). Ce classement « moyen » révèle une autre réalité : les élèves les plus performants ont de meilleurs résultats, alors que les résultats des élèves les moins performants ne cessent de diminuer, faisant de la France « la championne des inégalités » (Battaglia & Collas 2013). De plus, les études épidémiologiques montrent que le pourcentage d'élèves en grande difficulté de lecture atteint 30 % dans les milieux socio-économiques défavorisés (Billard *et al.* 2009 ; Fluss *et al.* 2009).

Les prises en charge des difficultés de lecture au sein de l'école semblent trop insuffisantes pour réduire le pourcentage élevé d'élèves (20-40 %) en difficulté de compréhension d'un texte à l'entrée du collège (CEDRE 2015). Quant aux prises en charge en dehors de l'école, elles se résument souvent à un bilan et des séances d'orthophonie à raison de 30 minutes par semaine. L'efficacité de cette prise en charge ne fait pas l'objet d'évaluations systématiques. Face à l'ampleur

1. Nous remercions Édouard Alavoine, Jean-Patrice Albrand, Liliane Sprenger-Charolles et Alain Desrochers pour leurs commentaires, suggestions et leur aide précieuse tout au long de ce projet. Julie Lassault bénéficie d'un contrat doctoral du Programme Investissements d'Avenir Action « Innovation numérique pour l'excellence éducative », volet « Espaces de formation, de recherche et d'animation numérique » (e-FRAN, projet LEMON). La préparation de cet article a bénéficié d'un financement de l'Agence National de la Recherche [ANR-13-APPR-0003], du soutien du Labex BLRI [ANR-11-LABX-0036], de l'Institut Convergence ILCB [ANR-16-CONV-0002] et de l'Initiative d'Excellence d'Aix-Marseille Université [A*MIDEX].

du problème, le numérique peut offrir aujourd'hui des solutions « de première intention » au sein de l'école.

2. APPRENTISSAGE DE LA LECTURE ET OUTILS NUMÉRIQUES

2.1. Apprentissage de la lecture

L'Observatoire National de la Lecture définit l'apprentissage de la lecture comme la capacité à « développer des habiletés dans deux domaines : l'identification des mots écrits, et le traitement du sens, pour la compréhension des textes » (ONL 1998). Cette capacité suppose des habiletés cognitives, sociales et linguistiques particulières, qui se développent depuis le plus jeune âge (Fayol & Morais 2004).

Pour comprendre un texte, l'élève doit d'abord être capable de décoder, *i.e.* d'établir les relations qui existent entre ce qu'il voit (graphèmes, syllabes, mots, etc.) et ce qu'il entend (phonèmes, syllabes, mots, etc.). C'est ce mécanisme qui va lui permettre de déchiffrer des mots qu'il n'aura jamais rencontrés, mais dont il connait déjà la forme phonologique et la signification avant l'apprentissage de la lecture, grâce à l'acquisition du langage oral (Ziegler, Perry & Zorzi 2014). La mise en route de ce mécanisme de décodage nécessite en premier lieu un apprentissage explicite des règles et associations simples et fréquentes permettant de transformer les graphèmes en phonèmes, mécanisme qui, d'après le modèle J. C. Ziegler, C. Perry et M. Zorzi (2014) devient progressivement automatique. Selon ce modèle, l'entrainement des processus de décodage est fondamental dans l'apprentissage de la lecture, ce qui fait aujourd'hui consensus (cf. le rapport de la conférence du consensus ; CNESCO 2016).

En parallèle, des études ont mis en évidence l'importance, au moment de l'apprentissage du décodage, du langage oral et, plus particulièrement, de la capacité à identifier et à manipuler les unités phonologiques de la langue, c'est ce que l'on appelle la « conscience phonologique » (pour une synthèse en français, v. Sprenger-Charolles & Colé 2013). L'élève apprend au fur et à mesure à reconnaitre les graphèmes qui composent les mots et prend conscience des unités phonologiques qui les composent.

Selon le système orthographique de la langue dans laquelle un élève apprend à lire, le décodage est efficace plus ou moins rapidement (v. Ziegler 2018 dans ce numéro). Lorsqu'il est mis en place, c'est le moment où l'élève peut entrer dans le cercle vertueux de « l'auto-apprentissage » (Share 1995, 2004). En effet, à chaque décodage réussi, *i.e.* à chaque fois que le mot décodé active en mémoire la bonne forme phonologique, le mécanisme de décodage est renforcé. C'est principalement à ce moment précis que l'utilisation du numérique prend tout son sens, car l'automatisation du décodage demande du temps, de la supervision individuelle et de très nombreuses répétitions.

2.2. Outils numériques d'aide à la lecture

Quels que soient l'âge, le support ou la compétence visée, l'idée qu'un apprentissage ludique est efficace n'est pas nouvelle. L'introduction des nouvelles technologies a permis d'offrir de nouveaux supports pédagogiques (Abt 1987), et l'usage de jeux numériques pour l'apprentissage ne cesse d'augmenter depuis une vingtaine d'années dans de nombreux domaines tels que l'histoire, les mathématiques ou la lecture. Sous l'appellation « jeu numérique », on retrouve à la fois les jeux numériques de divertissement (jeux vidéo) utilisés en contexte éducatif (p. ex. l'utilisation du jeu *Minecraft* en école d'architecture) et les jeux numériques sérieux éducatifs, développés dans une intention pédagogique (Romero 2016).

Les outils numériques permettent de proposer des procédures standardisées, de recueillir précisément les données de chaque utilisateur (le temps de réponse, les erreurs) afin de suivre de près leur évolution. Les propriétés mises en avant par ces outils sont, par exemple, la haute définition graphique, la qualité des stimulations auditives, le retour immédiat sur les performances, l'apprentissage par exploration et interaction active et l'aspect ludique qui permet de maintenir l'attention et la motivation des élèves (Mioduser, Tur-Kaspa & Leitner 2000). Bien que les premières études sur les jeux numériques aient été conduites dans les années 70 aux États-Unis (Ball 1978), celles présentant une évaluation de l'efficacité de ces dispositifs sont apparues plus tardivement (De Aguilera & Mendiz Noguero 2003). En effet, peu d'outils ont été évalués systématiquement dans des études randomisées et contrôlées, sans doute parce que cette méthodologie est assez lourde à mettre en place (Girard, Écalle & Magnan 2013).

Les premiers jeux sérieux ayant fait l'objet de validations scientifiques rigoureuses portaient sur un entrainement des habiletés phonologiques assisté par ordinateur. Par exemple, B. W. Wise, J. Ring et R. K. Olson (1999) ont comparé les effets de trois entrainements phonologiques d'une durée de 40 heures (articulation de phonèmes, manipulation de phonèmes, les deux) chez des enfants à l'école primaire (grade 2 à 5). Ils ont observé des bénéfices significatifs en décodage et en conscience phonémique par rapport à un groupe contrôle sans entrainement informatisé. Chez les enfants d'âge préscolaire, des études randomisées ont été réalisées pour tester l'efficacité d'un entrainement des habilités phonologiques à l'aide de deux jeux d'aventures, *Daisy Quest* et *Daisy Castle* (Foster *et al.* 1994). Les auteurs ont trouvé une augmentation importante (+1 écart-type) des habiletés phonologiques par rapport à un groupe contrôle. Des effets similaires ont été obtenus par T. A. Baker et J. K. Torgesen (1995) pour des enfants à risque de dyslexie après 8 heures d'entrainement phonologique, par rapport à un entrainement contrôle, de même nature, mais dans un autre domaine (les mathématiques). Enfin, en utilisant encore *Daisy Quest* et *Daisy Castle*, C. J. Lonigan *et al.* (2003) ont entrainé des enfants d'âge préscolaire à risque de troubles de la lecture dans 7 compétences phonologiques (p. ex. reconnaitre des mots qui riment ou qui partagent le même phonème au début, au milieu ou à la fin d'un mot). Après un entrainement de 8 semaines (à raison de

4 séances de 15-20 minutes par semaine), les enfants du groupe expérimental ont obtenu de meilleurs scores dans différentes tâches phonologiques (tâche de rime, suppression de phonème) que le groupe contrôle, mais ces enfants n'ont pas été suivis pour savoir si ces effets se généralisaient à la lecture.

Un autre type d'entrainement, également basé sur le développement de la sensibilité phonologique, a été proposé à des enfants américains faibles lecteurs, âgés de 6-7 ans (Macaruso, Hook & McCabe 2006). Cet entrainement, qui comportait de nombreux exercices phonologiques ainsi que d'autres portant sur des stratégies de décodage basées sur le début du mot (l'attaque), a produit des effets significatifs sur le décodage et la reconnaissance des mots, mais uniquement pour les élèves les plus faibles.

En France, l'un des premiers logiciels d'entrainement à la lecture – *Play-On* – a été développé par L. Danon-Boileau et D. Barbier (2001). Ce logiciel porte essentiellement sur la discrimination auditive et visuelle de paires minimales du type /ba/ *vs* /pa/ (voisement). L'écran représente un terrain de basket avec deux paniers (un pour chaque son ; p. ex. /ba/ et /pa/). Un ballon est lancé au milieu de l'écran : l'enfant entend un son (p. ex. /ba/) et doit, à ce moment-là, envoyer le ballon dans le bon panier. Un second jeu reprend la même disposition, mais à chaque panier correspond soit « Oui » soit « Non » ; au jeté de ballon, l'enfant entend un mot suivi d'une syllabe, il doit envoyer le ballon vers le panier « Oui » lorsque la syllabe reprend la fin du mot (*bateau – to*) et vers le panier « Non » lorsque ce n'est pas le cas (*jardin – don*). Ce jeu a été testé auprès de 14 enfants dyslexiques âgés de 8 à 12 ans (Magnan *et al.* 2004). L'entrainement a duré 10 heures réparties sur 5 semaines, à raison de 2 séances par jour de 15 minutes, 4 jours par semaine. Les enfants ont été répartis en 2 groupes de façon aléatoire, l'un bénéficiant de l'entrainement, l'autre non. À l'issue des 5 semaines, les groupes ont été inversés. Les résultats ont mis en évidence un bénéfice significatif de l'entrainement sur la tâche d'identification de mots « Timé 2 » (Écalle 2003).

Un second logiciel d'entrainement de la conscience phonologique *Chassymo* a fait l'objet d'une étude sur 16 enfants faibles lecteurs en classe de Cours Préparatoire (CP) répartis en 2 groupes (Écalle, Magnan & Jabouley 2010). Dans ce jeu, l'enfant entend et voit une syllabe *(gi)*, puis, il entend un mot contenant cette syllabe *(gilet)*, il doit alors retrouver le bon emplacement de la syllabe dans le mot. L'entrainement a duré une dizaine d'heures à raison de 2 séances de 15 minutes par jour, 4 jours par semaine pendant 5 semaines. Les résultats ont permis de mettre en évidence un bénéfice de l'entrainement sur les performances en décodage et en reconnaissance de mots écrits des élèves du groupe expérimental.

3. *GRAPHOLEARN*

Parmi les logiciels d'entrainement à la lecture, c'est surtout *GraphoLearn* qui a fait l'objet des développements et évaluations les plus importants. *GraphoLearn* est un logiciel d'entrainement audio-visuel à la lecture, basé sur la présentation simultanée de stimuli auditifs et de choix orthographiques (Richardson & Lyytinen 2014). Créé en Finlande par H. Lyytinen et son équipe de l'Université de Jyväskylä il y a plus d'une dizaine d'années, d'abord sur ordinateurs et désormais sur tablettes ou tout autre support numérique, il a depuis été adapté dans de nombreux pays à travers le monde.

Le principe du jeu repose sur la présentation simultanée de phonèmes et de graphèmes. L'enfant est équipé d'un casque où il entend les sons, en même temps plusieurs propositions de réponses sont présentées et il a pour tâche de choisir la réponse qui est correcte, *i.e.* celle qui correspond à ce qu'il vient d'entendre. S'il donne la réponse exacte, il passe à l'essai suivant ; en revanche, s'il se trompe, les propositions incorrectes s'effacent, il réentend alors le son et doit choisir la bonne proposition, qui est alors la seule visible. Cette présentation en simultané permet d'apprendre le lien systématique entre graphèmes et phonèmes, lien qui est à la base du décodage et qui est indispensable pour l'apprentissage de la lecture (Ziegler & Goswami 2005). La répétition massive permet, quant à elle, d'automatiser ce processus nécessaire pour la reconnaissance instantanée de mots. Parallèlement, *GraphoLearn*, comme *Play-On*, travaille sur la discrimination phonémique, souvent déficitaire chez les enfants dyslexiques (Serniclaes *et al.* 2001 ; Ziegler *et al.* 2009).

La première étude de validation de *GraphoLearn* a été conduite en 2010 sur 166 enfants finlandais de 7 ans (Saine *et al.* 2011). Au cours des deux premières semaines du grade 1 (qui correspond au CP en France), différentes compétences ont été évaluées (connaissance des lettres, conscience phonologique et dénomination rapide). Deux cohortes d'élèves ont été suivies, l'une bénéficiant de l'entrainement avec *GraphoLearn*, l'autre suivant un entrainement sans logiciel. L'intervention a été réalisée sur une période de 28 semaines, à raison de 4 séances hebdomadaires de 45 minutes (l'entrainement a été donc plus intensif que dans les études présentées précédemment). Les résultats ont montré que les enfants les plus en difficulté ont effectué des progrès significatifs en décodage et atteignaient un niveau de lecture correspondant à la fin de la deuxième année de scolarisation.

Au cours des années, la version finlandaise de *GraphoLearn* a été adaptée dans de nombreuses langues. Comme indiqué dans le Tableau 1, certaines versions ont été testées et validées, alors que d'autres sont encore en expérimentation.

Tableau 1 : Versions de *GraphoLearn* adaptées dans différents pays [+ = versions testées et validées]

Versions	Pays	Validation scientifique
GraphoLearn German	Autriche	
GraphoLearn Portuguese	Brésil	
GraphoLearn Canadian English	Canada	
GraphoLearn Chile	Chili	+
GraphoLearn Pinyin	Chine	+
GraphoLearn Mandarin Chinese	Chine	
GraphoLearn Greek	Chypre	
GraphoLearn Eesti	Estonie	
GraphoLearn US English Rime	États-Unis	+
GraphoLearn US Spanish	États-Unis	+
GraphoLearn French	France	+
GraphoLearn Greek	Grèce	+
GraphoLearn UK Phoneme	Hong-Kong	
GraphoLearn Hungarian	Hongrie	
GraphoLearn Bahasa	Indonésie	+
GraphoLearn Irish English	Irlande	
GraphoLearn Arabic	Israël	
GraphoLearn Kikuyu	Kenya	+
GraphoLearn Swahili	Kenya & Tanzanie	+
GraphoLearn Afrikaans	Namibie	
GraphoLearn Norwegian	Norvège	
GraphoLearn Spanish Peru	Pérou	
GraphoLearn Polish	Pologne	+
GraphoLearn Portuguese	Portugal	+
GraphoLearn UKRime	Royaume-Uni	+
GraphoLearn Russian	Russie	
GraphoLearn Swedish	Suède	+
GraphoLearn German	Suisse	+
GraphoLearn Zhuyin	Taïwan	
GraphoLearn Tonga	Zambie	
GraphoLearn ChiNyanja	Zambie	+

4. VERSION FRANÇAISE DE *GRAPHOLEARN*

L'adaptation française de *GraphoLearn* a tenu compte des caractéristiques principales de l'orthographe de cette langue afin de proposer une progression adaptée. Elle repose sur la base de données Manulex-MorphO (Peereman, Sprenger-Charolles & Messaoud-Galusi 2013), issue de Manulex (Lété, Sprenger-Charolles & Colé 2004) qui recense environ 2 000 000 formes non lemmatisées (l'ensemble des formes lexicales, incluant celles avec des marques morphologiques) contenues dans 54 manuels scolaires de primaire. La base de données Manulex-MorphO fournit la fréquence et la consistance des correspondances graphophonologiques (CGPh) et tient compte de la morphologie (marques de flexions et de dérivations ; Peereman & Sprenger-Charolles 2018, dans ce numéro). La progression proposée dans la version française de *GraphoLearn* repose sur la fréquence et la consistance des CGPh (les CGPh les plus fréquentes et les plus stables sont présentés dès le début). Les marques de flexions et les supports de dérivations portés par les lettres muettes en fin de mots sont introduits plus tard dans le jeu.

La progression « optimale » se découpe en plusieurs séquences de difficulté croissante. Chaque séquence est composée d'une dizaine de niveaux présentés sous forme de mini jeux (cf. Fig. 1). Un mini jeu dure environ 2 à 3 minutes et une séquence environ 20 à 30 minutes. À chaque début et fin de séquence, 2 niveaux identiques sont présentés aux élèves, ce sont des niveaux de pré- et post-tests qui permettent de quantifier le progrès de l'élève à l'issue de chaque séquence. Les 67 séquences proposées sont fixes et la progression dans le jeu ne tient pas compte des difficultés de l'élève à part le fait que, s'il n'atteint pas 80 % de réponses correctes, il doit recommencer le niveau (avec un nombre maximum de répétitions bloqué à 5 fois).

Au sein du jeu, on trouve des séquences spécifiques portant, d'une part, sur des graphèmes pouvant entrainer des confusions visuelles (p. ex. <u>-<n>) ou phonémiques (p. ex. /t/-/d/), d'autre part, sur ceux qui ont une prononciation dépendant du contexte (<c> et <g>) et, enfin, sur les marques morphologiques nominales (<e> et <s> dans *amies*) et verbales (<s> dans *tu chantes*). Afin de ne pas induire la mémorisation de formes orthographiques erronées, les distracteurs sont, le plus souvent (environ 90 %), des mots correctement écrits, mais qui peuvent être plus ou moins proches de la cible (p. ex. *équipe/épique*). Au fur et à mesure de la progression, les jeux de discrimination sont complétés par des jeux de construction de mots et de phrases.

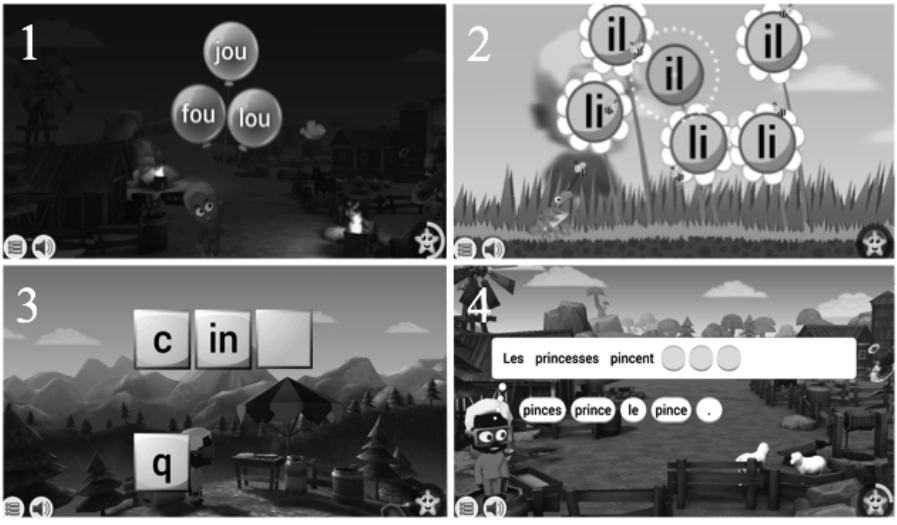


Figure 1 : Quatre exemples de niveaux de la version française de *GraphoLearn* :
1. Niveau classique (l'enfant entend la syllabe /lu/ et choisit l'une des options) ;
2. Niveau de confusion visuelle (il entend le mot « il » et doit le retrouver 3 fois) ;
3. Niveau de formation de mot (il entend le mot « cinq » et doit le former) ;
4. Niveau de formation de phrase (il entend la phrase « Les princesses pincent le prince. » et il doit la construire)

Une première évaluation portant sur la validation de l'outil a été réalisée sur 34 enfants de CP et 35 enfants de CE1 à risque de dyslexie (Ruiz *et al.* 2017). Après une phase de pré-tests (décodage, vocabulaire, etc.), les élèves étaient répartis en 2 groupes (intervention *vs* contrôle). Le protocole était tel qu'un groupe recevait l'intervention pendant 5 semaines (à raison de 3 séances de 20 minutes par session, 3 fois par semaine) pendant que l'autre travaillait sur un logiciel de mathématiques ; à l'issue de ces 5 semaines, les groupes ont été inversés. Les résultats ont permis de montrer une progression en lecture de mots (isolés et en contexte), plus importante après avoir suivi l'intervention qu'après un entrainement numérique aux mathématiques. Ils apportent une première validation de *GraphoLearn* pour les élèves à risque de dyslexie. Une expérimentation longitudinale à grande échelle est actuellement en cours dans 36 classes de CP (soit près de 500 enfants) en réseau d'éducation prioritaire (cf. www.grapholearn.fr).

