Using verbal fluency tasks to investigate the lexicon in Greek-speaking children with literacy and language disorders

Maria Mengisidou

Thesis submitted for the degree of Doctor of Philosophy

Department of Psychology and Human Development

Institute of Education

University College London

London, May 2019

I, Maria Mengisidou, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Word count (exclusive of list of references and the appendix): 72, 495 words

Signature

[redacted]

Abstract

In this thesis, semantic and phonological fluency tasks were used to investigate the lexicon in sixty-six children with dyslexia and/or DLD (hereafter DDLD group) aged 7-12 years and in 83 typically-developing (TD) children aged 6-12 years, all monolingual Greek speakers. In semantic and phonological fluency tasks, responses are often produced in clusters of semantically- or phonologically-related items, respectively (e.g. “cat-dog” is a semantic cluster; “*fl*ag-*fl*ower” is a phonological cluster). Once the retrieval of items within a cluster slows down, children tend to switch to another cluster.

In both groups, productivity in semantic and phonological fluency tasks correlated strongly with the number of clusters and the number of switches, but not with average cluster size. Regression analyses showed that the DDLD group retrieved significantly fewer correct items in semantic and phonological fluency tasks compared to the TD group, but average semantic and phonological cluster size did not differ significantly in the two groups. Furthermore, the two groups did not differ significantly on the number of correct designs generated in the design fluency task.

Poorer semantic fluency performance in children with DDLD is attributed to slower retrieval processes while children’s semantic structure is intact, as proposed by the Slow-Retrieval Model. Consistent with the Deficient Phonological Access Hypothesis, children with DDLD showed impaired explicit access but intact implicit access to phonological representations. For both verbal fluency categories, slower retrieval processes originating from deficient access to intact semantic and phonological representations, and also inferior language and literacy skills, explain poorer verbal fluency performance in children with dyslexia and/or DLD. The specificity of DDLD children’s verbal fluency deficit is supported by evidence showing that children with DDLD showed poorer semantic and phonological fluency performance relative to their TD peers even after design fluency performance was controlled. The underlying causes of slow lexical retrieval still need further investigation.

Impact statement

Children with dyslexia and/or Development Language Disorder (DLD) have been reported to have lexical difficulties. Two models attempting to account for lexical difficulties, the Structure-Loss Model and the Retrieval-Slowing Model, have been proposed on the basis of adult data and have been tested only in adults. The current study is the first developmental study designed to tease apart the Poor Lexical-Semantic Structure Model (adapted from the Structure-Loss Model) and the Slow-Retrieval Model (adapted from the Retrieval-Slowing Model). Semantic fluency tasks (e.g. "Name as many animals as you can in one minute") were used to investigate the organization of the lexicon in a group of Greek children with dyslexia and/or DLD relative to a group of typically developing (TD) children. The findings support the Slow-Retrieval Model in that children with dyslexia and/or DLD have intact lexical-semantic representations but access to these representations is impaired, resulting in slower retrieval processes of lexical items from the mental lexicon.

Further, this study adds to the theoretical debate on the locus of the phonological deficit in dyslexia and DLD. Phonological fluency tasks (e.g. "Name as many words as you can beginning with the letter 'f' in one minute") are lexical tasks which can be ideally used for the purpose of this study because they do not involve metalinguistic awareness skills and reading or spelling ability. The outcomes do not support the view that the phonological deficit in dyslexia and DLD lies in degraded phonological representations. The outcomes do, however, support the view that the phonological deficit in dyslexia and DLD lies in deficient explicit access to (intact) phonological representations.

From a scientific point of view, insight into the models and hypotheses accounting for poorer verbal fluency performance may inform theory and theory can inform treatment and training of children with dyslexia and DLD in clinical and educational settings. As a first step towards this direction, intervention studies designed to improve children's retrieval processes are needed to investigate any potential gains on productivity in semantic and phonological fluency tasks.

Table of Contents

Abstract 3

Impact statement 4

List of tables 10

List of figures 12

Acknowledgments 14

Abbreviations 17

Structure of the thesis 18

Chapter 1a. General background to the thesis 21

1a.1. General overview of the thesis aims and key questions 21

1a.2. Definitions of Developmental Dyslexia and Developmental Language Disorder (DLD) 28

1a.3. Language difficulties in children with dyslexia and DLD 30

1a.4. Comorbidity between dyslexia and DLD 33

Chapter 1b. Hypothesis-driven research questions 38

1b.1. The Structure-Loss Model and the Retrieval-Slowing Model 38

1b.2. The Poor Lexical-Semantic Structure Model and the Slow-Retrieval Model 42

1b.3. The role of semantic skills in reading ability 46

1b.4. The role of phonological processing skills in reading ability 48

1b.4.a. Phonological awareness. 48

1b.4.b. Phonological short-term memory (STM). 49

1b.4.c. Rapid automatic naming (RAN). 50

1a.5. Spelling accuracy performance in children with dyslexia and DLD 52

1b.6. The Phonological Theory and the phonological deficit 56

1b.7. The Degraded Phonological Representations Hypothesis 59

1b.8. The Deficient Phonological Access Hypothesis 62

Chapter 2. Fluency tasks 68

2.1. Semantic fluency 68

2.1.1. Semantic switching, clustering and cluster size 68

2.2. Phonological fluency 70

2.2.1. Phonological switching, clustering and cluster size 70

2.3. Semantic and phonological fluency tasks: Same or different? 72

2.4. Effects of switching and clustering behaviour on semantic and phonological fluency performance 75

2.5. Effects of age, level of intelligence, and gender on semantic and phonological fluency performance 76

2.5.1. Evidence from TD children and healthy adults 76

2.5.1.a. The effect of age. 76

2.5.1.b. The effect of the level of intelligence. 76

2.5.1.c. The effect of gender. 77

2.6. Automatic versus controlled processing in verbal fluency tasks 77
2.7. Semantic and phonological fluency performance in clinical groups 78
2.7.1. Evidence from children with dyslexia 78
2.7.1.a. Poorer semantic and phonological fluency performance. 78
2.7.1.b. Poorer phonological fluency but similar semantic fluency performance. 79
2.7.2. Evidence from children with DLD 81
2.7.3. Evidence from children with word-finding difficulties (WFDs) 82
2.7.4. Evidence from adults with dyslexia and/or DLD 82
2.7.5. Summary: Semantic and phonological fluency performance in clinical groups 83
2.8. Switching, clustering and cluster size in clinical groups 84
2.8.1. Evidence from children with dyslexia 84
2.8.2. Evidence from children with DLD 84
2.8.3. Evidence from adults with dyslexia and/or DLD 85
2.9. Correlation, network and computational modelling methodologies in semantic fluency tasks in clinical groups 86
2.9.1. Evidence from children with Cochlear Implants (CIs) 87
2.9.2. Evidence from adults with Mild Cognitive Impairment (MCI) and Alzheimer's Disease (AD) 87
2.9.3. Summary: Switching, clustering and cluster size in clinical groups 88
2.10. Automatic versus controlled processing in clinical groups 89
2.11. Effects of language, literacy and executive function (EF) skills on semantic and phonological fluency performance 90
2.11.1. The effect of language and literacy skills 90
2.11.2. The effect of EFs 92
2.12. Verbal fluency and executive functioning (EF) 93
2.13. Design fluency 93
2.13.1. Summary: Design fluency performance in clinical groups 95
2.14. Research questions 96
2.15. Predictions 97

Chapter 3. Methods 100
3.1. Ethical approval 100
3.2. Participants 100
3.2.1. Children with Dyslexia and/or Developmental Language Disorder (DDLD) 100
3.2.1.a. Initial selection criteria for the DDLD group. 100
3.2.1.b. Fine-grained selection criteria for the DDLD group. 106
3.2.2. Typically Developing (TD) Children 107
3.2.2.a. Initial selection criteria for the TD group. 108
3.2.2.b. Fine-grained selection criteria for the TD group. 108
3.3. Groups' general descriptives 108
3.4. Descriptive measures 109
3.4.1. Nonverbal IQ (NVIQ) task 111
3.4.1.a. Raven's Coloured Progressive Matrices (CPM). 111
3.4.2. Language ability tasks 111

3.4.2.a. WISC Similarities and Vocabulary subtasks. 111

3.4.2.b. Peabody Picture Vocabulary Test-Revised (PPVT-R). 112

3.4.2.c. Diagnostic Verbal Intelligence (DVIQ) Test's syntax comprehension and sentence repetition subtasks. 112

3.4.3. Literacy ability tasks 113

3.4.3.a. L'Alouette task. 113

3.4.3.b. Reading Test Alpha's reading accuracy and text-reading fluency subtasks (for children in Grade 3 onwards). 115

3.4.3.c. Test of Detection and Investigation of Reading Difficulties' syllable and nonword reading subtasks (for children in Grades 1 and 2). 116

3.5. Groups' descriptives of language, literacy and phonological tasks 116

3.5.1. Statistical methods. 116

3.5.2. Language tasks. 117

3.5.3. Literacy tasks. 120

3.5.4. Phonological tasks. 124

3.5.5. Summary of descriptives of language, literacy and phonological tasks 128

3.5.5.a. Group differences on language, literacy and phonological tasks. 128

3.5.5.b. Task limitations. 129

3.6. Experimental measures 131

3.6.1. Verbal fluency tasks 131

3.6.1.a. Semantic fluency tasks. 131

3.6.1.b. Phonological fluency tasks. 132

3.6.1.c. Coding of verbal fluency responses 133

3.6.1.d. Semantic clusters. 135

3.6.1.e. Computational modelling. 136

3.6.1.f. Phonological clusters. 138

3.6.2. The design fluency task 140

3.6.3. Phonological ability tasks 142

3.6.3.a. Phoneme deletion tasks. 142

3.6.3.b. Nonword repetition (NWR) task. 143

3.6.3.c. Rapid automatic naming (RAN) task. 143

3.6.4. The spelling-to-dictation task 144

3.6.4.a. Error classification 146

3.6.4.b. Phonological errors. 147

3.6.4.c. Grammatical errors. 149

3.6.4.d. Orthographic errors. 150

3.7. Assessment procedure 154

3.8. Calculation of statistical power 154

3.9. Threshold of statistical significance 156

Chapter 4. Results 158

4.1. Groups' Performance and Group Differences on Semantic Fluency Tasks 158

4.1.1. Statistical methods. 159

4.1.2. Regression analyses on the semantic fluency variables 159

4.1.3. Computational analysis for the semantic category of animals 162

4.1.4. Associations between the number of correct responses with the number of switches, the number of clusters and average cluster size in the TD and DDLD groups 162

4.1.5. Proportional scores of types of incorrect responses in semantic fluency categories 163

4.1.6. Error ratio in semantic fluency categories 163

4.1.7. Summary 164

4.2. Groups' Performance and Group Differences on Phonological Fluency Tasks 165

4.2.1. Statistical methods. 166

4.2.2. Regression analyses on the phonological fluency variables 166

4.2.3. Associations between the number of correct responses with the number of switches, the number of clusters and average cluster size in the TD and DDLD groups 168

4.2.4. Proportional scores of types of incorrect responses in phonological fluency categories 168

4.2.5. Error ratio in phonological fluency categories 169

4.2.6. Summary 169

4.3. Groups' Performance and Group Differences on Phonological Tasks 170

4.3.1. Statistical methods. 170

4.3.2. Phoneme deletion tasks 171

4.3.3. Nonword repetition (NWR) task 173

4.3.4. Rapid automatic naming (RAN) task 174

4.3.5. Summary 176

4.4. Groups' Performance and Group Differences on Types of Spelling Errors 177

4.4.1. Statistical methods. 177

4.4.2. Associations between the number of total words spelled with phonological, grammatical and orthographic errors 178

4.4.3. Group comparisons of phonological, grammatical and orthographic errors 178

4.4.6. Summary 184

4.5. Relation of language and literacy skills with semantic and phonological fluency 184

4.5.1. Statistical methods. 186

4.5.2. Associations between language and literacy measures with semantic and phonological fluency in the overall sample 187

4.5.5. Explanatory factor analysis 189

4.5.6. Regression analyses in the overall sample 192

4.5.7. Regression analyses by subgroup 193

4.5.8. Relation of language and literacy skills with automatic and controlled processing in semantic and phonological fluency categories 195

4.5.9. Summary 196

4.6. Groups' Performance and Group Differences on the Design Fluency Task 197

4.6.1. Statistical methods. 197

4.6.2. Regression analyses on the design fluency variables 197

4.6.3. Summary 201

4.7. Results meeting the old and the new threshold of statistical significance 201

4.7.1. Results meeting the new threshold of statistical significance 202

4.7.2. Results meeting the old threshold of statistical significance 202
4.8. Summary of the Results 203

Chapter 5. Discussion 213

5.1. Objectives of the study 213
5.2. Research question 1. What is the structure of the lexicon in children with DDLD compared to TD children? Is poorer semantic fluency performance in children with DDLD better explained by impoverished semantic structure or slower retrieval processes? 215
5.3.a. Research question 2a. Where does the phonological deficit in children with DDLD lie? Is poorer phonological fluency performance in children with DDLD better explained by degraded phonological representations or by deficient explicit access to (intact) phonological representations? 220
5.3.b. Research question 2b. Which hypothesis better characterises the locus of the phonological deficit in children with DDLD in phonological tasks—namely, phoneme deletion and RAN tasks, and a spelling-to-dictation task: The Degraded Phonological Representations Hypothesis or the Deficient Phonological Access Hypothesis? 225
5.4. Research question 4. Does semantic and phonological fluency performance relate to children's language and literacy skills? 230
5.5. Research question 5. How specific is the verbal fluency deficit in children with DDLD: Does it extend to a nonverbal task (design fluency)? 232
5.6. Theoretical implications 234
5.7. Strengths, limitations and further directions 236
5.8. Conclusions 241

References 244
Footnotes 259
Appendix 260

List of tables

Table 2.1. Findings of semantic and phonological fluency performance between different clinical groups and appropriate controls

Table 2.2. Findings of semantic and/or phonological switching, clustering and cluster size between different clinical groups and appropriate controls

Table 2.3. Findings of design fluency performance between different clinical groups and controls

Table 3.1. The loadings onto components 1 and 2 for each task generated by the Principal Component Analysis (PCA) with oblique rotation in the DDLD group

Table 3.2. General descriptives of the number of children, the gender distribution, mean age, age range and Raven's CPM performance in the DDLD group and the TD group

Table 3.3. Means (*SD*s) and 95% *CI*s (lower line) on language tasks for the DDLD and the TD group

Table 3.4. Percentage of variance (R^2) on language tasks explained by age in months and NVIQ entered in Step 1 and proportion of variance (ΔR^2) explained by group entered in Step 2 in the overall sample

Table 3.5. Means (*SD*s) and 95% *CI*s (lower line) on literacy measures for the DDLD and the TD group

Table 3.6. Percentage of variance (R^2) on literacy tasks explained by age in months and NVIQ entered in Step 1 and proportion of variance (ΔR^2) explained by group entered in Step 2 in the overall sample

Table 3.7. Means (*SD*s) and 95% *CI*s (lower line) on phonological measures for the DDLD and the TD group

Table 3.8. Percentage of variance (R^2) on phonological measures explained by age in months and NVIQ entered in Step 1 and proportion of variance (ΔR^2) explained by group entered in Step 2 in the overall sample

Table 3.9. Proportion of variance accounted for by group for language, literacy and phonological measures and *p*-values of significance after controlling for age in months and NVIQ

Table 3.10. An example for animal fluency data preparation of the first participating child

Table 3.11. Examples of minor subcategories of phonological errors using the word 'άλογο' <horse>

Table 3.12. Examples of phonological errors where the phonological errors are shown in bold and omitted grapheme is shown with _ in each word

Table 3.13. Examples of minor subcategories of grammatical errors using the words 'άλογο' <horse>, 'αυτός' <he>, 'είναι' <is> and 'ξεφυλλίζοντας' <flipping through>

Table 3.14. Examples of minor subcategories of orthographic errors using the words 'ποτίζω' <water>, 'δίχτυ' <net>, δανείζω <loan>, αυτός <he>, 'χείμαρρος' <torrent> and φωτισμένος <lit>

Table 3.15. The total number of error opportunities for the three major categories of errors in the first 30 words, in the subsequent 30 words and in total 60 words of the spelling-to-dictation task

Table 3.16. A priori power analyses computing required overall sample sizes

Table 4.1.1. Linear regression analyses conducted on the semantic fluency variables

Table 4.1.2. Types of incorrect responses in proportional scores in the semantic condition in the TD and DDLD groups

Table 4.2.1. Linear regression analyses conducted on the phonological fluency variables

Table 4.2.2. Types of incorrect responses in proportional scores in the phonological condition in the TD and DDLD groups
Table 4.3.1. Means (*SD*s) and *CI*s (lower line) of RAN measures in the DDLD group and the TD group
Table 4.3.2. Means (*SD*s), *CI*s and group comparisons of the number of correctly repeated nonwords in terms of the number of syllables in the DDLD group and the TD group
Table 4.3.3. Percentage of variance (R^2) on RAN measures explained by age in months and NVIQ entered in Step 1 and proportion of variance (ΔR^2) explained by group entered in Step 2 in the overall sample
Table 4.4.1. Mean (*SD*s) proportions of phonological, grammatical and orthographic errors and 95% *CI*s (lower line) in the DDLD group and the TD group
Table 4.5.1. Partial correlations (controlling for age) among all language measures in the overall sample
Table 4.5.2. Partial correlations (controlling for age) among literacy measures in the overall sample
Table 4.5.3. Partial correlations (controlling for age) between semantic fluency (number of correct responses) and language and literacy tasks in the overall sample, in the DDLD group and in the TD group
Table 4.5.4. Partial correlations (controlling for age) between phonological fluency (number of correct responses) and language and literacy tasks in the overall sample, in the DDLD group and in the TD group
Table 4.5.5. Percentage of variance in semantic and phonological fluency explained by age in months and NVIQ entered in Step 1 and proportion of variance explained by the LangLit variable entered in Step 2 in the overall sample
Table 4.5.6. Percentage of variance in semantic and phonological fluency explained by age in months and NVIQ entered in Step 1 and proportion of variance explained by the LangLit variable entered in Step 2 in the DDLD group and in the TD group
Table 4.5.7. Percentage of variance in the first 15 s and the subsequent 45 s in semantic and phonological categories explained by age in months and NVIQ entered in Step 1 and proportion of variance explained by the LangLit variable entered in Step 2 in the overall sample
Table 4.6.1. Linear regression analyses conducted on the design fluency variables
Tables 4.8.1. and 4.8.2. Partial correlations (controlling for age) between word productivity with the number of switches, the number of clusters and average cluster size in both verbal fluency categories in the TD and DDLD groups

List of figures

Figure 1a.1. Bishop and Snowling's (2004) two-dimensional model for the relationship between dyslexia and poor comprehension
Figure 1a.2. Phonological processing skills and phonological representations in control children and in children with dyslexia, DLD (referred to as SLI in the Figure), or DLD *plus* dyslexia
Figure 1b.1. A schematic diagram showing the typical adult state on the left, a storage deficit shown in the middle and an access deficit shown on the right
Figure 1b.2. A hypothesised network of semantic connections for TD children
Figure 1b.3. A hypothesised network of semantic connections in children with DDLD according to the Slow-Retrieval Model
Figure 1b.4. A hypothesised network of semantic connections in children with DDLD according to the Poor Lexical-Semantic Structure Model
Figure 1b.5. A causal model of dyslexia as a disorder originating from a phonological deficit
Figure 3.1. Number of words read correctly and words read incorrectly in raw scores, in addition to the time spent for text reading in the TD group, plotted against age in months
Figure 3.2. Mean performance on language tasks in the DDLD group in *z* scores
Figures 3.3. and 3.4. Scatterplots showing performance in sentence repetition and WISC Vocabulary in raw scores in the TD and DDLD groups, plotted against age in months
Figure 3.5. Mean performance on literacy tasks in the DDLD group in *z* scores
Figures 3.6. and 3.7. Scatterplots showing performance in spelling and l'Alouette in raw scores in the TD and DDLD groups, plotted against age in months
Figure 3.8. Mean performance on phonological measures in the DDLD group in *z* scores
Figures 3.9. and 3.10. Scatterplots showing accuracy performance in NWR and RAN in raw scores in the TD and DDLD groups, plotted against age in months
Figures 3.11. and 3.12. Scatterplots showing performance in phoneme deletion of CVC items and syllable reading in raw scores in the TD and DDLD groups, plotted against age in months
Figure 4.1.1. Scatterplot showing the number of correct responses in semantic categories in the DDLD group and the TD group, plotted against age in months
Figure 4.2.1. Scatterplot showing the number of correct responses in phonological categories in the DDLD group and the TD group, plotted against age in months
Figure 4.3.1. Accuracy performance in *z* scores in phoneme deletion tasks in the DDLD group
Figure 4.3.2. Time performance in *z* scores in phoneme deletion tasks in the DDLD group
Figures 4.4.1., 4.4.2. and 4.4.3. The distribution of phonological, grammatical and orthographic errors in the TD group and in the DDLD group
Figure 4.5.1. Scree plot showing the number of components generated by the Principal Component Analysis (PCA) with oblique rotation in the overall sample and each component's eigenvalue
Figure 4.5.2. The LangLit variable in *z* scores, plotted against age in months, in the DDLD group and the TD group
Figure 4.6.1. Scatterplot showing the number of correct designs produced in the DDLD group and the TD group, plotted against age in months

Figure 4.8.1. Raw scores (bars represent *SD*s) of the number of correct responses (semantic fluency), the number of switches, the number of clusters and average cluster size in semantic fluency categories in the two groups
Figure 4.8.2. Raw scores (bars represent *SD*s) of the number of correct responses (phonological fluency), the number of switches, the number of clusters and average cluster size in phonological fluency categories in the two groups
Figure 4.8.3. Raw scores of accuracy (bars represent *SD*s) for phoneme deletion tasks of CVCVCV items, CVC items and CCV items
Figure 4.8.4. Raw scores of accuracy (bars represent *SD*s) for 3-, 4-, 5- and 6-syllable items of the NWR task in the two groups
Figure 4.8.5. Raw scores of speed in sec (bars represent *SD*s) for the RAN task in the two groups
Figure 4.8.6. Raw scores (bars represent *SD*s) of the number of phonological and semantic errors in the RAN task in the two groups
Figure 4.8.7. Proportional scores (bars represent *SD*s) of the number of phonological, grammatical and orthographic errors in the two groups
Figure 4.8.8. Raw scores (bars represent *SD*s) of the number of correct designs (design fluency) in the two groups

Acknowledgments

I would like to express my gratitude to my supervisor Professor Chloë Marshall, for your encouragement and ever insightful input to the whole process. You provided me your complete attention, academic and constructive supervision and guidance. Thank you for your understanding, for recognising my needs throughout the whole process, for the meetings we had, for building a warm and supportive relationship between yourself and your students and for engaging us in the academic life. And, thank you for our continuous collaboration towards our publications. I would also like to express my gratitude to Associate Professor Dr. Stavroula Stavrakaki, Aristotle University of Thessaloniki, who acted as my second supervisor, for accepting to be my second supervisor, for your knowledge, your feedback and for your support during my period in Thessaloniki while I was collecting my data. I have been very lucky having you both as my supervisors. Thank you for giving me the opportunity to study for a PhD degree in England and for helping me throughout so that this thesis is now an entity.

I would like to thank Professor Lucy Henry, City University, who was my external examiner, and Professor Jackie Masterson, who was my internal examiner, for accepting to be examiners of my thesis. I very much appreciate your careful reading of my thesis and your helpful suggestions for improvement in my viva voce exam I had in December 2018. I am thankful to Senior Lecturer Dr. Roberto Filippi too who acted as the internal reader of my thesis.

I am grateful to everyone in the Department of Psychology and Human Development at the Faculty of Institute of Education, University College London. I am especially thankful to Professor Julie Dockrell and Associate Professor Dr. Yvonne Griffiths who were my examiners for my upgrade from MPhil to PhD, for your constructive feedback which helped me to better conceptualise my thesis. I am also especially thankful to my friend Dr. Nelly Joye, for being a real friend to me, for your politeness, your encouragement, your knowledge and the time we spent together during our studies. I have been very lucky to meet you and have you around. I also thank Dr. Jaime Balladares

Hernández, for helping me when I asked for your help, for your willingness to do so and for sharing your experiences of your studies with us. I thank all the tutors at University College London who support PhD students by offering courses to improve doctoral skills, for their enthusiasm to transfer research skills to doctoral students. I would also like to thank Associate Professor Dr. Eddy Davelaar, Birkbeck University, for offering us, my supervisor and myself, insight into his work on computational modelling techniques, a new way for analysing my semantic fluency data, and his willingness to collaborate with us towards a publication on semantic fluency data using his new metric. I thank Associate Professor Dr. Christian Hennig, University of Bologna, for his willingness to provide me statistical advice in order to analyse my spelling data in the future.

I thank the Hellenic Ministry of Education, Research and Religious Affairs for the ethical approval of this study. My sincere thanks also to Dr. Ioanna Talli, Speech and Language Therapist, for your kindness, your goodwill and your help during the time I spent in the school you work in. Thank you for allowing me to use certain tasks you had adapted into the Greek language, for hosting me in your Class and for helping me to recruit children for my clinical sample. I also thank all the headteachers, teachers, teachers of inclusive education, speech and language therapists and parents, for their trust and their contribution to data collection.

Furthermore, I thank the Editor, Professor Ludovic Ferrand, Centre National de la Recherche Scientifique (CNRS), and the two reviewers, Dr. Norbert Maïonchi-Pino, Laboratoire de Psychologie Sociale et Cognitive (LAPSCO) and Professor Franck Ramus, Centre National de la Recherche Scientifique (CNRS), for their constructive feedback of our recently (April, 2019) published paper in *Frontiers in Psychology* entitled "Deficient Explicit Access to Phonological Representations Explains Phonological Fluency Difficulties in Greek Children with Dyslexia and/or Developmental Language Disorder" (doi: 10.3389/fpsyg.2019.00638). I also thank, for their useful comments on content, the two anonymous reviewers of our paper entitled "Semantic fluency difficulties in Developmental Dyslexia and Developmental Language Disorder (DLD): Poor semantic structure of the lexicon or slower retrieval processes?", which was submitted in summer 2018. The work of the thesis was also presented as a poster in the *UCL Institute of Education, University College London Doctoral Poster Conference* in London on 23rd March 2018 entitled "Using verbal fluency tasks to investigate the lexicon in Greek-speaking children with literacy and language disorders" (doi: 10.13140/RG.2.2.13745.28006), as well as in the *1st Society for Research on Learning Disorders Conference* in Padova on June 7th 2019 entitled "Semantic and phonological fluency difficulties in children with dyslexia and/or Developmental Language Disorder: How their lexical retrieval difficulties are explained?" (doi: 10.13140/RG.2.2.21387.54565).

Presence is important. I thank my friends in London with whom I spent very good moments, walking around the city, discovering new places, taking photos and sharing thoughts and dreams. I am especially thankful to my family living in Greece. Even though England is more than two thousand kilometres away from Greece, my parents have been always present for me.

This thesis would not be an entity without the very wonderful children who very generously offered me their time along with their smile. I especially thank the children of my clinical sample for whom, I know, this might have been a demanding task. I hope that this thesis deserves your time and your smiles.

I hope that this thesis will be read pleasantly by the reader.

Abbreviations

AD: Alzheimer’s Disease
ADHD: Attention Deficit Hyperactivity Disorder
APA: American Psychiatric Association
ASD: Autism Spectrum Disorder
BSL: British Sign language
CI: Confidence Interval
CIs: Cochlear implants
CPM: Raven’s Coloured Progressive Matrices
CV: Consonant Vowel
DDLD: Dyslexia and/or Developmental Language Disorder
DLD: Developmental Language Disorder
DSM: Diagnostic and Statistical Manual of Mental Disorders
DVIQ: Diagnostic Verbal Intelligence
EF: Executive Functioning
EVALEC: Evaluation de la Lecture
ICD: International Classification of Diseases
LangLit: Language and Literacy
M: Mean
MANOVA: Multivariate Analysis of Variance
MCI: Mild Cognitive Impairment
NEPSY: Neuropsychological Assessment
NH: Normal Hearing
NVIQ: Nonverbal IQ
NWR: Nonword Repetition
PCA: Principal Component Analysis
PhAB: Phonological Assessment Battery
Phonological STM: Phonological Short-Term Memory
PPVT-R: Peabody Picture Vocabulary Test-Revised
RAN: Rapid Automatic Naming
SD: Standard Deviation
SLI: Specific Language Impairment
TD: Typically-developing/Typical Development
TR: Total Responses
WHO: World Health Organization
WISC: Wechsler Intelligence Scale for Children
WFD: Word-finding Difficulties

Structure of the thesis

This thesis is organised in five chapters. In Chapter 1a, a general overview of the thesis aims and also the research questions is presented followed by a section on recent definitions of Developmental Dyslexia and Developmental Language Disorder (DLD), a section describing language difficulties in children with dyslexia and DLD, and a section on the comorbidity of dyslexia and DLD. In Chapter 1b, the theoretical models and hypotheses are presented in order to define the theoretical background in which this study should be placed. The *Structure-Loss Model* and the *Slow-Retrieval Model* are two models accounting for lexical difficulties as exemplified in semantic fluency tasks in adults. A developmental perspective is adopted in the current study, however, for the two models, with the Poor Lexical-Semantic Structure Model and the Slow-Retrieval Model being considered in order to investigate which of the two models better explains DDLD children's retrieval difficulties in semantic fluency tasks. The presentation of the two theoretical models accounting for lexical-semantic difficulties is followed by two sections, one discussing the role of semantic and phonological processing skills in reading ability, and another discussing spelling in children with dyslexia and DLD. In the same chapter are also reviewed the *Phonological Theory* of dyslexia and the phonological deficit which is characteristic of dyslexia. The *Degraded Phonological Representations Hypothesis* and the *Deficient Phonological Access Hypothesis* are then presented as a theoretical framework on the basis that the current study investigates which of these two phonological hypotheses better explains the locus of the phonological deficit in dyslexia and DLD using phonological fluency tasks, three phonological tasks and a spelling-to-dictation task.

Chapter 2 presents the research literature on the verbal fluency tasks used in this study, namely, semantic and phonological fluency tasks, and what these tasks measure. Evidence is reviewed originating from TD children and healthy adults showing the effects of switching and clustering behaviour on semantic and phonological fluency performance, in addition to the effects of age, level of intelligence and gender. Evidence is also reviewed regarding automatic versus controlled processing in verbal fluency tasks. Evidence is then reviewed originating from children with dyslexia, DLD, word-finding difficulties (WFDs), and adults with dyslexia and/or DLD showing group differences between all those clinical groups and appropriate controls on semantic and phonological fluency performance. With respect to switching, clustering and cluster size in verbal fluency tasks in clinical groups, evidence is reviewed from children with dyslexia and DLD and adults with dyslexia and/or DLD. Correlation, network and computational modelling methodologies used to analyse data from verbal fluency tasks are presented as alternative methodologies offering insight into the structure of the semantic network. Automatic versus controlled processing in verbal fluency tasks in clinical groups is then presented. The chapter continues with a presentation of the effects of language, literacy and executive functions on semantic and phonological fluency performance, followed by a presentation of design fluency with evidence originating from TD children, children and adults with dyslexia and children with DLD being provided. The chapter ends with the study's research questions and predictions.

Chapter 3 presents the methods used to answer the study's research questions. Two groups were tested in a cross-sectional design using measures of verbal fluency, language, literacy, phonology, and nonverbal fluency: a group of TD children aged 6-12 years and a group of children with dyslexia and/or DLD (hereafter DDLD group) aged 7-12 years, all monolingual Greek speakers. Chapter 3 provides information on the methods and criteria of selection of the participants, and nonverbal IQ (NVIQ), language, literacy and phonological abilities of the two groups. Besides a detailed description of the participants, this chapter also describes what type of tasks were used to examine NVIQ, language, literacy and phonological abilities, and verbal and nonverbal fluency, and which were used as descriptive measures and which of these were used as experimental measures. The procedures of assessment, calculation of statistical power and threshold of statistical significance are also presented.

Chapter 4 presents first the results on group differences on semantic fluency tasks followed by group differences on phonological fluency tasks. The chapter continues with presenting regression analyses showing first the contribution of children's language and literacy skills to semantic and phonological fluency performance and then to automatic (first 15 s of the test period in verbal fluency tasks) and controlled processing (subsequent 45 s of the test period in verbal fluency tasks) in the two verbal fluency conditions. The results on group differences on the design fluency task are then presented. There is also a section presenting results meeting a less stringent (i.e. $p < .05$) and a more stringent (i.e. $p < .005$) threshold of statistical significance, in addition to a summary of the results.

In the final chapter, Chapter 5, the results of the study are discussed in relation to the study's objectives and research questions, and in relation to the research literature. To aid the reader, there is first a recap of the study's objectives and research questions. This is followed by a specification of the research questions with respect to semantic and phonological fluency tasks and the models and hypotheses accounting for poorer semantic and phonological fluency performance in children with DDLD. The results of phonological tasks and the spelling-to-dictation task in relation to the two phonological hypotheses of dyslexia and DLD are then discussed. The contribution of children's language and literacy skills to semantic and phonological fluency performance is also discussed, in addition to the contribution of children's language and literacy skills to automatic and controlled processing in the two verbal fluency conditions. Next, a discussion of the design fluency task used is presented. The findings are linked to relevant findings in the literature throughout Chapter 5. The chapter concludes with a discussion of theoretical implications, followed by a section of strengths, limitations, further directions and conclusions of the study.

Chapter 1a. General background to the thesis

1a.1. General overview of the thesis aims and key questions

This overview outlines the overall thesis aims, research questions and general approach. The overview also emphasises what is novel about the current research. Dyslexia and Developmental Language Disorder (DLD), hereafter DDLD, are two neurodevelopmental disorders which affect, respectively, the development of literacy and oral language skills. Lexical organization has been less well-studied in children with dyslexia and DLD than other components of language, such as phonology, morphology and syntax. In the current study, lexical organization was tested using three semantic and three phonological fluency categories in sixty-six children with dyslexia and/or DLD, combined in one group, the DDLD group, aged 7-12 years and 83 TD children aged 6-12 years, all monolingual Greek speakers.

In the light of evidence that children with DDLD show poor semantic fluency performance compared to age-matched TD children (as reviewed in Chapter 2), the first aim of the current study was to investigate whether the lexical retrieval difficulties that children with DDLD display in semantic fluency tasks can be attributed to the semantic structure of their lexicon being poor or to items being retrieved more slowly despite the semantic structure being intact. Semantic fluency tasks require people to produce as many words as they can which belong to certain categories, such as ‘animals’, in a limited period of time (e.g. 60 s). Further, in semantic fluency tasks, responses are often produced in clusters of semantically-related items (e.g. “cat-dog” is a semantic cluster of ‘pets’).

The current study was designed to test two theoretical models that could potentially account for retrieval difficulties in semantic fluency tasks. The two models considered, the Structure-Loss Model and the Slow-Retrieval Model, were initially developed based on adult data and to date they have been tested only in adults. The current study is therefore the first developmental study designed to test two theoretical models accounting for lexical difficulties in semantic fluency tasks. The two models were named the *Poor Lexical-Semantic Structure Model* and the *Slow-Retrieval Model*. Both models predict that children with DDLD will retrieve fewer items. However, while the Poor Lexical-Semantic Structure Model predicts a less sophisticated network of connections between items in the lexicon in children with DDLD, as evidenced by smaller clusters of related items, the Slow-Retrieval Model predicts intact inter-item associations in the lexicon, as evidenced by clusters being of a similar size in the two groups.

Previous studies have reported that word productivity in semantic fluency tasks is related to the number of switches and cluster number rather than average cluster size (e.g. Marshall, Rowley, Mason, Herman, & Morgan, 2013; Marshall et al., 2018). The basic fact that responses are clustered suggests that the lexical items are organised in subcategories. Whether it is the number of clusters retrieved or the size of clusters retrieved that drives productivity is a separate question.

That lexical items are organised in subcategories is suggested just by the existence of clusters. Furthermore, switching is considered a measure of executive functions (EFs) (e.g. Kavé, Kigel, & Kochva, 2008; Troyer, 2000), and children move from one subcategory to another, that is, they switch among subcategories, in order to retrieve as many lexical items as possible. This is the first study which will investigate semantic clustering and switching in Greek children with DDLD.

Further, verbal fluency tasks are governed by certain rules and require children to inhibit certain responses. Children should inhibit inappropriate responses that come readily to mind, i.e., responses that have already been produced in the sequence to avoid repeated responses, and out-of-category responses (e.g. the item ‘balcony’ in the category of ‘objects from around the house’) in order to abide by the rules of the task. It is predicted that any impairments in EFs might result in an increased number of incorrect responses, with a high proportion of errors suggesting difficulties in word search and retrieval processes, and impaired executive control over semantic search and retrieval strategies. The groups' number of incorrect responses, and a related measure, the error ratio, will be therefore compared.

Overall, this study set out to answer the following main research question about semantic fluency in Greek children with DDLD:

- What is the structure of the lexicon in children with DDLD compared to TD children? Is poorer semantic fluency performance in children with DDLD better explained by impoverished semantic structure or slower retrieval processes?
- Do cluster number and/or cluster size drive productivity in semantic fluency tasks in TD children and children with DDLD?

The second aim of the current study was to add to the theoretical debate on the locus of the phonological deficit in children with dyslexia and DLD. The phonological deficit is evident in three interrelated but distinct phonological processing skills—namely, phonological awareness, phonological short-term memory (STM) and rapid automatic naming (RAN) skills (Goswami, 2003). All these different abilities that have been considered relate to reading ability and involve phonological processing skills and the involvement of phonological representations. Children with dyslexia and DLD perform poorly compared to TD children on tasks tapping the three above-mentioned phonological skills (see for dyslexia: Wagner & Torgesen, 1987; see for DLD: Ramus, Marshall, Rosen, & van der Lely, 2013). However, there has been some debate as to whether their phonological deficit arises directly from an impairment in phonological representations (referring to the abstracted way that speech sounds of a particular language are represented in the brain), or instead from deficient access to (intact) phonological representations. The latter view reflects a central distinction in the literature between explicit and implicit phonology. Performance on phonological tasks tapping explicit phonology is fostered by learning to read, and therefore cause and effect is difficult to disentangle. Phonological tasks tapping explicit knowledge are those tasks which require metalinguistic awareness defined as an explicit awareness or knowledge about the structure and properties of a language (e.g. delete /s/ from the nonword ‘ston’ to pronounce ‘ton’). Explicit skills, however, contrast with the implicit skills that are automatically involved in phonological tasks and which invoke only a minimal level of metalinguistic awareness. Phonological tasks tapping implicit knowledge can be used even with preschool children who have not yet acquired reading.

Based on evidence that children with dyslexia perform poorly on a range of phonological tasks, the leading view on dyslexia for many years regarding the locus of the phonological deficit has been that phonological representations are degraded (i.e. less robust and distinct), and that this primary representational deficit impacts upon higher-level phonological processing skills, and ultimately, upon reading development. This view is called the Degraded Phonological Representations Hypothesis (e.g. Goswami, 2000; Leong, Hämäläinen, Soltész, & Goswami, 2011; Ziegler & Goswami, 2005). The concept of degraded phonological representations implies that in the course of development, children with dyslexia have experienced difficulties in establishing representations of phonological units adequately robust and distinct for the recognition and production of words.

There is an alternative view, however, explaining the origin of the phonological deficit observed in children with dyslexia. The Deficient Phonological Access Hypothesis (Ramus & Szenkovits, 2008) argues that the phonological deficit in dyslexia is evident only under certain task demands, namely, tasks requiring explicit manipulation of speech sounds, loading phonological STM, or requiring speeded access to phonological representations. According to this hypothesis, phonological representations of people with dyslexia are intact, but hard to access because of the involvement of the afore-mentioned processes, which are required to explicitly access phonological representations, processes which are deficient in dyslexia. Therefore, it is only by using phonological tasks with minimal processing demands that the quality of phonological representations themselves can be assessed (Ramus et al., 2013). Most of the studies supporting a phonological access deficit have been conducted in adults. Adopting a developmental perspective, however, allows one to test what is perhaps the most valid criticism of the Deficient Phonological Access Hypothesis: the possibility that adults with dyslexia have degraded phonological representations in childhood, but these representations have recovered in adulthood (e.g. Goswami, 2003).

The two phonological hypotheses presented were developed on the basis of the profile of people with dyslexia (children and adults). However, the two hypotheses also apply to DLD, given that many children with DLD have phonological difficulties linked to reading difficulties similar to those seen in children diagnosed with dyslexia (e.g. Bishop, MacDonald, Bird, & Hayiou-Thomas, 2009; Catts, 1993; Hulme & Snowling, 2009; Kamhi & Catts, 1986). Children with DLD have been reported to have degraded phonological representations and poorer phonological awareness, phonological STM and RAN skills relative to their TD peers (Ramus et al., 2013). It should be noted that considering these deficits, some theories of DLD propose that phonological representations mediate the relationship between impairments in auditory temporal processing and reading, and as such, in the case that there is no deficit in phonological representations *per se* in DLD, the causal chain underpinning auditory temporal processing theories would be weakened.

The current study's contribution to this debate is to test the Degraded Phonological Representations Hypothesis and the Deficient Phonological Access Hypothesis, using just one task – phonological fluency – which requires both explicit and implicit access to phonological representations. Phonological fluency tasks are lexical-retrieval tasks requiring children to produce as many words as they can which begin with particular letters, usually in a 60 s test period. Successful performance on phonological fluency tasks requires the search of the mental lexicon for words on the basis of their phonology. Importantly, the phonological fluency task measures two different aspects of access to phonological representations, namely, explicit access, as evidenced by the number of correct responses retrieved, and implicit access, as evidenced by the size of clusters produced. Both hypotheses predict that children with DDLD will retrieve fewer items than TD peers in the phonological fluency task. However, while the Degraded Phonological Representations

Hypothesis predicts smaller clusters of phonologically-related items in children with DDLD, the Deficient Phonological Access Hypothesis predicts that the two groups will not differ in cluster size.

Overall, the study addresses the following main research questions about phonological fluency in Greek children with DDLD:

- Where does the phonological deficit in children with DDLD lie? Is poorer phonological fluency performance in children with DDLD better explained by degraded phonological representations or by deficient explicit access to (intact) phonological representations?
- Do cluster number and/or cluster size drive productivity in phonological fluency tasks in TD children and children with DDLD?

The locus of the phonological deficit in children with DDLD is also investigated using (i) three phonological tasks, namely, phoneme deletion, nonword repetition (NWR) and RAN, and (ii) children's types of spelling errors in dictation which are assigned to three major categories of errors, namely, phonological, grammatical and orthographic errors. The aim is to investigate whether the phonological deficit in children with DDLD is better explained by the Degraded Phonological Representations Hypothesis or the Deficient Phonological Access Hypothesis. The groups' (i) accuracy in the three phoneme deletion tasks with items varying in syllabic length and syllabic complexity, (ii) the NWR accuracy in the different items varying in syllabic length, (iii) the time spent on naming items in the RAN task and the number of phonological errors found in the RAN task, and (iv) the proportion (proportionally to the total number of word spelled) of phonological (grapheme-to-phoneme) spelling errors in dictation are the variables of interest, and are therefore compared. With respect to the three phoneme deletion tasks, accuracy and speed performance are investigated in the phoneme deletion task of monosyllable items with simple CVC syllable structure, monosyllable items with complex CCV syllable structure, and trisyllable items with simple CVCVCV syllable structure. It is hypothesised that short nonwords with a simple syllable structure do not load phonological STM, and therefore that the two groups will show similar accuracy performance in this task; however, it is hypothesised that the DDLD group will show significantly poorer performance in the phoneme deletion task of trisyllable items with simple CVCVCV syllable structure and of monosyllable items with complex CCV syllable structure relative to the TD group. This is explained by the fact that long nonwords with three syllables or nonwords with complex syllable structure load phonological STM. It should be noted that phonological STM capacity is not actually measured, however, a NWR measure consisting of 3-, 4-, 5-, and 6-syllable nonwords is used to investigate whether the DDLD group shows phonological STM deficits relative to the TD group. With respect to the RAN task, slower naming performance is explained by the fact that a phonological access deficit in the DDLD group renders performance on tasks requiring speeded access to phonological representations particularly slow. Phonologically accurate performance is explained by intact access to phonological representations in the RAN task.

With respect to the spelling task, the objective of this study is to examine what specific spelling errors in dictation are made by Greek children with DDLD and whether the same errors are made by their TD peers. The ultimate objective, however, which is related to the two prominent phonological hypotheses considered, is to investigate which of the two prominent phonological hypotheses presented above better characterises the locus of the phonological deficit in Greek children with DDLD. To this end, it will be investigated whether children with DDLD differ from TD children on the proportional number of phonological spelling errors. The Degraded Phonological Representations Hypothesis predicts that qualitative analysis of spelling errors will reveal that the DDLD group produces a higher proportion of phonological spelling errors than the TD group. This is explained by inaccurate phonological representations in the DDLD group. The Deficient Phonological Access Hypothesis, however, predicts that qualitative analysis of spelling errors will reveal that in the DDLD group the majority of spelling errors will be phonologically correct. This is explained by accurate phonological representations but inappropriate orthographic encoding of words using grapheme-to-phoneme mappings that are inappropriate for a particular context in the DDLD group.

Another issue emerging from the two phonological hypotheses and the two lexical-semantic models presented above is the specificity of the verbal fluency deficit, be it semantic or phonological, in dyslexia and DLD. The design fluency task used in the current study measures visuospatial executive skills by assessing a child's ability to generate nonsense designs under time constraints and restricted design conditions. It is therefore a similarly-structured task to verbal fluency tasks, but it does not require phonological or semantic representations, or phonological processing or semantic processing skills. Both phonological hypotheses and both lexical-semantic models predict that children with DDLD will generate a similar number of correct designs in the design fluency task compared to TD children, advocating a 'modular' deficit within the language system which affects the verbal domain, whilst the nonverbal domain is unaffected. It is hypothesized that if a generalised slower processing speed in children with DDLD accounts for their lower verbal fluency performance, they will also have poorer design fluency performance compared to their TD peers; however, if only difficulties with verbal processing skills underlie poorer verbal fluency performance, the two groups will show similar design fluency performance. Thus, the following research question was also addressed. How specific is the verbal fluency deficit in children with DDLD: Does it extend to a nonverbal task (design fluency)? Furthermore, in order to test the specificity of the verbal fluency deficit in children with DDLD, design fluency performance will be used as a covariate in the analysis investigating group differences in semantic and phonological fluency performance.

1a.2. Definitions of Developmental Dyslexia and Developmental Language Disorder (DLD)

In the latest edition of the International Classification of Diseases (ICD)-11 (World Health Organization [WHO], 2018), developmental dyslexia (hereafter dyslexia) is included under the term 'developmental learning disorder with impairment in reading' characterised by “significant and persistent difficulties in learning academic skills related to reading, such as word reading accuracy, reading fluency, and reading comprehension. The individual’s performance in reading is markedly below what would be expected for chronological age and level of intellectual functioning and results in significant impairment in the individual’s academic or occupational functioning” (retrieved from: https://icd.who.int/dev11/l-m/en#/http://id.who.int/icd/entity/1008636089).

With respect to comorbidities, dyslexia commonly co-occurs with Developmental Language Disorder (DLD), defined as a neurodevelopmental disorder of communication in which the child’s language development falls well behind that of other children of the same age for no apparent reason that could account for problems with understanding or producing spoken language (e.g. Bishop & Norbury, 2008; Botting, 2014; Williams & Lind, 2013). Norbury et al. (2016) reported that DLD in the UK affects an estimated 7.58% in children aged 4 years 9 months to 5 years 10 months. Bishop, Snowling, Thompson, and Greenhalgh (2016) gathered an international group of experts, the CATALISE consortium, who agreed on the term ‘Developmental Language Disorder’ (DLD) to replace the term ‘Specific Language Impairment’ (SLI) when the child has receptive or expressive language problems that affect every day functioning and when language disorder is not part of a broader developmental condition, such as Autism Spectrum Disorder (ASD), or a known condition, such as brain injury and sensori-neural hearing loss. The term DLD is used by WHO (2018) in the latest ICD-11 in which DLD is characterized by

“persistent difficulties in the acquisition, understanding, production or use of language (spoken or signed), that arise during the developmental period, typically during early childhood, and cause significant limitations in the individual’s ability to communicate. The individual’s ability to understand, produce or use language is markedly below what would be expected given the individual’s age and level of intellectual functioning. The language deficits are not explained by another neurodevelopmental disorder or a sensory impairment or neurological condition, including the effects of brain injury or infection” (retrieved from:

https://icd.who.int/dev11/lm/en#/http://id.who.int/icd/entity/862918022).

The label DLD has been embraced by some researchers (Joye, 2018; Joye, Broc, Olive, & Dockrell, 2019; Mengisidou & Marshall, 2019) since it was first proposed by Bishop et al. (2016). However, some researchers are more ambivalent about this label (see Discussion chapter for details).

1a.3. Language difficulties in children with dyslexia and DLD

Most persons with DLD have restricted expressive and receptive language skills. This is reflected well in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; American Psychiatric Association [APA], 2013) criteria for language disorder and the equivalent criteria in the ICD-11 (WHO, 2018), which stress the importance for the diagnosis of language disorder to be based on standardised administered measures of receptive and expressive aspects of language.

Children with DLD whose deficits concern grammar in particular are reported to have problems both in producing and comprehending syntactically complex sentences (van der Lely, 1998). An overall impression therefore of a child experiencing DLD is that their oral language is like that of a much younger child as they may use, for example, simple language with limited vocabulary and words in short, ungrammatical strings; e.g. “Me go there”, rather than “I went there” (Bishop, 2006). Indeed, children with DLD may produce short, simple sentences, and they may have receptive language difficulties which affect language understanding. As such, they may be confused by longer sentences, comprehending only a few words of a sentence such as “If you bring your swimsuit tomorrow, we can go to the pool after lunch” (Bishop, 2007), or by complex sentence structures so that they may fail to understand someone’s intended meaning, for example, the passive sentence “The elephant was pushed by the boy” (Bishop, 1982). In addition, children with DLD may omit grammatical suffixes, such as the past tense ending in English regular verbs (i.e. -ed), or the third person singular ending in English verbs (i.e. -s), and may produce sentences with bare stem errors, such as “Yesterday, I play_ in the park”, or “My brother like_ chocolate”. They may also produce sentences such as “Who Marge saw someone?”, or “Yesterday I fall over” (van der Lely & Marshall, 2011).

Aside from grammatical deficits, children with DLD have been reported to have a phonological impairment exemplified in a NWR task varying in syllabic and/or metrical complexity and in particular problems with repeating nonwords containing consonant clusters (Gathercole & Baddeley, 1990; Marshall, Harris, & van der Lely, 2003; Marshall & van der Lely, 2009). The NWR task is widely considered to be a measure of phonological STM.

Children with DLD also experience lexical problems (Marshall, 2014) manifesting themselves as word-finding difficulties (WFDs); that is, they know the meaning of a word but they cannot remember it and they cannot access it readily. They might also have lower vocabulary knowledge, as measured by a test assessing receptive vocabulary, such as the Peabody Picture Vocabulary Test (Dunn & Dunn, 2007) or the British Picture Vocabulary Scale (Dunn, Douglas, & Dunn, 2009).

In the Greek language, children with DLD have been reported to show a similar profile to English-speaking children with DLD. For example, they have difficulty acquiring subject-verb agreement and grammatical morphemes (Stavrakaki, 2005), show difficulty with relative clauses and wh-questions (Stavrakaki, 2001; Stavrakaki, Chrysomallis, & Petraki, 2011), and perform poorly relative to children with dyslexia and TD children on tasks measuring listening and reading comprehension skills (Talli, Sprenger-Charolles, & Stavrakaki, 2015).