5. CONCLUSION

Grâce à l'accès de plus en plus facile au numérique au sein de l'école, les outils numériques offrent de réelles pistes « de première intention », *i.e.* dès le plus jeune âge et au sein de l'école, pour réduire les difficultés en lecture des élèves. En

effet, les outils numériques n'ont pas comme vocation de se substituer aux activités pédagogiques de l'enseignant, mais ils peuvent être complémentaires, particulièrement dans les tâches d'apprentissage (comme le décodage) qui nécessitent une supervision individuelle ainsi que, pour s'automatiser, un grand nombre de répétitions.

Le numérique peut également offrir des solutions pour les pays en voie de développement où l'accès à l'éducation reste parfois difficile. Cependant, même dans les pays développés, ces outils peuvent jouer un rôle important, notamment dans une approche de « réponse à l'intervention », qui consiste à proposer des entrainements ciblés à des enfants avant qu'ils ne se trouvent en situation d'échec, et de tester s'ils sont en mesure d'atteindre les cibles attendues grâce à l'intervention (Desrochers, Laplante & Brodeur 2016).

Comme nous l'avons déjà signalé, à ce jour, *GraphoLearn* a été traduit dans une trentaine de langues ; cependant, toutes les versions n'ont pas été validées avec un protocole rigoureux (choix aléatoire des écoles, répartition aléatoire des élèves aux conditions, utilisation d'un logiciel contrôle). Puisqu'il existe une forte corrélation négative entre la taille de l'échantillon et la taille d'effet des interventions, les études avec peu de sujets ont tendance à surestimer les effets (Cheung & Slavin 2016). C'est pourquoi, il est aujourd'hui crucial de tester l'efficacité des outils numériques à grande échelle, dans des études randomisées et contrôlées.

Références bibliographiques

Abt C. C. (1987), *Serious games*, Washington (DC), University Press of America.

Ball H. G. (1978), "Telegames teach more than you think", *Audiovisual Instruction* 23 (5), 24-26.

Barker T. A. & Torgesen J. K. (1995), "An evaluation of computer-assisted instruction in phonological awareness with below average readers", *Journal of Educational Computing Research* 13 (1), 89-103.

Battaglia M. & Collas A. (2013), « Classement PISA : la France championne des inégalités scolaires », *Le Monde*, 03-12-2013.

Billard C. *et al*. (2009), "Deficits in reading acquisition in primary school: Cognitive, social, and behavioral factors studied in a sample of 1062 children", *Revue d'Épidémiologie et de Santé Publique* 57 (3), 191-203.

Cedre (2015), « Cedre 2015. Nouvelle évaluation en fin de collège : compétences langagières et littératie. Évaluation des acquis des élèves », *Note d'information* 21, juillet 2016, Direction de l'Évaluation, de la Prospective et de la Performance (DEPP).

Cheung A. C. & Slavin R. E. (2016), "How methodological features affect effect sizes in Education", *Educational Researcher* 45(5), 283-292.

Cnesco (2016), *Conférence de consensus : « Lire, comprendre, apprendre : comment soutenir le développement de compétences en lecture ? »*, mars 2016. [www.cnesco.fr/events/event/lecture/]

DANON-BOILEAU L. & BARBIER D. (2001), *Play-On : un logiciel d'entraînement à la lecture*, Paris, Audivimédia.

DE AGUILERA M. & MENDIZ NOGUERO A. (2003), "Video games and education: (Education in the face of a 'parallel school')", *Computers in Entertainment* 1 (1), 10-14.

DESROCHERS A., LAPLANTE L. & BRODEUR M. (2016), « Le modèle de la réponse à l'intervention et la prévention des difficultés d'apprentissage de la lecture au préscolaire et au primaire », dans M.-F. Morin, D. Alarmagot & C. Gonçalves (éds), *Perspectives actuelles sur l'apprentissage de la lecture et de l'écriture*, Sherbrooke, Éditions de l'Université de Sherbrooke, 290-314.

ÉCALLE J. (2003), « Développement des processus d'identification de mots écrits : une étude transversale entre 6 et 8 ans », *Rééducation Orthophonique* 213, 77-96.

ÉCALLE J., MAGNAN A. & JABOULEY D. (2010), *Chassymo : un logiciel d'aide au traitement syllabique*, Châteauroux, Adeprio Diffusion.

FAYOL M. & MORAIS J. (2004), « La lecture et son apprentissage » dans *L'évolution de l'enseignement de la lecture en France, depuis dix ans*, Paris, ONL, 13-59.

FLUSS J. *et al.* (2009), "Poor reading in French elementary school: The interplay of cognitive, behavioral, and socioeconomic factors", *Journal of Developmental and Behavioral Pediatrics* 30 (3), 206-216.

FOSTER K. C. *et al.* (1994), "Computer administered instruction in phonological awareness: Evaluation of the DaisyQuest program", *Journal of Research & Development in Education* 27 (2), 126-137.

GIRARD C., ÉCALLE J. & MAGNAN A. (2013), "Serious games as new educational tools: How effective are they? A meta-analysis of recent studies", *Journal of Computer Assisted Learning* 29 (3), 207-219.

LÉTÉ B., SPRENGER-CHAROLLES L. & COLÉ P. (2004), "MANULEX: A grade-level lexical database from French elementary school readers", *Behavior Research Methods, Instruments, and Computers* 36 (1), 156-166.

LONIGAN C. J. *et al.* (2003), "A computer-assisted instruction phonological sensitivity program for preschool children at-risk for reading problems", *Journal of Early Intervention* 25 (4), 248-262.

MACARUSO P., HOOK P. E. & MCCABE R. (2006), "The efficacy of computer-based supplementary phonics programs for advancing reading skills in at-risk elementary students », *Journal of Research in Reading* 29 (2), 162-172.

MAGNAN A. *et al.* (2004), "The effects of an audio-visual training program in dyslexic children", *Dyslexia* 10 (2), 131-140.

MIODUSER D., TUR-KASPA H. & LEITNER I. (2000), "The learning value of computer-based instruction of early reading skills", *Journal of Computer Assisted Learning* 16 (1), 54-63.

OCDE (2014), *PISA 2012 Results: What Students Know and Can Do – Student Performance in Mathematics, Reading and Science*, vol. I, PISA, OECD Publishing.

ONL (1998), *Apprendre à lire*, Paris, CNDP & Odile Jacob.

PEEREMAN R. & SPRENGER-CHAROLLES L. (2018), « Manulex-MorphO, une base de données sur l'orthographe du français intégrant les morpho-phonogrammes », *Langue française* 199. (ce volume)

PEEREMAN R., SPRENGER-CHAROLLES L. & MESSAOUD-GALUSI S. (2013), "The contribution of morphology to the consistency of spelling-to-sound relations: A quantitative analysis based on French elementary school readers", *L'Année psychologique* 113 (1), 3-33.

RICHARDSON U. & LYYTINEN H. (2014), "The GraphoLearn method: The theoretical and methodological background of the technology-enhanced learning environment for learning to read", *Human Technology* 10 (1), 39-60.

ROMERO M. (2016), *Jeux numériques et apprentissages*, Paris, Éditions JFD.

RUIZ J.-P. *et al.* (2017), « GraphoGame : un outil numérique pour enfants en difficultés d'apprentissage de la lecture », *Approche Neuropsychologique des Apprentissages chez l'Enfant (A.N.A.E.)* 148, 333-343.

SAINE N. L. *et al.* (2011), "Computer-assisted remedial reading intervention for school beginners at risk for reading disability", *Child Development* 82 (3), 1013-1028.

SERNICLAES W. *et al.* (2001), "Perceptual discrimination of speech sounds in developmental dyslexia", *Journal of Speech, Language, and Hearing Research* 44 (2), 384-399.

SHARE D. L. (1995), "Phonological recoding and self-teaching: *Sine qua non* of reading acquisition", *Cognition* 55 (2), 151-218.

SHARE D. L. (2004), "Orthographic learning at a glance: On the time course and developmental onset of self-teaching", *Journal of Experimental Child Psychology* 87 (4), 267-298.

SPRENGER-CHAROLLES L. & COLÉ P. (2013[2]), *Lecture et dyslexie : approche cognitive*, Paris, Dunod.

WISE B. W., RING J. & OLSON R. K. (1999), "Training phonological awareness with and without explicit attention to articulation", *Journal of Experimental Child Psychology* 72 (4), 271-304.

ZIEGLER J. C. (2018), « Différences inter-linguistiques dans l'apprentissage de la lecture », *Langue française* 199. (ce volume)

ZIEGLER J. C. & GOSWAMI U. (2005), "Reading acquisition, developmental dyslexia and skilled reading across languages: A psycholinguistic grain size theory", *Psychological Bulletin* 131 (1), 3-29.

ZIEGLER J. C., PERRY C. & ZORZI M. (2014), "Modelling reading development through phonological decoding and self-teaching: Implications for dyslexia", *Philosophical Transactions of the Royal Society of London. B: Biological Sciences* 369 (1634), 20120397.

ZIEGLER J. C. *et al.* (2009), "Speech-perception-in-noise deficits in dyslexia", *Developmental Science* 12 (5), 732-745.

A.N.A.E., 2017; 148; 333-343

GraphoGame : un outil numérique pour enfants en difficultés d'apprentissage de la lecture

J.-P. RUIZ*, J. LASSAULT*, L. SPRENGER-CHAROLLES*, U.A RICHARDSON, H. LYYTINEN***, J.C. ZIEGLER***

* Aix-Marseille Université et CNRS, Laboratoire de psychologie cognitive (UMR 7290), Fédération de recherche 3C, Marseille, France. Auteur de correspondance : Johannes Ziegler, Laboratoire de psychologie cognitive, Aix-Marseille Université, 3, place Victor-Hugo, Bât 9, Case D, 13331 Marseille Cedex 3, France. Tél. : +33413550997.
Email : Johannes.ziegler@univ-amu.fr
** University of Jyväskylä, Centre for Applied Language Studies, Jyväskylä, Finlande.
*** UNESCO Chair on Inclusive Literacy Learning for All, University of Jyväskylä, Jyväskylä, Finlande.

RÉSUMÉ : GraphoGame : un outil numérique pour enfants en difficultés d'apprentissage de la lecture
Un enseignement systématique des relations graphophonologiques (le décodage) est nécessaire pour apprendre à lire. Des outils numériques, tel que GraphoGame, un logiciel audiovisuel finlandais, peuvent faciliter la mise en place et l'automatisation du décodage. L'article présente une adaptation en français de GraphoGame sur tablette qui tient compte de l'orthographe de cette langue. Un premier essai de validation, qui a été effectuée auprès d'enfants de CP et CE1 à risques de dyslexie, a permis de constater une progression en lecture de mots plus importante après l'entraînement GraphoGame qu'après un entraînement non-informatisé ou un entraînement mathématique informatisé. Ces premiers résultats suggèrent que les outils du type GraphoGame présentent un intérêt pour les enfants à risque de dyslexie.
Mots clés : *Dyslexie – Remédiation – GraphoGame – Numérique – École.*

SUMMARY: ***GraphoGame: a digital tool for children with difficulties in learning-to-read***
The systematic teaching of spelling-to-sound relations (decoding) is necessary for learning to read. Digital tools such as GraphoGame, a Finnish audio-visual reading game, can facilitate the initiation and automatization of the decoding process. This article presents a French adaptation of GraphoGame for smartphones and tablets. This adaptation implements a psycholinguistic progression that takes into account orthographic, phonological and lexical properties of the French language. A randomized control trial, which was conducted with 1st and 2nd grade children at risk for dyslexia, showed greater benefits in the reading of words after GraphoGame training than after a non-computerized training or a computerized math training. These first results suggest that computerized tools, such as GraphoGame, are beneficial for children at risk for dyslexia.
Key words: *Dyslexia – Intervention – GraphoGame – Digital tools – School.*

RESUMEN: ***El GraphoGame: un instrumento numérico para niños con dificultades de aprendizaje de la lectura***
La enseñanza sistemática de la relaciones entre ortografía y sonido (decodificación) es necesaria para aprender a leer. Las herramientas digitales como GraphoGame, un programa de lectura audiovisual en finlandés, pueden facilitar la iniciación y automatización del proceso de decodificación. Este artículo presenta una adaptación francesa de GraphoGame para tabletas teniendo en cuenta las propiedades ortográficas, fonológicas y lexicales del francés. Una primera validación con niños de primer y segundo grado de Primaria con riesgo de dislexia, mostró que un entrenamiento de lectura con GraphoGame representa un mayor beneficio para la lectura de palabras que un entrenamiento de lectura no informatizado o un entrenamiento matemático informatizado. Estos resultados iniciales sugieren que las herramientas informatizadas, como GraphoGame, son beneficiosas para niños con riesgo de dislexia.
Palabras clave: *Dislexia – Intervención – GraphoGame – Herramientas digitales – Escuela.*

Pour citer cet article : RUIZ, J.-P., LASSAULT, J., SPRENGER-CHAROLLES, L., RICHARDSON, U.A, LYYTINEN, H. & ZIEGLER, J.C. (2017). GraphoGame : un outil numérique pour enfants en difficultés d'apprentissage de la lecture. *A.N.A.E., 148*, 333-343.

© copyright Anae

INTRODUCTION

La lecture est un processus complexe qui fait appel à au moins deux composantes (Aaron, Joshi, Gooden & Bentum, 2008). D'une part, l'identification des mots écrits qui est un processus spécifique à la lecture. D'autre part, la capacité à effectuer un traitement sémantique pour la compréhension de ce qui est lu. La compréhension, qui s'appuie sur plusieurs habilités (traitements sémantique, syntaxique, morphologique), n'est pas spécifique à la lecture. L'identification du mot écrit nécessite un apprentissage fondamental. Le décodage graphophonologique, à voix haute, en constitue l'entrée, et représente un prérequis indispensable pour pouvoir ensuite accéder à une reconnaissance automatique des mots écrits (Share, 1995).

L'apprentissage de la lecture pose de sérieux problèmes pour de nombreux enfants. En France, à l'école primaire, environ 250 000 enfants souffrent de problèmes d'apprentissage de la lecture. En 2015, 40 % des élèves étaient en difficulté à la sortie de l'école primaire : ils n'étaient pas en capacité d'identifier le sujet principal d'un texte, de comprendre des informations implicites et de lier deux informations explicitées séparées dans le texte (CEDRE, 2015). L'illettrisme chez les adultes est évalué en France à 3,1 millions d'individus, et les récentes études de l'OCDE (OECD, 2014) montrent que 40,5 % des élèves de 15 ans ne maîtrisent pas la lecture ; 21,5 % sont même en grande difficulté. Les écarts de niveau entre les élèves les plus performants et les moins performants sont très importants. De la même façon, l'INSERM (2007) a montré que 20 à 25 % des élèves d'une tranche d'âge sont en échec scolaire. Pourtant, le niveau d'éducation est un facteur essentiel puisqu'il prédit le mieux l'état de santé d'une population (INSERM, 2007).

Les conséquences de ces difficultés peuvent être dramatiques et ont un coût personnel et sociétal (Beddington *et al*., 2008). Il semble donc primordial d'intervenir en cas de difficultés au plus tôt au sein de l'école auprès de ces enfants dits « à risque de dyslexie ». L'utilisation d'outils pédagogiques informatisés peut être l'une des réponses pour venir renforcer l'enseignement classique (Ecalle, Magnan, Bouchafa & Gombert, 2009 ; Girard, Ecalle & Magnan, 2013 ; Magnan, Ecalle, Veuillet & Collet, 2004 ; Potocki, Magnan & Ecalle, 2015). Ces dernières années, plusieurs outils ont été développés dont GraphoGame (GG). GG est un outil numérique d'entraînement à la lecture développé en Finlande par Heikki Lyytinen et ses collaborateurs de l'université de Jyväskylä (Ojanen *et al*., 2015 ; Richardson & Lyytinen, 2014). Il repose sur la présentation simultanée et répétée d'unités linguistiques en modalité auditive et visuelle. Notre article présente l'adaptation de cet outil à la langue française et une première validation à petite échelle lors d'une expérimentation auprès d'enfants, scolarisés au CP et au CE1. L'originalité de cette étude réside donc dans le fait de pouvoir tester l'outil GG en première intention à l'école, c'est-à-dire auprès d'enfants à « risque de dyslexie » pendant le temps d'activité pédagogique complémentaire.

L'APPRENTISSAGE DE LA LECTURE

La lecture est une invention culturelle récente qui nécessite tout d'abord un apprentissage explicite des règles de correspondance entre des formes écrites, les graphèmes, et leur forme sonore, les phonèmes : c'est le décodage. Ce décodage, à voix haute, permet à l'enfant de lire des mots écrits même s'il ne les a jamais vus avant, mais dont il connaît déjà la forme phonologique et la signification (Share, 1995 ; Ziegler & Goswami, 2006). La conclusion d'une méta-analyse du *National Reading Panel* (*National Institute of Child Health and Human Development*, 2000) montre l'importance du décodage en phase initiale de l'apprentissage de la lecture pour devenir un bon lecteur. Cet apprentissage est plus ou moins long en fonction des relations entre graphème et phonème (Ziegler & Goswami, 2005). Ainsi lorsque cette relation entre graphème et phonème est régulière, comme par exemple en finnois, chaque graphème correspond à un seul phonème et l'apprentissage de la lecture se fait en quelques semaines. En revanche, lorsque cette relation est moins régulière comme en anglais, le même graphème peut se prononcer différemment (comme 'a' dans '*cat*', '*lady*', '*cake*', '*car*') et l'apprentissage peut prendre quelques années (Landerl, Wimmer & Frith, 1997). Le français est une langue ayant un niveau de difficulté intermédiaire, relativement régulière dans le sens de la lecture mais beaucoup moins régulière dans le sens de la transcription orthographique (Peereman, Lété & Sprenger-Charolles, 2007 ; Ziegler, Jacobs & Stone, 1996).

Dans une écriture régulière, l'aisance du décodage se manifeste très concrètement par le temps de lecture du mot. Chez un lecteur débutant, ce temps est très long et directement lié à la longueur du mot et mobilise toutes les ressources cognitives de l'enfant (Zoccolotti *et al*., 2005). L'automatisation est donc primordiale pour libérer des ressources nécessaires pour effectuer le traitement sémantique. C'est un processus progressif (Sprenger-Charolles, Siegel, Bechennec & Serniclaes, 2003) qui se caractérise par le passage d'un mode sériel, où les mots sont lus graphème par graphème, à un mode parallèle de la lecture, où tous les graphèmes sont traités de façon simultanée (Ans, Carbonnel & Valdois, 1998 ; New, Ferrand, Pallier & Brysbaert, 2006 ; Perry, Ziegler & Zorzi, 2007). Ce passage vers une lecture fluide et instantanée résulte de plusieurs mécanismes :

- un décodage de plus en plus performant permettant un accès rapide à la forme phonologique des mots (Ziegler, Perry & Zorzi, 2014),
- un traitement orthographique de plus en plus efficace permettant un accès rapide à la forme orthographique des mots (Grainger, Dufau & Ziegler, 2016 ; Ziegler, Bertrand, Lété & Grainger, 2014),
- un traitement des unités morphologiques porteuses de sens qui facilite l'accès à la signification des mots (Beyersmann, Grainger, Casalis & Ziegler, 2015 ; Cavalli, Colé, *et al*., 2016).

© copyright Anae

La lecture quotidienne pendant l'enfance est le principal facteur permettant de parvenir à cette automatisation à l'adolescence (Cunningham & Stanovich, 1997) : chaque fois que l'enfant décode un mot, il renforce la forme orthographique, la forme phonologique et les couplages entre les deux (couplage orthographe/phonologie et couplage phonologie/orthographe). C'est donc l'apprentissage explicite du décodage, puis l'utilisation répétée de ces connaissances pendant la lecture (auto-apprentissage) qui permet à l'enfant de devenir un lecteur fluide et autonome (Cunningham, Perry, Stanovich & Share, 2002 ; Share, 1995 ; Ziegler, Perry *et al*., 2014).

LA DYSLEXIE

Lors de ce processus d'apprentissage de la lecture, un enfant peut rencontrer des difficultés pour différentes raisons (Gabrieli, 2009 ; Norton, Beach & Gabrieli, 2015 ; Ramus, 2001 ; Stein, 2014 ; Ziegler *et al*., 2008) (pour une revue voir Habib & Ziegler, 2016). La dyslexie développementale serait l'extrémité d'un continuum des difficultés de lecture. Elle se traduit par un retard de lecture important malgré une stimulation normale, une intelligence normale, l'absence de difficultés sensorielles ou cognitives. Selon les différentes sources, sa prévalence en France varie entre 3,5 % et 8 % (Billard *et al*., 2009 ; Habib, 2014 ; Sprenger-Charolles & Colé, 2003) ; elle peut atteindre des niveaux plus élevés dans les milieux socio-économiques défavorisés (Fluss *et al*., 2008).