The Greek language has a shallow orthography, which means that it is characterized by consistent grapheme-to-phoneme mappings (Seymour, Aro, & Erskine, 2003), estimated to be 95% consistent for reading and 80% consistent for spelling (Protopapas & Vlahou, 2009). Considering this high level of orthographic consistency, it is not surprising that reading difficulties are evident primarily in poor reading fluency rather than poor reading accuracy (Nikolopoulos, Goulandris, & Snowling, 2003). Poor reading fluency in turn is associated with poor performance on phonological awareness and rapid automatic naming tasks (Nikolopoulos, Goulandris, Hulme, & Snowling, 2006; Protopapas, Altani, & Georgiou, 2013; Protopapas, Fakou, Drakopoulou, Skaloumbakas, & Mouzaki, 2013). Having said that, reading accuracy difficulties are evident in children with dyslexia even in Grade 7 (Protopapas & Skaloumbakas, 2007; Protopapas, Skaloumbakas, & Bali, 2008; Protopapas, Simos, Sideridis, & Mouzaki, 2012). With respect to phonological difficulties, children with dyslexia and DLD have been reported to show phonological deficits in tasks measuring phonological awareness, phonological short-term memory, and rapid automatic naming skills (e.g. Diamanti, Goulandris, Campbell, & Protopapas, 2018; Spanoudis, Papadopoulos, & Spyrou, 2018; Talli, Sprenger-Charolles, & Stavrakaki, 2016). There is also evidence that relatively easy tasks for assessing phonological awareness, such as phoneme segmentation and phoneme deletion tasks, show ceiling effects by the end of Grade 1, and are not therefore able to reveal children's phonological difficulties (Papadopoulos, Spanoudis, & Kendeou, 2009; Papadopoulos, Kendeou, & Spanoudis, 2012). However, more demanding phoneme deletion tasks, when stimuli comprise polysyllabic nonwords with consonant clusters, can reveal group differences in 3rd and 4th Graders (Protopapas et al., 2008), and in children with dyslexia through secondary education (Anastasiou & Protopapas, 2015; Protopapas & Skaloumbakas, 2007). With respect to phonological skills in DLD, children with DLD aged 8-12 years are reported to show poorer phonological short-term, working and long-term memory skills relative to their TD peers (Spanoudis & Natsopoulos, 2011). Further research is needed to investigate the profile of Greek children with dyslexia and DLD in semantic fluency tasks that assess lexical organization and patterns of lexical retrieval, and in phonological fluency tasks that assess the quality of phonological representations and explicit and implicit access to them.

1a.4. Comorbidity between dyslexia and DLD

Dyslexia co-occurs with DLD with an overlap of approximately 50 percent, and accordingly, the probability of showing dyslexia is much higher in children diagnosed with DLD than in those without DLD (e.g. Bishop & Snowling, 2004; Eisenmajer, Ross, & Pratt, 2005; McArthur, Hogben, Edwards, Heath, & Mengler, 2000; Messaoud-Galusi & Marshall, 2010; Nash, Hulme, Gooch, & Snowling, 2013; Pennington & Bishop, 2009; van der Lely & Marshall, 2010). At the behavioural level, McArthur et al. (2000) reported that 55% of the children in their study who had initially been classified as having dyslexia also had oral language difficulties, and 51% of the children initially identified as having DLD had a reading disability. Thus, 53% of their 212 participants met the criteria for both dyslexia and DLD. Evidence suggests that neither the specificity of deficits in children with dyslexia and DLD can be assumed (as nearly 50% of the children with dyslexia and DLD have deficits in *both* written and oral language) nor the comorbidity between dyslexia and DLD (as nearly 50% of the children with dyslexia and DLD have *either* written- *or* oral language-specific deficits).

At the cognitive level, Bishop and Snowling (2004) argued that conceptualising two aspects of spoken language - namely, phonological language skills and broader oral language skills - helps elucidate the relationship between dyslexia and DLD, as shown in Figure 1a.1. This model proposes that TD readers have both good phonological and language comprehension skills. Children with dyslexia have poor phonological processing skills and good language comprehension skills, while poor comprehenders show the reverse profile, namely, they have good phonological processing skills but poor language comprehension skills. Children with DLD have poor language comprehension skills and some of them also have poor phonological processing skills, or dyslexia, that is, they have DLD *plus* dyslexia.

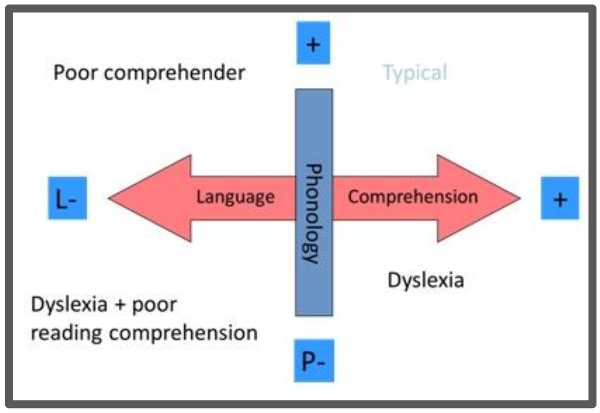


Figure 1a.1. Bishop and Snowling's (2004) two-dimensional model for the relationship between dyslexia and poor comprehension (from Snowling, 2017)

According to many of the proposed models attempting to account for the relationship between dyslexia and DLD, the phonological deficit underlies the overlap between the two conditions. Indeed, it has been reported that children with DLD have similar phonological difficulties as those with dyslexia (e.g. Bishop, McDonald, Bird, & Hayiou-Thomas, 2009; Brooks & Kempe, 2012; Hulme & Snowling, 2009; Kamhi & Catts, 1986). Nevertheless, even though there is a lot of overlap between the two conditions, there are children who have either dyslexia or DLD but not both. In a study designed to investigate why the overlap is not complete, Bishop et al. (2009) investigated what characterizes children who learn to read and write despite DLD. They reported that tasks measuring speed of lexical retrieval were a key factor. Children with DLD who can develop their decoding skills are children who show a better performance on a rapid serial naming task requiring speeded lexical access compared to children with DLD who had poor decoding skills. As such, according to Bishop et al. (2009) study, children with DLD only and children with DLD *plus* dyslexia seem not to have exactly the same deficits in phonological processing skills but that they can be differentiated in a task assessing speeded lexical access.

Similarly, in another study designed to investigate why the overlap between dyslexia and DLD is not complete, Ramus et al. (2013) used a wide battery of phonological and nonphonological language tasks to investigate whether children with dyslexia and children with DLD show the same pattern of phonological difficulties. Factor analysis revealed evidence for a possible dissociation between explicit and implicit phonology in a group of sixty-five 8-12 year-old English children with dyslexia and/or DLD. Explicit phonology was measured by phonological tasks that really do require phonological processing skills and implicit phonology was measured by tasks that demand phonological representations themselves. Ramus et al. (2013) argued that the phonological deficit in dyslexia may not lie in deficient phonological representations themselves—as only a subset (one-third) of the children with dyslexia had a deficit in phonological representations—but rather in some cognitive skills involved in complex phonological tasks that also involve explicit skills. Rhyme, spoonerisms, rapid digit naming, digit span, phonemic categorization, and prosodic output tasks loaded on the phonological processing skills component. NWR, NW discrimination, picture-word matching, and articulation tasks loaded on the phonological representations component. As Figure 1a.2. shows, there is a strong positive association between phonological processing skills and phonological representations. It is also evident from Figure 1a.2. that the majority of children with dyslexia show poor phonological processing skills but few of them showed a deficit in phonological representations. Most of the children with dyslexia *plus* DLD showed both poor phonological processing skills and representations. As such, Ramus et al. (2013) proposed that the profile of dyslexia and DLD is different. DLD is related to deficits in both phonological representations and phonological processing skills, whereas dyslexia is related to deficits in the skills that operate on phonological representations and not the representations themselves. In Figure 1a.2. lines correspond to a -1.5-SD threshold (Ramus et al., 2013). It should be noted that the study of Ramus et al. (2013) is discussed in more detail in the subsequent sections as it is crucially relevant to the main hypotheses considered in this thesis.

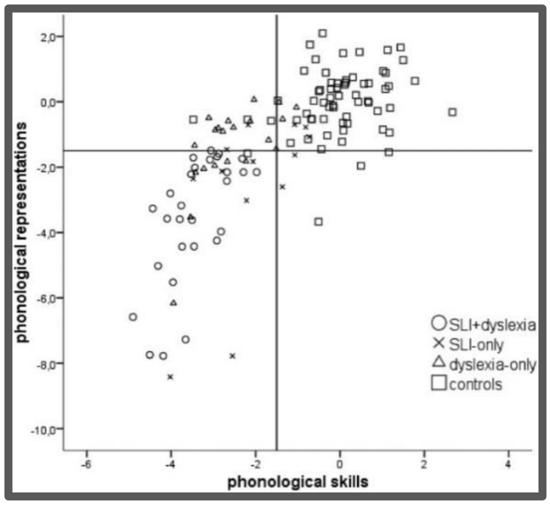


Figure 1a.2. Phonological processing skills and phonological representations in control children and in children with dyslexia, DLD (referred to as SLI in the Figure), or DLD *plus* dyslexia (from Ramus et al., 2013, p. 639)

Two hypotheses have been presented above to account for the relationship between dyslexia and DLD. Snowling (2014) argued that recent findings originating from studies of children at family risk of dyslexia are consistent with dyslexia being a language-learning impairment, or a subtype of DLD. This is based on evidence showing that reading difficulties originate from two separable causes: poor phonological processing skills and poor language skills. Snowling argues that studies focused on developmental trajectories showed that if poor language problems resolve early, then the phonological deficit which is at the core of language learning impairments will solely affect written language skills, with the affected child showing the dyslexia profile. This is one hypothesis. Another hypothesis has been proposed by Ramus et al. (2013) claiming that children with DLD have impaired phonological representations resulting in severe and persistent language problems, and children with dyslexia show deficient access to (intact) phonological representations. An access deficit would mean that children's phonological representations developed normally, but for some reason, which is not yet evident in the literature, are less accessible. As Snowling (2014) argues, this reduced accessibility would result in reading difficulties because it would make grapheme-to-phoneme mappings, defined as the rules determining how to pronounce each grapheme, more difficult to learn.

Chapter 1b. Hypothesis-driven research questions

1b.1. The Structure-Loss Model and the Retrieval-Slowing Model

In a seminal publication, Collins and Loftus (1975) claimed that words are stored in a mental lexicon which is organized in a semantic network. Decades of research since then have investigated the structure of this network. One characteristic feature is considerable local structure in the form of clusters of words that are highly interconnected by semantic relatedness (e.g. Hills, Maouene, Maouene, Sheya, & Smith, 2009; Rogers & McClelland, 2008). Individuals differ in this connectivity, and these individual differences are associated with individual differences in language development. For example, Beckage, Smith, and Hills (2011) investigated the structure of the semantic networks in children aged 15 to 36 months who had either typical vocabulary size for their age or whose vocabulary size was small relative to their age (i.e. children who were late talkers). They found that the semantic networks of typical language learners had a greater degree of connectivity and clustering compared to late talkers, even when their overall vocabulary size was the same.

The semantic network is important for many aspects of language processing, and different experimental tasks can be used to investigate its structure. One such task – which can be used across the age span, from children to older adults, and is widely administered in clinical settings – is semantic fluency. Two models, namely the Structure-Loss Model and the Retrieval-Slowing Model, account for lexical retrieval difficulties in semantic fluency tasks. The Structure-Loss Model states that lexical difficulties can be accounted for by the loss of semantic structure, and that effective word retrieval is facilitated by inter-item associations between that word and other, semantically-related words (Rohrer, Wixted, Salmon, & Butters, 1995). The category of foods, for example, subsumes subcategories, such as fruits and vegetables, which in turn subsume exemplars, such as apple and orange, and spinach and cabbage. According to this model therefore, evidence for the loss of semantic structure originates from one's inability to cluster words around subcategories when performing a semantic fluency task.

Evidence (in Chertkow & Bub, 1990) for an impoverished semantic structure originates from people with Alzheimer's Disease (AD), in that people with AD have been found to produce statistically smaller clusters in the category of 'objects found in a supermarket' (around 2 items per subcategory) compared to healthy controls, and to switch between subcategories less often than controls. It has also been found that people with AD produced more subcategory responses (e.g. fruits) than specific exemplars (e.g. apple). Chertkow and Bub (1990) assessed how many words out of a total of 130 words people with AD knew, and they divided the 130 words into two groups for each participant: intact and degraded. They then asked participants to produce category exemplars, and they found that about 90% of the intact words and only 12% of the degraded words were produced by people with AD. That the same words could not be accessed either by probing or during the verbal fluency task is consistent with the Structure-Loss Model.

The Retrieval-Slowing Model states that lexical difficulties can be accounted for by an access deficit, while the associative networks within semantic memory are preserved. According to this model, lower performance in a semantic fluency task implies that the participant has intact semantic structure, but their lower performance is attributable to slower retrieval times or a slowing of retrieval processes (Rohrer et al., 1995). In Nebes, Boiler and Holland's (1986) study, participants heard a sentence in which the final word was missing and were asked to complete the sentence choosing a word that could make sense. The easier sentences had fewer acceptable final words ("Most cats see well at _") than the more difficult, less restrictive sentences ("In the distance they heard a _"). At each level of difficulty, the mean response-time for people with AD was significantly slower than controls, and mean response-time increased as a function of sentence-completion difficulty to the same degree for both groups. These findings were interpreted as evidence for the preservation of associative networks within semantic memory.

As Mirman and Britt (2014) argued, the term ‘access’ is used to describe: “the set of phenomena that are distinguished from storage deficits” (p. 12). In Mirman and Britt’s (2014) own words, “the central premise of access deficits is that the knowledge itself is intact, but access is ineffective, inefficient or inconsistent” (p. 2). By ‘access deficits’ therefore is meant that lexical-semantic representations are intact but access to these representations is impaired. In their paper, they use a cartoon to schematize the distinction between ‘storage’ and ‘access’ deficits. As shown in Figure 1b.1., a storage deficit (shown in the middle) means that the container that stores words is missing words, whereas an access deficit (shown on the right) means that the container is not missing any words but the mechanism for retrieving, or accessing, these words is problematic, or less effective. The typical adult state is shown on the left where in the container there is both a full collection of words and an effective mechanism for accessing them.

Figure 1b.1. A schematic diagram showing the typical adult state on the left, a storage deficit shown in the middle and an access deficit shown on the right

Troster et al. (1995) tested the two models, that is, the Structure-Loss Model and the Slow-Retrieval Model, in a sample of people with intractable temporal lobe epilepsy using the *supermarket* fluency test (“Name as many items found in a supermarket as possible”). They reported that people with epilepsy produced fewer words, and a significantly higher ratio of category labels than exemplars (e.g. ‘fruit’ or ‘vegetable’ instead of ‘orange’ or ‘broccoli’) than controls. They argued that this finding possibly suggests that the semantic memory network is disrupted, that is, people with epilepsy might have a smaller network, or it might be that the connections between different semantic concepts are eroded.

Recently, Lenio et al. (2016) also investigated whether lexical difficulties can be better accounted for by the Structure-Loss Model or the Slow-Retrieval Model in adults drawn from a larger study of patient-reported outcome measures in older adults. The evidence for Structure-Loss before the study of Lenio et al. (2016) was based on the number of correct responses produced, and in the case that the number of responses was significantly lower for the clinical group compared to the control group, it was interpreted as evidence for an impaired network of semantic associations. There was no response time data available, however. The evidence for retrieval slowing on the other hand was based on response times but without any available evidence originating from semantic fluency tasks.

In order to tease apart the two hypotheses, Lenio et al. (2016) used a procedure called 'detrending'. Detrended time scores control for the varying speeds of retrieval across individuals, offering an insight into the structure of the semantic memory retrieval process without the confounding effects of retrieval slowing. The researchers defined intra-cluster response times as the time between responses within a cluster (e.g. the response time between the first and the second response of a cluster), and inter-cluster response times as the time between last response of a cluster and the first response of the next cluster. They argued that consistent with the Structure-Loss Model, the semantic deficits would be evident in slower intra-cluster response times (participants will spend a longer time to produce the same number of animals in a cluster due to an impaired network of semantic associations), while inter-cluster response times would be unaffected (participants will spend similar amounts of time switching from one subcategory to another one). Consistent with the Slow-Retrieval Model, the semantic deficits would be evident both in intra- and inter-cluster response times as low performance is accounted for by a general slow speed in retrieving items, and an intact network of semantic associations. They found that by analysing the raw data (i.e. prior to detrending procedure), lexical difficulties were attributable to a breakdown in the associative networks. Specifically, people who showed a low performance spent more time in producing items within a cluster, with this finding being interpreted as evidence for a difficulty to make connections between semantically-related words, or for a difficulty to produce clustering quickly. Those low performers did not show slower inter-cluster times, however. By analysing detrended time scores though, the researchers revealed that those showing a low performance had quicker inter-cluster times but exhausted more quickly as well. Thus, while the raw data appeared to support the Structure-Loss Model, the detrended data showed an effect better explained by the Slow-Retrieval Model.

Moreover, the raw and the detrended inter-cluster times stand in contrast to the notion that impaired semantic fluency is associated with an inability to switch between semantic clusters quickly, as previously measured by switch number. Although it has been suggested that the number of switches in a semantic fluency task could be used as an index of EFs (Bertola et al., 2014), Lenio et al.'s study suggested that by using a detrending procedure in their sample, estimating the number of switches in semantic fluency tasks does not offer full insight into one's recall of items from semantic memory, but that estimating timing in the analysis does. Lenio et al. (2016) concluded that producing fewer responses in semantic fluency tasks is not deterministic to its cause but rather multifactorial as retrieval slowing, structure-loss and impaired EFs all appear to play a role.

1b.2. The Poor Lexical-Semantic Structure Model and the Slow-Retrieval Model

The two models presented above have been proposed on the basis of adult data and have been tested only in adults. There is currently no developmental model accounting for lexical difficulties in children as they are manifested in semantic fluency tasks. Using therefore a model which was initially developed to account for lexical difficulties in adults to account for lexical difficulties in children implies that the model should be adapted developmentally. As such, developmentally, the concept of the loss of semantic structure argued by the Structure-Loss Model should be replaced by the concept of poor semantic structure of the network in dyslexia and DLD, implying that in the course of development, children with dyslexia and DLD have experienced difficulties in establishing lexical-semantic representations. For the purpose of the current study, this model is called the Poor Lexical-Semantic Structure Model. Further, in the current study, the Retrieval-Slowing Model is called instead the Slow-Retrieval Model. In the context of adult samples, 'slowing' means that retrieval speeds were once normal but have slowed down. In developmental disorders, however, the model implies that retrieval speeds were never normal. Poorer semantic fluency performance will therefore be interpreted by using the two models presented above: The Poor Lexical-Semantic Structure Model claiming that poorer semantic fluency can be attributed to children's poor lexical-semantic structure and the Slow-Retrieval Model claiming that poorer semantic fluency performance can be attributed to slow retrieval processes while lexical-semantic representations are intact.

Both models predict that children with DDLD will show lexical retrieval difficulty exemplified in poorer semantic fluency and in the production of fewer clusters compared to TD children. The pattern of retrieval, however, is predicted to be different for the two models. The Poor Lexical-Semantic Structure Model predicts that children with DDLD will cluster their responses around subcategories of a smaller average cluster size than TD children, while the Slow-Retrieval Model predicts a similarly-sized average cluster in the two groups.

The assumption is that in a sophisticated network of semantic connections, as illustrated in Figure 1b.2. for TD children, individuals are able to cluster their responses around subcategories based on the inter-item associations between items whose semantic representations partly overlap. In the 'animals' category, for example, an individual could produce a cluster of 'lion-tiger-giraffe-elephant' since all four items might be connected to each other under the subcategory of 'safari animals' in their semantic network. The richer those semantic connections are – in terms of their number and their strength – the easier it will be for individuals to retrieve items belonging to a subcategory, and therefore the greater the number of items produced belonging to that subcategory (i.e. cluster size) will be. Thus, lexical retrieval is facilitated by the inter-item associations between words, and clustering behaviour supports this argument. Even if, according to the Slow-Retrieval Model in Figure 1b.3., lexical retrieval is slower in children with DDLD compared to TD children, cluster size should not differ between the two groups (although the number of items retrieved overall in the task will differ). According to the Poor Lexical-Semantic Structure Model, however, in an impoverished network of semantic connections, as illustrated in Figure 1b.4., individuals with DDLD are predicted to have fewer and weaker semantic connections between semantically-related items, which will result in clusters of a smaller size being produced and therefore the retrieval of fewer items during the fluency task overall. The size of clusters retrieved during the semantic fluency task is therefore considered to reflect the structure of the semantic lexicon, and the two models make different predictions with respect to cluster size. The legend for Figures 1b.2.-1b.4. attempts to explain more clearly how these predictions arise from the models.

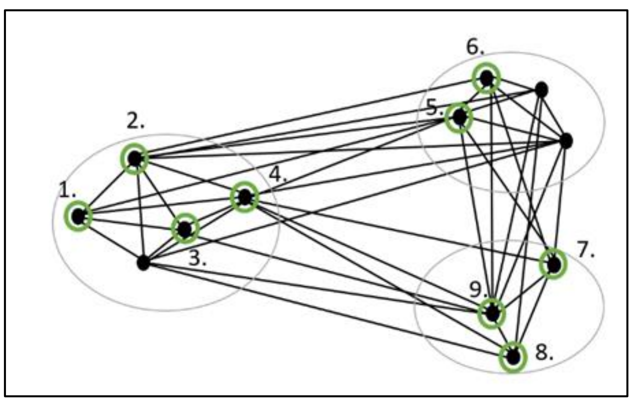


Figure 1b.2. A hypothesised network of semantic connections for TD children

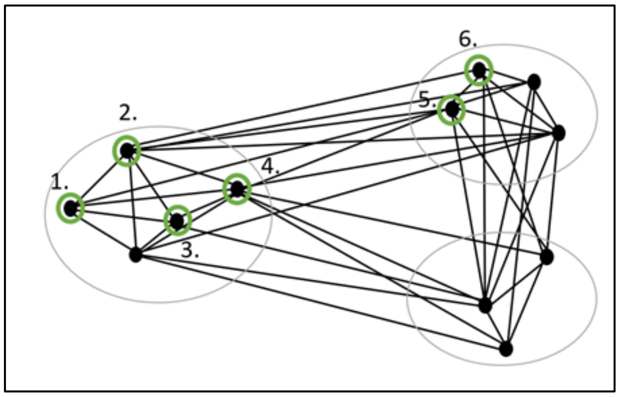


Figure 1b.3. A hypothesised network of semantic connections in children with DDLD according to the Slow-Retrieval Model

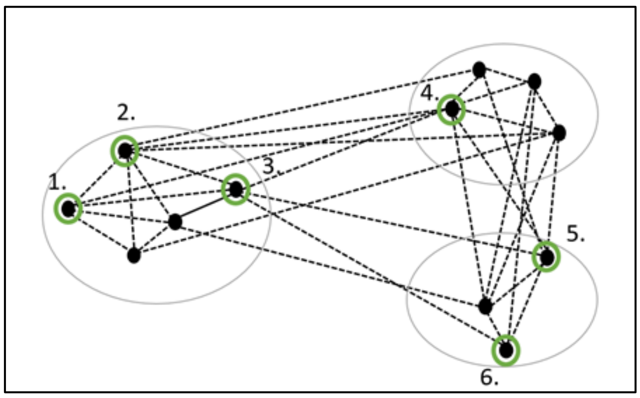


Figure 1b.4. A hypothesised network of semantic connections in children with DDLD according to the Poor Lexical-Semantic Structure Model

Figures 1b.2.-1b.4. Hypothesised semantic networks and lexical retrieval during the semantic fluency task in 1b.2. typical development, 1b.3. children with DDLD according to the Slow-Retrieval Model, and 1b.4. children with DDLD according to the Poor Lexical-Semantic Structure Model. In all Figures, words are the nodes of the network, indicated by black dots. Lines between words indicate the semantic associations between words. Shorter lines indicate that words share more semantic features and are therefore closer together in semantic space. Clustering, as indicated by the grey circles, emerges from this architecture. The lines are dotted rather than solid in Figure 1b.4., to represent the hypothesised weakness of the associations, and there are fewer lines to indicate that there are fewer associations between words. The "direction of travel" through this semantic network during the course of the semantic fluency task is shown by the green numbered circles, which indicate the order in which words are retrieved. Words retrieved from within clusters give rise to the clustering of responses that can be identified in the child's spoken output. Importantly for the predictions of the two models, clustering occurs in both Figure 1b.3. and 1b.4., but the average size of clusters is smaller in Figure 1b.4. ((3+2)/2=2.5) compared to Figures 1b.2. ((4+2+3)/3=3) and 1b.3. ((4+2)/2=3). This is because the poorer semantic structure results in children being less likely to retrieve a word from within the same cluster as the previous word than they would if the semantic structure was developmentally-appropriate: the overall result is the retrieval of fewer items than is typical. Meanwhile, fewer items are retrieved in Figure 1b.3. compared to Figure 1b.2. because retrieval, while following an age-appropriate pattern retrieval, is slower than is typical.

1b.3. The role of semantic skills in reading ability

Nation (2017) suggests that there is a close relationship between reading and semantics in that semantic representations also have an effect on word reading skills as any word has a phonological form, an orthographic form, but also a meaning. According to the *Lexical Legacy Hypothesis* (Nation, 2017), the development of word reading is achieved via the experience of words in diverse and meaningful language environments: because reading experience allows a reader to read words in different semantic contexts, it leads to a rich and nuanced database about a word and its connections to other words. This hypothesis therefore states that word knowledge is based on lexical co-occurrence in the sense that a word is known as it is related in meaning with other words.

Joseph and Nation (2018) tested children aged 10-11 years, and asked whether encountering new words in semantically diverse contexts can help children to acquire new words more than encountering novel words in semantically uniform contexts, as predicted by the Lexical Legacy Hypothesis. Children read sentences containing novel words as their eye movements were monitored. The authors concluded that incidental reading helps children to acquire new words and that semantic diversity of the text does not seem to lead to an improvement in learning in a direct way but rather in an indirect way. The argument is that children might have benefited indirectly from semantic diversity in that they improved their learning of the orthographic form and the meaning of new words using contextual informativeness and their reading comprehension skills. Children who used contextual informativeness and children who had good reading comprehension skills substantially reduced the reading times over the experiment. However, the study failed to find support for the Lexical Legacy Hypothesis (Nation, 2017), as presented above, with Joseph and Nation (2018) claiming that contextual experience in a word's lexical history leads to differences in lexical quality emerging as a consequence of learning.

Evidence from intervention studies also supports the view that semantics plays a significant role in reading ability. Best (2005) in her intervention study of five children with WFDs aged 6 years 10 months-10 years 7 months, reported that children with WFDs can improve their naming skills after a therapy focused on strengthening links from meaning to form. Perfetti (2007) proposed the *Lexical Quality Hypothesis* claiming that variations in the quality of word representations have consequences for reading ability. High lexical quality includes well-specified and partly redundant representations of form (orthography and phonology) and flexible representations of meaning, allowing for rapid and reliable meaning retrieval. Recently, Dyson, Best, Solity, and Hulme (2017), in an intervention study based on the Lexical Quality Hypothesis, showed that children could show benefits in their ability to learn to read after a 4-week programme focusing on improving their ability to pronounce words correctly and also to read and define the words. The researchers provided evidence that it is through access to the meaning of the word after pronouncing the word correctly that improves children's ability to learn to read.

From this overview, it appears that there is a relationship between children's word knowledge and reading ability. In the current study, it will be investigated whether word productivity in semantic and phonological fluency tasks is predicted by children's language and literacy skills. If this is the case, and children's word productivity in semantic and phonological fluency tasks is predicted by language and literacy skills, poorer semantic and phonological fluency performance in children with DDLD will be partly attributed to DDLD children's inferior language and literacy skills. In line with the Lexical Quality Hypothesis (Perfetti, 2007), this implies the impact of well-specified orthographic and phonological representations of words on rapid lexical retrieval in children with DDLD (see also section "4.5. Relation of language and literacy skills with semantic and phonological fluency" in the Results chapter for the argument just presented above and how the argument is linked to the research question).

1b.4. The role of phonological processing skills in reading ability

As presented in the general overview of the thesis, phonological ability includes three interrelated but distinctive phonological processing skills—namely, phonological awareness, phonological STM and RAN (Goswami, 2003). All these different abilities that have been considered in relation to reading ability require phonological processing skills and the involvement of phonological representations. In this section, the three abilities are presented in detail.

1b.4.a. Phonological awareness. In order to become proficient decoders, children must develop an explicit knowledge that spoken words can be segmented into individual components of speech sounds, or phonemes, and combinations of speech sounds, namely, syllables, onsets and rimes; i.e. phonological awareness (Hulme, Snowling, Caravolas, & Carroll, 2005; Ziegler & Goswami, 2005). Importantly, the ability to link graphemes with their corresponding phonemes in an alphabetic script requires phoneme awareness, defined as an explicit awareness that words are made up of phonemes, and this is the most accurate predictor of later reading skills (Melby-Lervåg, Lyster, & Hulme, 2012).

Phonological awareness deficits have also been associated with reading difficulties in DLD (e.g. Bird, Bishop, & Freeman, 1995; Briscoe, Bishop, & Norbury, 2001; Nathan, Stackhouse, Goulandris, & Snowling, 2004). Van Alphen et al. (2004) studied children with DLD, children with a familial risk for dyslexia and age-matched controls (from 36 to 42 months) assessing grammatical morphology, speech perception, phonological processing, and phoneme awareness. The children at risk for dyslexia scored worse than the controls but better than children with DLD, which led the researchers to suggest that children at risk for dyslexia resemble children with DLD, but that they are less severely impaired.

1b.4.b. Phonological short-term memory (STM). A frequently-used measure of phonological STM is the NWR task (Gathercole & Baddeley, 1996), which reflects the ability to encode, store and retrieve novel phonological representations. In this task, children are instructed to attempt to repeat nonwords, such as *favéli*, *munolivoura* or *tirsatabito*, that vary in word length (number of syllables), and/or syllabic structure (Sprenger-Charolles, Colé, Béchennec, & Kipffer-Piquard, 2005), or even in terms of word position and stress of consonant clusters (Marshall & van der Lely, 2009). Phonological STM is considered to be another requisite skill for decoding. In word reading via decoding, children have to blend the phonemic units resulting from the decoding process and store these units in their STM.

However, Melby-Lervåg et al.'s (2012) meta-analytic study reported that although phonological STM was a reliable correlate of individual differences in children's word reading skills, only phoneme awareness was a unique predictor of word reading skills after controlling for the effects of the two other predictors, namely, rhyme awareness and verbal STM; further, data revealed a stronger association between word reading and phoneme awareness than between the former and phonological STM. Nation and Hulme (2011) showed that learning to read at 6 years of age influenced the development of NWR at 7 years, even after controlling for the effects of oral language skills, phonological awareness and earlier NWR, consistent with the notion that spoken and written language are reciprocally related (Hulme & Snowling, 2014); however, NWR was not a longitudinal predictor of reading development. These findings are interpreted on the basis that, as children acquire reading, orthographic information is assimilated by the language system, and this information influences on-line performance when children have to repeat novel words.

Poor NWR performance in children with DLD is well-established in the literature (Gathercole, Willis, Baddeley, & Emslie, 2007). NWR, along with other measures, has been reported to be an excellent behavioural marker of DLD, as it can discriminate with high accuracy children who received a diagnosis of DLD from TD children (Conti-Ramsden, Botting, & Faragher, 2001; Hesketh & Conti-Ramsden, 2013). Kalnak, Peyrard-Janvid, Forssberg, and Sahlén (2014) reported that the task can distinguish well children with DLD from children without DLD, with a large effect size being evident and with 90.2% sensitivity and 97.7% specificity at a cut-off level of -2 *SD*s for binary scoring of nonwords. Effect sizes revealed that the longer the nonwords, the bigger the gap between children with and without DLD. This is a very well-replicated finding (e.g. Barry, Yasin, & Bishop, 2007; Gathercole & Baddeley, 1996; Graf Estes, Evans, & Else-Quest, 2007; Loucas, Baird, Simonoff, & Slonims, 2016). Poor phonological STM skills have also been consistently found in dyslexia (e.g. Catts, Adlof, Hogan, & Weismer, 2005; de Bree, Rispens, & Gerrits, 2007; Gathercole & Baddeley, 1990; Kamhi & Catts, 1986).

1b.4.c. Rapid automatic naming (RAN). RAN tasks, such as picture or colour naming, require the child to name from a printed page familiar visual symbols (e.g. pictures, colours), presented many times on the page, as quickly and as accurately as possible. Georgiou and Parrila (2013) review evidence for an association between the speed with which people perform in RAN tasks and their reading ability, and an association between a slow naming speed and reading difficulties. There are many hypotheses attempting to account for this association. Recently, Georgiou, Ghazyani, and

Parrila (2018) tested university students with dyslexia, and reported that slower times in RAN in the disordered group were accounted for by lexical access impairments, serial processing and articulation, and not by impaired anchoring (namely, the inability of adults with dyslexia to form a perceptual anchor in the RAN task based on repeated items). According to the *Perceptual Anchor Theory* of dyslexia, children with dyslexia show deficits in phonological processing skills not because they have phonological impairments, but because they are unable to form a perceptual anchor in tasks that rely on a small set of repeated stimuli. The anchoring deficit predicts that with a large open set, the performance of people with dyslexia in RAN will not differ from that of controls. Georgiou et al. (2018) did not find greater difficulties in RAN tasks involving a small set of repeated stimuli (five different items repeated 16 times each) than in RAN tasks involving a large set of less frequently repeated stimuli (20 different items repeated only four times each). This study supports therefore the view that the RAN task measures phonological processing skills and challenges the Perceptual Anchor Theory of dyslexia (see also Di Filippo, Zoccolotti, & Ziegler, 2008).

With respect to DLD, Bishop et al. (2009) found that RAN was the strongest predictor of DLD versus dyslexia. Bishop et al. (2009) also reported that performance in the RAN task at 9 years was the strongest predictor to divide the children with DLD into groups on the basis of literacy achievement. The researchers concluded that RAN is not related to oral language ability, but that it is related to reading ability and that good RAN skills protect the child against reading difficulties, even when oral language skills are impaired. Catts (1993), who examined children with speech and/or language problems, reported that they had significantly lower First- and Second-Grade reading scores than TD children. For these children, preschool measures of phonological awareness and RAN turned out to be better predictors for word recognition in Grade 1 and Grade 2 than measures of receptive and expressive language ability. Further, Vandewalle, Boets, Ghesquiere, and Zink (2010) examined longitudinally phonological and early literacy development of 18 Dutch children with DLD from age 5 to age 7 and compared them with their TD peers. They found that RAN skills (and not phonological awareness or phonological STM skills) measured in kindergarten were strongly correlated with reading and spelling performance in both groups, and that only children with DLD who additionally failed on RAN in kindergarten developed reading and spelling difficulties at the end of Grade 1. Their findings are consistent with the findings of previous studies (e.g. Bishop et al., 2009; Brizzolara et al., 2006; Catts, 1993). Thus, the important role of RAN skills in the development of reading is demonstrated both in languages using shallow orthographies, such as Dutch, and in languages using deep orthographies, such as English.

1a.5. Spelling accuracy performance in children with dyslexia and DLD

Accurate spelling requires several years of formal schooling. Children with dyslexia and DLD show difficulty, however, in achieving age-appropriate spelling skills (e.g. for dyslexia: Protopapas et al., 2013; for DLD: Joye et al., 2019). In the light of direct evidence that spelling accuracy is moderated by orthographic consistency (Marinelli, Romani, Burani, & Zoccolotti, 2015), the objective of the current study is to use a spelling task to examine what specific spelling errors in dictation are made by Greek children with DDLD and whether the same errors are made by their TD peers. The ultimate objective, however, which is related to the two prominent phonological hypotheses considered, is to investigate whether the Degraded Phonological Representations Hypothesis or the Deficient Phonological Access Hypothesis (reviewed later below) better characterises the locus of the phonological deficit in Greek children with DDLD. To this end, it will be investigated whether children with DDLD differ from TD children on the number of phonological spelling errors.

In alphabetic writing systems, spelling is based on the relationships between speech sounds (phonemes) and letters/letter combinations (graphemes). It is not surprising therefore that across languages varying in orthographic consistency, spelling performance is predicted by a range of tasks assessing phonological processes, such as phonological awareness, phonological STM, and RAN (e.g. Caravolas et al., 2012; Diamanti et al., 2017; Georgiou, Parrila, & Papadopoulos, 2008; Moll et al., 2014; Ziegler et al., 2010). Greek has a shallow orthography, which means that it is characterised by consistent grapheme-to-phoneme mappings (Seymour et al., 2003). However, grapheme-to-phoneme mappings are more consistent than phoneme-to-grapheme mappings. This indicates that there are alternative ways to spell a single speech sound (for example, phoneme /i/ can be spelled with five graphemes [ι, η, υ, ει, and οι] but all five graphemes spell the phoneme /i/), and this results in phonologically-plausible spelling errors. In fact, the Greek language is estimated to be 95% consistent for reading but only 80% consistent for spelling (Protopapas & Vlahou, 2009). It would therefore seem reasonable that given that there is less consistency in spelling than in reading, spelling is more challenging a task than reading; it hinges to a greater extent on children's phonological processing skills (Cassar, Treiman, Moats, Pollo, & Kessler, 2005), and concomitantly, poor phonological processing skills might show a greater impact on spelling than on reading accuracy (Protopapas et al., 2013). This is particularly so for children with dyslexia and DLD showing an underlying phonological deficit (e.g. Diamanti et al., 2018; Hulme & Snowling, 2009; Ramus et al., 2013; Saksida et al., 2016; Talli et al., 2016; Vellutino, Fletcher, Snowling, & Scanlon, 2004). Less consistent spellings, however, can benefit from an insight into words' morphological and orthographic structure.

An alphabetic writing system not only reflects how graphemes map into phonemes, but also grammatical morphemes represent what part of speech a certain word belongs to. In Greek, for example, /i/ sounds in words such as 'ποτ**ί**ζω' (water), 'μαγ**ι**κός' (magical), and 'δράσ**η**' (action) can be spelled correctly if one is aware that the first word is a verb, the second is an adjective, and the third word is a noun, but can be misspelled if one tries to spell them solely by mapping graphemes onto phonemes, ignoring the part of speech the word belongs to. For example, sounding out the word 'ποτίζω' might yield the following spelling errors: 'ποτήζω', 'ποτύζω', 'ποτείζω', 'ποτοίζω'. Grammatical knowledge is crucial therefore for the correct spelling of inflectional suffixes. In fact, spelling ability is predicted by morphological awareness (Diamanti et al., 2017; Grigorakis &

Manolitsis, 2012; Pittas & Nunes, 2014). In Greek, for example, verbs ending with the morpheme -*izo*- are written with the letter 'ι' (this is a rule, even though there are few exceptions), feminine nouns ending with the sound -*i*- are written with the letter 'η' (again this is a rule, even though there are few exceptions). Thus, knowledge of the inflectional type is required for correct spelling of verb suffixes, nouns, adjectives, pronouns, etc.

Importantly, the unique role of morphological awareness in the development of spelling ability, after controlling for the effect of phonological awareness and RAN, has been recently demonstrated in second Grade Greek spellers (Desrochers, Manolitsis, Gaudreau, & Georgiou, 2017). It has also been reported that morphological richness of a language has an impact on the spelling profile of children with dyslexia. Diamanti et al. (2018) directly compared English and Greek children with dyslexia. They reported that even though Greek has a richer morphology than English, English children with dyslexia were less able to apply morphological knowledge to spell inflectional suffixes correctly compared to Greek children. The researchers concluded that Greek children have more opportunities to learn how to spell suffixes compared to English children. How a word is spelled might not be predicted either by a sounding out strategy or by its morphological structure, but it might reflect its origin or etymology. In Greek, spellings such as *λ* and *κ* for the phoneme /l/ and /k/in words like 'ξεφυλλίζοντας' and 'εκκλησία' reveal the origins in Ancient Greek. Orthographic spelling therefore reflects children's knowledge of word stems.

With respect to the types of spelling errors observed in Greek children with dyslexia, studies have revealed that despite a persistent spelling difficulty, the number of phonological errors is negligible (Nikolopoulos et al., 2003; Niolaki & Masterson, 2013; Porpodas, 1999; Protopapas et al., 2013). For example, Porpodas (1999) reported that the proportion of word spelling accuracy of first Graders suffering poor reading and spelling skills was 25% while the proportion of their nonword spelling phonological accuracy was 88%. However, despite the fact that Greek children with dyslexia make a negligible number of phonological errors, research has shown that they produce significantly more phonologically incorrect words than TD children, with their greater difficulty being observed in the level of the morphological and orthographic structure of words, as reported by Protopapas et al. (2013). Protopapas et al. (2013) followed a systematic, fine-grained, approach to error classification in Greek TD children and children with dyslexia, providing one of the most precise spelling analyses in the dyslexia literature to date. They assessed children of Grades 3-4 and 7 in a spelling-to-dictation task and in a passage spelling task. They reported that both groups made primarily grammatical errors followed by orthographic errors, while phonological errors were negligible. Group comparisons revealed that children with dyslexia produced significantly more phonological, grammatical and orthographic errors than TD children. The researchers concluded that spelling errors of children with dyslexia indicate a persistent difficulty with internalizing regularities of the Greek orthographic lexicon, including derivational, inflectional and word families. Diamanti, Goulandris, Stuart, and Campbell (2014) also found weaknesses in how Greek children with dyslexia applied morphological knowledge to correctly spell word suffixes.

With respect to the types of spelling errors produced by children with DLD, it has been reported that English children with DLD aged 9-10 years showed difficulties with phonological processes and difficulties in applying derivational morphological rules in their spelling compared to age- and language-matched children aged 6-8 years (Critten, Connelly, Dockrell, & Walter, 2014); however, no difference was found in accuracy and error patterns for inflectional morphemes. To date, however, there is no study investigating the types of spelling errors in Greek children with DLD.

What predicts spelling accuracy in dyslexia and DLD is another area of exploration. The impact of phonological and reading skills on the spelling profiles of children with DLD has been confirmed in a recent meta-analytic study which included 984 children with DLD (Joye et al., 2019). In the absence, however, of substantial evidence investigating what processes underlie poor spelling performance in more consistent orthographies than English, with a richer morphology, researchers are not able to identify whether difficulties in nonphonological processing skills may also have an impact on spelling accuracy (Joye et al., 2019). The Greek orthography is ideal towards this investigation given that it is a more consistent and morphologically richer orthography than English.

In sum, in Greek, in order for a word to be spelt accurately in dictation, requires not only phonological processing skills but also grammatical and orthographic processing skills. A few studies have analysed spelling errors in Greek children with dyslexia, but no study has investigated spelling errors in DLD. Our contribution is to fill this gap with the study aiming to analyse spelling errors using a dictation task in a large sample of Greek children with dyslexia and/or DLD, combined in the DDLD group, and next, to examine whether the same errors were found in TD children. Types of spelling errors will inform us about which processes are problematic in Greek children with DDLD and which processes function age-appropriately. In the light of evidence reviewed, phonological errors reflect knowledge of grapheme-to-phoneme mappings, grammatical errors reflect knowledge of inflectional morphology, and orthographic errors reflect knowledge of word stems.

1b.6. The Phonological Theory and the phonological deficit

According to *three-level framework* (Morton & Frith, 1995) presented in Figure 1b.5., reading disability in dyslexia (unlike in general learning disability) implies a specific cognitive deficit (phonological deficit) that underlies a particular pattern of symptoms in dyslexia (e.g. difficulties in reading and spelling words and decoding nonwords). In addition, the model implies complex causal links that link this cognitive deficit on the one hand to behaviour and on the other hand to the brain and that account for dyslexia as a disorder with neurocognitive origin (Frith, 1999). In other words, it is an underlying cognitive cause centred on phonology that renders dyslexia a *specific* disorder (Frith, 1992; Frith & Happé, 1998). Environment is a transversal concept because it may interact with other factors at either the biological, cognitive or behavioural level. This model acknowledges that neurodevelopmental disorders are dynamic such that the behavioural manifestations of dyslexia change as a function of different contexts (e.g. the orthography of the language).

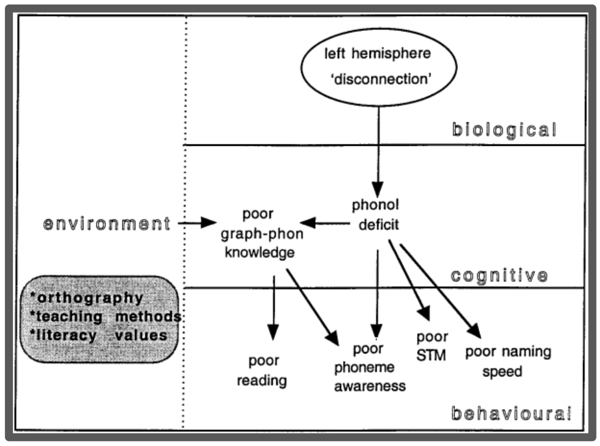


Figure 1b.5. A causal model of dyslexia as a disorder originating from a phonological deficit (from Frith, 1999, p. 203)

As the initial stages of reading development are characterised by learning how graphemes (novel orthographic codes) map onto their corresponding sounds (pre-existing phonological codes), referred to as *phonological decoding*, it is not surprising that the consensus view for many years has been that dyslexia is the behavioural outcome of an underlying phonological deficit. This view, which has received substantial empirical support from a range of experimental studies (e.g. Gabrieli, 2009; Grigorenko, 2001; Hulme & Snowling, 2009; Ramus, 2004a, 2013; Ramus et al., 2003; Ramus & Altarelli, 2014; Rice & Brooks, 2004; Snowling, 2000; Snowling & Hulme, 2008; Varvara, Varuzza, Sorrentino, Vicari, & Menghini, 2014; Vellutino et al., 2004; White et al., 2006a), maintains that the phonological deficit in dyslexia lies in the ability to represent or recall speech sounds, or phonological representations, with this resulting in the condition described above (e.g. Ramus, 2001, 2010).

The consensus is that the phonological deficit in dyslexia across all languages so far studied makes itself manifest in three main dimensions that rely on the efficient functioning of the phonological system. Children with dyslexia demonstrate difficulties in three areas of phonological processing skills, which are related to phonological decoding—namely, in those tasks that require the manipulation of phonological representations at the different 'grain' sizes of syllable, onset and phoneme (such as deletion tasks), the retention of verbal material in STM (such as nonwords in NWR tasks), and access to phonological representations efficiently and quickly (such as high frequency word forms in RAN tasks) (Snowling, 2000). Longitudinal studies show that performance on these tasks predicts individual differences in reading skills in both TD children and children with dyslexia (Boets et al., 2010).

Ramus and his collaborators (2003), in their seminal work, tested 16 English adult people with dyslexia and 16 controls, using a 10-hour battery of psychometric, phonological, auditory, visual and motor skills to investigate whether any single deficit, that is, phonological, auditory, visual or motor, can account for dyslexia. They reported that all 16 people with dyslexia had a phonological deficit, 10 had an auditory deficit, 4 had a motor defect and 2 had a visual defect. Interestingly, 5 people with dyslexia had a phonological deficit in the absence of any sensorimotor deficit. Ramus et al. (2003) concluded that the phonological deficit is, in their own words, “a sufficient cause” of dyslexia (p. 861), even though not a necessary one in the sense that there might be other causes that can lead to the genesis of dyslexia.

Similar findings have been obtained with children with dyslexia (e.g. Kronbichler, Hutzler, & Wimmer, 2002; White et al., 2006b). Recently, Saksida et al. (2016) assessed a group of 164 children with dyslexia and 118 TD children aged 8-13 years to investigate whether a phonological deficit, a visual stress deficit, or a visual attention span deficit better explains dyslexia. They reported that most participating children with dyslexia had a phonological deficit (up to 92.1% of the clinical group) as measured by the level of children’s accuracy, speed, or both on phonological tasks, and that all of the children who had a visual attention span deficit (28.1% of the clinical group) also had a co-occurring phonological deficit. In addition, a large amount of variance in literacy skills was predicted by phonological variables, while visual variables did not explain any additional variance in literacy skills. They also reported that children who had both phonological and visual deficits were no worse in reading skills compared to children who had only a phonological deficit, with this finding supporting the hypothesis of the phonological deficit in dyslexia. Children with a visual stress deficit were estimated to be 5.5% of the clinical group and 8.5% of the control group. This study showed that there are only few children with dyslexia who do not have a phonological deficit and also that there are children with dyslexia whose reading problems can be accounted for by visual hypotheses of dyslexia as they display visual stress or visual attentional deficits.

Overall, evidence originating from such studies focusing not only on the phonological domain but also on the auditory, visual and motor domains within the same individuals and using a broad test battery of tasks in each modality, clearly places the proximal cause of dyslexia within the phonological language system, at least for the majority of people with dyslexia. The review so far supports the view that the Phonological Theory advocates a 'modular' deficit within the language system: it predicts normal nonverbal ability as well as verbal ability and semantic functioning.

1b.7. The Degraded Phonological Representations Hypothesis

While the supporters of the Phonological Theory agree on the central and causal role of phonology in dyslexia, they disagree on the precise locus of the phonological deficit (Ramus, 2003; Ramus & Szenkovits, 2008; Snowling, 2000). This thesis aims to test and contrast the two prominent phonological hypotheses that have been claimed to explain the underlying cause of the phonological deficit in dyslexia—namely, the Degraded Phonological Representations Hypothesis and the Deficient Phonological Access Hypothesis.

Prior to presenting these hypotheses, however, a reasonable question is the following: Why is it that children with dyslexia and DLD show poor performance in phonological tasks and how is poor phonological ability linked to reading difficulties? Phonemes are distinct cognitive categories imposed by our phonological system upon a gradient acoustic space (Ladefoged, 2001). For example, in the case of the words 'bat' and 'pat', voice onset time is a gradient cue signalling the difference between voiced /b/ and voiceless /p/ in English. In order for spoken word recognition (e.g. the recognition of the word 'bat') to proceed successfully, phonological representations must be robust (i.e. all /b/ sounds must be assigned to the same phoneme category) and distinct (i.e. /b/ sounds must be distinguished from /p/ sounds). However, whilst phonetic forms can be identified at a level of perception, recognition of phonological units (e.g. phonemes) involves additional cognitive processes such as categorization (Ladefoged, 2001). It follows that phonological representations are a way of storing the sound sequences that make up words in an abstracted form. Further, it is generally assumed that their internal structure corresponds to the hierarchy of units developed in Phonological Theory: word, syllable, onset-rime, and phoneme (Harris, 1994). Conscious knowledge of this hierarchical internal structure is known as phonological awareness. It follows that lack of specification at any of these levels during the execution of a phonological awareness task could adversely affect task performance. Moreover, as the initial stages of reading development are characterized by learning how graphemes (i.e. letters and group of letters) map onto their corresponding sounds, it is not surprising that the consensus view for many years has been that dyslexia is the behavioural outcome of an underlying phonological deficit.

The most commonly accepted hypothesis regarding the phonological deficit in dyslexia is that the phonological representations are somehow degraded (i.e. less robust, less distinct), and that this primary representational deficit impacts upon higher-level phonological processing skills and reading development. This view is termed the Degraded Phonological Representations Hypothesis, that is, dyslexia is associated with impaired neural representation of the sound structure of words (Goswami, 2000; Snowling, 2000; Ziegler & Goswami, 2005). One of the first steps when the child learns to read is to learn the mappings between graphemes and phonemes. It seems logical, therefore, that if a child has a difficulty understanding what units speech is made of and difficulty paying attention to those units and remembering them, it will be more difficult for them to learn the associations between graphemes and phonemes. This is the beginning of an explanation of reading disability through the phonological deficit.

The concept of degraded phonological representations in dyslexia implies that in the course of development, children with dyslexia have experienced difficulties in establishing representations of phonemes (such as the /b/ in <bat>) adequate for the recognition and production of words. In addition, it is assumed that these difficulties are subtle, as they become obvious only when the affected child has to use their phonological representations for tasks such as reflecting on the internal sound structure of words or in acquiring the alphabetic principle. The Degraded Phonological Representations Hypothesis proposes that children with dyslexia have a difficulty in establishing, storing and consequently retrieving phonological representations to the same degree of detail as their TD peers (Goswami, 2000; Ramus et al., 2003; Snowling, 2000; Ziegler & Goswami, 2005). The resulting phonological representations have been variously referred to as 'fuzzy', 'indistinct', or 'underspecified' and this lack of distinctness or segmental specificity has been held responsible for the range of difficulties with phonological processing skills observed on the aforementioned phonological tasks.