Si l'origine génétique et neurobiologique est largement admise (Norton *et al*., 2015), plusieurs hypothèses explicatives sont encore débattues et ce numéro donne un excellent aperçu de la diversité de ces explications. Sans aucun doute, l'une des hypothèses largement partagée concerne la théorie phonologique qui selon les travaux expérimentaux montre un double déficit (Sprenger-Charolles & Colé, 2013 ; Sprenger-Charolles, Siegel, Jimenez & Ziegler, 2011). Le premier serait relié aux compétences d'analyse phonémique et de mémoire phonologique à court terme qui entraveraient le décodage et l'assemblage des unités phonémiques. Le second serait relié à l'accès lexical qui entraverait la récupération rapide des informations phonologiques. Avec un traitement phonologique perturbé, le mécanisme du décodage ne peut se mettre en place efficacement ce qui rend l'apprentissage de la lecture et l'automatisation extrêmement difficiles (pour des simulations de ce déficit voir Ziegler, Perry *et al*., 2014).

Une deuxième théorie concerne les traitements visuels et visuo-attentionnels qui peuvent être déficitaires chez certains enfants dyslexiques (Bosse, Tainturier & Valdois, 2007 ; Stein, 2014 ; Vidyasagar & Pammer, 2010). Il est évident qu'une lecture fluide nécessite un traitement orthographique efficace. Si l'empan visuo-attentionnel est trop réduit, le traitement parallèle des lettres devient impossible tout comme l'exploitation des informations parafovéales (Bosse *et al*., 2007 ; mais voir Ziegler, Pech-Georgel, Dufau & Grainger, 2010). Si le codage de la position des lettres est déficitaire (Collis, Kohnen & Kinoshita, 2013) possiblement dû à des déficits de la voie magnocellulaire (*e.g*. Stein, 2014), le mécanisme de décodage est perturbé et l'apprentissage orthographique est fortement affecté (pour des simulations de ce déficit voir Ziegler, Perry, *et al*., 2014). Si le traitement visuo-attentionnel est déficitaire, l'enfant peut souffrir d'un encombrement perceptif plus important (Zorzi *et al*., 2012) et/ou d'un guidage oculomoteur perturbé. Bien que la question de l'indépendance et de la prévalence de ces déficits fasse toujours débat (Saksida *et al*., 2016), il semble aujourd'hui clair que les causes de la dyslexie sont multifactorielles (Menghini *et al*., 2010 ; Pennington, 2006 ; van Bergen, van der Leij & de Jong, 2014).

GRAPHOGAME

GraphoGame (GG) est un logiciel d'entraînement audio-visuel qui présente des stimuli auditifs et des choix orthographiques (voir *figure 1*) à différentes tailles d'unités (phonème, syllabe, rimes, mots, phrases). À ce titre, le logiciel renforce tout d'abord la perception auditive, l'accès aux représentations phonologiques et la discrimination phonémique et phonologique. Puis il travaille la mise en relation entre les unités phonologiques et orthographiques (le décodage). Enfin, il renforce directement le traitement orthographique et le codage de la position des lettres. Grâce à la répétition massive, le logiciel a pour but de stabiliser le codage orthographique et phonologique et de renforcer l'automatisation du couplage entre ces deux codes. GG s'attaque donc directement aux domaines qui sont souvent déficitaires chez les enfants dyslexiques ou chez les enfants à risque.

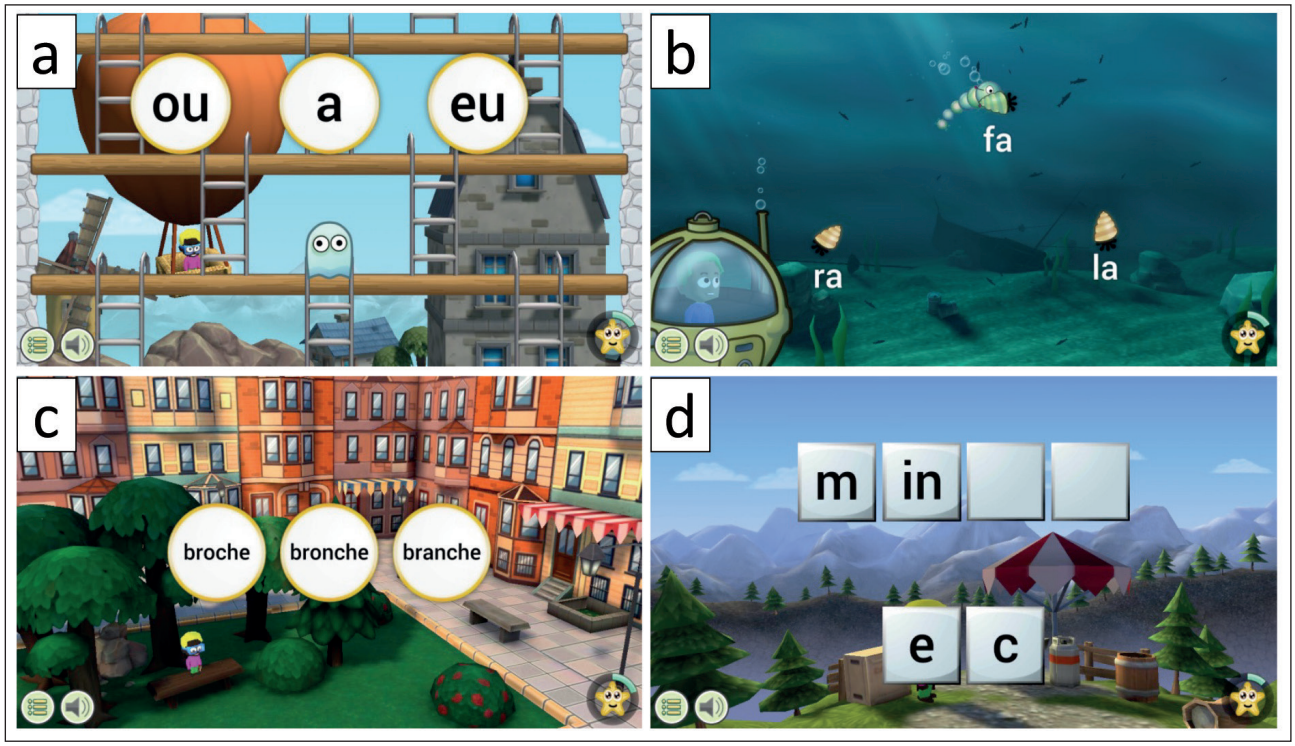
L'efficacité de la version finlandaise de GG a déjà été mise en évidence en Finlande auprès d'enfant âgés de 7 ans considérés à « risque de dyslexie » (Saine, Lerkkanen, Ahonen, Tolvanen & Lyytinen, 2010). Les enfants ayant suivi l'entraînement GG ont montré des progrès significatifs en décodage et en orthographe par rapport à un groupe contrôle ayant suivi une remédiation classique sans le logiciel. Des résultats similaires ont été obtenus pour l'anglais (Kyle, Kujala, Richardson, Lyytinen & Goswami, 2013).

Adaptation de Graphogame à la langue française

L'adaptation française de GG a tenu compte des caractéristiques principales de l'orthographe du français pour établir une progression systématique :
- la consistance des correspondances entre les graphèmes et les phonèmes (dans le sens de la lecture, 85 % des correspondances graphème-phonème sont prévisibles) ;
- la fréquence des mots rencontrés même s'ils sont irréguliers (par exemple, présentation assez rapide des mots outils comme les articles, les pronoms, les prépositions) ;
- les règles de variation contextuelle (par exemple le « s » en début de mot, entre deux voyelles, en fin de mot) ;
- les lettres muettes de fin de mots : certaines de ces lettres sont des « morpho-phonogrammes » de deux types : les uns

© copyright Anae

Figure 1. Quatre niveaux différents de la version française de GraphoGame. a) l'enfant entend le phonème /a/ et doit choisir l'une des options proposées ; b) l'enfant entend la syllabe /fa/ et doit choisir l'une des options ; c) l'enfant entend le mot « bronche » et doit choisir parmi trois options : « branche », « bronche » et « broche » ; d) l'enfant entend le mot « mince » et doit remettre les lettres dans le bon ordre.



marquent les flexions nominales (le genre et le nombre, *e* et *s* dans « amies ») et verbales (le *s* dans « tu chantes » et *ent* dans « ils mangent »). D'autres lettres muettes sont des supports de dérivation (le *t* de « toit/toiture ») qui sont aussi, dans certains cas, des supports de flexion (le *d* dans « grand/grande/grandeur »).

L'élaboration de cette progression s'est faite à partir de la base de données Manulex-Morpho (Peereman, Sprenger-Charolles & Messaoud-Galusi, 2013), issue de Manulex (Lété, Sprenger-Charolles & Colé, 2004). Manulex est une base de données créée à partir de 54 manuels scolaires du primaire. Cette base fournit les fréquences d'occurrence des mots écrits à 3 niveaux d'expertise de la lecture : le CP où se construit le lexique de l'enfant sur la base de la médiation phonologique, le CE1 où se construit le lexique orthographique par automatisation progressive de la reconnaissance du mot écrit et le cycle 3 (CM1-CM2) où se consolide et s'enrichit le stock lexical par exposition répétée à l'écrit. Cette base contient 1 900 000 mots dont 49 000 formes différentes (avec les marques morphologiques de flexion et bien sûr, toutes les formes dérivées) et seulement 24 000 mots sans formes fléchies.

À partir de cette base de données, a été développé Manulex-Infra (Peereman *et al.*, 2007) pour décrire les caractéristiques infralexicales des mots puis Manulex-Morpho (Peereman *et al.*, 2013). Dans cette base, qui a été utilisée pour construire la progression de GG (voir l'article de Sprenger-Charolles dans ce numéro), les morphophonogrammes font partie des unités infralexicales prises en compte dans les analyses de la fréquence et de la consistance des correspondances graphème-phonème et phonème-graphème. Par exemple, en utilisant Manulex-Morpho, nous pouvons connaître la fréquence du graphème s par rapport aux autres graphèmes du français. De la même manière, nous pouvons aussi connaître la régularité des correspondances graphème-phonème pour le graphème s, à savoir, combien de fois il se lit /s/ comme dans « sol », /z/ comme dans « rose » et, en fin de mots, combien de fois il se prononce (comme dans « ours ») ou est muet. Dans ce dernier cas, le plus fréquent, il est souvent une marque morphologique, comme le s qui est une flexion nominale (amis) ou verbale (tu chantes) ou une marque de flexion/dérivation (gris) ; il peut aussi n'avoir aucune valeur spécifique comme dans « alors ».

Description du fonctionnement de GG

L'enfant est équipé d'un casque dans lequel il entend un stimulus auditif. En fonction du type de jeu (voir ci-dessous), plusieurs propositions sont présentées visuellement et l'enfant doit identifier celles qui correspondent au stimulus auditif et cliquer dessus pour valider son choix. Si l'enfant donne la bonne réponse, un autre stimulus lui est proposé. Lorsque l'enfant se trompe, la correction est automatique et visible, et le stimulus auditif lui est à nouveau présenté tandis que la syllabe réapparaît, marquée d'une couleur. Le logiciel enregistre les réponses et adapte le passage des niveaux en fonction du taux de réussite. Si le taux de réussite est insuffisant (< 85 %), l'enfant doit recommencer le niveau. Chaque stimulus apparaît plusieurs fois au cours d'un niveau. GG enregistre également les temps de réponse.

© copyright Anae

Le contenu (*i.e.*, la progression) est organisé en une vingtaine d'unités pédagogiques ou séquences (par exemple, la première séquence propose les voyelles simples *a, o, ou*). Une séquence dure entre 15 et 20 minutes selon la vitesse de l'enfant et le taux d'erreur. La progression est faite de telle sorte que chaque séquence comprendra uniquement des mots comportant des relations graphophonologiques déjà vus précédemment. Chaque séquence est subdivisée en une dizaine de niveaux (ou *levels*) d'une 1 à 2 minutes (10 à 15 essais). Dans la présente version du logiciel, l'ordre des séquences est fixé d'avance. Le choix des problèmes ne dépend donc pas du niveau de performance des élèves. En revanche, le passage au prochain niveau se fait en fonction du taux de réussite. Donc un élève avec un faible niveau de lecture restera plus longtemps dans les niveaux élémentaires qu'un élève avec un bon niveau de lecture.

À chaque niveau, un exercice différent sous forme de jeu est proposé à l'enfant dans un environnement graphique renouvelé pour maintenir la motivation. Les différents jeux sont :
- les jeux classiques (plusieurs environnements possibles, cf. *figure 1a et 1b*) où il s'agit pour l'enfant d'identifier le graphème, la syllabe ou encore le mot qui correspond à une prononciation donnée ;
- les jeux à choix multiple où il s'agit pour l'enfant de rechercher tous les graphèmes, les syllabes ou encore les mots qui correspondent à une prononciation donnée ;
- les jeux de discrimination auditive et visuelle ;
- les jeux de formation de mots et de phrases où il s'agit pour l'enfant de remettre dans l'ordre une suite de graphèmes (pour un mot) ou une suite de mots (pour une phrase) en fonction de ce qu'il a entendu (cf. *figure 1c et 1d*).

Les premier et dernier niveau de chaque séquence servent de prétest et post-test « interne » permettant d'évaluer si l'enfant a progressé au sein du jeu. Ces niveaux sont identiques et proposent un contenu représentatif de la séquence. Ces niveaux ne sont pas répétés même si l'enfant est en dessous du seuil de réussite.

ÉTUDE DE VALIDATION

Population

Au CP, 34 enfants issus de deux écoles élémentaires du département des Alpes de Haute-Provence ont participé à l'étude. Les enfants ont été choisis à partir des signalements des enseignants (au retour de la Toussaint pour les CP démarrant l'expérimentation début janvier). Au CE1, 40 enfants issus de 6 écoles du département des Alpes de Haute-Provence ont été sélectionnés à la fin du CP pour un démarrage de l'expérimentation en septembre. Les données de 5 enfants du CE1 ont été éliminées pour différentes raisons (abandon, blocage au même niveau tout au long du jeu, problèmes techniques) ramenant le nombre total de participants à 35. Les élèves ont été assignés à un groupe de façon aléatoire. Le groupe 1 commençait avec GG tandis le groupe 2 effectuait un entraînement Math (CE1) ou aucun entraînement sur tablette (CP). Les groupes ont été inversés après 5 semaines. Le *tableau 1* résume la répartition des élèves par groupe et l'appariement des groupes en termes de nombre d'enfants, âge, sexe ainsi que les résultats cognitifs et langagiers. L'étude a été validée par l'Inspecteur d'académie, directeur académique des services de l'Éducation nationale (IA-DASEN). Une lettre d'information a été transmise aux parents afin d'avoir leur autorisation écrite pour inclure leur enfant dans l'étude.

Les tests d'évaluation

Les tests du CP

Les matrices progressives de Raven (PM 47, Raven, 1976).
Pour évaluer le niveau cognitif non verbal des enfants, les matrices progressives de Raven couleur (qui testent le raisonnement analogique) ont été choisies (PM 47). Ce test, qui est l'un des plus sensibles pour évaluer le facteur général d'intelligence (Prabhakaran, Smith, Desmond, Glover & Gabrieli, 1997), se compose de 3 séries de 12 items proposées aux enfants sans limite de temps. Les normes ont été établies sur une population française de 1 064 enfants âgés de 4 à 11 ans et demi. Aucun indice de fiabilité (alpha de Cronbach) n'est donné pour ce test en français.

Test de vocabulaire actif et passif (TVAP 5-8 ans Deltour & Hupkens, 1984)
Pour évaluer le niveau d'intelligence verbale (et écarter un trouble spécifique du langage oral avec retentissement

Tableau 1. Composition des groupes ayant participé à l'expérimentation et résultats au Test de vocabulaire actif et passif (TVAP), matrices progressives de Raven (PM 47) et tests de lecture (CP, lecture en 1 minute ; CE1, Alouette).

	CP			CE1		
	Groupe 1	**Groupe 2**	**Total**	**Groupe 1**	**Groupe 2**	**Total**
Nombre	19	15	34	18	17	35
Âge (mois)	77.3	74.9	76.0	87.9	89.1	88.5
Sexe (G/F)	10/9	8/7	18/16	11/7	7/10	18/17
TVAP	39.5	43.0	41.0	46.5	44.7	45.6
PM47	22.1	22.4	22.2	24.9	25.6	25.2
Lecture (M/min)	12.4	13.9	13.0	12.1	13.4	12.8

© copyright Anae

sur la lecture), nous avons choisi un test de vocabulaire (TVAP 5-8 ans). Nous avons évalué uniquement le niveau de vocabulaire en réception, c'est-à-dire la capacité à associer un mot prononcé par l'expérimentateur avec la bonne image parmi 6 représentées : l'image correcte, 1 proche et 4 distracteurs. Par exemple, pour « château », l'image d'un château (choix correct, 2 points), d'une grande maison (choix proche, 1 point), d'un gâteau, d'un chat, d'une tour de contrôle et d'un roi. Aucun indice de fiabilité (alpha de Cronbach) n'est donné pour ce test.

Test de lecture en 1 min (Gentaz, Sprenger-Charolles, Theurel & Colé, 2013)
Le test utilisé au CP pour évaluer le niveau de lecture au cours des 10 semaines d'entraînement avec GG est un test de lecture de mots familiers et de mots inventés pour le début du CP. Il consiste à lire à haute voix le plus de mots possibles en 1 minute pour chaque catégorie.

Les mots familiers (35, 5 par ligne) sont fréquents et figurent parmi les 1 000 premiers mots des manuels du premier grade (Lété *et al.*, 2004). Ils sont courts, mono- ou bi-syllabiques, réguliers sur le plan des correspondances graphème-phonème (mais 5 d'entre eux ont un graphème dont la prononciation dépend d'une règle contextuelle, *c* ou *g*) et ils comportent peu de groupes consonantiques

Les mots inventés ou pseudo-mots sont appariés aux mots fréquents en longueur, en structure syllabique et en difficulté orthographique (30 mots, 5 par ligne). Ils sont tous réguliers sur le plan des correspondances graphème-phonème, mais comme pour les mots, 5 d'entre eux ont un graphème dont la prononciation dépend d'une règle contextuelle (*c* ou *g*). Le coefficient de fidélité pour ce test (*Cronbach alpha*) est de .92.
Lorsque l'enfant a lu moins de 4 mots correctement sur les 10 premiers mots d'une catégorie, l'épreuve est arrêtée. Ce test a été passé à trois reprises lors des phases de prétest, à 5 semaines (mi-test) et à la fin des 10 semaines d'entraînement (post-test).

Le test de l'Alouette (Lefavrais, 1965)
Ce test a été utilisé pour évaluer le niveau de lecture au CE1. Il propose un texte asémantique de 265 mots avec une présence accrue de mots peu fréquents obligeant l'enfant à mettre en œuvre la procédure de décodage graphophonologique plutôt que ses connaissances lexicales. Ce test est très sensible aux difficultés de lecture (Bertrand, Fluss, Billard & Ziegler, 2010 ; Cavalli *et al.*, 2017).

Méthode d'expérimentation

Nous avons choisi de comparer deux groupes d'entraînement randomisés afin d'éviter un biais de sélection : un groupe d'entraînement à la lecture sur GG et un groupe contrôle (APC classique au CP, mathématiques au CE1). L'utilisation d'un groupe contrôle répond à deux objectifs. Il permet d'une part de neutraliser d'autres facteurs pouvant intervenir dans l'évolution du niveau de lecture des élèves, comme par exemple les apprentissages scolaires ou un désir accru de bien faire, lié au fait de participer à cette étude. Chaque groupe d'enfants a été tour à tour groupe GG et groupe contrôle. Le design de cette étude permet ainsi à chaque enfant d'être son « propre contrôle ».

L'évolution du niveau de lecture des élèves a été mesurée grâce aux tests de lecture étalonnés décrits précédemment, mais aussi grâce aux données issue de GG. Nous pourrons ainsi comparer les résultats obtenus à partir de ces données internes et externes à GG.

Groupe GG versus Groupe contrôle

Pour les élèves de CE1, le groupe contrôle s'entraînait avec un logiciel de mathématiques *Toké Maths* (http://www.tokemaths.com). Pendant une première période de 5 semaines, le 1er groupe utilisait GG et le 2e Toké Maths. Pendant la seconde période de 5 semaines, les 2 groupes étaient inversés. Pour les élèves de CP, nous n'avons pas trouvé de logiciel de mathématiques adapté. Il a donc été décidé de laisser le groupe contrôle en soutien scolaire (APC) classique avec l'enseignant. Au cours des prétests, tous les tests étaient proposés aux enfants de CE1 ou CP. Au cours des mi-tests et post-tests, seuls les tests de lecture étaient présentés aux enfants. L'ordre de passation des tests ainsi que les modalités de passation ont été strictement les mêmes pour les expérimentateurs.

Les entraînements ont été réalisés dans le cadre des temps d'APC. La durée des entraînements était d'1 heure par semaine à raison de 3 fois 20 min ou 2 fois 30 min (5 heures en total).