There is convincing evidence for this hypothesis from longitudinal studies indicating that people with dyslexia do show early phonological impairments, even before going to school or before receiving any reading instruction (e.g. Boets et al., 2010; Gallagher, Frith, & Snowling, 2000; Lundberg, 2002; Lyytinen et al., 2004; Snowling, Gallagher, & Frith, 2003). This is particularly important given that competence in reading also influences phonological abilities, thus possibly questioning the causal relationship between these two in dyslexia.

Boada and Pennington (2006) tested the segmentation hypothesis. This hypothesis proposes that poorly specified phonological representations themselves that compromise reading acquisition is the primary cause of reading difficulties in dyslexia (Snowling, 2000; Snowling & Hulme, 1994). Boada and Pennington used tasks that measure implicit phonological representations, such as lexical gating, priming and syllable similarity, in children aged 11-13 years identified as reading disordered to test whether implicit phonological representations are impaired in dyslexia. The assumption of using implicit phonological tasks is that these tasks invoke only a minimal level of metalinguistic ability. They reported that reading disordered children made more syllable than phoneme structure confusion errors in the syllable similarity task, and this finding was interpreted as evidence for less specified (not phonemically organized) phonological representations in children with dyslexia. The same interpretation of data comes from Elbro, Borstrøm, and Petersen (1998) who argued that children with dyslexia have less-distinct phonological representations.

One hypothesis proposed by Goswami et al. (2002) is that phonological difficulties, and the associated reading difficulties, originate from an underlying difficulty with supra-segmental (or prosodic) attributes of the speech stream, and in particular with speech prosody. Goswami et al. (2002) showed that speech intelligibility was affected by fluctuations in the energy of the speech soundwave produced as people speak, called amplitude modulations. Children with dyslexia were less sensitive to slow changes in amplitude in comparison to typically reading controls, and in particular they were less sensitive to changes in amplitude rise time defined as the amplitude envelope onset (children with dyslexia aged 11 needed 100 ms to perceive a change compared with about 50 ms that controls aged 9 needed). Behaviourally, rise time is closely associated with the perceptual experience of speech rhythm and stress. Goswami et al. (2002) reported that individual differences in sensitivity to the shape of amplitude modulation accounted for 25% of the variance in reading and spelling acquisition even after controlling for individual differences in age, NVIQ and vocabulary in children with dyslexia. In a related study, adults with dyslexia performed poorly in stress discrimination tasks, indicating a persistence of stress perception deficits throughout the life span (Leong et al., 2011).

1b.8. The Deficient Phonological Access Hypothesis

The fact that phonological representations are degraded in people with dyslexia has mostly been inferred indirectly on the basis of behavioural evidence. Indeed, despite the widespread acceptance of the Degraded Phonological Representations Hypothesis, Ramus and Szenkovits (2008) argued in an influential review paper that the data from individuals with dyslexia are actually inconsistent with that hypothesis. They instead proposed an alternative hypothesis arguing that the phonological representations of people with dyslexia are perfectly intact, but hard to access. Ramus and Szenkovits (2008) reviewed the dyslexia literature aiming to define what precisely underlies the phonological deficit, and argued that the phonological deficit is evident only under certain task demands; for example, phonological tasks that yield reliably poor performance of people with dyslexia either require explicit manipulation of speech sounds (phonological awareness tasks), or load STM (phonological STM tasks), or require speeded access (lexical retrieval tasks).

This finding of less accessible phonological representations in dyslexia has since been replicated by studies showing that the poor specification of phonological representations has the most detrimental impact when phonological representations are employed in demanding phonological tasks that involve some additional cognitive processes in order for successful completion from perception and STM through to working memory and even EFs that influence task performance (Ramus & Ahissar, 2012; Soroli, Szenkovits, & Ramus, 2010; Szenkovits & Ramus, 2005; Szenkovits, Darma, Darcy, & Ramus, 2016).

Specifically, it is certainly the case that certain phonological awareness tasks, such as phoneme deletion tasks, require greater involvement of additional cognitive processes than other tasks, such as rime judgment tasks. The former task is harder than the latter task, and one of the reasons is that phoneme deletion tasks require greater involvement of cognitive processes, as manipulation of a specified phonological unit and not just judgment about the phonological structure of words is needed. As such, although classic tasks of phonological processing skills might claim to measure whether children can manipulate, retrieve, or access speech sounds, in reality, these tasks are not process pure. Ramus et al. (2010) acknowledge this fact and argue that:

"indeed, all linguistic data are behavioural data. Linguistic representations are hidden in the brain, and can never be assessed directly by the experimenter … behavioural data are always collected using a task. All tasks involve multiple levels of representation and processing" (p. 332).

This might be better illustrated in the following example, indicating that measuring phonological processes is complex. In a classic phonological awareness task, the phoneme deletion task, a child might be instructed to delete /s/ from the nonword 'ston' and pronounce the resulting word. In order to correctly respond to this task the child must first:

- detect the nonword 'ston' and the phoneme /s/ to delete
- maintain both the nonword and the phoneme in working memory
- segment the nonword 'ston' in order to separate the /s/ from the other sounds in the nonword, and then delete the /s/
- blend the remaining sound units /t/ /o/ /n/
- produce a verbal response of the nonword *ton*.

The argument therefore is as follows. Whenever high demands are placed upon phonological representations, the phonological deficit observed in people with dyslexia is due to poorly specified phonological representations themselves, and concomitantly, because more phonological material needs to be specified, or due to a greater involvement of processes that require phonological access or load phonological STM and that they are deficient in dyslexia. Data that would allow adjudication between the two views remain scarce, however (Ramus & Ahissar, 2012). This is because it is almost impossible in practice to design experimental tasks tapping phonological representations without involving access to these representations and vice versa.

The next paragraphs outline key studies at the cognitive and at the biological level of explanation, which have been described as reflecting deficient phonological access deficits in children and adults with dyslexia. Szenkovits and Ramus (2005) reported that French adult people with dyslexia had a substantial impairment in the input pathway of lexical and sub-lexical phonological representations and that they failed to discriminate and repeat verbal material whenever the STM capacity was substantial. With these findings in mind, Soroli et al. (2010) tested the perception and production of foreign (i.e. non-native) phonological (speech) contrasts by French adult people with dyslexia and controls in different conditions varying in the load of STM. In order to test the contribution of STM capacity, the researchers conducted auditory discrimination (measuring perception) and repetition (measuring perception and production). The results showed that there were no group differences when participants were asked to discriminate or repeat single nonwords. Group differences were evident when participants were asked to discriminate or repeat sequences of two or three nonwords, and particularly so for the lexical stress. Differences between people with dyslexia and controls were only apparent when STM capacity increased. Nonetheless, one possibility for the affected individuals' intact performance on phonological tasks used to minimise effects of task demands might be that such tasks are not able to identify the affected people's phonological difficulties, difficulties that otherwise exist. Another possibility is of course that such tasks can reliably distinguish the affected-non-affected differences whenever these exist.

Ramus et al. (2013) tested a group of sixty-five 8-12-year-old English children with dyslexia and/or DLD using a wide battery of phonological and language tasks. They provided, by using factor analysis, evidence for a possible distinction between explicit phonology, measured by phonological tasks that really require phonological processing skills, and implicit phonology, measured by tasks that demand phonological representations themselves. Ramus and collaborators argue that the phonological deficit in dyslexia may not lie in deficient phonological representations themselves—as only a subset (one-third) of the children with dyslexia had a deficit in phonological representations—but rather in some cognitive skills involved in complex phonological tasks that also involve explicit skills. Ramus et al.'s factor analytic study, using more sophisticated statistics, a larger group of participants, and a larger test battery, offers the most convincing empirical evidence at the cognitive level of explanation for a deficient phonological access in children with dyslexia and not for degraded phonological representations themselves, so far.

In sum, it is well-established that children with DDLD perform poorly on phonological tasks compared to TD children. However, there has been some debate as to whether their phonological deficit arises directly from an impairment in phonological representations, or instead from deficient access to (intact) phonological representations. The current study tested the Degraded Phonological Representations Hypothesis and the Deficient Phonological Access Hypothesis using three phonological tasks, namely, phoneme deletion, NWR and RAN tasks. Using phoneme deletion tasks, the Degraded Phonological Representations Hypothesis predicts that the TD group will outperform the DDLD group in accuracy and speed performance in all three phoneme deletion tasks. This is explained by impaired phonological representations in the DDLD group. In contrast, the Deficient Phonological Access Hypothesis predicts that accuracy and speed performance in the phoneme deletion task of monosyllable items with simple CVC syllable structure will be equivalent for the two groups. This is explained by the fact that short nonwords with a simple syllable structure do not load children's phonological STM. The Deficient Phonological Access Hypothesis also predicts that accuracy and speed performance in the phoneme deletion task of trisyllable items with simple CVCVCV syllable structure and in the phoneme deletion task of monosyllable items with complex CCV syllable structure will be poorer for the DDLD group than the TD group. This is explained by the fact that long nonwords with three syllables and nonwords with complex syllable structure load phonological STM. It should be noted that phonological STM capacity is not actually measured, however, a NWR measure consisting of 3-, 4-, 5-, and 6-syllable nonwords will be used to investigate whether the DDLD group shows phonological STM deficits relative to the TD group.

Using the RAN task, the Degraded Phonological Representations Hypothesis predicts that the DDLD group will make phonological errors in their picture naming. Phonologically inaccurate performance is explained by inaccurate phonological representations. The Deficient Phonological

Access Hypothesis, however, predicts that the DDLD group will name pictures significantly slower than the TD group but that the two groups will not differ on phonological accuracy. Slower naming performance is explained by the fact that the phonological access deficit in the DDLD group renders performance on tasks requiring speeded access to phonological representations particularly slow. Phonologically accurate performance is explained by intact access to phonological representations in the RAN task not requiring metalinguistic manipulation.

Chapter 2. Fluency tasks

What are fluency tasks and what do they measure? The aim of this chapter is to present the two main experimental tasks of the current study, namely, semantic and phonological fluency tasks, in addition to a design fluency task. Semantic and phonological fluency tasks have been used by other researchers to investigate lexical organisation and lexical retrieval processes. Design fluency tasks have been used by other researchers to investigate visuospatial skills. In this chapter therefore, relevant studies which used the three fluency tasks are presented.

2.1. Semantic fluency

Semantic fluency tasks require children to produce as many words as they can belonging to certain categories, such as 'animals', in a 60 s period. The task therefore requires the search of the mental lexicon for words on the basis of their meaning, offering an important window into lexical organization (i.e. how words are stored) and lexical retrieval processes (i.e. how words are accessed).

2.1.1. Semantic switching, clustering and cluster size

What drives productivity in semantic fluency tasks? Although the most common performance measures in verbal fluency categories are the number of total and correct responses, other analyses such as the number and size of clusters and the number of switches between clusters can be carried out in order to investigate what drives verbal fluency performance. Troyer, Moscovitch, and Winocur (1997) introduced clustering and switching as two components of verbal fluency performance. Words are often produced in clusters of semantically-related words. For example, "cat-dog" is a cluster of 'pets'. Semantic clusters provide a measure of how words are stored in the mental lexicon (lexical organization) on the basis that semantic similarity (or overlap) in successively produced responses might aid word retrieval. In the example given above, the retrieval of *cat* might facilitate the retrieval of *dog* because their semantic representations partly overlap: both animals are pets, for example. Semantic clusters are classified as either conventional or contextual (or thematic/schematic) subcategories. ‘Fruits’ and ‘vegetables’ are two examples of conventional subcategories within the superordinate category of foods. In contextual (or thematic/schematic) clusters, items are clustered based around the context (or theme/schema) in which they typically occur—for example, foods grouped according to the meal at which they are usually eaten (e.g. breakfast foods—milk, cereal, bread, butter). In the current study, both contextual and conventional subcategories were used. Once a semantic subcategory is “exhausted”, people switch to another subcategory (e.g. from ‘pets’ to ‘fish’). It is not that people always use up a category completely - there are likely to be items still to be retrieved. There might be items hard to retrieve, however, and as people know that the task is timed, they switch to another item, from another subcategory, which is more easily retrieved.

As such, good performers search mentally for subcategories, and then produce words within an identified subcategory as described above (i.e. clustering process). Once a subcategory is exhausted, it is most efficient to quickly move to another subcategory or cluster. This process is referred to as switching. It has been reported that both of these processes are a product of strategic searching and cognitive flexibility. Troyer et al. (1997) argued that switching involves EFs to a greater extent than clustering. EFs are defined as higher order cognitive processes that encompass skills necessary for purposeful, goal-directed behaviour and are essential to the ability to respond to novel and unfamiliar situations. According to Troyer (2000), switching requires strategic search, conscious control and cognitive flexibility to shift between representational sets.

2.2. Phonological fluency

Phonological fluency tasks are explicit word-retrieval tasks requiring children to produce as many words as they can beginning with particular letters in a 60 s period. The task therefore requires the search of the mental lexicon for words on the basis of their phonology, offering, like semantic fluency tasks, an important window into lexical organization and lexical retrieval processes.

2.2.1. Phonological switching, clustering and cluster size

Producing words starting with particular letters would suggest that one has representations of those words in which an initial phoneme is distinct, or segmented, from the rest of the word form. Phonological fluency tasks are therefore important for tapping children's conscious, or explicit, access to phonological representations. Words are often produced in clusters of phonologically-related words. For example, "***fl***ag-***fl***ower" is a phonological cluster since the two words share the initial two phonemes ('***fl***'). This phonological clustering provides a more implicit measure of the quality of children's phonological representations on the basis that it is the phonological similarity (or overlap) in successively produced responses that might aid word retrieval. In the example given above, the retrieval of *flag* might facilitate the retrieval of *flower* because their phonological representations partly overlap. Theoretically, phonological clustering at the word onset can be explained by the Cohort Model (Marslen-Wilson, 1984). According to the Cohort Model, an initial phoneme of a word (e.g. 'f') is used to activate the set of all words in the lexicon that have the same initial phoneme (e.g. fun, flag, flower). This set of words is called a 'cohort' in the model. As more words are retrieved over the test period, words can be eliminated from the cohort due to new phonological information added. For example, the word 'flag' might activate the word 'flower' since both words share the initial two phonemes. In phonological fluency tasks thereby, this results in phonological clustering. Given the limited time of the test period, once lexical retrieval within a cluster slows down, individuals tend to switch to another cluster (e.g. from "flag-flower" to "freefriend"). Switching allows them a more rapid retrieval of lexical items from the mental lexicon. Both clustering and switching strategies show a strong positive correlation with the number of correct items retrieved in phonological fluency tasks (e.g. Kosmidis, Vlahou, Panagiotaki, & Kiosseoglou, 2004). Overall, successful performance on phonological fluency tasks requires the search of the mental lexicon for words on the basis of their phonology. Importantly, the phonological fluency task measures two different aspects of access to phonological representations, namely explicit access to phonological representations, as evidenced by the number of correct responses retrieved, and implicit access to phonological representations, as evidenced by the size of clusters produced.

As presented in the previous paragraph, producing words starting with particular letters would suggest that one has distinct or segmented phonological representations. Indeed, John, Rajashekhar, and Guddattu (2016) assessed a number of 1,015 Indian children aged 5-15 years on phonological fluency tasks. The researchers reported which were the most often produced phonological subcategories. Successively retrieved words which shared the first two letters or the last letters were the two most common types of subcategories observed. They also reported that third Graders produced the following subcategories: words sharing the first two syllables, words sharing the first two consonants with differing vowels, words differing only by a vowel or consonant sound, words having the same number of syllables and words differing only in the last part. Likewise, Nash and Snowling (2008) found that the most common types of phonological subcategories in English children were the following: words that shared the same beginning consonant and vowel (e.g. tea-teacher), those differing only with a vowel (e.g. tail-tall; ball-bill), words sharing a consonant and a vowel with an intervening consonant (e.g. tea-tree) and words that shared a syllable (e.g. tooth-toothpaste; book-bookcase).

Both hypotheses of the phonological deficit in children with DDLD predict that children with DDLD will retrieve fewer items than TD peers in the phonological fluency task. However, while the Degraded Phonological Representations Hypothesis predicts smaller clusters of phonologically-related items in children with DDLD, the Deficient Phonological Access Hypothesis predicts that the two groups will not differ in cluster size. It is assumed that TD children have age-appropriate explicit and implicit access to phonological representations. In this context, the more robust and distinct children's phonological representations are, the easier it will be for them to retrieve items belonging to a phonological category. In the context of typical phonological representations, the easier it will also be for them to produce a phonological cluster, such as '***st***ar-***st***are-***st***reet-***st***rong', since all four items share the initial two phonemes, and therefore the greater the number of items produced belonging to that subcategory (i.e. the greater the cluster size) will be. Even if, according to the Deficient Phonological Access Hypothesis, a phonological access deficit is evident in children with DDLD, cluster size should not differ between the two groups. According to the Degraded Phonological Representations Hypothesis, however, if phonological representations are less robust and distinct, individuals with DDLD are predicted to have difficulty in retrieving words in clusters, which will result in smaller clusters being produced. The size of clusters is therefore considered to be an implicit, and therefore more direct, measure of access to phonological representations, and the two hypotheses make different predictions with respect to cluster size.

2.3. Semantic and phonological fluency tasks: Same or different?

Do semantic and phonological fluency tasks measure the same or different cognitive processes? Word productivity in semantic categories is reliably greater than word productivity in phonological categories (e.g. Arán-Filippetti & Allegri, 2011; Hazin et al., 2016; Hurks et al., 2006; Kosmidis et al., 2004; Marshall, Rowley, & Atkinson, 2014). Nevertheless, word productivity in both tasks is moderately or strongly correlated (in Ardila, Ostrosky-Solis, & Bernal, 2006; Matute, Rosselli, Ardila, & Morales 2004). Unsworth, Spillers, and Brewer (2011) found using exploratory factor analysis that productivity in semantic and phonological fluency tasks was accounted for by a single factor, suggesting that the two tasks measure to a great extent the same processes.

In an adult study, Vonberg, Ehlen, Fromm, and Klostermann (2014) asked 42 participants to produce as many words beginning with the letter *S* as possible in two minutes, aiming to investigate whether word productivity in phonological fluency tasks is performed by pursuing semantic or phonological search strategies. They defined clusters temporally rather than on the basis of the semantics or phonology of the responses. They argued that even in a condition where participants were asked to search words in the lexicon by using a phonological strategy, content and sound-related information interacted with each other, and therefore that the different aspects of lexical information cannot be retrieved independently from each other. The argument is that a semantic search is the default search strategy given how the lexicon is organised and that participants cannot scan the mental lexicon only under the premise of phonological word features, leaving semantic information aside. However, inspection of the results reveals very low semantic and phonological relatedness between items in their temporally-defined clusters. This suggests therefore that the findings from Vonberg et al.'s study might be better interpreted as evidence showing that even if semantics is involved in retrieving items in phonological conditions, this does not imply that phonology is not involved. In another adult study, Woods, Wyma, Herron, and Yund (2016a) also revealed that the sequence of words retrieved in phonological conditions was influenced by semantics.

Likewise, in a child study, John et al. (2016) reported some overlap in the processes involved in the two conditions. They argued that the phonological condition predominantly taps grapho-phoneme processes, even though semantic processes are involved too. They based this argument on the finding that in the phonological condition, task-discrepant clustering was evident in less than 5% of the children in higher Grades. Task-discrepant clusters were defined as semantic clusters produced in the phonological condition, that is, words which shared phonological characteristics, but which were also related in meaning.

Evidence originating from clinical populations also offered insight into whether semantic and phonological fluency tasks measure the same or different cognitive processes. Marshall et al. (2014) reported that phonological fluency is particularly hard in deaf signers using British Sign Language (BSL). The researchers argued that since there is no orthography for sign languages, signers show reduced phonological awareness and fewer opportunities to engage in metaphonological activities. As such, it seems that phonological fluency tasks involve phonological processing skills to a greater extent than semantic fluency tasks.

Smith-Spark, Henry, Messer, and Zięcik (2017) offered another interpretation for poorer phonological than semantic fluency performance in adults with dyslexia. They argued that worse phonological than semantic fluency performance could be attributed to the fact that phonological fluency tasks place higher demands upon EFs than semantic fluency tasks, and therefore not difficulties with phonological processing skills but increased EF demands might result in poorer phonological than semantic fluency performance. Their claim that phonological fluency placed higher demands upon EFs than semantic fluency was supported, according to the researchers, by the finding that short-form IQ score was a stronger predictor of phonological than semantic fluency performance. This finding was consistent with the view that phonological fluency tasks reflect cognitive complexity to a greater extent than semantic fluency tasks, as was previously reported by Ardila et al. (2006).

In sum, some researchers found evidence that even when participants are asked to retrieve lexical items in phonological categories, semantic influences are evident in that content and sound-related information interact with each other. These recent findings contradict the view for a componential organization of lexical knowledge, namely, that the semantic component is independent of phonological and orthographic form knowledge (Shelton & Caramazza, 1999). However, the fact that semantic processes are involved in phonological fluency tasks does not imply that phonological processes are not crucial for retrieving items in phonological fluency conditions.

Indeed, researchers have found that phonological fluency tasks predominantly tap grapho-phoneme processes and that performance in those tasks is particularly poor for those who show reduced phonological awareness and fewer opportunities to engage in metaphonological activities. It has also been reported that phonological fluency tasks involve EFs to a greater extent than semantic fluency tasks.

2.4. Effects of switching and clustering behaviour on semantic and phonological fluency performance

Both switching and clustering behaviour strongly associate with word productivity in semantic and phonological categories (e.g. Arán-Filippetti & Allegri, 2011; Kosmidis et al., 2004). With respect to the semantic condition, Kosmidis et al. (2004), in a study with healthy Greek adults, showed that the total number of responses correlated weakly with cluster size and strongly with the number of switches. This indicates that as the size of clusters and the number of switches increased, so did the total number of responses. Arán-Filippetti and Allegri's (2011) child study reported that the number of clusters explained 52% of the variance, the number of switches explained an additional 18% and cluster size explained an additional 16%. In another child study, Resch, Martens and Hurks (2014) tested 225 Dutch 4-6-year-old children using the semantic category of animals. The researchers found that semantic fluency correlated with the number of switches, the number of clusters and cluster size.

With respect to the phonological condition, Kosmidis et al.'s (2004) study with Greek adults showed that the total number of responses correlated moderately with cluster size and strongly with the number of switches. Arán-Filippetti and Allegri's (2011) child study reported that the number of switches explained 84% of the variance, the number of clusters explained an additional 10% and cluster size explained an additional 6%. Hence, in adult and child studies alike, productivity in semantic and phonological fluency tasks is driven mainly by switching the number of clusters retrieved, and to a lesser extent by cluster size.

2.5. Effects of age, level of intelligence, and gender on semantic and phonological fluency performance

It has been reported that semantic and phonological fluency performance is influenced by age and level of intelligence, with older children and those with higher scores on tests of intelligence performing better in both types of verbal fluency task. Mixed findings have been reported for the effect of gender on verbal fluency performance.

2.5.1. Evidence from TD children and healthy adults

2.5.1.a. The effect of age. With respect to semantic and phonological conditions, previous studies have generally shown that the number of correct items increases significantly as a function of age (e.g. Chami et al., 2018; Cohen, Morgan, Vaughn, Riccio, & Hall, 1999; Klenberg, Korkman, & Lahti-Nuuttila, 2001; Korkman, Kemp, & Kirk, 2001; Nieto, Galtier, Barroso, & Espinosa, 2008; Pastor-Cerezuela, Fernández-Andrés, Feo-Álvarez, & González-Sala, 2016; Resch et al., 2014; Sauzéon, Lestage, Raboutet, N'Kaoua, & Claverie, 2004). For example, Hurks et al. (2010) assessed a large sample of Dutch children from first to ninth Grade, and found significant grade-related differences in semantic fluency performance until Grade 7.

2.5.1.b. The effect of the level of intelligence. Regard, Strauss, and Knapp (1982) showed that aside from age, level of intelligence had a moderate effect on verbal fluency performance in TD children from Grades 5 to 7, as indexed by WISC Vocabulary and Block Design subtasks. Weak correlations have been reported between semantic fluency and verbal IQ and full-scale IQ (in Ardila et al., 2006). Resch *et al.* (2014) also reported a moderate association between Raven's Coloured Progressive Matrices (CPM; Raven, 2008) and semantic fluency in 4-6-year-old children.

2.5.1.c. The effect of gender. Variable effects of gender on verbal fluency performance have been reported: many studies have failed to find significant gender differences (e.g. Arán-Filippetti & Allegri, 2011; Barry, Bates, & Labouvie, 2008; Hazin et al., 2016; Hurks et al., 2006; 2010), while others have reported a main effect of gender, with girls outperforming boys (Klenberg et al., 2001), or with boys outperforming girls using, for example, the category of 'brands of cars' (Zarino, Crespi, Launi, & Casarotti, 2014). Gender differences might be attributed to the fact that males and females are more familiar with certain semantic categories. It appears that males produce more responses using the categories of 'cars' and 'tools' and females produce more responses using the category of 'fruits', while no gender differences have been found for the category of 'animals' (in Woods et al., 2016a).

2.6. Automatic versus controlled processing in verbal fluency tasks

Response output rate tends to decline over the 60 s test period, especially after the initial 15 s have elapsed (e.g. Crowe, 1998; Henry, Messer, & Nash, 2015; Hurks et al., 2006; Smith-Spark et al., 2017). Crowe (1998) showed that in semantic and phonological fluency categories, adults produced more and more frequent words in the first 15 s, and as the time period on task increased, word productivity and word frequency decreased. The author concluded that words accessed early on in the task are words of high frequency and once these words are retrieved one tends to retrieve less frequent words, with this resulting in lower word productivity. Likewise, Hurks et al. (2006) reported that children retrieved more correct responses in the first 15 s of the test period than in the remaining 3 intervals of the test period, with their performance decreasing from 4.93 responses in the first quartile to 2.66 in the second quartile to 2.06 in the third quartile to 1.65 in the fourth quartile.