RÉSULTATS

L'analyse des logs a montré que le temps effectif passé sur les exercices de GG était en moyenne de 2 heures 50 minutes au CP et de 2 heures 37 minutes au CE1.

Progression au sein de GG

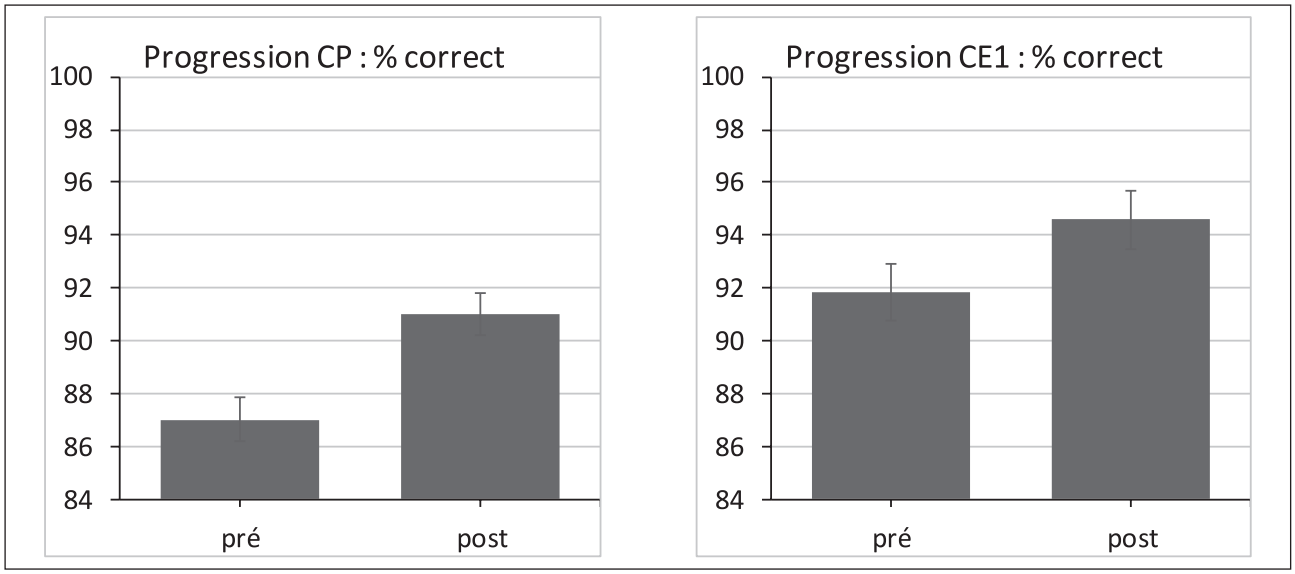
La progression des élèves au sein du jeu GG a été évaluée en comparant le pourcentage de bonnes réponses du premier niveau de chaque séquence (prétest) au pourcentage de bonnes réponses du dernier niveau (post-test) de chaque séquence. Ces niveaux étaient identiques permettant de mesurer la progression « immédiate » au sein du jeu. Comme le montre la *figure 2*, les enfants du CP ont progressé en moyenne de 4 % entre le prétest et le post-test de chaque séquence ($t(33) = 4.53$; $p < .0001$). Au CE1, ils ont progressé en moyenne de 2,75 % entre les prétests et les post-tests de chaque séquence ($t(35) = 3.8$; $p < .001$).

Progression dans les tests de lecture

La progression des élèves en fonction du type d'entraînement a été évaluée par une analyse de variance (ANOVA)

© copyright Anae

Figure 2. Progression (% correct) au sein du jeu GG entre le premier et le dernier niveau de chaque séquence pour les enfants du CP (à gauche) et du CE1 (à droite). Les barres d'erreurs correspondent à des erreurs standards intra-sujets.



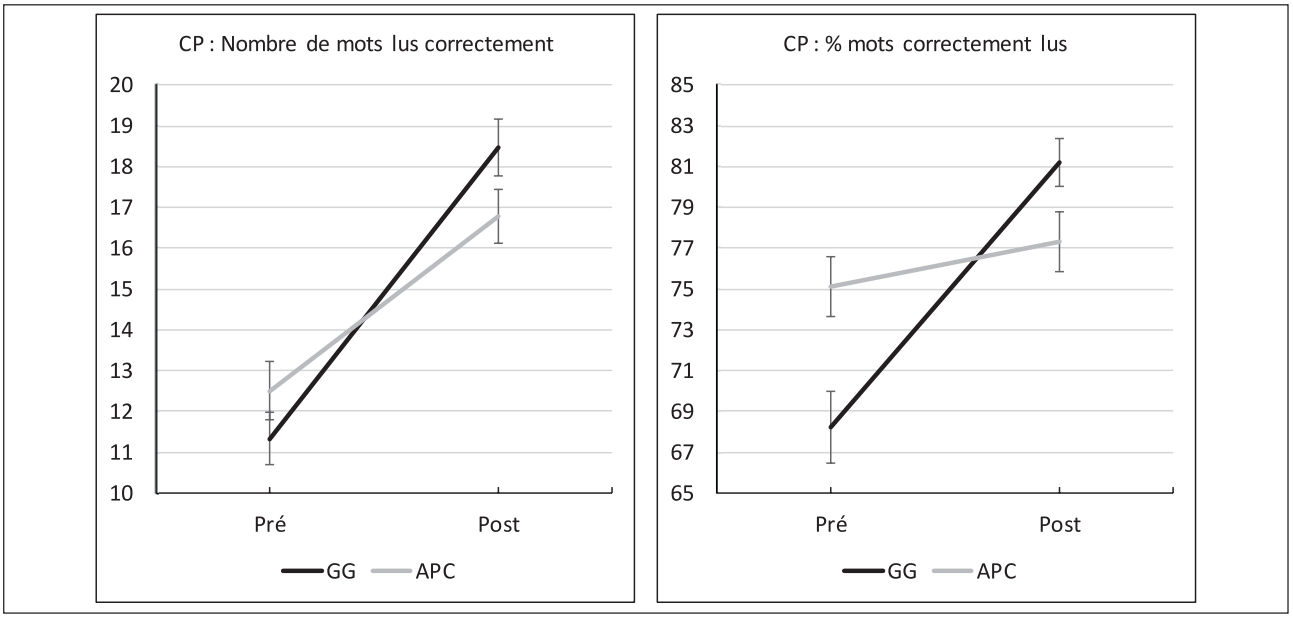
avec deux facteurs intra-sujets, le facteur Progression (prétest versus post-test) et le facteur Type d'intervention (GG versus APC classique pour le CP ; GG versus Math pour le CE1). Un effet spécifique de l'entraînement se manifeste par une interaction significative entre le facteur Progression et le facteur Type d'intervention.

Résultats CP

Les résultats du test de lecture (lecture en 1 minute) réalisé avant et après l'entraînement sont présentés dans la *figure 3* pour la fluence (nombre de mots correctement lus en 1 minute) et la précision (% de mots correctement lus). Pour la fluence, l'ANOVA montre un effet du facteur Progression ($F(1,33) = 94.4$, $p < .0001$) reflétant le fait que l'ensemble des enfants ont progressé entre le prétest et le post-test. Plus important, l'interaction entre le facteur Progression et le facteur Type d'intervention est significative ($F(1,33) = 12.8$, $p < .001$) indiquant que les enfants ont davantage progressé après les séances GG qu'après un entraînement APC classique. Pour la précision, l'ANOVA montre également un effet significatif du facteur Progression ($F(1,33) = 32.4$, $p < .001$) et une interaction significative entre le facteur Progression et le facteur Type d'intervention ($F(1,33) = 8.2$, $p < .01$). Cette interaction reflète le fait que la précision de lecture était meilleure après les séances de GG (+ 13.01 %) qu'après l'APC classique (+ 2.23 %).

En ce qui concerne la lecture de pseudomots au CP, l'ANOVA montre un effet significatif du facteur Progression pour la vitesse ($F(1,33) = 75.8$, $p < .0001$) et la précision ($F(1,33) = 24,2$, $p < .0001$) mais l'interaction entre le facteur Progression et le facteur Type d'Intervention n'est pas significative ni pour la fluence ni pour la précision ($F < 1$).

Figure 3. Nombre de mots familiers correctement lus en 1 minute (gauche) et pourcentage de mots correctement lus (droite) au prétest et post-test du test de lecture en 1 minute au CP. Les barres d'erreurs correspondent à des erreurs standards intra-sujets.



© copyright Anae

Résultats CE1

Les résultats du test de l'Alouette passés avant et après l'entraînement sont présentés dans la *figure 4* pour la fluence (nombre de mots correctement lus en 3 minutes) et la précision (% de mots correctement lus). L'analyse de la fluence montre que l'ensemble des enfants ont progressé entre le prétest et le post-test avec un effet significatif du facteur Progression ($F(1,34) = 116.2$, $p < .0001$). L'interaction entre le facteur Progression et le facteur Type d'intervention est également significative ($F(1,34) = 7.6$, $p < .01$) indiquant que les élèves ont davantage progressé après l'entraînement GG qu'après l'entraînement mathématique. Pour la précision, l'ANOVA montre un effet significatif du facteur Progression ($F(1,34) = 13.4$, $p < .001$) mais l'interaction entre le facteur Progression et le facteur Type d'intervention n'était pas significative ($F < 1$).

Profil des enfants « répondants » et « résistants » à l'intervention

Pour analyser les profils des enfants « répondants » et « résistants » à l'intervention (Fuchs & Fuchs, 2006) au CP, nous avons créé deux groupes en fonction du taux de progression moyen (13 %) : un avec 17 enfants « répondants » qui ont obtenu des scores de progression au-dessus de cette moyenne ; l'autre avec 17 enfants « résistants » qui ont obtenu des scores de progression en dessous de cette moyenne. Pour identifier des variables qui expliqueraient pourquoi certains enfants répondent mieux à l'intervention que d'autres, nous avons comparé les 2 groupes sur les variables suivantes : âge, niveau de vocabulaire (TVAP), intelligence non-verbale (PM47), niveau de lecture initial, progression au sein de GG, temps effectif joué et nombre de niveaux joués. Comme le montre le *tableau 2*, les « répondants » (++) avaient un niveau de lecture initial significativement plus faible que les « résistants » ($p < .0001$). De façon tendancielle mais non significative ($p < .10$), les enfants « répondants » avaient un niveau de vocabulaire légèrement supérieur et une meilleure progression de jeu.

Une analyse similaire a été effectuée pour les enfants du CE1. Nous avons regroupé les enfants en fonction de leur score de vitesse obtenu à l'Alouette avec un premier groupe de 10 enfants « répondants » ayant progressé de plus de 20 points (++) et un second groupe de 10 enfants « résistants » ayant progressé de moins de 10 points (--). Le profil de ces enfants a été étudié en fonction des variables Âge, TVAP, PM47, niveau de lecture initial, temps effectif joué, nombre de niveaux joués et progression au sein de GG (cf. *tableau 3*). Il en résulte que les enfants répondants ont un meilleur niveau de vocabulaire mais aucune différence en termes d'âge, PM47, ou progression au sein de GG.

Figure 4. Nombre de mots correctement lus en 3 minutes (gauche) et pourcentage de mots correctement lus (droite) au prétest et post-test du test de l'Alouette CE1. Les barres d'erreurs correspondent à des erreurs standards intra-sujets.

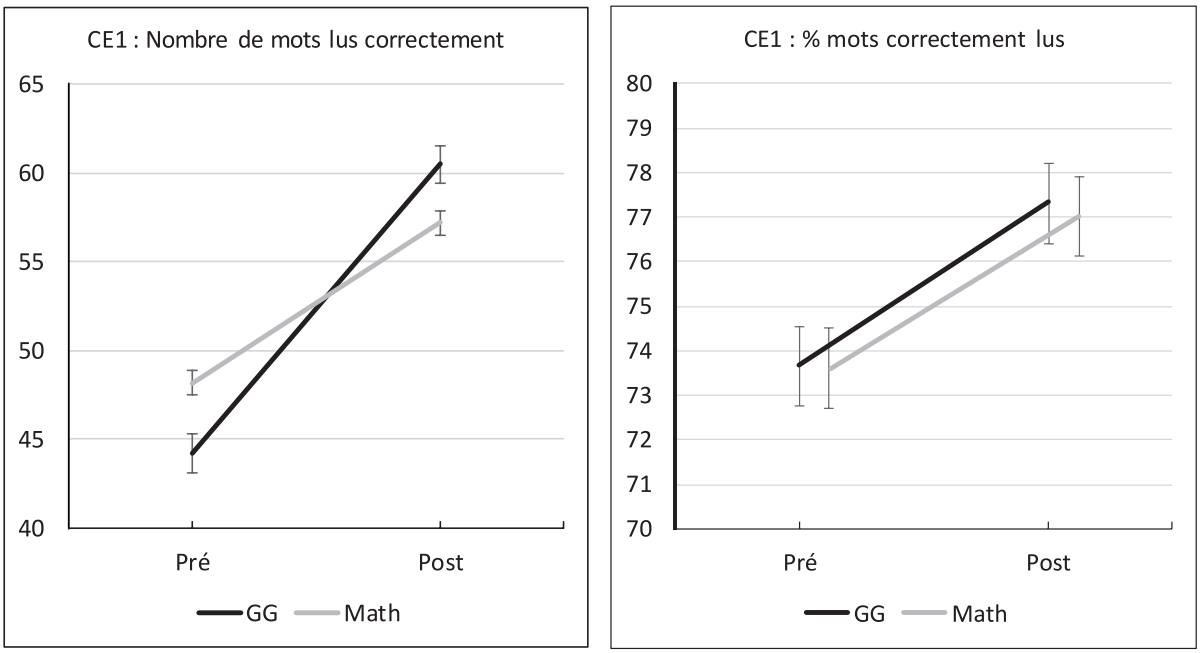


Tableau 2. Comparaison des enfants du CP ayant progressé plus que la moyenne (++, « répondants ») avec les enfants ayant progressé moins que la moyenne (--, « résistants »).

Groupe	Âge	TVAP	PM47	Niveau de lecture initial	Progression au sein de GG (%)	Temps de jeu effectif	Nombre de niveaux joués
--	76,29	39,88	22,82	79,52	3,62	2 h 57	191,29
++	71,94	42,24	21,65	56,89	4,14	2 h 43	175,76
Test t				***			

*** = p < .0001

© copyright Anae

Tableau 3. Profil des enfants du CE1 ayant progressé plus de 20 points après les séances GG (++, « répondants ») versus ceux qui ont progressé moins de 10 points (--, « résistants »).

Groupe	Âge	TVAP	PM47	Niveau de lecture initial	Progression au sein de GG (%)	Temps de jeu effectif	Nombre de niveaux joués
--	91,4	44,3	25,3	35,5	2 h 32	161	2,6
++	88,4	48,6	24,9	40,2	2 h 47	159	2,8
Test t		*					

* = p < .05

DISCUSSION

L'objectif de cette étude consistait à tester et valider par un protocole contrôlé et randomisé l'efficacité de l'adaptation française de GG auprès d'enfants « à risques » de dyslexie scolarisés au CP et au CE1 et participant à une prise en charge complémentaire au sein de l'école. Les résultats obtenus permettent d'affirmer que GG a apporté un bénéfice spécifique pour ces élèves. En effet, pour les enfants du CP, les résultats du test de lecture en 1 minute, proposé avant et après chaque phase d'entraînement, montrent que la vitesse et la précision de lecture de mots familiers ont progressé plus après la phase GG qu'après la phase APC classique. Ces bénéfices ont été obtenus après seulement 5 semaines d'entraînement (durée de 2 h 50). En revanche, aucun effet spécifique de l'entraînement n'a été obtenu pour la lecture de mots inventés (pseudo-mots). Un entraînement avec GG génère donc un effet de transfert à la lecture à voix haute de mots familiers mais pas de pseudo-mots. Il est possible que le transfert à la lecture de pseudo-mots nécessite un entraînement plus intense.

Les résultats du CE1 montrent également des effets d'entraînement spécifique à GG pour la vitesse de la lecture dans le test de l'Alouette mais pas sur la précision de la lecture : les enfants ont lu significativement plus de mots après les phases d'entraînement GG qu'après les phases d'entraînement *TokéMath*. Ces effets ont été obtenus après seulement 5 semaines d'entraînement, soit 2 h 37. Il convient de noter que les 2 tests de lecture impliquent la lecture à voix haute tandis que GG n'entraîne pas explicitement la lecture à voix haute. Il est donc possible que les effets de transfert de l'entraînement aient été plus importants dans un test de lecture silencieuse basé sur reconnaissance visuelle des mots, tel que *Timé2* (Ecalle, 2003), ou dans un test de choix orthographique (Siegel, Share & Geva, 1995).

L'analyse des profils des enfants « répondants » versus « résistants » à l'intervention montre l'importance de deux variables. Tout d'abord, au CP, les enfants ayant un niveau de lecture initial plus faible ont plus bénéficié de l'entraînement avec GG que les enfants ayant un niveau de lecture initial plus élevé. Cela suggère que les outils informatiques tels que GG sont particulièrement bénéfiques pour les enfants en grande difficulté de lecture. Deuxièmement, dans le groupe CE1, les enfants « répondants » avaient un niveau de vocabulaire plus élevé que les enfants « résistants ». Ce résultat suggère que les compétences langagières, indexées ici par le niveau de vocabulaire, puissent agir comme un facteur de protection comme cela a été montré dans des études récentes chez les dyslexiques adultes (Cavalli, Casalis *et al.*, 2016). Par exemple, les élèves qui ont un vocabulaire plus étendu ont plus de chance de reconnaître (ou de reconnaître plus rapidement) les mots qu'ils décodent parce qu'ils en ont une représentation dans leur lexique phonologique. Ils pourraient mieux profiter de GG parce que les conditions préalables à la formation d'une représentation orthographique sont réunies.

Cette étude présente certaines limitations qu'il convient de souligner. Premièrement, le taux d'entraînement semble trop faible par rapport aux études qui ont rapporté des effets d'entraînement très impressionnants tant au niveau comportemental que cérébral. Par exemple, dans l'étude de Eden *et al.* (2004), les adultes dyslexiques ont bénéficié d'un entraînement de 112 heures (3 heures par jour pendant 8 semaines). Dans l'étude de Simos *et al.* (2002), les enfants dyslexiques ont eu un entraînement de 80 heures (1-2 heures par jour pendant 8 semaines). Enfin, dans l'étude de Franceschini *et al.* (2013), les enfants dyslexiques ont été entraînés pendant 12 heures. Nous nous attendons à des effets d'entraînement plus robustes avec un entraînement plus important. Deuxièmement, nous n'avons pas pu suivre ces enfants pour savoir si les effets d'entraînement persistent dans la durée et lesquels de ces enfants « à risque » ont été diagnostiqués dyslexiques par la suite. En effet, la persistance des effets d'entraînements dans la durée est un enjeu majeur et peu d'études d'entraînement ont étudié les effets d'entraînement à long terme (Elbro & Petersen, 2004). Troisièmement, comme discuté plus haut, nous avons uniquement utilisé des tests de lecture à voix haute. Or, nous pouvons nous attendre à des effets plus importants sur la lecture silencieuse (Ziegler, Bertrand *et al.*, 2014). Enfin, nous n'avons pas pu vérifier si l'entraînement a un effet bénéfique sur la compréhension de textes lors de la lecture (Gough &Tunmer, 1986).

En conclusion, cette première validation de GG pour les enfants à risque en contexte scolaire est assez encourageante. L'outil, adopté par les enseignants comme par les élèves, a permis d'obtenir des effets positifs sur la vitesse et la précision de la lecture des mots familiers au CP et la vitesse de la lecture au CE1. Ces résultats ont été obtenus après seulement 5 semaines d'entraînement en raison d'1 heure par semaine. Il est probable qu'on puisse obtenir de meilleures progressions avec une période d'utilisation plus longue. Plus que jamais, dans une volonté de promouvoir une approche éducative fondée sur la preuve (Davies, 1999), des études randomisées et contrôlées sont

© copyright Anae

nécessaires pour évaluer et quantifier l'apport de ces nouveaux outils numériques dans un contexte scolaire (Ecalle, Kleinsz & Magnan, 2013 ; Girard *et al*., 2013).

REMERCIEMENTS

Nous remercions Nathalie Fina et Christine Garcia d'avoir réalisé cette étude dans le contexte de leur mémoire d'orthophonie et Sylvie Rouillard, Édouard Alavoine et Jean-Patrice Albrand pour leur aide technique tout au long de ce projet. Nos remerciements vont également aux équipes pédagogiques et aux enfants des écoles de la circonscription de Sisteron Sud dans les Alpes de Haute Provence. Cette recherche a été financée par un contrat ANR [ANR- 13-APPR-0003]. Cette recherche a également bénéficié du support du Labex BLRI (ANR-11-LABX-0036) et de l'Institut Convergence ILCB (ANR-16-CONV-0002) et de l'initiative d'excellence d'Aix-Marseille Université (A*MIDEX).

RÉFÉRENCES

AARON, P.G., JOSHI, R.M., GOODEN, R. & BENTUM, K.E. (2008). Diagnosis and treatment of reading disabilities based on the component model of reading: an alternative to the discrepancy model of LD. *Journal of Learning Disabilities, 41* (1), 67-84.

ANS, B., CARBONNEL, S. & VALDOIS, S. (1998). A connectionist multiple-trace memory model for polysyllabic word reading. *Psychological Review, 105* (4), 678-723.

BEDDINGTON, J., COOPER, C. L., FIELD, J., GOSWAMI, U., HUPPERT, F.A., JENKINS, R. & THOMAS, S. M. (2008). The mental wealth of nations. *Nature, 455* (7216), 1057-1060.

BERTRAND, D., FLUSS, J., BILLARD, C. & ZIEGLER, J.C. (2010). Efficacité, sensibilité, spécificité : comparaison de différents tests de lecture [Efficiency, sensitivity, specificity: Comparison of different reading tests]. *Année psychologique, 110*, 299-320.

BEYERSMANN, E., GRAINGER, J., CASALIS, S. & ZIEGLER, J.C. (2015). Effects of reading proficiency on embedded stem priming in primary school children. *Journal of Experimental Child Psychology, 139*, 115-126.

BILLARD, C., FLUSS, J., DUCOT, B., BRICOUT, L., RICHARD, G., ECALLE, J., …, ZIEGLER, J. (2009). Deficits in reading acquisition in primary school: cognitive, social and behavioral factors studied in a sample of 1062 children. *Revue Epidemioligue et Santé Publique, 57* (3), 191-203.

BILLARD, C., FLUSS, J., RICHARD, G., ZIEGLER, J.C., ECALLE, J., MAGNAN, A. & SPIRA, A. (2007). *Résultats préliminaires d'une étude épidémiologique transversale des apprentissages en lecture, orthographe et calcul au CE1*. Retrieved from Paris:

BOSSE, M.L., TAINTURIER, M.J. & VALDOIS, S. (2007). Developmental dyslexia: the visual attention span deficit hypothesis. *Cognition, 104* (2), 198-230.

CAVALLI, E., CASALIS, S., EL AHMADI, A., ZIRA, M., PORACCHIA-GEORGE, F. & COLÉ, P. (2016). Vocabulary skills are well developed in university students with dyslexia: Evidence from multiple case studies. *Research Developmental Disabilites, 51-52*, 89-102.

CAVALLI, E., COLÉ, P., BADIER, J.M., ZIELINSKI, C., CHANOINE, V. & ZIEGLER, J.C. (2016). Spatiotemporal Dynamics of Morphological Processing in Visual Word Recognition. *Journal of Cognitive Neuroscience, 28* (8), 1228-1242.

CAVALLI, E., COLÉ, P., LELOUP, G., PORACCHIA-GEORGE, F., SPRENGER-CHAROLLES, L. & EL AHMADI, A. (2017). Screening for Dyslexia in French-Speaking University Students: An Evaluation of the Detection Accuracy of the Alouette Test. *Journal of Learning Disabilities*, 22219417704637.

CEDRE. (2015). *Nouvelle évaluation en fin de collège : compétences langagières et littératie. Évaluation des acquis des élèves*. Direction de l'evaluation, de la prospective et de la performance (DEPP).

COLLIS, N.L., KOHNEN, S. & KINOSHITA, S. (2013). The role of visual spatial attention in adult developmental dyslexia. *Quarterly Journal of Experimental Psychology (Hove), 66* (2), 245-260.

CUNNINGHAM, A.E., PERRY, K.E., STANOVICH, K.E. & SHARE, D.L. (2002). Orthographic learning during reading: Examining the role of self-teaching. *Journal of Experimental Child Psychology, 82* (3), 185-199.

CUNNINGHAM, A.E. & STANOVICH, K.E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33* (6), 934-945.

DAVIES, P. (1999). What is Evidence-based Education? *British Journal of Educational Studies, 47* (2), 108-121.

DELTOUR, J.J. & HUPKENS, D. (1984). *Test de vocabulaire actif et passif pour enfants (5 à 8 ans) (TVAP)*. Braine-le-Château : éditions de l'Application des techniques modernes.

ECALLE, J. (2003). *Timé-2 : test d'identification de mots écrits pour enfants de 6 à 8 ans*. Paris : Centre de psychologie appliquée.

ECALLE, J., KLEINSZ, N. & MAGNAN, A. (2013). Computer-assisted learning in young poor readers: The effect of grapho-syllabic training on the development of word reading and reading comprehension. *Computers in Human Behavior, 29* (4), 1368-1376.

ECALLE, J., MAGNAN, A., BOUCHAFA, H. & GOMBERT, J.E. (2009). Computer-based training with ortho-phonological units in dyslexic children: new investigations. *Dyslexia, 15* (3), 218-238.

EDEN, G.F., JONES, K.M., CAPPELL, K., GAREAU, L., WOOD, F.B., ZEFFIRO, T.A. & FLOWERS, D. L. (2004). Neural Changes following Remediation in Adult Developmental Dyslexia. *Neuron, 44* (3), 411-422.

ELBRO, C. & PETERSEN, D. K. (2004). Long-Term Effects of Phoneme Awareness and Letter Sound Training: An Intervention Study With Children at Risk for Dyslexia. *Journal of Educational Psychology, 96* (4), 660-670.

FLUSS, J., ZIEGLER, J.C., ECALLE, J., MAGNAN, A., WARSZAWSKI, J., DUCOT, B. & BILLARD, C. (2008). Prevalence of reading disabilities in early elementary school: Impact of socioeconomic environment on reading development in 3 different educational zones. *Archives de pédiatrie, 15* (6), 1049-1057.

FRANCESCHINI, S., GORI, S., RUFFINO, M., VIOLA, S., MOLTENI, M. & FACOETTI, A. (2013). Action Video Games Make Dyslexic Children Read Better. *Current Biology, 23* (6), 462-466.

FUCHS, D. & FUCHS, L.S. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly, 41* (1), 93-99.

GABRIELI, J.D. (2009). Dyslexia: a new synergy between education and cognitive neuroscience. *Science, 325* (5938), 280-283.

GENTAZ, É., SPRENGER-CHAROLLES, L., THEUREL, A. & COLÉ, P. (2013). Reading comprehension in a large cohort of French first graders from low socio-economic status families: a 7-month longitudinal study. *PLoS ONE, 8* (11), e78608.

GIRARD, C., ECALLE, J. & MAGNAN, A. (2013). Serious games as new educational tools: how effective are they? A meta-analysis of recent studies. *Journal of Computer Assisted Learning, 29* (3), 207-219.

GOUGH, P.B. & TUNMER, W.E. (1986). Decoding, Reading, and Reading Disability. *Remedial and Special Education, 7* (1), 6-10.

GRAINGER, J., DUFAU, S. & ZIEGLER, J.C. (2016). A Vision of Reading. *Trends in Cognitive Sciences*, *20* (3), 171-179.

HABIB, M. (2014). *La Constellation des dys : bases neurologiques de l'apprentissage et de ses troubles*. Paris-Bruxelles : DeBoek.

HABIB, M. & ZIEGLER, J.C. (2016). Dyslexie et troubles apparentés : une revue critique de 15 ans de recherche scientifique « Perspectives thérapeutiques ». In S. Pinto & L. Sato (Eds.), *Traité de neurolinguistique* (pp. 329-336). Louvain-la-Neuve : De Boeck Supérieur.

INSERM (2007). *Dyslexie, dysorthographie et dyscalculie. Bilan des données scientifiques*. Éditions INSERM.

KYLE, F., KUJALA, J., RICHARDSON, U., LYYTINEN, H. & GOSWAMI, U. (2013). Assessing the Effectiveness of Two Theoretically Motivated Computer-Assisted Reading Interventions in the United Kingdom: GG Rime and GG Phoneme. *Reading Research Quarterly, 49*, 59-74.

LANDERL, K., WIMMER, H. & FRITH, U. (1997). The impact of orthographic consistency on dyslexia: A German-English comparison. *Cognition, 63* (3), 315-334.

© copyright Anae

LEFAVRAIS, J. (1965). *Test de l'Alouette (révision 2005)*. Paris : ECPA.

LETE, B., SPRENGER-CHAROLLES, L. & COLÉ, P. (2004). MANULEX: a grade-level lexical database from French elementary school readers. *Behavior Research Methods Instruments & Computers, 36* (1), 156-166.

MAGNAN, A., ECALLE, J., VEUILLET, E. & COLLET, L. (2004). The effects of an audio-visual training program in dyslexic children. *Dyslexia, 10* (2), 131-140.

MENGHINI, D., FINZI, A., BENASSI, M., BOLZANI, R., FACOETTI, A., GIOVAGNOLI, S. & VICARI, S. (2010). Different underlying neurocognitive deficits in developmental dyslexia: a comparative study. *Neuropsychologia, 48* (4), 863-872.

NATIONAL INSTITUTE OF CHILD HEALTH AND HUMAN DEVELOPMENT (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction (NIH Publication No. 00-4769)*. Washington DC: US Government Printing Office.

NEW, B., FERRAND, L., PALLIER, C. & BRYSBAERT, M. (2006). Reexamining the word length effect in visual word recognition: new evidence from the English Lexicon Project. *Psychonomic Bulletin & Review, 13* (1), 45-52.

NORTON, E.S., BEACH, S.D. & GABRIELI, J.D. (2015). Neurobiology of dyslexia. *Current Opinions in Neuobiology , 30*, 73-78.

OECD (2014). *PISA 2012 Results: What Students Know and Can Do – Student Performance in Mathematics, Reading and Science* (Vol. I, Revised edition, February 2014). PISA: OECD Publishing.

OJANEN, E., RONIMUS, M., AHONEN, T., CHANSA-KABALI, T., FEBRUARY, P., JERE-FOLOTIYA, J. & LYYTINEN, H. (2015). GraphoGame - a catalyst for multi-level promotion of literacy in diverse contexts. *Frontiers in Psychology, 6*, 671.

PEEREMAN, R., LETE, B. & SPRENGER-CHAROLLES, L. (2007). Manulex-infra: distributional characteristics of grapheme-phoneme mappings, and infralexical and lexical units in child-directed written material. *Behavioral Research Methods, 39* (3), 579-589.

PEEREMAN, R., SPRENGER-CHAROLLES, L. & MESSAOUD-GALUSI, S. (2013). The contribution of morphology to the consistency of spelling-to-sound relations: A quatitative analysis based on French elementary school readers. *Année psychologique, 113*, 3-33.

PENNINGTON, B.F. (2006). From single to multiple deficit models of developmental disorders. *Cognition, 101* (2), 385-413.

PERRY, C., ZIEGLER, J.C. & ZORZI, M. (2007). Nested incremental modeling in the development of computational theories: the CDP+ model of reading aloud. *Psychological Review, 114* (2), 273-315.

POTOCKI, A., MAGNAN, A. & ECALLE, J. (2015). Computerized trainings in four groups of struggling readers: Specific effects on word reading and comprehension. *Research in Developmental Disabilities, 45-46*, 83-92.

PRABHAKARAN, V., SMITH, J.A., DESMOND, J.E., GLOVER, G.H. & GABRIELI, J.D. (1997). Neural substrates of fluid reasoning: an fMRI study of neocortical activation during performance of the Raven's Progressive Matrices Test. *Cognitive Psychology, 33* (1), 43-63.

RAMUS, F. (2001). Dyslexia - Talk of two theories. *Nature, 412* (6845), 393-395.

RAVEN, J.C. (1976). *Standard progressive matrices: Sets A, B, C, D & E*. Oxford: Psychologists Press.

RICHARDSON, U. & LYYTINEN, H. (2014). The Graphogame method: the theoretical and methodological background of the technology-enhanced learning environment for learning to read. *Human Technology, 10*, 39-60.

SAINE, N.L., LERKKANEN, M.K., AHONEN, T., TOLVANEN, A. & LYYTINEN, H. (2010). Computer-assisted remedial reading intervention for school beginners at risk for reading disability. *Child Development, 82* (3), 1013-1028.

SAKSIDA, A., IANNUZZI, S., BOGLIOTTI, C., CHAIX, Y., DEMONET, J. F., BRICOUT, L. & RAMUS, F. (2016). Phonological skills, visual attention span, and visual stress in developmental dyslexia. *Developmental Psychology, 52* (10), 1503-1516.

SHARE, D.L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55*(2), 151-218.

SIEGEL, L.S., SHARE, D. & GEVA, E. (1995). Evidence for superior orthographic skills in dyslexics. *Psychological Science, 6* (4), 250-254.

SIMOS, P.G., FLETCHER, J.M., BERGMAN, E., BREIER, J.I., FOORMAN, B.R., CASTILLO, E.M. & PAPANICOLAOU, A.C. (2002). Dyslexia-specific brain activation profile becomes normal following successful remedial training. *Neurology, 58* (8), 1203-1213.

SPRENGER-CHAROLLES, L. & COLÉ, P. (2003). *Lecture et dyslexie : approche cognitive*. Paris : Dunod.

SPRENGER-CHAROLLES, L. & COLÉ, P. (2013). *Lecture et dyslexie : approche cognitive (2e édition)*. Paris : Dunod.

SPRENGER-CHAROLLES, L., SIEGEL, L. S., BECHENNEC, D. & SERNICLAES, W. (2003). Development of phonological and orthographic processing in reading aloud, in silent reading, and in spelling: A four-year longitudinal study. J*ournal of Experimental Child Psychology, 84* (3), 194-217.

SPRENGER-CHAROLLES, L., SIEGEL, L.S., JIMENEZ, J.E. & ZIEGLER, J.C. (2011). Prevalence and Reliability of Phonological, Surface, and Mixed Profiles in Dyslexia: A Review of Studies Conducted in Languages Varying in Orthographic Depth. *Scientific Studies of Reading, 15* (6), 498-521.

STEIN, J. (2014). Dyslexia: the Role of Vision and Visual Attention. *Current Developmental Disorders Reports, 1* (4), 267-280.

VAN BERGEN, E., VAN DER LEIJ, A. & DE JONG, P.F. (2014). The intergenerational multiple deficit model and the case of dyslexia. *Frontiers in human neuroscience, 8* (346).

VIDYASAGAR, T.R. & PAMMER, K. (2010). Dyslexia: a deficit in visuo-spatial attention, not in phonological processing. *Trends in Cognitive Sciences, 14* (2), 57-63.

ZIEGLER, J.C., BERTRAND, D., LÉTÉ, B. & GRAINGER, J. (2014). Orthographic and phonological contributions to reading development: Tracking developmental trajectories using masked priming. *Developmental Psychology, 50* (4), 1026-1036.

ZIEGLER, J.C., CASTEL, C., PECH-GEORGEL, C., GEORGE, F., ALARIO, F.X. & PERRY, C. (2008). Developmental dyslexia and the dual route model of reading: Simulating individual differences and subtypes. *Cognition, 107*, 151-178.

ZIEGLER, J.C. & GOSWAMI, U. (2006). Becoming literate in different languages: similar problems, different solutions. *Developmental Science, 9* (5), 429-436.

ZIEGLER, J.C., JACOBS, A.M. & STONE, G.O. (1996). Statistical analysis of the bidirectional inconsistency of spelling and sound in French. *Behavior Research Methods, Instruments & Computers, 28* (4), 504-515.

ZIEGLER, J.C., PECH-GEORGEL, C., DUFAU, S. & GRAINGER, J. (2010). Rapid processing of letters, digits, and symbols: What purely visual-attentional deficit in developmental dyslexia? *Developmental Science, 13*, F8-F14.

ZIEGLER, J.C., PERRY, C. & ZORZI, M. (2014). Modelling reading development through phonological decoding and self-teaching: Implications for dyslexia. *Philosophical Transactions of the Royal Society B, 369*, 20120397.

ZOCCOLOTTI, P., DE LUCA, M., DI PACE, E., GASPERINI, F., JUDICA, A. & SPINELLI, D. (2005). Word length effect in early reading and in developmental dyslexia. *Brain and Language, 93* (3), 369-373.

ZORZI, M., BARBIERO, C., FACOETTI, A., LONCIARI, I., CARROZZI, M., MONTICO, M. & ZIEGLER, J.C. (2012). Extra-large letter spacing improves reading in dyslexia. *Proceedings of the National Academy of Sciences, 109* (28), 11455-11459.

© copyright Anae

Positive effects of GraphoGame reading intervention in French: the importance of initial reading level and engagement

Julie Lassault[1], Liliane Sprenger-Charolles[1], Jean-Patrice Albrand[1], Edouard Alavoine[1], Ulla Richardson[2], Heikki Lyytinen[2,3], Johannes C. Ziegler[1]

1 Aix-Marseille Univ, CNRS, Laboratoire de Psychologie Cognitive (UMR 7290), France

2 University of Jyväskylä, Centre for Applied Language Studies, Jyväskylä, Finland

3 UNESCO Chair on Inclusive Literacy Learning for All, University of Jyväskylä, Jyväskylä, Finland

Author Note

Johannes C. Ziegler https://orcid.org/0000-0002-2061-5729

Correspondence should be addressed to Johannes C. Ziegler, Laboratoire de Psychologie Cognitive, Aix-Marseille University, 3, place Victor Hugo, 13003 Marseille. France. Email: Johannes.Ziegler@univ-amu.fr

Abstract

The aim of the present study was to test a French version of GraphoGame (GG) that was developed to cover the entire 1st grade reading program based on a theoretically optimal progression. GG intervention was tested in a large school sample of grade 1 readers (N=921) from socioeconomically disadvantaged neighborhoods. GG was assessed against an equally engaging math intervention. Teacher and school effects were kept constant by using the math intervention in the same schools with the same teachers one year later. Progress in reading was assessed with 4 standardized tasks that measured phoneme suppression, reading aloud of single words and text and word recognition before and after the intervention. Results were analyzed using hierarchical mixed effect models, which took into account classes as random factors along with two important moderator variables: interindividual differences of the children in pre-test measures and within-game engagement. We found that GG had a positive effect in the four reading tasks (effect sizes between .11 and .26). Most importantly, the main effect of the GG intervention interacted with initial pre-test level and engagement. That is, initially weak students benefitted to a greater extent from GG intervention than initially good students and students with high levels of engagement benefitted more from GG intervention than students with low levels of engagement. Together, the results argue against "all size fits all solutions" and suggest that effect sizes are underestimated if initial reading level and within-game variables, such as engagement, are not taken into account.

Key words: Reading, Dyslexia, Intervention, Digital, Classroom, Serious Games

Positive effects of GraphoGame reading intervention in French: the importance of initial reading level and engagement

Introduction

International reading assessments (PISA) continue to show that, on average across OECD countries, 23% of the 15-year old students fail to acquire the technical skills to read (see OECD, 2019). These students are unable to identify the main idea in a text of moderate length and unable to find information based on explicit criteria. In almost all countries that participated in PISA 2018, students who were socio-economically disadvantaged were less likely to attain the minimum level of proficiency in reading. The strength of the relationship between a student's socio-economic status and his or her performance varies greatly across countries and was particularly strong in France, Hungary, Luxembourg, Peru, the Philippines, Romania and the Slovak Republic (OECD, 2019). This result is in line with previous epidemiological studies, which showed that French pupils from socially-disadvantaged neighborhoods were almost 10 times more likely to encounter reading difficulties than pupils from socially-advantaged neighborhoods (Fluss et al., 2009). In PISA 2018, the gap in reading performance in France (but also in Germany, Hungary, Israel, Peru and the Slovak Republic) between the 10% most socio-economically advantaged and the 10% most disadvantaged students was over 170 score points – the equivalent of well over four years of schooling.

There is an increasing consensus that many children are at risk for reading difficulties from their first day of school (Loeb & Bassok, 2007) and that early and focused interventions for at-risk children within the school setting is the way to go about reducing them (Castles, Rastle, & Nation, 2018; Foorman, Breier, & Fletcher, 2003; Foorman & Torgesen, 2001; Hulme, Snowling, West, Lervåg, & Melby-Lervåg, 2020; Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001; Ziegler, Perry, & Zorzi, 2020). The recommendations of OECD (2019) also state that interventions should occur early with a special focus on socio-economically disadvantaged students and/or schools. There are indeed some examples of fairly successful early interventions (e.g., oral language training) that produce sustainable effects especially when implementation quality is high (for review, see Hulme et al., 2020).

In the context of early focused intervention for reading within the school context, educational technology, and in particular computer-assisted instruction (CAI), held the promise to support the learning process in unpreceded ways because of its potential to provide individualized and adaptive support (Cheung & Slavin, 2013). However, one of the first studies that systematically evaluated five first-grade reading software products in 43 schools (2600 students) in the US found no significant effect of CAI on reading outcomes (Dynarski et al., 2007). In a follow-up study, the effect was even negative for low-achieving students (Campuzano et al., 2009). In their review of 14 studies, Slavin, Lake, Davis,

and Madden (2011) found only very small effects (Effect Size, ES = 0.09) of educational technology programs for struggling elementary readers. More recently, the synthesis of 20 CAI studies for elementary school reading showed somewhat larger but still small effects (ES = 0.14). Thus, by and large, there is deceivingly little evidence that educational technology programs have the expected effect of improving reading skills for at-risk children in primary school. One of the reasons might have to do with the students' motivation. Indeed, in a meta-analysis of 31 studies, Wouters, van Nimwegen, van Oostendorp, and van der Spek (2013) found that, contrary to common beliefs, serious games were not more motivating than conventional instruction methods.