Once the relatively automatic access to prototypical items has been exhausted, more controlled, effortful searching is required in verbal fluency tasks (Hurks et al., 2006). Automatic processing refers to word productivity in the early stages of the test period, and controlled processing refers to word productivity in later stages of the test period. Hurks et al. (2010) investigated the developmental trajectories of automatic and controlled processing using semantic categories in children aged 6 to 15 years of age, and reported that controlled processing was established 2 years later (Grades 7-8) than automatic processing. In Henry et al.'s (2015) study of semantic and phonological fluency, the number of lexical items produced in the first 15 s (quartile 1) showed a stronger positive correlation with executive-loaded working memory than the number of lexical items produced in the subsequent 45 s (quartiles 2, 3 and 4). This finding was interpreted as revealing a qualitative difference between the first quartile and the subsequent quartiles.

2.7. Semantic and phonological fluency performance in clinical groups

2.7.1. Evidence from children with dyslexia

Findings for children with dyslexia are mixed. Some studies have reported that children with dyslexia perform poorly both on semantic and phonological fluency. There is another line of evidence showing that children with dyslexia perform poorly on semantic fluency but similarly to controls on phonological fluency.

2.7.1.a. Poorer semantic and phonological fluency performance. The studies presented in this section tested children with dyslexia in different orthographies, and across orthographies children with dyslexia showed significantly poorer semantic and phonological fluency performance compared to age-matched TD children. Levin's (1990) early study assessed 20 children with dyslexia who were selected on the basis of average intelligence, 1.3 years below grade level on reading and 1.5 grades below grade level on the California Achievement Test, but within grade level performance on maths and spelling, alongside 20 TD children. Verbal fluency used the categories of 'proper names', 'foods' and 'words beginning with the letter *F*'. Significant differences were found between children with dyslexia and TD children in all three categories. Levin also reported that in phonological fluency, children with dyslexia showed more out-of-category responses compared to TD children. Similar findings for poorer semantic and phonological fluency performance in children with dyslexia have been reported by other studies (e.g. Menghini et al., 2010; Moura, Simões, & Pereira, 2015; Plaza, Cohen, & Chevrie-Muller, 2002; Reiter, Tucha, & Lange, 2005; Varvara et al., 2014). Reiter *et al.* (2005) reported that the differences in semantic and phonological fluency performance represented large and medium effect sizes (respectively, Cohen's d = .81 and .48), while Moura *et al.* (2015) reported large effect sizes in semantic and phonological fluency performance (respectively, partial eta squared (η_p^2) = .11 and .13) even after controlling for general intellectual ability.

2.7.1.b. Poorer phonological fluency but similar semantic fluency performance. All the studies discussed in this section show poorer performance in children with dyslexia compared to age-matched TD children on phonological fluency tasks but a similar performance. Frith, Landerl, and Frith (1995) measured the time it took 12-year-old children with dyslexia and their age-matched peers to produce ten exemplars from the semantic category of animals and ten exemplars from the phonological category of the sound /S/. They also measured the number of repeated words, which was a type of error. Children with dyslexia repeated more words in both conditions compared to TD children but the differences were nonsignificant. Importantly, they found a nonsignificant difference between the two groups in the time in which ten exemplars from the semantic category of animals were produced, but a significant difference between the two groups in the phonological condition, with the children with dyslexia requiring longer to retrieve ten words. Further, with respect to phonological fluency, the researchers showed that one can rely on orthographic representations, in addition to phonological representations, when retrieving words from the lexicon. Indeed, the spelling of the sound /S/ is S or C, and 11 out of 19 children with dyslexia produced words beginning with C compared to 6 out of 19 controls. According to the researchers, this finding might suggest that the two groups utilize different strategies.

Frith et al. (1995) were the first researchers to explore performance in children with dyslexia at a finer-grained level than just recording the number of responses produced in semantic and phonological conditions. The researchers asked children to produce items in three more successive trials for each condition and measured the number of new words produced on these three trials. Children with dyslexia and TD children did not differ in the number of new words produced on trials 2 to 4 of either the semantic or the phonological condition. The researchers estimated a word pool size for each condition reflecting the number of total words from which the children could retrieve words to tease apart if group differences on phonological fluency performance could be accounted for by a smaller number of words from which children could retrieve words or by lexical difficulties. Even though TD children showed a larger estimated word pool size in the phonological condition than children with dyslexia, this difference was not significant.

Frith et al. (1995) therefore argued that children with dyslexia had a similar sized lexicon compared to TD children, and that group differences in phonological fluency could be accounted for by a difficulty that children with dyslexia had in accessing these words by their initial phoneme. This study therefore showed that even though children with dyslexia might have the same number of words in their lexicon compared to TD children, when they are asked to retrieve as many words as possible in a phonological fluency task, they show a lower performance than TD children. Frith et al. interpreted this finding as revealing that word retrieval difficulties in children with dyslexia can be accounted for by access difficulties, not by an impaired underlying organization of the lexicon.

Moreover, the finding that children with dyslexia did not differ from TD children in semantic fluency but differed in phonological fluency could be interpreted as revealing that lexical difficulties in children with dyslexia might originate from a phonological processing or phonological representational deficit rather than a semantic processing or semantic representational deficit. Other studies have also reported significant group differences between children with dyslexia and controls on phonological fluency but nonsignificant group differences on semantic fluency (e.g. Brosnan et al., 2002; Landerl, Fussenegger, Moll, & Willburger, 2009; Marzocchi et al., 2008).

2.7.2. Evidence from children with DLD

Studies which have assessed English children with DLD on semantic and phonological fluency are consistent in terms of finding group differences, which contrasts with the previous section on dyslexia, where the findings are much more mixed. Children with DLD show significantly poorer semantic and phonological fluency performance than TD children.

Weckerly, Wulfeck, and Reilly (2001) found that children with DLD produced significantly fewer correct responses in semantic categories compared to TD children (~18 versus 21 responses). Children with DLD produced about 8 correct responses in phonological categories and TD children about 11 correct responses, with the difference being significant. Henry, Messer, and Nash (2012) assessed children with DLD, low language functioning and controls. They reported that after controlling for age, NVIQ and verbal IQ, 39% of the variance in semantic and phonological fluency scores was significantly accounted for by the dummy-coded group variable of ‘DLD vs typical group’. Even after controlling for age, NVIQ and verbal IQ, there was an effect of group on verbal fluency performance. Likewise, Henry et al. (2015) found that children with DLD produced significantly fewer responses in semantic and phonological conditions compared to TD children. Marshall et al. (2013) compared TD deaf children who used BSL with deaf children who had DLD in their BSL. Children’s sign-finding difficulties were evident in particular in the first 15 s of the test period, where the TD group outperformed the DLD group (8 versus 6 signs), and also in some sign-finding errors which were not evident in the TD group.

2.7.3. Evidence from children with word-finding difficulties (WFDs)

There is also noteworthy evidence originating from children with WFDs. Messer, Dockrell, and Murphy (2004) assessed semantic, phonological and rhyme fluency in children with WFDs. They found that only 6 and 10% of the 7- and 9-year-old children scored within one *SD* of the mean on semantic fluency. On phonological fluency, however, 20% of the 7- and 9-year-old children scored within one *SD* of the mean, and on rhyme fluency the scores were 27 and 69% at 7 and 9 years, respectively. This study suggests that children with WFDs have greater difficulties with semantic representations and/or semantic processing skills than phonological representations and/or phonological processing skills relative to TD children.

2.7.4. Evidence from adults with dyslexia and/or DLD

In the absence of a reading age-matched control group, it cannot be adjudicated whether dyslexic children's poor performance in phonological fluency tasks is a consequence rather than a cause of poor reading. Frith et al. (1995) tested eight highly educated and well-compensated adults with dyslexia in that they had a reading level in the average range but still showed significantly poorer reading and spelling abilities compared to controls and also signs of phonological impairment. Results showed that adults with dyslexia produced on average 26.1 words for three semantic categories compared to 23.8 words produced by controls, and on average 11.2 words for three phonological categories compared to 14.3 words produced by controls. There was a significant group by task interaction: adult people with dyslexia differed from controls only in phonological fluency but not in semantic fluency. Frith et al. argued that these results suggest that for this group of highly educated and well-compensated adults with dyslexia, group differences in phonological conditions are not consequential of their poorer reading experience but are attributable instead to an underlying impaired phonological system in dyslexia.

Smith-Spark et al. (2017) tested university students with dyslexia and controls in semantic and phonological fluency to investigate whether dyslexia can account for any group differences in the two verbal fluency conditions. Consistent with the findings of Frith et al. (1995), they found that after controlling for short-form IQ, adults with dyslexia performed significantly more poorly on phonological fluency, but nonsignificant group differences were found on semantic fluency. The findings originating from the two adult studies in dyslexia are consistent in that adults with dyslexia differ from controls on phonological fluency tasks but not on semantic fluency tasks. However, Hall, McGregor, and Oleson (2017) tested a group of adults with dyslexia and/or DLD, and reported lower overall semantic fluency than controls (24 versus 27 words), with a small effect size being found (η_p^2 = .04).

2.7.5. Summary: Semantic and phonological fluency performance in clinical groups

Table 2.1. summarises the evidence reviewed above. Overall, it is evident that mixed findings have been reported for children with dyslexia. For children with DLD, findings appear to be consistent revealing significantly lower semantic and phonological fluency performance compared to TD children.

Table 2.1. Findings of semantic and phonological fluency performance between different clinical groups and appropriate controls

Groups	Semantic fluency	Phonological fluency
Children with dyslexia[1]	< TD	< TD
Children with dyslexia[2]	= TD	< TD
Children with dyslexia[3]	= TD	= TD
Children with DLD[4]	< TD	< TD
Children with DLD[5]	< TD	< TD
Children with DLD[6]	< TD	< TD
Deaf children with DLD[7]	< TD	-
Children with WFDs[8]	< TD	< TD
Adults with dyslexia[9]	= controls	< controls
Adults with dyslexia[10]	= controls	< controls
Adults with dyslexia and/or DLD[11]	< controls	-

Notes: coloured are significant results showing a poorer semantic and/or phonological fluency performance between different clinical groups and appropriate controls; [1]Levin (1990), Menghini et al. (2010), Moura et al. (2015), Plaza et al. (2002), Reiter et al. (2005), Varvara et al. (2014); [2]Brosnan et al. (2002), Griffiths (1991), Landerl et al. (2009), Marzocchi et al. (2008); [3]Landerl et al. (2009); [4]Weckerly et al. (2001); [5]Henry et al. (2012); [6]Henry et al. (2015); [7]Marshall et al. (2013); [8]Messer et al. (2004); [9]Frith et al. (1995); [10]Smith-Spark et al. (2017); [11]Hall et al. (2017).

2.8. Switching, clustering and cluster size in clinical groups

2.8.1. Evidence from children with dyslexia

Research on switching and clustering within verbal fluency tasks in children with dyslexia is almost entirely lacking. To the best of the author's knowledge, the only published study which investigated switching and clustering in children with dyslexia is by Mielnik, Łockiewicz, and Bogdanowicz (2015). They reported that Polish children with dyslexia switched fewer times on a semantic fluency task and produced fewer clusters compared to controls. Nonsignificant group differences were found in the phonological fluency task with respect to the number of switches, however.

2.8.2. Evidence from children with DLD

Weckerly et al. (2001) found that children with DLD did not differ on cluster size from TD children. TD children, however, significantly outperformed DLD children in the number of switches and the number of clusters. The authors' interpretation of these findings was that the organization of the lexicon did not differ between the two groups, and that the groups were equally efficient in terms of their access to semantic knowledge. They further reported no evidence for a condition (semantic, phonological) by group effect, which according to the researchers suggests that children with DLD showed general processing difficulties, that is, not specific to the phonological aspects of language.

Henry et al. (2015) found that children with DLD produced fewer switches in semantic and phonological categories compared to controls. In the semantic categories, a similar cluster size was evident between the two groups, while in the phonological categories, children with DLD produced a marginally smaller cluster size than controls. Marshall et al. (2013), who compared TD deaf children who used BSL with deaf children who had DLD in their BSL, reported that the number of clusters and the size of clusters in semantic fluency tasks did not differ between the two groups. Marshall *et al.* interpreted these data as indicating that the organization of the lexicon in deaf signers with DLD is not different from that of TD signers. However, given that significantly fewer signs were retrieved from the DLD group of signers compared to the TD group of signers, an access deficit was proposed to account for sign-finding difficulties in deaf signers with DLD. This access deficit could originate either from a slower access to the lemma, suggesting a semantic processing or a semantic representations deficit, or from a slower or unsuccessful retrieval of the phonological form, suggesting a phonological processing or a phonological representations deficit, with the exact locus not yet being clear (Marshall, 2014).

2.8.3. Evidence from adults with dyslexia and/or DLD

Smith-Spark et al. (2017) tested English adults with dyslexia. They found that after controlling for short-form IQ, adults with dyslexia did not differ from controls on semantic switching and on the number of words per semantic cluster. In phonological fluency, group predicted the number of switches but did not predict the number of words per phonological cluster. Smith-Spark et al.

argued that this finding is consistent with Frith et al.'s (1995) argument that group differences in phonological fluency in their sample could be accounted for by the difficulty of participants with dyslexia with accessing words based on their initial phoneme and not by differences in vocabulary size. Hall et al. (2017) found, using semantic categories, that adults with dyslexia and/or DLD produced a significantly smaller cluster size and embedded cluster ratio than controls but the two groups did not differ on cluster switch ratio. What an embedded cluster ratio was can be explained as follows. In the category of food, for example, 'dairy' is a cluster, and within 'dairy', *ice creams* are embedded. Thus, in the following sequence of words, 'milk, cheese, *chocolate ice cream, vanilla ice cream, strawberry ice cream*, yogurt', a cluster of 'dairy' can be identified, but it can be also identified an embedded cluster of 'ice cream' within the larger 'dairy' cluster. An embedded cluster ratio was the ratio of embedded clusters to total clusters of three or more items. Cluster switch ratio can be defined as the ratio of cluster switches to hard switches, where a hard switch was a transition between a response in a cluster and a response not in a cluster, or vice versa, or a switch between each item in a series of unclustered items, and a cluster switch was a transition between two adjacent clusters. Embedded clusters and cluster switches were expressed as ratios to correct for differences in the number of clusters that each person named.

2.9. Correlation, network and computational modelling methodologies in semantic fluency tasks in clinical groups

Correlation, network and computational modelling can be used as alternative methodologies providing insight into the structure of children's semantic networks. How semantic memory is organized, that is, which lexical items are close in semantic space to others can be investigated using computational network tools. In this thesis, the first study using computational modelling to investigate the semantic network of children with DDLD will be conducted, by comparing the semantic network of these children with the semantic network of TD children. To this end, the results of the semantic category of animals will be analysed. Analysing the semantic network of the two groups allows one to investigate whether the structure of the semantic network is similar in the two groups or not. This investigation offers an analysis of the data which is independent of the coding scheme that the experimenter followed to code the data into clusters. Findings originating from two studies which used correlation, network and computational modelling methodologies are discussed below.

2.9.1. Evidence from children with Cochlear Implants (CIs)

Kenett et al. (2013) examined, using correlation and network methodologies, the semantic memory organization in children with CIs compared to normal hearing (NH) children. Correlation methodologies refer to a correlation score being observed among words belonging to the same subcategory, implying that if the word ‘dog’ is retrieved from the subcategory of ‘household pets’, then it is likely that the word ‘cat’ will also be retrieved. The researchers provided evidence for a less developed semantic network structure using the semantic category of animals in children with CIs compared to NH children in that the semantic network of the former was more condensed and less spread out than that of the latter.

2.9.2. Evidence from adults with Mild Cognitive Impairment (MCI) and Alzheimer’s Disease (AD)

Davelaar, Vrontissis, and Fullgrabe (2015) used computational modelling, and found a significant difference between healthy adults and adults with MCI using the computed metric, which suggests that adults with MCI showed compromised semantic memory compared to healthy adults. The metric was computed as follows. Produced sequences were combined into a single network by overlaying the individual paths. The resulting network was directed and weighted. Each individual’s sequence had a likelihood of being produced by that reference network. The likelihood of an individual’s sequence was the product of the frequency of each word and the conditional probability of all the steps in the sequence. A maximum likely sequence was constructed as a reference path to control for the influence of sequence length. A ratio score was then computed for each individual and was the metric of interest in the study.

2.9.3. Summary: Switching, clustering and cluster size in clinical groups

From the studies reviewed so far, it is evident that most of our knowledge of how children with dyslexia and DLD retrieve items during semantic and phonological fluency comes from English. There is very little research from other languages. The current study aims to fill this gap in the developmental literature by testing patterns of lexical retrieval in semantic and phonological fluency tasks in Greek-speaking children with dyslexia and/or DLD. Given that there are still related improvements in semantic and phonological fluency until mid-adolescence (Hurks et al., 2010), semantic and phonological fluency tasks should be sufficiently sensitive to differentiate amongst primary school-aged children. Table 2.2. summarises the evidence reviewed above.

Table 2.2. Findings of semantic and/or phonological switching, clustering and cluster size between different clinical groups and appropriate controls

Groups	Switching	Clustering	Cluster size
Children with dyslexia[1]	Semantic: < TD;	Semantic: < TD	-
	Phonological: = TD	-	-
Adults with dyslexia[2]	Semantic: = TD;	-	Semantic: = TD;
	Phonological: < TD	-	Phonological: = TD
Children with DLD[3]	Semantic: < TD;	Semantic: < TD;	Semantic: = TD;
	Phonological: < TD	Phonological: < TD	Phonological: = TD
Children with DLD[4]	Semantic: < TD;	-	Semantic: = TD;
	Phonological: < TD	-	Phonological: < TD
Deaf children with DLD[5]	-	Semantic: = TD	Semantic: = TD
	-	-	-
Adults with DDLD[6]	Semantic: = controls	Semantic: < controls	Semantic: < controls
	Network methodology		
Children with CIs[7]	Semantic: < controls		
	Computed metric		
Adults with MCI[8]	Semantic: < controls		

Notes: coloured are significant results showing a smaller number of switching, clustering and cluster size between different clinical groups and appropriate controls; [1]Mielnik et al. (2015); [2]Smith-Spark et al. (2017); [3]Weckerly et al. (2001); [4]Henry et al. (2015); [5]Marshall et al. (2013); [6]Hall et al. (2017); [7]Kenett et al. (2013); [8]Davelaar et al. (2015).

2.10. Automatic versus controlled processing in clinical groups

Evidence from children with DLD, adults with dyslexia and adults with dyslexia and/or DLD is reviewed hereafter. Marshall et al. (2013) compared a group of deaf children with a group of deaf signing children with DLD, and found that the latter group produced significantly fewer responses belonging to the categories of ‘animals’ and ‘food’ in the first 15 s of the test period compared to the former group. The authors argued that this finding suggests that deaf signers with DLD show slower access to the lexicon. It is not clear yet whether slower access to the lexicon is due to slower access to the lemma, or to slower retrieval of the phonological form (Marshall et al., 2013). In Henry et al.’s (2015) study, children with DLD showed weaker performance at all four quartiles of the semantic and phonological tasks, including the more automatic first quartile. The authors argued that these data show weaknesses in automatic processing (lexical items retrieved in quartile 1), in addition to controlled processing (lexical items retrieved in quartiles 2, 3 and 4), suggesting performance limitations for easily accessed items at the beginning of the task but also for items requiring more demanding search efforts later in the task.

In Smith-Spark et al.'s (2017) adult study of university students, group predicted the number of total responses in the first, second and fourth quartile in phonological fluency tasks but did not predict any quartile in semantic fluency tasks, after controlling for short-form IQ. In another adult study, however, Hall et al. (2017) reported that adults with dyslexia and/or DLD showed a similar change over the course of the test period in that they were not less fluent than controls after the initial 15 s of the test period in semantic fluency. Given that they found no evidence for difficulties in controlled processing, they interpreted this finding as evidence that EFs did not show any effect on semantic fluency performance in adults with dyslexia and/or DLD. It remains to be investigated whether children with and without DDLD differ in automatic processing, in controlled processing, or in both. It is hypothesised that if automatic processing reflects the lexico-semantic structure and controlled processing reflects EFs, children's language and literacy skills will have a greater effect on automatic than controlled processing in the regression models.

2.11. Effects of language, literacy and executive function (EF) skills on semantic and phonological fluency performance

2.11.1. The effect of language and literacy skills

It has been reported that semantic fluency is related to language measures, namely, naming, repetition, comprehension and phonological fluency (in Ardila et al., 2006). Studies have also shown that greater productivity in semantic and phonological fluency is associated with better performance on language measures assessing a child's ability to define words and identify similarities across sets of words, as recently reported by Henry et al. (2015) who assessed English children with DLD and TD. Henry et al. found that better language ability supports performance in earlier parts of the semantic fluency task, when items are more readily available, but ceases to be important during the more effortful searching required in later parts of the task; however, better language ability supports performance both in earlier and later parts of the phonological fluency task. The finding that language ability was a stronger predictor of phonological than semantic fluency in Henry et al.'s study is not consistent with Luo, Luk, and Bialystok's (2010) study, however, which revealed that language ability is a more important predictor of semantic than phonological fluency. They argued that this is because more integrated semantic knowledge was needed for the semantic task and because language ability was relevant throughout the semantic fluency task. In an adult study, Whiteside et al. (2016) found that both semantic and phonological fluency tasks loaded on to a language factor, with the researchers arguing that language processing is of critical importance for both verbal fluency conditions.

Aside from the role of language ability, literacy skills have also been found to play a role in verbal fluency performance. Indirect evidence for the effect of literacy skills on verbal fluency performance originates from the study of Riva, Nichelli, and Devoti (2000) who tested children aged 5-11 years. An important finding of their study was that verbal fluency performance increased linearly from first Grade to fifth Grade, with the most significant increase observed between first and second Graders. The authors argued that this is because at that time formal teaching begins, and children begin to know the components of language. Riva et al. (2000) therefore proposed an association between the development of the ability to organize and retrieve words according to phonological categories and reading skills. To the best of the author's knowledge, only one published study directly showed that in phonological conditions, the effect of reading ability was significant, with a small effect size (η^2 = .06) being found (Landerl et al., 2009).

In the current study, it will be investigated for the first time whether language and literacy skills predict word productivity in semantic and phonological fluency categories in Greek TD children and children with DDLD. Considering that children with DDLD show inferior language and literacy skills relative to their TD peers, it is predicted that word productivity in semantic and phonological fluency categories will be partly accounted for by children's language and literacy skills in the regression models.

2.11.2. The effect of EFs

Aside from language and literacy skills, greater productivity in verbal fluency categories is associated with better performance on measures of EFs too (Hall et al., 2017; Henry et al., 2015; Luo et al., 2010; Marshall et al., 2018; Shao, Janse, Visser, & Meyer, 2014; Unsworth et al., 2011). Luo et al. (2010) reported, however, that EFs are more important predictor of phonological than semantic fluency since the former task demands greater increased strategic search processes than the latter task. This is because semantic search is the default search strategy used in verbal fluency tasks (Vonberg et al., 2014).

Recently, Marshall et al. (2018) found that 26.2% of the variance in semantic fluency was predicted by age and NVIQ and 23.4% was predicted by expressive vocabulary and by a composite score of EFs (i.e. measures of visuospatial working memory, inhibition, planning, shifting and design fluency) in two groups of deaf and hearing children. The researchers concluded that semantic fluency can be used in deaf children who use BSL and spoken English, as an index of both vocabulary and EFs. In another population, university students with specific learning disabilities (dyslexia and/or DLD), Hall et al. (2017) found that lexical-semantic knowledge and EFs were related to semantic fluency performance and they argued that semantic fluency tasks can reveal difficulties in lexical-semantic knowledge and EFs.

2.12. Verbal fluency and executive functioning (EF)

Prior to turning to the design fluency task and relevant studies, and given that design fluency tasks have been argued to involve EFs, this section presents evidence suggesting difficulties with EFs in children with dyslexia and DLD. As well as reflecting lexical organization, verbal fluency tasks also require the use of word retrieval strategies, which rely on higher level thinking and reasoning skills, or EFs. EFs include: (i) executive-loaded working memory (i.e. requirement for concurrent remembering and processing of information); (ii) inhibition (i.e. suppression of possible yet inappropriate responses that come readily to mind); (iii) switching (between different mental sets); and (iv) fluency (i.e. generation of new responses according to a rule). For example, Henry et al. (2015) found that inhibition is related to the number of errors in the phonological condition since it significantly predicted the total number and proportional scores of errors. A high proportion of errors might be indicative of difficulties in word search and retrieval processes, caused by impaired executive control over semantic and phonological search and retrieval strategies. Thus, in the current study, the groups' number of incorrect responses, and a related measure, the error ratio, in verbal fluency categories will be therefore compared.

Thompson et al. (2015) assessed children at high risk of dyslexia, and reported that EFs at age 4½ improved the prediction of dyslexia. When it comes to DLD, Henry and Botting (2017) reported in their review paper that children with DLD show deficits both in verbal and nonverbal central executive resources of working memory. In a similar vein, Gooch, Hulme, Nash, and Snowling (2014) reported that preschool children with language impairment showed severe and persistent difficulties in EFs compared to children without language impairment.

2.13. Design fluency

Tests of design (nonverbal) fluency assess an individual's ability to generate geometric patterns and are argued to measure visuospatial EFs. Another issue emerging from the two phonological hypotheses and the two lexical-semantic models is the specificity of the fluency deficit to verbal material in dyslexia and DLD. For example, Protopapas (2014) argues that to establish the viability of any phonological hypothesis, one has to ensure that statistically poorer performance on tasks requiring phonological processing is accompanied by normal performance on similarly-structured tasks that do not involve phonological processing. The design fluency task used in this study measures visuospatial executive skills by assessing a child's ability to generate nonsense designs under time constraints and restricted design conditions. It is therefore a similarly-structured task to semantic and phonological fluency tasks without requiring, however, phonological or semantic representations and phonological or semantic processing skills.