GraphoGame intervention

A highly promising serious game, which had not been tested in any of the above-mentioned meta-analyses, is GraphoGame (GG, Richardson & Lyytinen, 2014). GG was initially developed in Finland in the context of dyslexia prevention (Lyytinen, Erskine, Hamalainen, Torppa, & Ronimus, 2015), but its potential as a tool for promoting reading acquisition and providing early intervention was quickly recognized (Lyytinen, Erskine, Kujala, Ojanen, & Richardson, 2009). Over the past years, GG has been implemented and tested in many languages, such as German (Brem et al., 2010; Huemer, Landerl, Aro, & Lyytinen, 2008), English (Kyle, Kujala, Richardson, Lyytinen, & Goswami, 2013), Portuguese (Carvalhais, Richardson, & Castro, 2018) or Spanish (Rosas, Escobar, Ramírez, Meneses, & Guajardo, 2017). The core idea of GG is in line with major theories of reading acquisition, which are all based on the idea that the initial steps of reading acquisition are all about learning the mapping between visual symbols (e.g., letters, graphemes) and their equivalent units in spoken language (Castles et al., 2018; Ehri, 1992; Goswami, 2002; Share, 1995; Ziegler & Goswami, 2005; Ziegler, Perry, & Zorzi, 2014). This mapping provides an extremely parsimonious and straightforward way to retrieve the spoken form and therefore the meaning of the thousands of words children have stored in their phonological lexicon prior to reading (Share, 1995).

In line with these theories, the key idea behind GG was to systematically introduce these mappings (starting small and consistent) and, most importantly, find a way to intensively train and automatize them. This is done in GG by presenting auditory stimuli and have children select or manipulate their orthographic equivalents in various games that are designed to be motivating to the child (for examples, see Figure 1). While the original Finnish version was a basic letter-sound game, which might be sufficient in a highly transparent writing system where literally all words can be correctly decoded on the basis of 20 letter-sound correspondences (see Landerl et al., 2013; Ziegler et al., 2010), the French but also English version of GG introduces and trains these mappings at various grain sizes including graphemes, syllables, rimes, whole words and even sentences (Kyle et al., 2013; Lassault & Ziegler, 2018).

The first validation study of GG was conducted in Finland with beginning readers who were at risk of reading difficulties (Saine, Lerkkanen, Ahonen, Tolvanen, & Lyytinen, 2010, 2011). The authors initially screened 166 children and then selected 50 of them (the lowest performing 30%) for the intervention study. All children received a regular phonics-based remedial training that was spread out over a 28-week period, with four weekly 45-minute sessions. Each session contained prereading activities, word segmentation, decoding and spelling activities and vocabulary training. For half of the children (N=25), which were selected randomly, the prereading activities of the regular phonics-based remedial training were replaced by a 15-minute GG session. The results showed that the additional GG intervention produced significant improvements over and above the regular phonics-based remedial training not only in letter knowledge, decoding, and accuracy, but also in fluency and spelling (the effect sizes varied between 0.22 and 2.08).

In contrast to the impressive results of Saine et al. (2010, 2011), a recent meta-analysis by McTigue, Solheim, Zimmer, and Uppstad (2020), which included 15 studies that used the GG intervention in 7 different languages (English, Polish, Spanish, German, Finnish, Portuguese, Kiswahili), came to a somewhat less optimistic conclusion. Across the 15 studies (19 comparisons), the authors reported a slightly negative mean effect size (g[1]= −0.02, 95% CI [−0.14, 0.09]) that was not significantly different from zero (p = .70). The authors came to the devastating conclusion that "as a word-reading intervention, there is no evidence that GG produced growth in students' word reading… although students often learned from GG, their learning did not typically surpass that of control groups" (McTigue et al., 2020, p. 60). However, the authors found a wide range of effect sizes (g = −1.07 to 1.55), which was taken to suggest that additional, contextual variables might moderate the intervention effects. They investigated the effects of four moderator variables, which were orthographic transparency of the writing system (shallow, moderate, deep), duration of the intervention (short, medium, long), type of control group (untreated control, math game, reading intervention) and the level of adult interaction (low, high). They found that only the level of adult interaction significantly moderated the effects. Studies with a high level of adult interaction showed significant effect sizes (g = .48).

Goal of the present study

The goal of the present study was four-fold. First, we wanted to test a French version of GG that was developed to cover the entire 1st grade reading program based on a theoretically optimal progression (Dehaene, 2011; Sprenger-Charolles, 2017), which took into account the frequency and consistency of grapheme-phoneme correspondences (GPCs) and the frequency of words (Peereman & Sprenger-Charolles, 2018). That is, it was structured in a way that the most consistent and frequent grapheme-phoneme correspondences were introduced progressively and they were then trained in the context of

[1] Hedges's g was calculated as the difference in gain (pretest to posttest) between the GG and control groups divided by the pooled standard deviation.

syllables, words and sentences that were made up of previously trained GPCs (Lassault & Ziegler, 2018). Thus, the French version of GG went well beyond teaching basic letter-sound correspondences[2]. Second, we wanted to test the GG intervention in a large school sample of beginning readers (1 grade) from socioeconomically disadvantaged neighborhoods because the need for early within-school intervention is particularly high in this part of the population and this is particularly true in France (Fluss et al., 2008; Fluss et al., 2009). Third, we wanted to assess the GG intervention against an equally engaging math intervention while keeping teacher and school effects constant. It is well known that teachers effects are the dominant factor affecting student academic gain (Sanders, Wright, & Horn, 1997). Finally, we wanted to investigate whether intervention gains were modulated by some moderator variables, such as the initial reading level of the child and within-game engagement. Surprisingly few previous studies included variables that allowed one to estimate students' engagement and motivation (McTigue et al., 2020). In our study, engagement was assessed through with-game variables. We actually used the highest-level obtained in the game because this measure was available for both GG and math intervention. This variable reflects a mixture between perseverance, success and motivation.

Methods

Participants

A total of 921 French participated in the study. They attended grade 1 in nine public schools (37 classes) that were located in socially disadvantaged neighborhoods (priority education areas, PEA) in the city of Marseille, south of France. The schools and classes were selected by the local school authorities on the basis of the following criteria: they had to have at least four classes in that grade, they had to be accessible by public transportation and they should not have been involved in other intervention programs.

In the first year of our study, 451 children (239 boys and 212 girls) took part in the GG intervention. Their average age was 6;4 years (range 5;10 – 8;1) at the time of pre-tests (T1), which was in November (~2 months after the beginning of the school year). In the second year, 470 (244 boys and 226 girls) took part in the math intervention. Their mean age was 6;4 years (range 5;10 – 8;1) at T1. The children came from the same 9 schools (36 classes). Most of the teachers participated in the study for two years and hence delivered the GG intervention in the first year and the math intervention in the second year. This is possible in France because teachers change classes every year (i.e., after teaching a grade 1 class in one year, a teacher will take on a new grade 1 class the next year).

[2] The French version of GG is freely available in France and French-speaking territories on google play and apple store. It can be downloaded in all other countries for a small fee.

Interventions

Reading Intervention using GG

GG is an audio-visual reading training software based on the simultaneous presentation of auditory stimuli and spelling choices embedded in a series of games (Richardson & Lyytinen, 2014). The children typically hear a phoneme, syllable word or sentence and, at the same time, several orthographic choices are presented on the screen (for examples, see Figure 1). In most games, the children have to choose amongst the alternatives the one that matches the spoken input. The number of alternatives increases adaptively as a function of a child's performance. Errors are indicated immediately, and children have to correct incorrect responses. To make sure that children find the correct answer and learn from the error, only the correct option is available immediately after an error occurred. If children make more than 25% errors in a given level, they cannot move on to the next level. However, after five repetitions of the same level, they are allowed to move on to the next level.

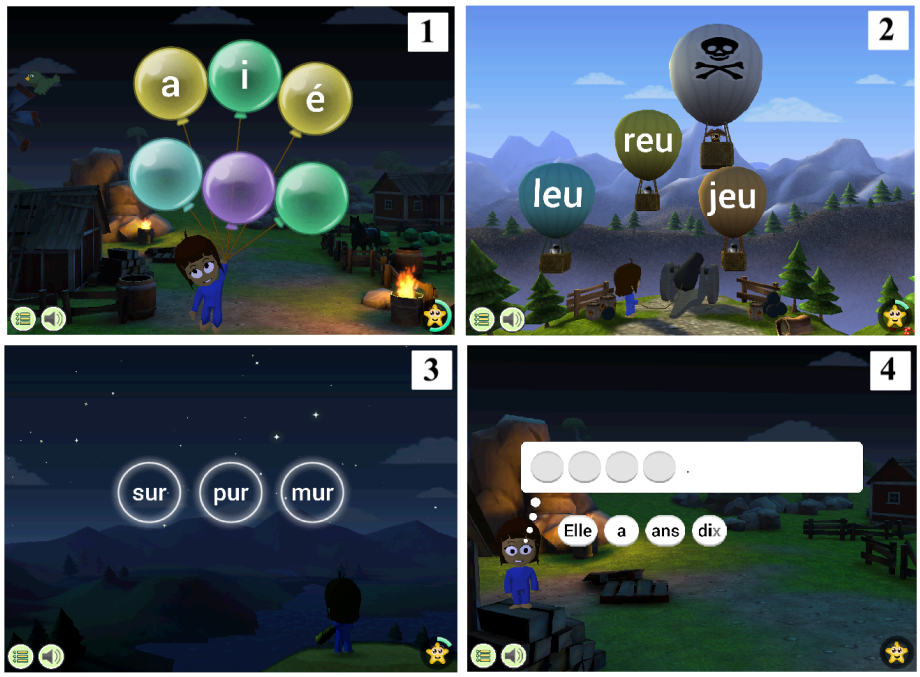


Figure 1: An illustration of four of the levels used in the the French version of GraphoGame: 1. Phoneme level; 2. Syllable level 3. Word level; 4. Sentence level

The content of the French version of GG was specifically developed for this school-based intervention study on the basis of the theoretically optimal progression initially described by Dehaene (2011) et further developed by Sprenger-Charolles (2017), which took into account the frequency and consistency of grapheme-phoneme correspondences (Peereman & Sprenger-Charolles, 2018) as they occurred in a grade-level lexical database of 54 elementary school French textbooks (MANULEX; Lété, Sprenger-Charolles, & Colé, 2004), which contained approximately 2,000,000 non-lemmatized forms. The progression is built such that the most frequent and consistent GPCs are introduced first and more

complex or context-sensitive correspondences are introduced later. Apart from a small number of highly frequent and potentially inconsistent sight words, all newly introduced words can be decoded on basis of previously learnt GPCs.

The theoretical progression was organized into 67pedagogical sessions (sequences), which were organized around a specific content, such as consistent vowels (a-i-o-u-e-eu-ou-ou-e) and consistent consonants (j-f-l-r) in sequence 1. Each sequence contained about 10 levels, which trained the specific content of the sequence across various games. Each game lasted about 2-3 minutes. Thus, each sequence would last about 15 to 30 minutes depending on the speed and accuracy of the child. A description of the 67 sequences can be found in Appendix A. Note that some sequences are repetitions of earlier ones, such that the same content was reviewed again from time to time. At the beginning and end of each sequence, we added a pre-test and post-test level that contained some typical items of each sequence to test the within-game progression for each sequence.

Math Intervention

We used a commercial application for training mathematical skills, *Fiete Math* (Ahoiii Entertainment), which was based on the principles of *Number Catcher* (Dotan, Dehaene, Huron, & Piazza, 2011), the follow-up version of *Number Race (Wilson et al., 2006).* This commercial application was preferred over Number Catcher because it contained many more training sessions. In *Fiete Math*, children use a shipload to calculate "containers" of numbers that can be put together or taken apart (see Figure 2). The tasks are simple: arriving ships have to be loaded appropriately. Like building blocks, the children can assemble the load into any shape and divide them in order to reach the right size (that changes on every trial). In order to keep the children motivated, they can unlock new lands and ports, win medals and larger ships. Each port has an arithmetic skill to work on: count (counting, recognize groups), split (divide numbers differently, get to know number space up to 20), value of 5th (calculate with base 5, decimal transition, structure of the decimal system), operating techniques (doubling and halving, exchange tasks, neighbour tasks, change in the opposite direction or in the same direction), addition and subtraction (logical digital suites from simple to difficult tasks, reverse tasks), free calculation (practice addition and subtraction tasks). The game contains over 1000 consecutive tasks from the number range up to 30.

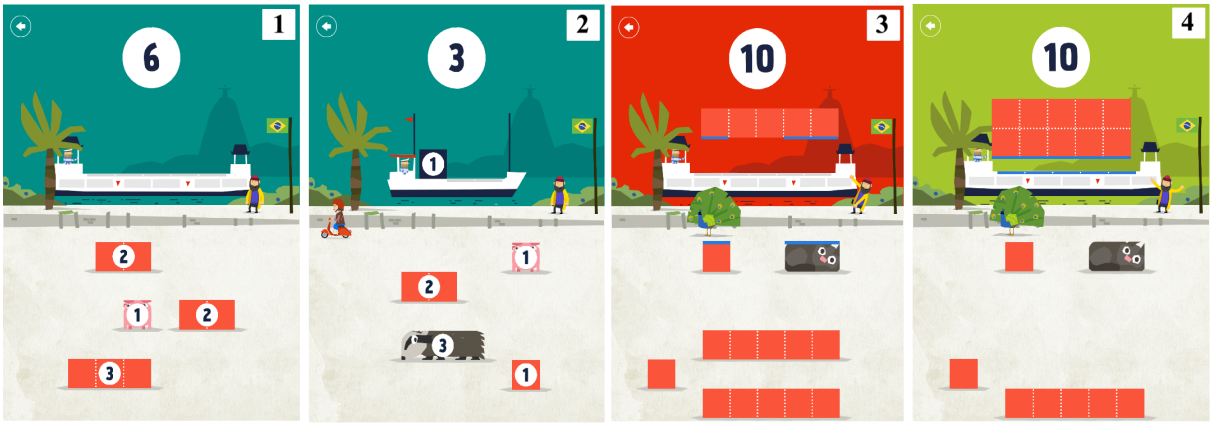


Figure 2: Illustrations of four different levels in Fiete Math: 1. Load an empty ship ; 2.Load a partially loaded ship ; 3. Feedback when the shipment is not accurate (red); 4. Feedback when the shipment is correct (green)

Assessment Tasks

Reading Fluency and Decoding Skills

Two tasks were used to assess reading fluency at T1 and T2. The first was the classic One-Minute Reading test, in which children had to read aloud a list of familiar words (max 35) and a list of pseudowords (max 30) in one minute each (Gentaz, Sprenger-Charolles, Theurel, & Colé, 2013). The score was the number of words read correctly in one minute. The second task was the classic standardized French reading test *Alouette* (Lefavrais, 2005), in which children have to read aloud a meaningless text. The score is the number of correctly read words in 3 minutes.

Word identification

We used a standardized word identification test, the *Timé 2* (Ecalle, 2003). This test is divided into 3 tasks (12 items each). In the first task, the examiner pronounces a target word and the student must find the target word among 5 written alternatives, which are the correct word, a homophone, a visually close pseudo word, a spelling neighbor or a pseudoword. In the second task, the student is presented with a picture and has to identify the corresponding word amongst five alternatives (same as in task 1). In the third task, the student is presented with a written word and has to find amongst four alternatives the one that is semantically associated with the target word (for example, for the target word, such as *ocean*, the alternatives could be *ship, shop, shp or shoe).*

Phonemic Awareness

Three tasks from the EVALEC battery (Sprenger-Charolles, Colé, Bechennec, & Kipffer-Piquard, 2005) were used to assess phoneme awareness. In the first task, the examiner pronounced a tri-syllabic pseudoword and children had to delete the first syllable (10 items). In the second task, the examiner pronounced a consonant-vowel-consonant pseudoword and the students had to delete the first

phoneme (12 items). In the third task, the examiner pronounced a consonant-consonant-vowel pseudoword and students had to delete the first phoneme (12 items). The final score was the total of correct responses.

Comprehension

The students performed an oral comprehension test at T1 and a written comprehension test at T2. The oral comprehension task contained 25 sentences from the *ECoSSe* (Lecocq, 1996), which is the French equivalent of the Test for Reception of Grammar (TROG). The examiner read a sentence aloud and child had to select the corresponding image out of four alternatives. Six sentences structures were used (active, pronoun, double negation, spatial terms, relative and passive). For the reading comprehension test, the procedure was similar except that instead of hearing the sentences, the students had to read them. The same 6 sentences structures were used for the 14 items of the test.

Vocabulary

A standardized receptive vocabulary test was used, the *TVAP* (Deltour & Hupkens, 1984). The test is composed by 30 items. The examiner pronounces a word and the student had to select the corresponding image out of six images. Each answer was scored on a scale of 0 to 2 (2 points were awarded for a correct answer, 1 point for the approximate answer and 0 for the others). The maximum possible score was therefore 60.

Non-verbal IQ

Raven's Progressive Matrices (Raven, 1976) were used to assess nonverbal intelligence. This version, for children aged between 4 and 11 years, is composed of 36 matrices (divided into 3 series A, AB and B).

Mathematical abilities

Three tasks were taken from the TEDI math evaluation battery : 6 pictorial operations where the examiner describe a picture and ask a question to the child (for example “there are two red balloons and three blue balloons. How many balloons are there in all?”); 8 classical additions and 4 incomplete additions. Another task of comparing pairs of numbers was assessed. On a tablet 36 pairs of numbers were presented one by one, children had to press as quickly as possible on the largest number of the pair. The list contained 4 pairs per number difference (from 1 to 9), with 2 correct answers on the right and 2 on the left. The digital support made possible to accurately collect the accuracy and time of the response (in ms).

Procedure

The children of the two cohorts (year 1, year 2) participated in a series of pre-tests that took place in November of grade 1 (T1). The post-tests took place in June of the following year (T2). At T1,

each child was tested for decoding, word identification, phonemic awareness, listening comprehension, vocabulary, non-verbal IQ and mathematical skills (see below). The order of tests was the same for all participants. The pre-tests were carried out in 2 separate sessions and lasted about 1 hour. At T2, the same tests were used, except for IQ and vocabulary, and listening comprehension was replaced by reading comprehension. Both assessments took place in a quiet room in the schools. The children were tested individually, except for the word identification assessment, which was done in the classroom.

The interventions took place between T1 and T2 (i.e., 5 months). Children were typically trained in groups of six, which in most cases corresponded to half of the class (footnote starting in 2018, the size of a class in PEAs in France has been reduced to 12 children per class). Each child was attributed an individual Android tablet, which remained at the school. The tablet only contained the GG application (year 1) or the math application (year 2). Each child used the GG and Math application approximately 20 minutes per day, 4 times per week, for a total of 16 weeks. The mean duration of the GG intervention was 14.73 h (+/- 3.51 h)

Statistical Analyses

The general analysis strategy was the following: Given that each individual *i* is nested in a class *j*, we used hierarchical linear mixed models to take class into account as a random effect using the IBM SPSS MIXED procedure (see Heck, Thomas & Tabata, 2009). For each critical variable (e.g., phoneme deletion scores, reading fluency), we predicted the post-test scores on a particular test by the pre-test-scores on the same test. This allows us to express the overall gain of a child on a given test as a function of his or her pre-test level.

$$Y_{ij}\,(\text{post}) = \beta_{0j} + \beta_1 x_{ij}(\text{pre}) + u_{0j} + \varepsilon_{ij} \qquad (1)$$

Next, we added intervention as a fixed effect because we were interested in finding out whether the scores on the post-test were differently affected by the two interventions (GG versus Math) when the initial level of the child (his/her score on the pre-test) was taken into account (see equation 2). We also added the interaction between intervention and pre-test scores because one could expect that the efficiency of the training depends on the initial level of the child. Interaction terms were only added if they significantly increased the model fit (log likelihood ratio) in a forward model selection procedure.

$$Y_{ij}\,(\text{post}) = \beta_{0j} + \beta_1 x_{ij}(\text{pre}) + \beta_2\,\text{intervention}_j + u_{0j} + \varepsilon_{ij} \qquad (2)$$

Next, we added student engagement (highest level obtained within the GG or math intervention) as another fixed effect because we expected that student engagement should affect the efficiency of the intervention (equation 3).

$$Y_{ij}\,(\text{post}) = \beta_{0j} + \beta_{1j} x_{ij}(\text{pre}) + \beta_2\,\text{intervention}_j + \beta_3 x_{ij}(\text{engagement}) + u_{0j} + \varepsilon_{ij} \qquad (3)$$

We also added the interaction between intervention * engagement and the triple interaction between intervention * pre-test * engagement but only if they significantly increased the model fits.

Results

The mean results for each dependent variable as a function of group (GG versus Math) and time points (T1 and T2) is presented in Table 1.

Table 1. Means (SDs) for each dependent variable as a function of group (GG versus Math) for the pre-test (T1) and the post-test

Tests	T 1 - Mean (SD)			T 2 - Mean (SD)		
	All	GG	Math	All	GG	Math
Phoneme Awareness						
Syllable suppression	6.12 (3.36)	6.34 (3.43)	5.91 (3.28)	7,08 (3,22)	7,65 (3,02)	6,54 (3,31)
Phoneme suppression (CVC)	5.25 (4.39)	4.80 (4.63)	5.68 (4.09)	8.40 (3.90)	8.55 (3.86)	8.2((3.94)
Phonemic suppression (CCV)	5.13 (3.75)	4.66 (4.03)	5.57 (3.41)	7.13 (3.63)	7.25 (3.81)	7.01 (3.46)
Reading Aloud skills						
1-minute Reading (Words)	9.59 (8.59)	7.94 (6.89)	11.18 (9.69)	36.87 (22.34)	35.38 (21.64)	38.30 (22.93)
1-minute Reading (Pseudowords)	9.50 (7.08)	7.92 (5.56)	11.01 (8.00)	26.36 (16.20)	26.93 (16.23)	29.72 (16.06)
3-minute Reading (Text. Alouette)	14.51 (17.21)	11.21 (12.67)	17.69 (20.16)	69.73 (42.93)	66.31 (40.85)	73.02 (44.64)
Word identification						
Oral Presentation	3.80 (2.19)	3.40 (2.05)	4.17 (2.26)	7.75 (2.57)	7.90 (2.60)	7.59 (2.54)
Visual Presentation	2.69 (1.79)	2.59 (1.67)	2.78 (1.89)	5.28 (2.35)	5.42 (2.41)	5.14 (2.28)
Semantic Associates	2.50 (1.51)	2.43 (1.53)	2.56 (1.50)	4.27 (2.37)	4.40 (2.42)	4.14 (2.31)
Comprehension						
Listening comprehension	18.41 (4.52)	18.36 (4.61)	18.45 (4.43)	-	-	-
Reading comprehension	-	-	-	8.31 (3.28)	7.98 (3.36)	8.63 (3.18)
Vocabulary	34.59 (8.92)	34.67 (10.04)	34.51 (7.70)	-	-	-
Nonverbal IQ	18.27 (4.93)	18.58 (5.15)	17.97 (4.70)	-	-	-
Mathematical abilities						
Pictural operations	5.11 (1.23)	5.14 (1.24)	5.07 (1.22)	5.45 (0.89)	5.45 (0.95)	5.45 (0.83)
Classic additions	4.78 (2.33)	4.48 (2.42)	5.07 (2.21)	6.35 (1.85)	6.09 (1.94)	6.60 (1.74)
Incomplete additions	1.79 (1.66)	1.63 (1.68)	1.93 (1.63)	2.92 (1.44)	2.62 (1.54)	3.20 (1.28)
Number comparison	32.07 (5.46)	31.64 (5.82)	32.48 (5.08)	33.60 (3.93)	33.62 (3.69)	33.58 (4.15)

As explained above, the data were analyzed using hierarchical mixed effect models (IBM SPSS MIXED) with subjects nested in classes as random effects. In the following sections, we present the ANOVA III results of the fixed effects and interactions, the parameter estimates of the final hierarchical mixed effect models, and the fitted data for the four critical tests that evaluated the effects of intervention outside the game: 1. phoneme suppression, 2. word recognition (Timé 3 test), 3. text reading fluency (Alouette test), and 4. single word reading (1-minute reading test). We also analyzed the effects of the math intervention compared to GG intervention of the math test. To get an idea about the effect size (ES) of the GG intervention, we calculated the standardized coefficients, which can be obtained by standardizing (i.e. M = 0; SD = 1) each variable before analysis (for an explanation why standardized coefficient are appropriate effect size measures in mixed effect models, see Lorah, 2018).

1. Phoneme Suppression

The final model included the fixed effects of intervention (GG versus Math), pre-test level and engagement of the intervention along with all interaction terms. The ANOVA III analyses showed significant fixed effects of intervention ($F(1, 290.1) = 23.27$, $p < .0001$), pre-test level ($F(1, 897.8) = 191.2$, $p < .0001$) and engagement ($F(1, 891.9) = 70.8$, $p < .0001$). There were significant two-way interactions between the effects of intervention and pre-test ($F(1, 897.8) = 13.4$, $p < .0001$) and intervention and engagement ($F(1, 891.9) = 4.04$, $p < .05$), as well as a triple interaction between intervention, pre-test level and engagement ($F(1, 865.3) = 9.35$, $p < .0001$). To appreciate the direction of the effects and interactions, it is important to look at the parameter estimates, which are shown in Table 2.

Table 2. Non-standardized parameter estimates of the final model predicting phoneme suppression post-test scores on the function of intervention, pre-test level and engagement of the training.

Parameters	Betas	t	Sig.
Random Effects			
Between-class	3.602	3.510†	0.000
Between-Students	10.026	17.101	0.000
Fixed Effects			
[Intervention=GG]	3.755	4.825	0.000
[Intervention=MATH]	0.000	.	.
Pre-test	0.484	11.712	0.000
Engagement	2.021	4.166	0.000
[Intervention=GG] * Pre-test	-0.202	-3.659	0.000
[Intervention=Math] * Pre-test	0.000	.	.
[Intervention=GG] * Engagement	1.451	2.224	0.026
[Intervention=Math] * Engagement	0.000	.	.
[Intervention=GG] * Engagement* Pre-test	-0.118	-3.781	0.000
[Intervention=Math] * Engagement * Pre-test	-0.079	-2.097	0.036

† *z*-value (Wald)

As can be seen in Table 2, GG intervention has a strong positive effect (+3.75 points) over and above Math intervention. The ES for the GG intervention was .26. The interaction between intervention and pre-test level suggests that children with initially weaker phoneme suppression skills benefit to a greater extent from GG intervention (-.20 points; ES = .23; see also Figure 3). The interaction between intervention and engagement shows that GG intervention benefits increased with the engagement of the intervention (+1.45 points; ES = .16). The triple interaction suggests that the biggest GG intervention benefits were obtained when the initial level of children was low and the engagement of the training was high (ES = .13). Figure 1 shows the model-fitted post-test scores as a function of the initial pre-test level for 3 levels of engagement: low engagement (1 SD below the mean engagement), average engagement, and high engagement (1 SD above the mean engagement).

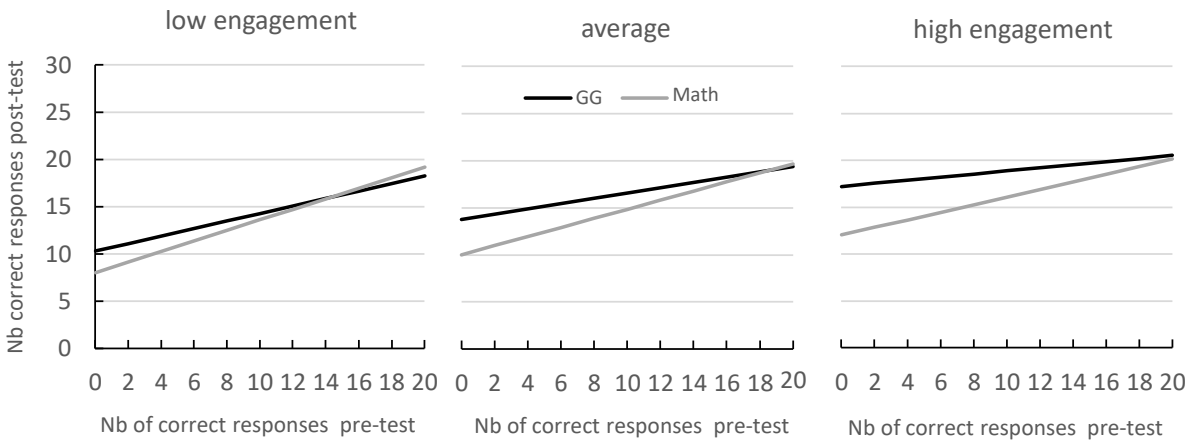


Figure 3. Model-fitted post-test scores on the phoneme suppression test as a function of intervention, pre-test level and the engagement of the training. Low engagement, 1 SD below the mean engagement; Average Engagement, mean engagement; High engagement, 1 SD above the mean engagement.

2. *Word recognition (Timé 3)*

The final model included the fixed effects of intervention, pre-test level, and engagement as well as the two two-way interactions. The analyses showed significant fixed effects of intervention (F1, 519.448) = 44.99, p < .0001), pre-test (F (1, 898.94) = 128.62, p < .0001), and engagement (F(1, 893.50) = 321.44, p < .0001). There were significant interactions between the effects of intervention and pre-test (F(1, 898.94) = 30.01, p < .0001) and the effects of intervention and engagement (F(1, 8993.51) = 57.65, p < .0001). The parameter estimates are shown in Table 3.

Table 3. Parameter estimates of the final model predicting word recognition post-test scores on the function of intervention, pre-test level and engagement of the training

Parameters	Beta	t	Sig.
Random Effects			
Between-class	1.255	2.600†	0.009
Between-Students	10.441	18.793	0.000
Fixed Effects			
[Intervention=GG]	5.439	6.708	0.000
[Intervention=MATH]	0.000		
Pre-test	0.667	13.254	0.000
Engagement	1.700	7.518	0.000
[Intervention=GG] * Pre-test	-0.434	-5.487	0.000
[Intervention=MATH] * Pre-test	0.000		
[Intervention=GG] * Engagement	2.499	7.593	0.000
[Intervention=MATH] * Engagement	0.000		

† *z*-value (Wald)

The results show a clear fixed effect of intervention with GG producing higher post-test word recognition scores (+5.4 points, ES = .25) than Math intervention everything else being equal (i.e., pre-reading level, engagement). Importantly, there was an interaction between intervention and initial word

recognition level, which reflects the fact that GG intervention was more beneficial for readers with poorer pre-test word recognition skills (β = -.0438; ES = .30; see Figure 4). Similarly, there was a significant interaction between the effects of intervention and engagement, which reflects the finding that GG intervention was more efficient (β = +2.5 points; ES = .40) when the engagement of the intervention was high. Figure 4 shows the model-fitted post-test scores as a function of initial word recognition level and 3 training intensities.

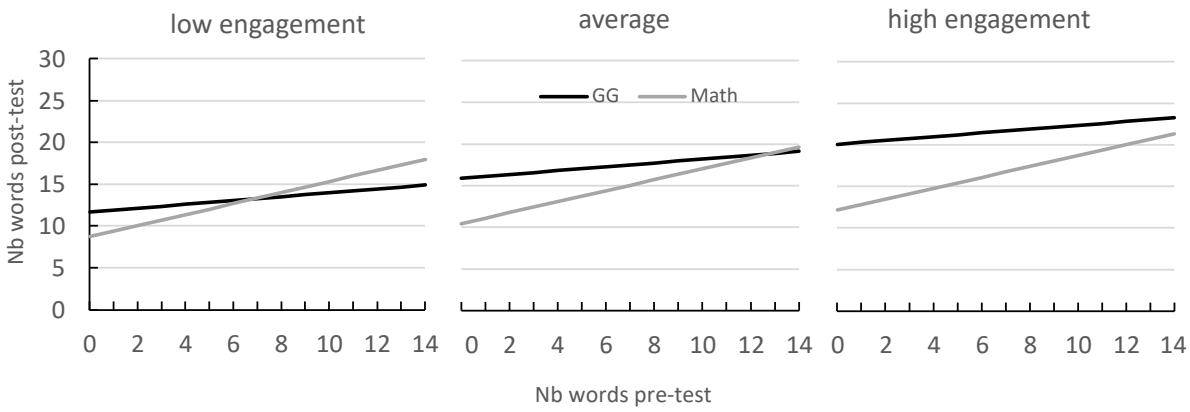


Figure 4. Model-fitted word recognition post-test scores as a function of intervention, pre-test level and the engagement of the training.

Reading aloud text (Alouette)

The final model included the fixed effects of intervention, pre-test level and engagement, the two-way interaction between intervention and engagement and the three way-interaction between intervention, engagement and pre-test level. The fixed effect of intervention failed to reach significance (F1 (1, 81.91) = 2.21, p = .141). However, there was a significant interaction between the effects of intervention and engagement (F(1, 897.5 = 20.8, p < .0001) and a significant triple interaction between the effects of intervention, engagement and pre-test level (F(1, 897.5) = 23.8, p < .0001). The fixed effects of pre-reading level and engagement were also significant (F(1, 896.7) = 300.8, p < .0001 and F(1, 897.1) = 228.03, p < .0001, respectively). The parameter estimates are shown in Table 4.

Table 4. Parameter estimates of the final model predicting text-reading post-test scores on the function of intervention, pre-test level and engagement of the training

Parameters	Betas	t	Sig.
Random Effects			
Between-class	63.460	2.708†	0.007
Between-Students	49.895	21.374	0.000
Fixed Effects			
[Intervention=GG]	4.268	1.488	0.141
[Intervention=MATH]	0.000	.	.
Pre-test	1.468	17.346	0.000
Engagement	14.478	7.777	0.000
[Intervention=GG] * Engagement	12.549	4.562	0.000
[Intervention=Math] * Engagement	0.000	.	.
[Intervention=GG] * Engagement* Pre-test	-0.564	-6.023	0.000
[Intervention=Math] * Engagement * Pre-test	-0.355	-4.408	0.000

† *z*-value (Wald)

As can be seen in Table 4, the main result in text reading improvement is a significant interaction between the effects of intervention and engagement suggesting that GG intervention resulted in higher post-test reading scores only when the engagement of the training was average or high (+12.5 scores, ES = .22 see also Figure 5). The triple interaction with pre-test reading level suggests that the intervention gains of GG over Math were the highest for readers with initially weak pre-test scores (poor readers) and decreased for readers for readers with high pre-test scores (good readers).

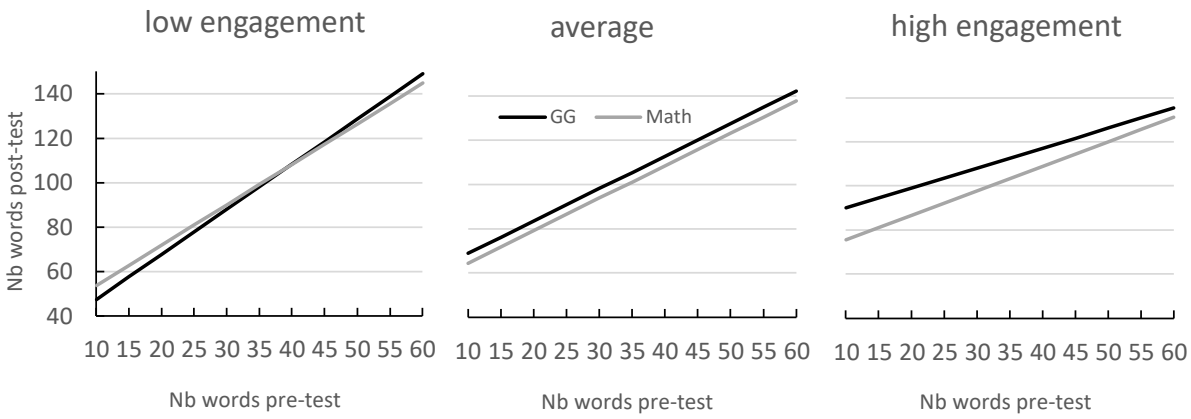


Figure 5. Model-fitted text-reading post-test scores (Alouette test) as a function of intervention, pre-test scores and the engagement of the training.

Single Word reading (1-minute reading test)

The final model included the fixed effects of intervention, pre-test level and engagement, the two-way interaction between intervention and engagement and the three way-interaction between intervention, engagement and pre-test level. The effect of intervention was not significant ($F(1, 93.5) = 2.85$, $p = .095$). However, there was a significant interaction between the effects of intervention and engagement ($F(1, 887.5) = 6.95$, $p < .01$) and a significant triple interaction between the effects of intervention, engagement and pre-test level ($F(2, 870.2) = 12.31$, $p < .0001$). The fixed effects of pre-reading level and engagement were also significant ($F(1, 899.8) = 290.6$, $p < .0001$ and $F(1, 886.4) = 149.6$, $p < .0001$, respectively). The parameter estimates are shown in Table 5.

Table 5. Parameter estimates of the final model predicting text-reading post-test scores on the function of intervention. pre-test level and engagement of the training

Parameters	Betas	t	Sig.
Random Effects			
Between-class	17.201	2.820†	0.005
Between-Students	23.624	17.746	0.000
Fixed Effects			
[Intervention=GG]	2.475	1.643	0.104
[Intervention=MATH]	0.000	.	.
Pre-test	1.475	17.008	0.000
Engagement	9.926	8.951	0.000
[Intervention=GG] * Engagement	4.481	2.784	0.005
[Intervention=Math] * Engagement	0.000	.	.
[Intervention=GG] * Engagement* Pre-test	-0.526	-6.050	0.000
[Intervention=Math] * Engagement * Pre-test	-0.471	-5.582	0.000

† *z*-value (Wald)

The main results of single word reading can be summarized as follows. Although the main effect of intervention failed to reach significance (ES = .11), there was a significant interaction between the effects of intervention and engagement, showing that high engagement training produced significant gains of GG over and above the Math intervention (β = +4.48 points; ES = .18). The triple interaction between the effects of intervention, pre-test level and engagement was also significant suggesting that high-engagement training for poor readers produced the largest gains of GG over Math intervention (see Figure 6).

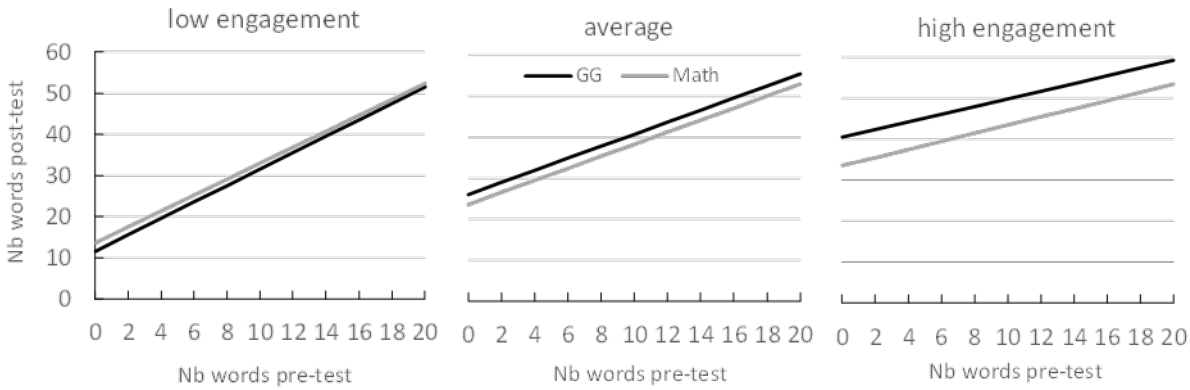


Figure 6. Model-fitted single-word reading post-test scores (1-minute reading test) as a function of intervention, pre-test level and the engagement of the training.

Math scores (TEDI evaluation battery)

Although the present study focused on evaluating the efficiency of GG, it is nevertheless interesting to investigate whether the Math intervention produced a positive training effect. We just followed the same analysis procedure with the math scores. The final model included the fixed effects

of intervention, pre-test level and engagement and the three way-interaction between intervention, engagement and pre-test level. The effect of intervention was significant ($F(1, 66.94) = 18.772$, $p < .001$), as well as the effects of pre-test level and engagement w ($F(1, 891.9) = 311.21$, $p < .0001$ and $F(1, 900.7) = 184.92$, $p < .0001$, respectively). There was also a significant significant triple interaction between the effects of intervention, engagement and pre-test level ($F(2, 883.7) = 47.62$, $p < .0001$). The parameter estimates are shown in Table 6.

Table 6. Parameter estimates of the final model predicting text-reading post-test scores on the function of intervention. pre-test level and engagement of the training

Parameters	Beta	t	Sig.
Random Effect			
Between-class	.563	3.265	<.001
Between-Students	10.577	37.848	<.001
Fixed Effects			
[Intervention=MATH]	1.012	4.333	<.001
[Intervention=GG]	.000		
Pre-test	.340	17.641	<.001
Engagement	2.813	13.599	<.001
[Intervention=MATH] * Pre-test * Engagement	-.168	-9.597	<.001
[Intervention=GG] * Pre-test * Engagement	-.145	-8.202	<.001

As can be seen in Table 6, the Math intervention had a strong positive effect over and above the GG intervention (+1.01 points, ES = .29). The triple interaction was significant suggesting that children with initially weaker math scores showed greater intervention benefits when engagement was high and that children with initially greater math scores showed larger benefits when engagement was low (ES = .21). The results are plotted in Figure 7.