Both phonological hypotheses and both lexical-semantic models predict that children with DDLD will generate a similar number of correct designs in the design fluency task compared to TD children. With respect to phonological hypotheses, this is because both hypotheses advocate a 'modular' deficit within the language system which affects the phonological domain, whilst the nonverbal domain is unaffected. However, given that empirical evidence shows that children with dyslexia and DLD demonstrate deficits not related only to the effective functioning of the phonological system (e.g. Gooch et al., 2014; Henry et al., 2012; Henry and Botting, 2017; Varvara et al., 2014), for the purpose of this study, further investigation in the nonverbal domain is needed. Moreover, as Messer and Dockrell (2006) have argued, in the context of children with word-finding difficulties, lexical-retrieval difficulties can be potentially caused by impairments in processing speed, amongst other proposed causes. It is hypothesized that if there is a slower processing speed in children with DDLD accounting for lower semantic and phonological fluency performance, lower design fluency performance would be also found in the DDLD group; however, if only verbal processing difficulties were to underlie poorer semantic and phonological fluency performance in children with DDLD, the two groups would show similar design fluency performance. Furthermore, in order to test the specificity of the fluency deficit to verbal material in children with DDLD, design fluency performance will be used as a covariate in the analyses investigating group differences in semantic and phonological fluency.

Existing research on design fluency in children with dyslexia is limited, and inconsistent findings have been reported, with one study reporting that the dyslexia group generated significantly fewer correct designs than the TD group (Griffiths, 1991), and another study reporting no group difference (Reiter et al., 2005). To my knowledge, only one published study used design fluency in children with DLD, and showed that the DLD group generated significantly fewer correct designs compared to the TD group (Henry et al., 2012). Recently, Smith-Spark et al. (2017) tested university students with dyslexia using a design switching task, in which participants were asked to switch alternately between empty and filled dots in each design. They reported that, after controlling for IQ, adults with dyslexia did not differ from controls on design fluency. The researchers argued that given that the two groups of participants did not differ on EFs, as measured with the design fluency task, phonological fluency deficits found in adults with dyslexia could not be attributed to difficulties with EFs but rather to phonological processing problems in dyslexia. There are no design fluency data originating from Greek children with dyslexia and DLD.

2.13.1. Summary: Design fluency performance in clinical groups

From the review so far, it is evident that inconsistent findings have been reported in relation to design fluency performance in children and adults with dyslexia or DLD compared to controls. Table 2.3. summarises the evidence reviewed above.

Table 2.3. Findings of design fluency performance between different clinical groups and controls

Groups	Design fluency
Children with dyslexia[1]	= TD
Children with dyslexia[2]	< TD
Adults with dyslexia[3]	= controls
Children with DLD[4]	< TD

Notes: [1]Griffiths, 1991; [2]Reiter et al. (2005); [3]Smith-Spark et al. (2017); [4]Henry et al. (2012).

2.14. Research questions

Research Question 1. Using analysis of semantic clustering behaviour and computational modelling analysis, what is the structure of the lexicon in children with DDLD compared to TD children? Is poorer semantic fluency performance in children with DDLD better explained by impoverished semantic structure, as proposed by the Poor Lexical-Semantic Structure Model, or slower retrieval processes, as proposed by the Slow-Retrieval Model?

Research Question 2a. Using analysis of phonological clustering behaviour, where does the phonological deficit in children with DDLD lie? Is poorer phonological fluency performance in children with DDLD better explained by degraded phonological representations, as proposed by the Degraded Phonological Representations Hypothesis, or by deficient explicit access to (intact) phonological representations, as proposed by the Deficient Phonological Access Hypothesis?

Research Question 2b. Which hypothesis better characterises the locus of the phonological deficit in children with DDLD in phonological tasks and a spelling-to-dictation task: The Degraded Phonological Representations Hypothesis or the Deficient Phonological Access Hypothesis?

Research Question 3. Do cluster number and/or cluster size drive productivity in semantic and phonological fluency tasks in TD children and children with DDLD?

Research Question 4. Does semantic and phonological fluency performance relate to children's language and literacy skills?

Research Question 5. How specific is the verbal fluency deficit in children with DDLD: Does it extend to a nonverbal task (design fluency)?

2.15. Predictions

Research Question 1. Using analysis of clustering behaviour in the semantic condition, the Poor Lexical-Semantic Structure Model predicts that the DDLD group will produce significantly fewer correct responses and a significantly smaller average cluster size than the TD group. This is explained by poor semantic structure in the DDLD group. In contrast, the Slow-Retrieval Model predicts that the DDLD group will produce significantly fewer correct responses than the TD group but that the two groups will not differ on average cluster size. This is explained by slow retrieval processes while lexical-semantic representations are intact in the DDLD group. Moreover, using computational modelling analysis, the Poor Lexical-Semantic Structure Model predicts that the DDLD group will produce a significantly larger ratio score than the TD group, suggesting structural differences in the semantic network of children with DDLD. In contrast, the Slow-Retrieval Model predicts that the two groups will not differ on the computed ratio score, suggesting an adequate semantic network but difficulties in accessing semantic information quickly and efficiently, that is, suggesting retrieval differences between the two groups.

Research Question 2a. Using analysis of clustering behaviour in the phonological condition, the Degraded Phonological Representations Hypothesis predicts that the DDLD group will produce significantly fewer correct responses and a significantly smaller average cluster size than the TD group. This is explained by impaired phonological representations in the DDLD group. In contrast, the Deficient Phonological Access Hypothesis predicts that the DDLD group will produce significantly fewer correct responses than the TD group but that the two groups will not differ on average cluster size. This is explained by impaired explicit access but intact implicit access to phonological representations in the DDLD group.

Research Question 2b. Using phoneme deletion tasks, the Degraded Phonological Representations Hypothesis predicts that the TD group will outperform the DDLD group in accuracy and speed performance in all three phoneme deletion tasks. This is explained by impaired phonological representations in the DDLD group. In contrast, the Deficient Phonological Access Hypothesis predicts that accuracy and speed performance in the phoneme deletion task of monosyllable items with simple CVC syllable structure will be equivalent for the two groups. This is explained by the fact that short nonwords with a simple syllable structure do not load phonological STM. The Deficient Phonological Access Hypothesis also predicts that accuracy and speed performance in the phoneme deletion task of trisyllable items with simple CVCVCV syllable structure and in the phoneme deletion task of monosyllable items with complex CCV syllable structure will be poorer for the DDLD group than the TD group. This is explained by the fact that long nonwords with three syllables or nonwords with complex syllable structure load phonological STM. It should be noted that phonological STM capacity is not actually measured, however, a NWR measure consisting of 3-, 4-, 5-, and 6-syllable nonwords will be used to investigate whether the DDLD group shows phonological STM deficits relative to the TD group.

Using the RAN task, the Degraded Phonological Representations Hypothesis predicts that the DDLD group will make phonological errors in their picture naming. Phonologically inaccurate performance is explained by inaccurate phonological representations. The Deficient Phonological Access Hypothesis, however, predicts that the DDLD group will name pictures significantly slower than the TD group but that the two groups will not differ on phonological accuracy. Slower naming performance is explained by the fact that the phonological access deficit in the DDLD group renders performance on tasks requiring speeded access to phonological representations particularly slow. Phonologically accurate performance is explained by intact access to phonological representations in the RAN task not requiring metalinguistic manipulation.

Using the spelling-to-dictation task, the Degraded Phonological Representations Hypothesis predicts that qualitative analysis of spelling errors will reveal that the DDLD group produces a higher proportion of phonological spelling errors than the TD group. This is explained by inaccurate phonological representations in the DDLD group. The Deficient Phonological Access Hypothesis, however, predicts that qualitative analysis of spelling errors will reveal a similar proportion of phonological spelling errors in the DDLD group relative to the TD group. This is explained by accurate phonological representations but inappropriate orthographic encoding of words using grapheme-to-phoneme mappings that are inappropriate for a particular context in the DDLD group.

Research Question 3. It is predicted that in the TD group, semantic and phonological fluency performance will be driven mainly by the production of more clusters and more switches between clusters rather than by the production of bigger clusters.

Research Question 4. It is predicted that some of the variance in semantic and phonological fluency performance will be accounted for by language and literacy measures after controlling for age in months and NVIQ in the analyses.

Research Question 5. In accordance with the two phonological hypotheses and the two lexical-semantic models, it is predicted that the DDLD group will not differ from the TD group on design fluency performance.

Chapter 3. Methods

3.1. Ethical approval

Ethical approval for the study was obtained from the Departmental Research Ethics Committee of the UCL Institute of Education, University College London, UK, and from the Hellenic Ministry of Education, Research and Religious Affairs (Reference no. Φ15/1049/160131/Δ1). Parents gave informed written consent on behalf of the participating children. Informed consent documents are available in the Appendix; see Information Sheet, Written Consent and Questionnaire (respectively, A, B and C) in the Greek language.

3.2. Participants

3.2.1. Children with Dyslexia and/or Developmental Language Disorder (DDLD)

Sixty-six monolingual Greek children aged 7-12 years, from second to sixth Grade, were included in this group (see Table 3.1. of general descriptives). Participants had a mean NVIQ of 96.74 (*SD* = 15.12) as measured using standard scores from the Greek standardization (Sideridis, Antoniou, Mouzaki, & Simos, 2015) of the Raven's Coloured Progressive Matrices (CPM; Raven, 2008).

3.2.1.a. Initial selection criteria for the DDLD group. This clinical sample was recruited from Inclusion Classes in the North of Greece, that is, units of Special Education attached to mainstream schools run by Special Education teachers who offer specialized educational programs in small groups for children with a range of special educational needs. The Special Education teacher of each school was contacted to identify children having a formal diagnosis of dyslexia and/or DLD. This means that children assigned to the DDLD group were selected on the basis of the diagnosis they had received. Descriptive measures assessing language, literacy and phonological processing skills were administered but were not used as selection criteria in this initial selection of participants. In accordance with ethical considerations, an information leaflet with details about the study, a consent form and a questionnaire designed for the purposes of this study were given to parents of potential participating children. This group was required to meet, according to parental responses to a questionnaire, the following inclusion criteria: (a) monolingual Greek speaker; (b) no current or prior history of hearing deficit (parents had to report no history of recurrent middle ear infections or known episodes within the preceding 12-month period); (c) visual acuity normal or corrected; (d) no absence from school for at least 3 months; (e) continuous attendance at Greek primary school (since the first Grade); (f) no history of neurological disease; and (g) no history of medication for any neurological, psychiatric, behavioural, or emotional disorder. These inclusion criteria were used on the basis that factors such as bilingualism, hearing or visual deficits, absence from school for a long period of time or a delayed start to schooling, and other medical conditions, can all affect children's performance on the tasks. These inclusion criteria have been used by other researchers too. Importantly, according to WHO (2011), a child to be diagnosed with 'developmental dyslexia' must present with significantly below average written language ability despite adequate schooling, normal cognitive capacity, and in the absence of any obvious sensory or neurological damage that might account for written language difficulties, such as hearing loss, or brain damage. It can be argued therefore that the current study followed the exclusion and inclusion criteria accompanying an official identification and diagnosis of developmental dyslexia.

In Greece, multidisciplinary teams identify children with dyslexia based primarily on the discrepancy (between general intellectual abilities and achievement) model (but see below discussion of this discrepancy as an exclusion criterion), following the definition of the National Joint Committee on Learning Disabilities (Al-Yagon et al., 2013). No formal definition exists for DLD, however. This means that for the identification of DLD, multidisciplinary teams use criteria based upon their theoretical orientation and operational definition (Al-Yagon et al., 2013). Moreover, experts in Greece often use umbrella terms for a diagnosis, and rather describe the child's condition. For research purposes, this makes the selection of participants a difficult task given that the testing battery is limited to certain tests and cannot be a diagnostic battery on its own. Given this issue, to be eligible for the DDLD group, children with dyslexia had to have received a diagnosis because of persistent and specific reading problems, such as 'developmental dyslexia' of course, but also 'specific disorder of reading', 'specific learning disorder (dyslexia)', 'signs of specific learning difficulty (dyslexia)', 'signs of dyslexia', 'specific difficulties in learning (dyslexia)', 'learning gaps with signs of dyslexia', or 'disorder of reading and syllabication'. Children who had received the following diagnoses were also considered: 'learning difficulties' because of poor decoding, reading fluency, or phonological performance, slow reading pace, and/or spelling errors; 'learning difficulties (features of specific type focused on reading and writing)'; 'learning difficulties in written language and in reading'. All those diagnostic terms are given in Footnotes in the Greek language in that order[1].

Likewise, to be eligible for the DDLD group, children with DLD had to have received a diagnosis because of persistent and specific language problems, such as 'specific language impairment', 'disorder in language expression and as such generalized learning difficulties', 'learning difficulties in the context of disorder of language expression', 'language disorder', 'disorder in the expression of language', 'disorder of expressive language', 'disorder of the expression and perception of language', or 'difficulty in written and expressive language'. Some of the descriptive diagnoses of DLD were the following: 'difficulties in the development of oral language (medium output and organization of speech, relatively poor vocabulary)'; 'difficulties in the development of oral and written language; difficulties in vocabulary, difficulties with structure and output of oral and written language'; 'difficulties in the development of learning skills, in oral language and in phonological system'; 'problems of language and speech'; 'mixed language disorder of receptive and expressive type'. All those diagnostic terms are given in Footnotes in the Greek language in that order[2].

Further, it is well known that there are currently no gold-standard assessments of diagnosing dyslexia and DLD with adequate psychometric properties, namely, valid and reliable assessments with diagnostic or prognostic value. Dockrell and Marshall (2015) argue that screening measures to date do not meet psychometric properties to identify language problems and also that the interpretation of assessments of language is challenged by a range of factors, from a child's socioeconomic and language status to hearing impairment and even characteristics of the assessment. The last of these factors is particularly relevant to the current study. In Greece, while there has been some progress in the development of psychometric materials over recent years (e.g. Sideridis, Mouzaki, Protopapas, & Simos, 2008; Vogindroukas, Protopapas, & Stavrakaki, 2009), standardised clinical tools for the diagnosis of dyslexia and DLD for preschool- and school-aged children are still lacking. Dyslexia is therefore often diagnosed on the basis of non-standardized measures of reading and spelling ability (Anastasiou & Polychronopoulou, 2009), and the same is also the case for DLD. This raises the issue of how accurately children with dyslexia, children with DLD, and children with dyslexia plus DLD can be differentiated; this might not be as easy as in studies of English-speaking children (e.g. Catts, Adlof, Hogan, & Weismer, 2005; Ramus et al., 2013).

Previous research in Greek has explored the overlap between dyslexia and DLD, and reported that dyslexia and DLD show common deficits in tasks measuring reading skills and reading-related phonological skills (Spanoudis et al., 2018; Talli et al., 2016), even though they do not completely overlap. With respect to semantic fluency in particular, evidence originates from the study of Hall e*t al.* (2017) who reported that participants with just a reading impairment were more fluent than participants with both language and reading impairments, while participants with language impairment did not differ from either subgroup. Hall *et al.*'s study therefore suggests that dyslexia and DLD should not be treated separately. In the light of this evidence, a Principal Component Analysis (PCA) with rotation (oblique) within the language and literacy skills of the children with dyslexia and/or DLD was carried out in order for the experimenter to determine whether there were separate loadings onto different components that might justify keeping the children with dyslexia and DLD separate. If there are different components for children's language and literacy skills, it can be argued that dividing the children into separate subgroups would be appropriate. If not, then this finding would offer strong evidence suggesting that combining children with dyslexia and/or DLD in one group is appropriate (Mengisidou & Marshall, 2019).

The dataset was suitable for the PCA: Kaiser-Meyer-Olkin Measure of Sampling Adequacy value was .787, meeting Kaiser's (1974) criterion for this value, Bartlett's Test of Sphericity value was significant ($p < .001$; Bartlett, 1954), and most of the intercorrelations observed among all seven measures of interest had a value of .30 and above. The PCA revealed that five language tasks (WISC Similarities, WISC Vocabulary, syntax comprehension, sentence repetition, and receptive vocabulary) and two literacy tasks (l'Alouette and spelling-to-dictation) used in the overall sample to profile children with dyslexia and/or DLD loaded onto component 1. Table 3.1. presents each task's contribution to components 1 and 2, which is expressed by its loading value. WISC Vocabulary, receptive vocabulary and WISC Similarities had the highest loadings onto the first component, while l'Alouette and sentence repetition had the lowest loadings onto this component.

Components 1 and 2 had an eigenvalue larger than 1, meeting Kaiser's (1974) criterion. The first component had, however, by far the largest eigenvalue of all seven components generated by the PCA. The second component had an eigenvalue of 1.2 and accounted for 18% of the variance in all measures, while the remaining components had an eigenvalue lower than 1, and as such, they were not considered further. Even though components 1 and 2 had an eigenvalue larger than 1, a one-factor solution was selected. This selection was based on the scree plot generated by the PCA illustrating a clear split between component 1 and the remaining components. The PCA was therefore repeated, and a one-factor solution was selected. This analysis revealed that component 1 had an eigenvalue of 3.4 and explained 49.68% of the variance in all seven measures. The result of the PCA suggests that it is appropriate to combine the children with dyslexia and/or DLD into a combined DDLD group. If the first component had loaded essentially on language variables and the second component on literacy variables (or the other way around), then this would have been strong evidence that language and literacy variables were two distinct sources of variance in this dataset. If this had been the case, it would have been a good reason to group the children with dyslexia and DLD separately. The PCA revealed that language and literacy variables did not load on different components, which suggests that it is appropriate to combine the children with dyslexia and/or DLD into a single DDLD group (Mengisidou & Marshall, 2019).

Table 3.1. The loadings onto components 1 and 2 for each task generated by the Principal Component Analysis (PCA) with oblique rotation in the DDLD group

Tasks	Component 1	Component 2
WISC Vocabulary	0.83	0.10
Receptive vocabulary	0.78	0.13
WISC Similarities	0.80	-0.02
Spelling-to-dictation	0.70	-0.52
Syntax comprehension	0.68	0.48
L'Alouette	0.60	-0.62
Sentence repetition	0.41	0.57

Further, the experimenter made no distinction between receptive and expressive language difficulties when recruiting children to the DLD group. The rationale was that there is no distinction between receptive and expressive language impairment in the definition of language disorder proposed by DSM-5 (APA, 2013). Moreover, Tomblin and Zhang's (2006) epidemiological study investigated whether performance on language tasks can reflect different dimensions of children's language ability. The researchers assessed children at kindergarten, second, fourth and eighth grade with standardised tasks of receptive vocabulary, expressive vocabulary, receptive sentence use and expressive sentence use. Factor analysis revealed that all four components of language 'loaded' on one single factor and therefore no evidence was found for the potential dimensionality of language ability into two modalities, receptive and expressive.

Moreover, in the sample of children with dyslexia in the current study, some children had co-occurring difficulties accompanying the diagnosis of dyslexia, such as Attention Deficit Hyperactivity Disorder (ADHD), dysgraphia, problems with time management, learning gaps, cognitive, emotional and social difficulties, hyperactivity, difficulty in using numbers, in performing calculations or in mathematical reasoning. Likewise, in the sample of children with DLD, some children had co-occurring difficulties accompanying the diagnosis of DLD, such as ADHD, articulation disorder, attention deficit, disorder of motor function, developmental disorder of motor skills, difficulties with visual, cognitive structure, separation anxiety, developmental disorder of speech, learning difficulties (in reading, writing and in maths), learning difficulties, difficulties with phonological awareness, grapho-motor difficulties, difficulties in visual-motor co-ordination, specific disorder in educational skills (in reading, writing and spelling), specific disorder in speech fluency (stuttering), persistent impulsivity, difficulty in information processing, difficulties of cognitive organization. Thirty out of sixty-six children had co-occurring difficulties accompanying the diagnosis of dyslexia and DLD. This is in line with the CATALISE consortium (Bishop et al., 2017) where there was an agreement that for DLD, additional disorders should be used as a descriptor rather than an exclusionary factor. Bishop (2017) argues that it is misleading to assume that co-occurring conditions are causes of language disorder, but that DLD should be distinguished from cases of language disorder associated with 'differentiating conditions' that have a known or likely biomedical origin, including brain injury, sensorineural hearing loss, genetic syndromes, intellectual disability and autism spectrum disorder. None of the children recruited to the current study had any of these conditions. Considering this conceptualization of DLD, it was not considered appropriate to control for co-occurring conditions in the analyses.

3.2.1.b. Fine-grained selection criteria for the DDLD group. More fine-grained selection criteria of the DDLD group were applied, since the identification of the DDLD sample on the basis of expert diagnosis resulted in a heterogeneous sample of children. The NVIQ task was a descriptive measure assessing children's nonverbal reasoning skills, however, performance on this task was used as a selection criterion. NVIQ affects performance on the tasks used, and as such, a low IQ score can function as a confounding factor accounting for children's poor performance. Furthermore, the NVIQ inclusion criterion for both the DDLD and TD groups was a score of 70 or above, following the CATALISE consortium (Bishop et al., 2017) and Norbury et al.'s (2016) population study which reported that children with a lower nonverbal ability (i.e. a standard score between 70 and 85) did not differ significantly in their language profile from children with an average nonverbal ability (i.e. a standard score > 85). Norbury et al.'s (2016) population study included language impaired children who scored on the NVIQ test within the normal range (> 85) but also those impaired children who had NVIQ scores between 70 and 85. They therefore considered children with language disorder irrespective of their intellectual level, and they reported that there was no real difference in terms of the nature of language impairment (in addition to learning aptitude or behaviour) between children who achieved a low and those who achieved a higher (i.e. within the normal range and above) NVIQ score. This is also in line with the CATALISE consortium (Bishop et al., 2017) where there was an agreement that for the diagnosis of DLD, NVIQ should no longer be used as an exclusionary criterion as long as there is no diagnosis of Intellectual Disability accompanying the child's language problems. Including in the clinical group children who had comorbid disorders and/or a NVIQ standardised score between 70 and 85 can be considered as a strength of the current study on the basis that the sample is more representative of the children seen in clinics.

3.2.2. Typically Developing (TD) Children

The sample of TD children was recruited from mainstream schools, and consists of eighty-three monolingual Greek children, aged 6-12 years, from first to sixth Grade (see Table 3.2. of general descriptives). Their mean NVIQ was 104.75 (*SD* = 12.94). A broader age range was included in order to investigate the typical pattern of development of semantic, phonological and design fluency.

3.2.2.a. Initial selection criteria for the TD group. The head teacher of each school was contacted to allow access to the school and identify monolingual Greek speakers, and then, in accordance with ethical considerations, an information leaflet with details about the study, a consent form and a questionnaire designed for the purposes of this study were given to parents of potential participating children.

3.2.2.b. Fine-grained selection criteria for the TD group. TD children were carefully selected, in part based on parents' answers to the questionnaire and in part based on the child's performance on the day of their assessment. TD children who achieved a percentile score of 10 or lower on a standard text-reading fluency measure, or substantial difficulties with the language and literacy tasks (e.g. the child had substantial difficulty to understand instructions and/or their response times to tasks were extremely slow) were excluded from the study. Nine children who had a percentile score of 10 or lower on a text-reading fluency measure, another child who also had substantial difficulties with the language and literacy tasks, and another one because of an uncorrected visual deficit were therefore excluded from the study. None of the TD children included in the study had a current or prior history of hearing or visual deficit, neurological disease, or medication for any neurological, psychiatric, or behavioural disorder, as reported in the section of *fine-grained selection criteria for the DDLD group* above. Although the inclusion criteria permitted TD children who had a NVIQ score as low as 70, none actually scored lower than 80.

3.3. Groups' general descriptives

Table 3.2. shows the number of children, the gender distribution, mean age, age range and Raven's CPM performance in the two groups.

Table 3.2. General descriptives of the number of children, the gender distribution, mean age, age range and Raven's CPM performance in the DDLD group and the TD group

	DDLD group	TD group	Group differences		
	M (*SD*)	*M* (*SD*)	*t*	*p*	η_p^2
N	66	83			
Gender ratio (m:f)	43:23	35:48			
Age in months	118.38 (16.92)	104.82 (21.55)	-4.30	< .001	.06
Age in months: range	88-146	75-148			
Age in years	9.51 (1.46)	8.37 (1.77)			
Age in years: range	7;04-12;02	6;03-12;04			
Raven's CPM standard score	96.74 (15.12)	104.75 (12.94)	3.48	.001	.07
Raven's CPM raw score	26.15 (4.95)	26.33 (5.60)	.21	.83	

The DDLD group was significantly older than the TD group, $t_{(147)} = -4.30$, $p < .001$, $\eta_p^2 = .06$. This means that group comparisons should consider age in months as a covariate variable in all analyses. Further, the TD group significantly outperformed the DDLD group on NVIQ, $t_{(147)} = 3.48$, $p = .001$, $\eta_p^2 = .07$, and therefore NVIQ was statistically controlled in analyses. There were 23 females and 43 males in the DDLD group, in line with the well-replicated finding that children with dyslexia and DLD are predominantly male (e.g. for dyslexia: Snowling, 2000; for DLD: Leonard, 1998).

3.4. Descriptive measures

NVIQ is the best predictor of achievement in Grades 6 and 9, when other predictors are also investigated, such as self-efficacy, school environment, parental education and sex (Guez, Panaïotis, Peyre, & Ramus, 2018). Concomitantly, there is a general assumption that a low cognitive ability results in poor language and/or reading ability. Thus, a NVIQ task standardised in the Greek language was used to exclude children with a low cognitive ability.