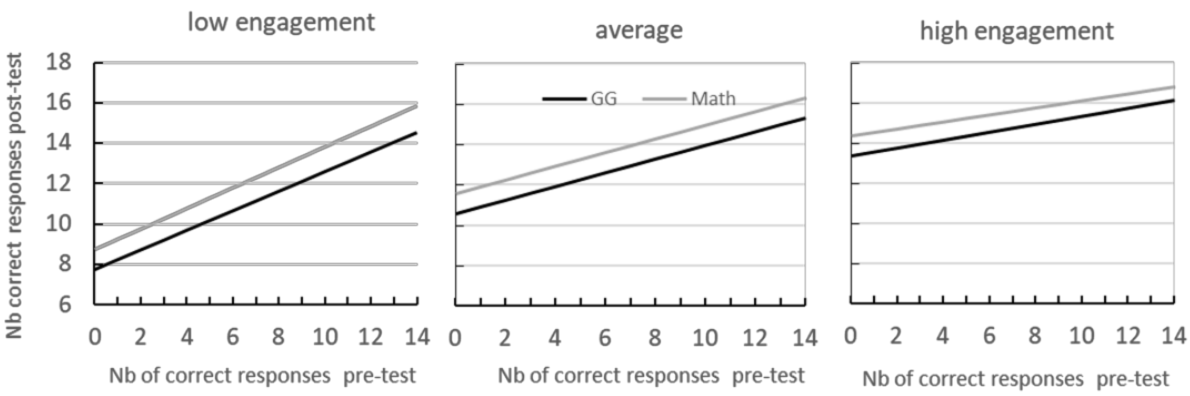


Figure 7. Model-fitted math post-test scores (Nb correct/20) as a function of intervention, pre-test level and the engagement of the training.

Discussion

The goal of the present study was four-fold. First, we wanted to test a French version of GG that was developed to cover the entire 1st grade reading program based on a theoretically optimal progression (Dehaene, 2011; Sprenger-Charolles, 2017). Second, we wanted to test the GG intervention in a large school sample of beginning readers (1 grade) from socioeconomically disadvantaged neighborhoods because the need for early within-school intervention is particularly high in this part of the population and this is particularly true in France (Fluss et al., 2008; Fluss et al., 2009). Third, we wanted to assess the GG intervention against an equally engaging math intervention while keeping teacher and school effects constant. Finally, we wanted to include some moderator variables, such as the initial pre-test level of the child and within-game engagement, to better understand why recent meta-analyses found weak or absent effects of GG intervention with the standard ANOVA-type analyses (McTigue et al., 2020).

We tested the effects of GG intervention, as compared to math intervention, on four classic standardized tests, one on phonemic awareness, one on visual word recognition skills, and two on reading fluency. The results showed positive effects of GG (as a main effect) over and above Math on phoneme suppression and word recognition (ES = .26 and .25, respectively). Importantly, for both tests, the effects of GG intervention were modulated by interindividual differences in phoneme awareness and word recognition skills at pre-test (ES=.20 and ES = .30, for phoneme awareness and word recognition, respectively), which reflect the fact that children with weak phoneme awareness or word recognition skills at pre-test benefitted more from GG intervention than children with stronger skills at pre-test. As concerns the two fluency measures (1 minute reading aloud and 3-minute text reading aloud), the main effect of GG intervention failed to reach significance but the gains in reading fluency as a consequence of intervention were significantly influenced by student's within-game enagagement, that is, fluency after GG intervention increased significantly than fluency after math intervention but only when the engagement was high (ES =.18 and .22, for single word and text reading fluency, respectively). For both fluency measures, the triple interaction between the effect sof intervention, pre-test-level and engagement was significant reflecting the fact that high engagement produced stronger GG intervention effect especially for the slow readers.

Although the assessment of the efficiency of the math intervention was not the main goal of the study, we nevertheless investigated whether the math intervention produced a significant increase in mathematical skills after the training over and-above the effects of the reading intervention. Indeed, on the mathematical ability post-test, we obtained significant math intervention effects and the effect size was comparable to that of GG intervention (i.e., ES = .29). Interestingly, the math intervention showed exactly the same triple interaction as the GG intervention, in which the intervention gains in mathematical abilities were stronger for children with initially weaker math skills especially when engagement was high.

Together, the strong interactions with the child's initial level and engagement in both the math and reading intervention highlight two important issues that are often neglected in classroom intervention studies. First, it is implicitly assumed that efficient interventions should benefit all children to a similar extent ("all size fits all"). However, our results clearly show that this is not the case. Instead, the interventions benefitted the most for children with initially weak skills. Second, it is also assumed that children are equally good and motivated in using the digital games. Again, our results clearly show that only those children with high within-game engagement show strong intervention benefits, especially for fluency. This is not only unsurprising because we would only expect intervention benefits for those

who properly played the game but it is also rather reassuring because it shows that the intervention games are directly related to the within-game behavior and thus is not an artefact of some other external variable that is not related to the game. More generally speaking, the effects of engagement are often discussed in terms of implementation fidelity or motivation and our results suggest that these variables need to be taken into account for a fair efficiency assessment of intervention games in classroom settings.

The present study has a number of limitations. The most obvious one is that we did not choose a strict randomized controlled trial, in which interventions were randomly attributed to classes or schools (xxx). We rather used a quasi-experimental design, in which the same teachers administered both interventions successively. This was done to reduce teacher effects and biases due to social comparisons between teachers who believed to be in the "good" treatment group (i.e., the GG intervention) versus the "bad" control group '(i.e., the math intervention). We have previously seen that teachers who believed to be in the control group (math intervention) increased the intensity of the reading instruction because they felt that their students were disadvantaged and that their teaching skills were being evaluated (Lassault, 2021). The second limitation is that we did not test the students one or two years later to see whether the intense GG training improved durably their reading performance. This longitudinal follow-up was initially planned and actually carried out for one of the groups but because of the corona-related lockdown of France in spring 2020, we could not obtain the follow-up measures for the other group. Thus, we could only compare the reading age of the GG-intervention group one year later to the norms on a standardized reading test that was administered at T1, T2 and T2+1year. While the reading age of the children in the GG group was 3 months below their chronological age at T1, their reading age was identical to their chronological age at T2 and this was also the case at the 1-year follow-up (see Lassault, 2021).

The present work opens up some interesting research avenues. The most important one is to use the large-scale data of several hundreds of students to render the training content and the progression much more adaptive using unsupervised learning algorithms that maximize the gains as a function of initial skills and engagement using within-game variables. Other perspectives include the development of games that train reading comprehension using short sentences and the use of automatic speech recognition software to train reading aloud, which are two dimensions that are currently not being trained.

Acknowledgments

This research has been supported by the French National Agency for Research (ANR) through the GraphoGame project (ANR-13-APPR-0003), the eFRAN Lemon project, the Institute of Convergence ILCB (ANR-16-CONV-0002) and the Excellence Initiative of Aix-Marseille University A*MIDEX (ANR-11-IDEX-0001-02). The authors thank all the participants and the teachers who accepted to use the digital applications in their classrooms, and the school authorities who supported this study. We thank Iivo and Michael and the GraphoGame Group Ldt for their collaboration in the development and diffusion of the project.

References

Brem, S., Bach, S., Kucian, K., Guttorm, T. K., Martin, E., Lyytinen, H., . . . Richardson, U. (2010). Brain sensitivity to print emerges when children learn letter-speech sound correspondences. *Proceedings of the National Academy of Sciences, 107*(17), 7939-7944. doi:10.1073/pnas.0904402107

Carvalhais, L., Richardson, U., & Castro, S. L. (2018). Computer-Assisted Reading and Spelling Intervention with Graphogame Fluent Portuguese. In Á. Rocha, H. Adeli, L. Reis, & S. Costanzo (Eds.), *Trends and Advances in Information Systems and Technologies. WorldCIST'18 2018. Advances in Intelligent Systems and Computing* (Vol. 746). Cham: Sptringer.

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. doi:10.1177/1529100618772271

Dehaene, S. (2011). *Apprendre à lire: Des sciences cognitives à la salle de classe*. Paris: Odile Jacob.

Deltour, J. J., & Hupkens, D. (1984). *Test de Vocabulaire actif et passif pour enfants (5 à 8 ans) (TVAP)*. Braine-le-Château: Editions de l'application des Techniques Modernes.

Dotan, D., Dehaene, S., Huron, C., & Piazza, M. (2011). The Number Catcher. Retrieved from http://www.thenumbercatcher.com/nc/home.php

Dynarski, M., Agodini, R., Heaviside, S., Novak, T., Carey, N., Campuzano, L., . . . Sussex, W. (2007). *Effectiveness of Reading and Mathematics Software Products: Findings from the First Student Cohort*. Research report - Report number NCEE 2007-4005. Retrieved from https://telearn.archives-ouvertes.fr/hal-00190019

Ecalle, J. (2003). *Timé-2 : test d'identification de mots écrits pour enfants de 6 à 8 ans*. PAris: Edition du Centre de Psychologie Appliquée.

Ehri, L. C. (1992). Reconceptualizing the development of sight word reading and its relationship to recoding. In P. B. Gough, L. E. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 105-143). Hillsdale, NJ: LEA.

Fluss, J., Ziegler, J., Ecalle, J., Magnan, A., Warszawski, J., Ducot, B., . . . Billard, C. (2008). Prevalence of reading disabilities in early elementary school: Impact of socioeconomic environment on reading development in 3 different educational zones. *Archives de Pédiatrie, 15*(6), 1049-1057. doi:10.1016/j.arcped.2008.02.012

Fluss, J., Ziegler, J. C., Warszawski, J., Ducot, B., Richard, G., & Billard, C. (2009). Poor Reading in French Elementary School: The Interplay of Cognitive, Behavioral, and Socioeconomic Factors. *Journal of Developmental and Behavioral Pediatrics, 30*(3), 206-216.

Foorman, B. R., Breier, J. I., & Fletcher, J. M. (2003). Interventions aimed at improving reading success: An evidence-based approach. *Developmental Neuropsychology, 24*(2-3), 613-639.

Foorman, B. R., & Torgesen, J. (2001). Critical Elements of Classroom and Small-Group Instruction Promote Reading Success in All Children. *16*(4), 203-212. doi:10.1111/0938-8982.00020

Gentaz, E., Sprenger-Charolles, L., Theurel, A., & Colé, P. (2013). Reading comprehension in a large cohort of French first graders from low socio-economic status families: a 7-month longitudinal study. *PLoS ONE, 8*(11), e78608. doi:10.1371/journal.pone.0078608

Goswami, U. (2002). Phonology, reading development and dyslexia: A cross-linguistic perspective. *Annals of Dyslexia, 52*, 141-163.

Huemer, S., Landerl, K., Aro, M., & Lyytinen, H. (2008). Training reading fluency among poor readers of German: many ways to the goal. *Annals of Dyslexia, 58*(2), 115-137. doi:10.1007/s11881-008-0017-2

Hulme, C., Snowling, M. J., West, G., Lervåg, A., & Melby-Lervåg, M. (2020). Children's Language Skills Can Be Improved: Lessons From Psychological Science for Educational Policy. *Current Directions in Psychological Science, 29*(4), 372-377. doi:10.1177/0963721420923684

Kyle, F., Kujala, J., Richardson, U., Lyytinen, H., & Goswami, U. (2013). Assessing the Effectiveness of Two Theoretically Motivated Computer-Assisted Reading Interventions in the United Kingdom: GG Rime and GG Phoneme. *Reading Research Quarterly, 49*, 59–74

Landerl, K., Ramus, F., Moll, K., Lyytinen, H., Leppanen, P. H., Lohvansuu, K., . . . Schulte-Korne, G. (2013). Predictors of developmental dyslexia in European orthographies with varying complexity. *J Child Psychol Psychiatry, 54*(6), 686-694. doi:10.1111/jcpp.12029

Lassault, J., & Ziegler, J. C. (2018). Les outils numériques d'aide à l'apprentissage de la lecture. *Langue Francaise, 119*.

Lecocq, P. (1996). *L'E.C.O.S.S.E*. Villeneuve d'Ascq, France: Presses Universitaires du Septentrion.

Lefavrais, J. (2005). *Test de l'Alouette (revised version)*. Paris: ECPA.

Lété, B., Sprenger-Charolles, L., & Colé, P. (2004). MANULEX: a grade-level lexical database from French elementary school readers. *Behavior Research Methods Instruments & Computers, 36*(1), 156-166. doi:10.3758/BF03195560

Lyytinen, H., Erskine, J., Hamalainen, J., Torppa, M., & Ronimus, M. (2015). Dyslexia-Early Identification and Prevention: Highlights from the Jyvaskyla Longitudinal Study of Dyslexia. *Curr Dev Disord Rep, 2*(4), 330-338. doi:10.1007/s40474-015-0067-1

Lyytinen, H., Erskine, J., Kujala, J., Ojanen, E., & Richardson, U. (2009). In search of a science-based application: A learning tool for reading acquisition. *Scandinavian Journal of Psychology, 50*(6), 668-675. doi:10.1111/j.1467-9450.2009.00791.x

McTigue, E. M., Solheim, O. J., Zimmer, W. K., & Uppstad, P. H. (2020). Critically Reviewing GraphoGame Across the World: Recommendations and Cautions for Research and Implementation of Computer-Assisted Instruction for Word-Reading Acquisition. *55*(1), 45-73. doi:10.1002/rrq.256

OECD. (2019). *PISA 2018 Results (Volume I)*.

Peereman, R., & Sprenger-Charolles, L. (2018). Manulex-MorphO, une base de données sur l'orthographe du français intégrant les morpho-phonogrammes. *199*(3), 99-109. doi:10.3917/lf.199.0099

Raven, J. C. (1976). *Standard progressive matrices: Sets A, B, C, D & E*. Oxford: Psychologists Press.

Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2001). How psychological science informs the teaching of reading. *Psychol Sci, 2*(2 Suppl), 31-74.

Richardson, U., & Lyytinen, H. (2014). The Graphogame method: the theoretical and methodological background of the technology-enhanced learning environment for learning to read. *Human Technology, 10*, 39-60. doi:doi: 10.17011/ht/urn.201405281859

Rosas, R., Escobar, J.-P., Ramírez, M.-P., Meneses, A., & Guajardo, A. (2017). Impact of a computer-based intervention in Chilean children at risk of manifesting reading difficulties / Impacto de una intervención basada en ordenador en niños chilenos con riesgo de manifestar dificultades lectoras. *Journal for the Study of Education and Development, 40*(1), 158-188. doi:10.1080/02103702.2016.1263451

Saine, N. L., Lerkkanen, M.-K., Ahonen, T., Tolvanen, A., & Lyytinen, H. (2010). Predicting word-level reading fluency outcomes in three contrastive groups: Remedial and computer-assisted remedial reading intervention, and mainstream instruction. *Learning and Individual Differences, 20*(5), 402-414. doi:https://doi.org/10.1016/j.lindif.2010.06.004

Saine, N. L., Lerkkanen, M.-K., Ahonen, T., Tolvanen, A., & Lyytinen, H. (2011). Computer-Assisted Remedial Reading Intervention for School Beginners at Risk for Reading Disability. *82*(3), 1013-1028. doi:10.1111/j.1467-8624.2011.01580.x

Sanders, W. L., Wright, S. P., & Horn, S. P. (1997). Teacher and Classroom Context Effects on Student Achievement: Implications for Teacher Evaluation. *Journal of Personnel Evaluation in Education, 11*(1), 57-67. doi:10.1023/A:1007999204543

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55*(2), 151-218. doi:10.1016/0010-0277(94)00645-2

Slavin, R. E., Lake, C., Davis, S., & Madden, N. A. (2011). Effective programs for struggling readers: A best-evidence synthesis. *Educational Research Review, 6*(1), 1-26. doi:https://doi.org/10.1016/j.edurev.2010.07.002

Sprenger-Charolles, L. (2017). Une progression pédagogique construite à partir de statistiques sur l'orthographe du français (d'après Manulex-Morpho): pour les lecteurs débutants et atypiques. *Approche Neuropsychologique des Apprentissages chez l'Enfant (ANAE), 148*, 247-256.

Sprenger-Charolles, L., Colé, P., Bechennec, D., & Kipffer-Piquard, A. (2005). French normative data on reading and related skills from EVALEC, a new computerized battery of tests. *Revue Européenne de Psychologie Appliquée, 55*, 157-186.

Wilson, A. J., Dehaene, S., Pinel, P., Revkin, S. K., Cohen, L., & Cohen, D. (2006). Principles underlying the design of "The Number Race", an adaptive computer game for remediation of dyscalculia. *Behavioral and Brain Functions, 2*(1), 19. doi:10.1186/1744-9081-2-19

Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology, 105*(2), 249-265. doi:10.1037/a0031311

Ziegler, J. C., Bertrand, D., Tóth, D., Csépe, V., Reis, A., Faísca, L., . . . Blomert, L. (2010). Orthographic depth and its impact on universal predictors of reading: A cross-language investigation. *Psychological Science, 21*(4), 551–559. doi:10.1177/0956797610363406

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychological Bulletin, 131*(1), 3-29. doi:10.1037/0033-2909.131.1.3

Ziegler, J. C., Perry, C., & Zorzi, M. (2014). Modelling reading development through phonological decoding and self-teaching: implications for dyslexia. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1634). doi:10.1098/rstb.2012.0397

Ziegler, J. C., Perry, C., & Zorzi, M. (2020). Learning to Read and Dyslexia: From Theory to Intervention Through Personalized Computational Models. *Current Directions in Psychological Science, 29*(3), 293-300. doi:10.1177/0963721420915873

Appendix

Sequences	Content
1.1	Vowel (a-i-o-u-e-eu-ou-ou-e) Consonant (j-f-l-r)
1.2	Repetition of sequence 1
1.3	Addition of 's' at the beginning
1.4	Repetition of sequence 1.3
1.5	Words Vowel-Consonant and Consonant-Vowel-Consonant
1.6	Silent letter and frequent words
2.1	Consonants (p-m-d-v)
2.2	Consonants (b-n-t-ch)
2.3	Repetition of sequences 2.1 and 2.2
2.4	Words Consonant-Consonant-Vowel
2.5	Repetition of sequence 2.4
2.6	Assessment of sequence 2
3.1	Visual discriminations (u-n,...)
3.2	Phonemic discrimination (t-d,...)
3.3	Mixed discrimination (p-b,...)
4.1	Nasal vowels (an-on)
4.2	Nasal vowels (in, un)
5.1	Summary of Module 1 - (Part 1)
5.2	Summary of Module 1 - (Part 2)
6.1	Vowel (oi) and nasal vowel (oin)
6.2	Vowel (ui)
6.3	Vowel (ai)
6.4	Vowel (ei)
6.5	Vowels (au-eau)
6.6	Silent consonant (h)
6.7	Vowels (e-è-ê)
7.1	Visual discriminations (u-n,...)
7.2	Phonemic discrimination (t-d,...)
7.3	Mixed discrimination (p-b,...)

8	Consonants (qu-k)
9.1	Contextual consonant (c - sound :/k/)
9.2	Contextual consonant (c - sound :/s/)
10.1	Consonant (z)
10.2	Contextual consonant (s - sound: /z/)
10.3	Contextual consonant (s, ss - sound : /s/)
11	Vowels (es-ez-er-and)
12.1	Contextual consonant (g - sound : /g/)
12.2	Contextual consonant (g - sound : /Z/)
13	Consonant (gn)
14.1	Nasal vowels (ant-and-ent)
14.2	Nasal vowels (om-im-am-am-em)
14.3	Nasal vowels (ain-ein-aim)
15.1	Vowel (e) in front of Double Consonants (ell-ett-err-eff-ess)
15.2	Vowel (e) in the middle of a word (ec-es-er)
16	Consonant (ph)
17.1	Vowel (i) in front of a Vowel (sound: /j/)
17.2	Vowel (i) in front of one or two l (sound : /j/)
18.1	Vowel (y) between two Consonants (sound: /i/)
18.2	Vowel (y) in front of a Vowel
19	Exceptions (ill - sound /il/)
20.1	Silent letter supporting the derivation
20.2	Final consonant(s), plural mark
21	Consonant (x)
22.1	Consonants (ç, sc)
22.2	Consonant (t - sound : /s/)
23.1	Exceptions: consonant (t) pronounced at the end of a word
23.2	Exceptions: consonants (s and c) pronounced at the end of a word
24	Consonant (ch - sound /k/)
25	Vowel (oeu)
26	Exceptions

27.1	Present verbs from the first group
27.2	Presents verbs from the first group (sentences)
27.3	Present tense of verb "to be"
27.4	Personal pronoun subject and verb "to be"
27.5	Present of the verb "to have"
27.6	Personal pronoun subject and verb "to have"
28	Assessment:gender, number and verb