Language, literacy and phonological processing skills were assessed using a wide range of tasks in order to profile the DDLD group's language, literacy and phonological difficulties. In the overall sample (*n* = 149), language skills were assessed with a widely used task of receptive vocabulary, in addition to tasks drawing upon a range of language processing skills, namely, verbal comprehension, syntax comprehension, and sentence repetition. Ziegler et al. (2010) argued that reading fluency performance, combining accuracy and speed, is a more sensitive index of reading difficulty than reading accuracy performance in consistent orthographies since reading accuracy rapidly approaches ceiling in these orthographies. Literacy skills were assessed with reading accuracy, text-reading fluency and spelling tasks. Reading accuracy and reading fluency are sensitive measures and can reveal difficulties in children who are reading in the Greek transparent orthography (Diamanti et al., 2018; Talli et al., 2016). Spelling performance is another sensitive index of reading difficulty in Greek (Porpodas, 1999; Protopapas & Skaloumbakas, 2007). Two literacy tasks were used in the overall sample, namely, text-reading fluency, as measured by l'Alouette, and spelling. However, given that there were no Greek standardised reading accuracy tasks designed for the age range of the study, first and second Graders (*n* = 47) were assessed with syllable and nonword reading tasks, and third to sixth Graders (*n* = 102) were assessed with reading accuracy and text-reading fluency tasks. In the overall sample, phonological processing skills were assessed with widely-used tasks assessing reading-related phonological processing skills, namely, phoneme deletion, NWR and RAN tasks, which reveal the typical phonological deficit in children with dyslexia and DLD (e.g. Diamanti, Goulandris, Stuart, Campbell, & Protopapas, 2018; Ramus et al., 2013). Further, phoneme deletion, NWR and RAN tasks account for significant amounts of variance in reading and spelling performance across orthographies, as has been reported by a number of large-scale (Moll et al., 2014; Ziegler et al., 2010) and small-scale (Caravolas, Volín, & Hulme, 2005) cross-linguistic studies in TD children in different orthographies. Landerl et al.'s (2013) large-scale cross-linguistic study also showed that phoneme deletion and RAN were the strongest predictors of dyslexia in 6 different languages varying in orthographic consistency. There is also evidence showing that in shallow orthographies, the phonological deficit can be identified most clearly using tasks that require implicit phonological processing such as phonological STM and RAN in that phoneme awareness deficits seem to be resolved by the end of the 2nd Grade (Landerl & Wimmer, 2000; Wimmer, 1996).

3.4.1. Nonverbal IQ (NVIQ) task

3.4.1.a. Raven's Coloured Progressive Matrices (CPM). The Greek standardization (Sideridis et al., 2015) of the Raven's CPM (Raven, 2008) was used. In this task, children were instructed to look carefully at a visual design with one missing part and to choose the image (from a choice of six) that fits the missing part of the design. In total, 1,042 Greek children aged between 4 years to 11 years and 11 months were selected by stratified random sampling for this standardization, with the task having good internal consistency (Cronbach's α = .867). One point was scored one point for each correct answer, and a total raw score was computed which was then converted to a standardised score.

3.4.2. Language ability tasks

3.4.2.a. WISC Similarities and Vocabulary subtasks. The Greek version (Georgas, Paraskevopoulos, Mpezevegkis, & Giannitsas, 1997) of the Wechsler Intelligence Scale for Children-III (WISC-III; Wechsler, 1991) was used. For the WISC Similarities subtask, children were instructed to identify how two words are alike ("How are piano and guitar alike?"). Responses scored two, one, or zero points, with the difference in scores reflecting the quality (accuracy and detail) of the response given (maximum score = 33). Internal consistency for the WISC Similarities subtask was computed using Cronbach's α; for 6- and 7-year-old children: α = .65, for 8-year-olds: α = .59, for 9-year-olds: α = .62, for 10-year-olds: α = .78, for 11-year-olds: α = .76, for 12-year-olds: α = .81. This subtask is also a good measure of general verbal ability, as it is highly correlated with the total scale (factor loadings: 74 for children aged 6 to 7 and 77 for those aged 8 to 10; Wechsler, 1991). For the WISC Vocabulary subtask, children were asked to listen carefully to some words and define them

(“What is a ‘clock’?”). Responses scored two, one, or zero points, with the difference in scores reflecting the quality of the definition given (maximum score = 33). Nation (2014) argues that even though a definition task (such as the WISC Vocabulary subtask) offers a more sensitive index of children’s word knowledge than a receptive vocabulary measure (such as the PPVT-R that follows), it also engages expressive skills and EFs, and therefore it cannot be deemed a ‘pure’ measure of children’s word knowledge. Internal consistency for the WISC Vocabulary subtask was computed using Cronbach’s α; for 6-year-old children: α = .62, for 7-year-old children: α = .68, for 8-year-olds: α = .76, for 9-year-olds: α = .81, for 10-year-olds: α = .84, for 11-year-olds: α = .83, for 12-year-olds: α = .81.

3.4.2.b. Peabody Picture Vocabulary Test-Revised (PPVT-R). The Greek non-standardised version (Simos, Sideridis, Protopapas, & Mouzaki, 2011) of the PPVT-R (Dunn & Dunn, 1981) was used as an estimate of receptive vocabulary. Children were provided orally with a word and had to decide which of the four pictures provided best represented its meaning. There was a possible total of 173 trials. Each response was scored with either one point (correct) or zero points (incorrect). The child’s score was the number of correctly selected pictures (maximum = 173). The task has good internal consistency (Cronbach’s α = .92-.98) and test-retest reliability (6 months; Pearson’s r = .65-.86).

3.4.2.c. Diagnostic Verbal Intelligence (DVIQ) Test’s syntax comprehension and sentence repetition subtasks. Syntax comprehension and sentence repetition subtasks of the DVIQ for school-age children (Stavrakaki & Tsimpli, 2000) were used. The syntax comprehension subtask assessed oral comprehension of syntax using 17 sentences. The child’s score was the number of correctly-selected pictures (maximum = 17). Children were instructed to listen carefully to a sentence and then to select from the response booklet the black and white picture (with a choice of either two or four alternatives) that best depicted the meaning of the sentence. These sentences begin very simply ('the boy gives her a card') and progress to more complex grammatical structures ('the man who is looking at the lady is embracing the old man'). Responses were scored 'online' as correct, one point, or incorrect, zero points. Sentence repetition used 10 sentences. Children heard a sentence and were asked to repeat it. Responses were scored 'offline' according to the manual, and three points were awarded for a completely correct response, two for a response with one error, one point for two errors, and zero points for three or more errors. The maximum possible score was 30.

3.4.3. Literacy ability tasks

3.4.3.a. L'Alouette task. L'Alouette's (Lefavrais, 1967) adaptation in Greek (Talli, 2010) was used to assess reading fluency (i.e. the speed of accurate reading). In France, l'Alouette is used in 6-12-year-old children as a screening test of reading difficulties evaluating accuracy and speed of text reading where semantic or pragmatic cues cannot be employed to aid word reading. Children were instructed to read as accurately and fast as possible a 271-word text bearing no meaning. The number of words read correctly within 3 min was recorded for each child. In its Greek adaptation, the text contains 179 high-frequency, 71 medium-frequency and 45 low-frequency words, most of which contain one or two syllables (respectively, 86 and 118 words), in addition to some words which have three and four syllables (respectively, 37 and 28 words) and just 2 words of five syllables. There is a proportion of misleading contextual information in the Greek version of the text with:

- words phonologically similar but semantically different to those that would be predicted by the context (e.g. the word 'κουτί' [box] instead of 'κουπί' [paddle] after the word 'ναύτες' [sailors]);

- slightly modified fixed expressions or slightly modified repetitions of an expression (e.g. ‘φεγγαράκι λαμπρό’ [bright moon] instead of ‘φεγγαράκι μου λαμπρό’ [my bright moon] coming from a traditional song which Greek children are taught at preschool age); and
- drawings including contextual errors (e.g. a drawing of a ‘δέμα’ [package] next to the written word ‘δέρμα’ [skin]).

More details about the Greek version of the text can be found in Talli (2010) and Talli et al. (2015). Text reading is discontinued after 3 minutes. Four measures were considered separately in this study for a more accurate presentation of children’s performance on this task, with the number of words read correctly being the explanatory variable used in further analyses. Figure 3.1. presents the number of words read correctly and the number of words read incorrectly in raw scores, in addition to the time spent for text reading in the TD group. It is evident that the number of words read correctly increased with age, that the number of words read incorrectly was small, and that most children spent the full 3 minutes on text reading, with only some older children being able to read the text in less than 3 minutes (with red shape fill in Figure 3.1.).

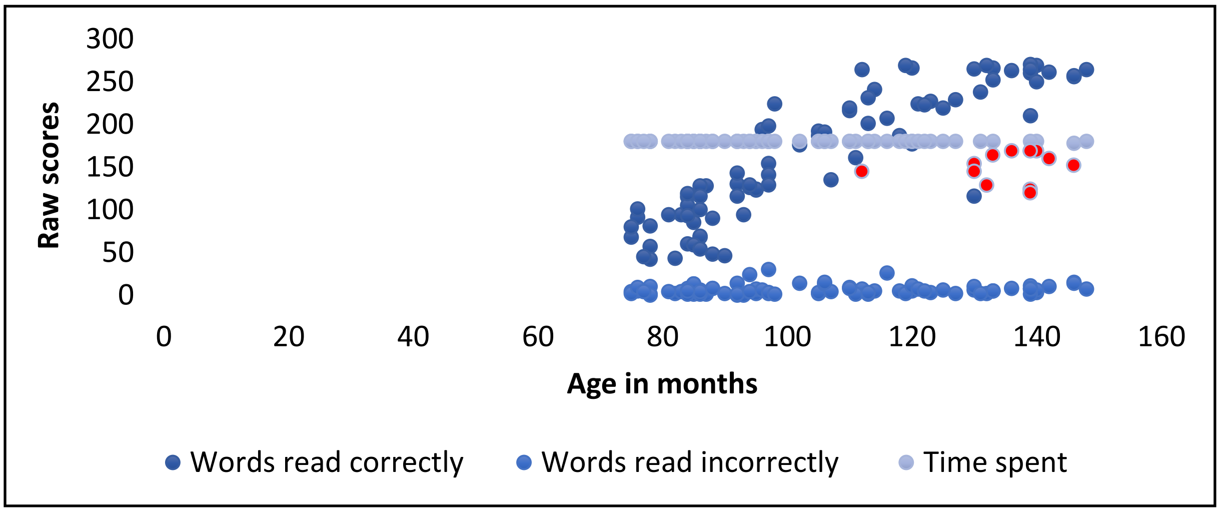


Figure 3.1. Number of words read correctly and words read incorrectly in raw scores, in addition to the time spent for text reading in the TD group, plotted against age in months

Note: with red shape fill are the children who read the entire text in less than 3 minutes.

3.4.3.b. Reading Test Alpha's reading accuracy and text-reading fluency subtasks (for children in Grade 3 onwards). The reading accuracy subtask of the Reading Test Alpha (Panteliadou & Antoniou, 2007) designed for children aged 8-15 years was used to assess reading accuracy in children in Grade 3 onwards. It includes three different subtasks assessing nonword reading, word reading and lexical decision. The first two subtasks consist of 24 nonwords (mean number of letters = 9.6, *SD* = 3.1) and 53 words (mean number of letters = 10.5, *SD* = 3.3). Words and nonwords are of increasing difficulty, according to the manual. The third subtask of Reading Test Alpha's reading accuracy subtask, the lexical decision subtask, consists of 16 nonwords (mean number of letters = 7.1, *SD* = 1.8) and 20 words (mean number of letters = 6.1, *SD* = 1.1) presented in arrays. Words and nonwords are intermixed and of increasing difficulty, according to the manual. Test-retest reliability for all three subtasks ranges between .74 and .87.

The children were asked to read as accurately as possible the presented nonwords and words, and also to report aloud only the real words of the lexical decision subtask. The number of nonwords and words read correctly, as well as the number of nonwords and words identified as such in the lexical decision subtask determines the reading accuracy score. The task has good test-retest reliability (r = .74-.87). The reading accuracy score was the number of words and nonwords read correctly (maximum = 77), alongside the number of words and nonwords identified as such in the lexical decision subtask (maximum = 36). The text-reading fluency task of the same test was used to assess reading fluency in children in Grade 3 onwards. It consists of a text of 279 words. The children were asked to read as accurately and fast as possible for a 60 s-period. The number of words read correctly within the time limit determines the reading fluency score. The task has good test-retest reliability (r = .74-.87). The reading fluency score was the number of words read correctly within the time limit.

3.4.3.c. Test of Detection and Investigation of Reading Difficulties' syllable and nonword reading subtasks (for children in Grades 1 and 2). The syllable reading and the nonword reading subtasks of the Test of Detection and Investigation of Reading Difficulties (Porpodas, 2007), or Test of DIRD for short, designed for first and second Graders were used. They consist of 24 syllables and 24 nonwords, respectively. There were three columns with 8 items each for the two subtasks. The difficulty of each column of syllables increases systematically from syllables without consonant clusters to syllables with clusters of two or three consonants, and the column of nonwords increases again systematically from nonwords of two syllables to nonwords of three or four syllables. Children were asked to read as accurately as possible syllables and nonwords, scored as correct (one point) or incorrect (zero points). A child's score for each subtask was the number of syllables and nonwords read correctly (maximum = 24, for each subtask).

3.5. Groups' descriptives of language, literacy and phonological tasks

DDLD and TD group's performance and group differences on every language, literacy and phonological variable are presented in this chapter to offer an indication of the profiles of children in the two groups.

3.5.1. Statistical methods. Groups' performance in raw scores for each language, literacy and phonological variable are presented as means and standard deviations (*SD*). *Z* scores are then presented for all language, literacy and phonological variables. *Z* scores were computed based on means and *SD*s of the TD group to illustrate the DDLD group's performance compared to the TD group's performance. Group differences on every language, literacy and phonological variable were tested by using regression analyses showing how much of the variance on performance in each language, literacy and phonological variable was accounted for by participant group, entered in Step 2, after controlling for age in months and NVIQ, entered in Step 1. In tables presenting results of regression analyses, Column 1 shows each explanatory variable of interest. Column 2 shows the amount of variance accounted for by age in months and NVIQ, and Column 3 shows the proportion of variance accounted for by the participant group. Columns 4, 5 and 6 show, respectively, *beta*-values (β) of age in months, NVIQ and group. Independent samples *t*-tests were used to compare statistically the two groups on age in months when needed. Paired samples *t*-tests were used to test within group differences between reading accuracy and reading fluency and between syllable reading and nonword reading. Univariate analyses of variance were used to reveal the effect sizes between the two groups for group differences between reading accuracy and reading fluency and between syllable reading and nonword reading.

3.5.2. Language tasks. Table 3.3. presents means (*SD*s) and 95% *CI*s (lower line) on language tasks for the DDLD and the TD group. The results showed that the TD group outperformed the DDLD group on every language task.

Table 3.3. Means (*SD*s) and 95% *CI*s (lower line) on language tasks for the DDLD and the TD group

Language tasks	*Max score*	DDLD group *M* (*SD*)	TD group *M* (*SD*)
Receptive vocabulary: PPVT-R	*173*	113.93 (18.24)	117.16 (19.81)
		(109.45-118.42)	(112.84-121.49)
Syntax comprehension: DVIQ Test	*17*	13.22 (2.25)	13.43 (2.49)
		(12.67-13.78)	(12.88-13.97)
Sentence repetition: DVIQ Test	*30*	23.84 (4.79)	26.77 (3.13)
		(22.66-25.02)	(26.08-27.45)
Verbal comprehension: WISC Similarities	*33*	9.33 (3.77)	10.20 (4.99)
		(8.40-10.26)	(9.11-11.29)
Verbal comprehension: WISC Vocabulary	*60*	17.00 (6.02)	19.85 (8.31)
		(15.51-18.48)	(18.03-21.67)

Notes: PPVT-R, Peabody Picture Vocabulary Test-Revised; DVIQ Test, Diagnostic Verbal Intelligence Test; WISC, Wechsler Intelligence Scale for Children.

Moreover, Figure 3.2. shows mean performance on language tasks in the DDLD group in *z* scores computed based on means and *SD*s of the TD group. From this Figure, it becomes apparent that the

DDLD group performed within -1 *SD* of the TD group's mean on all language tasks, with performance on sentence repetition being the lowest (-.93).

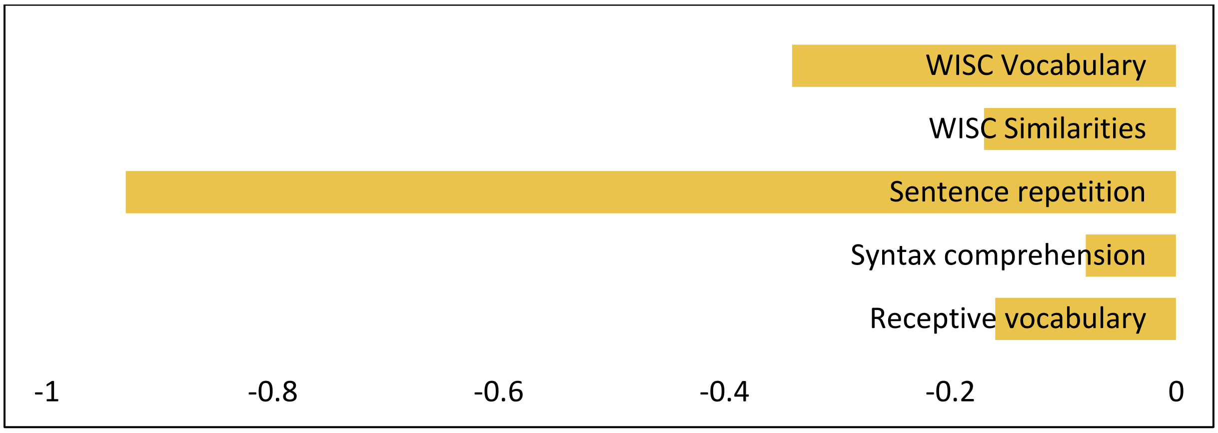


Figure 3.2. Mean performance on language tasks in the DDLD group in *z* scores

As shown in Table 3.4., regression analyses also revealed that age in months and NVIQ were significant predictor of almost all language tasks. Age in months was not a significant predictor of sentence repetition, however. Group was also a significant predictor of almost all language measures except for syntax comprehension.

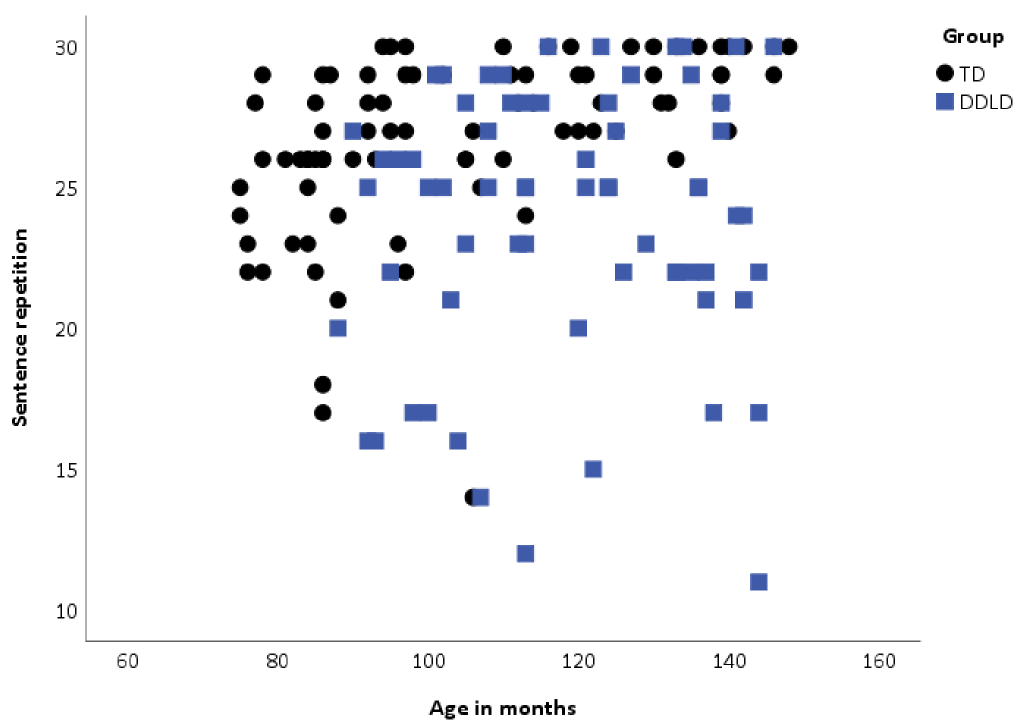
Table 3.4. Percentage of variance (R^2) on language tasks explained by age in months and NVIQ entered in Step 1 and proportion of variance (ΔR^2) explained by group entered in Step 2 in the overall sample

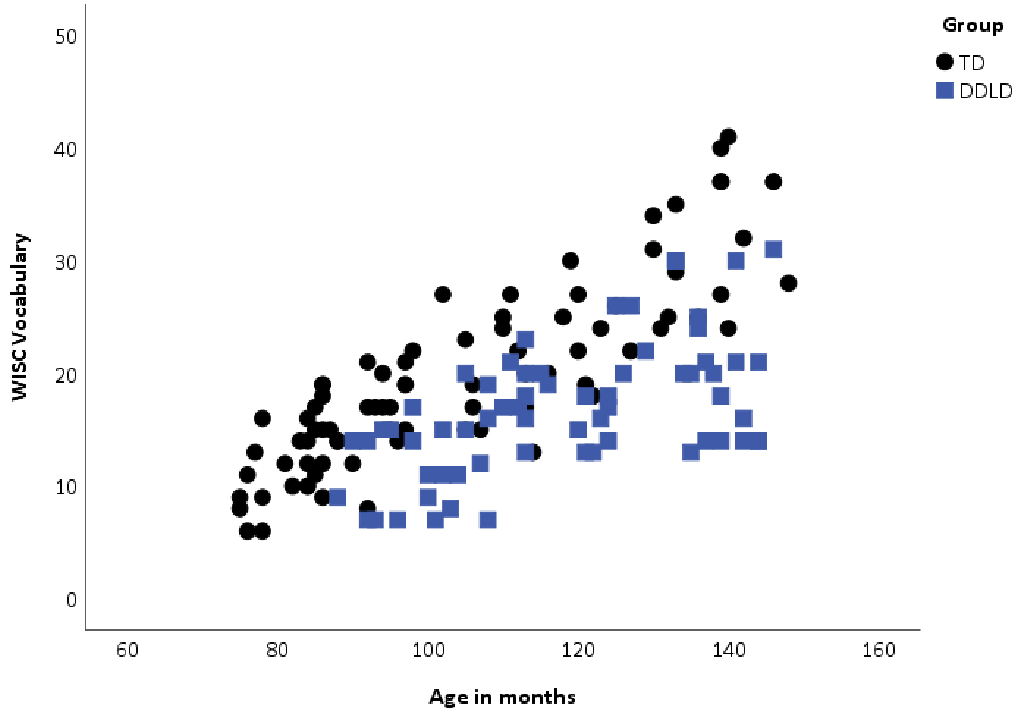
Explanatory variables	R^2	ΔR^2	β age	β NVIQ	β group
Receptive vocabulary	54.7***	.050***	.652***	.343***	-.248***
Syntax comprehension	38.7***	.008	.461***	.413***	-.098
Sentence repetition	9.1**	.118***	.121	.275**	-.380***
WISC Similarities	47.3***	.050***	.602***	.327***	-.248***
WISC Vocabulary	55.3***	.117***	.656***	.344***	-.378***

Notes: For each regression, age in months and NVIQ were entered in Step 1 and group was entered in Step 2. For Step 1, information is provided on the amount of variance accounted for by age in months and NVIQ (R^2), and for Step 2, information is provided on the proportion of variance accounted for by group (ΔR^2). The β-values for the three predictor variables are presented in Columns 4, 5 and 6. Significance values are given where they are relevant; $^{**}p < .01$, $^{***}p < .001$.

Checks indicated that Mahalanobis distances were less than 15 for all the children for all five language tasks, therefore all cases were included in regressions (Field, 2013). The regression model was significant for receptive vocabulary, $F_{(3, 145)} = 71.681$, $p < .001$, accounting for 57.7% of the variance, for syntax comprehension, $F_{(3, 145)} = 31.469$, $p < .001$, accounting for 39.4% of the variance, for sentence repetition, $F_{(3, 145)} = 12.759$, $p < .001$, accounting for 20.9% of the variance, for WISC Similarities, $F_{(3, 145)} = 53.005$, $p < .001$, accounting for 52.3% of the variance, and for WISC Vocabulary, $F_{(3, 145)} = 97.810$, $p < .001$, accounting for 66.9% of the variance.

It is evident from the analyses that of the language tasks, the largest amount of variance explained by group was for sentence repetition and WISC Vocabulary (respectively, 11.8 and 11.7%). Figures 3.3. and 3.4. show performance in sentence repetition and WISC Vocabulary in raw scores in the TD and DDLD groups, plotted against age in months. With regard to sentence repetition, it should be noted that many children in the TD group seem to be at ceiling in the task, and as such, the group difference may have been more pronounced if there had been more difficult sentences on this task.





Figures 3.3. and 3.4. Scatterplots showing performance in sentence repetition and WISC Vocabulary in raw scores in the TD and DDLD groups, plotted against age in months

3.5.3. Literacy tasks. Table 3.5. presents means (*SD*s) and 95% *CI*s (lower line) on literacy tasks for the DDLD and the TD group. The results showed that the TD group outperformed the DDLD group on every literacy task. In the Table below, the three sets of scores of word reading, nonword reading and lexical decision were subsumed in the reading accuracy score.

Table 3.5. Means (*SD*s) and 95% *CI*s (lower line) on literacy measures for the DDLD and the TD group

Literacy tasks	*Max score*	**DDLD group (*n* = 66)** *M* (*SD*)	**TD group (*n* = 83)** *M* (*SD*)
Text-reading fluency: L'Alouette, words read correctly	*271*	124.92 (51.12) (112.35-137.49)	162.16 (74.48) (145.90-178.43)
Text-reading fluency: L'Alouette, time spent (in sec)		179.54 (2.73) (178.87-180.21)	175.63 (12.44) (172.92-178.35)
Spelling ability: Spelling-to-dictation	*60*	18.31 (6.64) (16.68-19.95)	26.02 (13.91) (22.98-29.06)
		DDLD group (*n* = 56) *M* (*SD*)	**TD group (*n* = 46)** *M* (*SD*)
Reading accuracy: Reading Test Alpha	*113*	95.12 (12.00) (91.90-98.34)	105.00 (7.36) (102.81-107.18)
Word reading	*53*	44.75 (7.36) (42.77-46.72)	49.43 (4.02) (48.23-50.63)
Nonword reading	*24*	17.91 (4.12) (16.80-19.01)	20.93 (3.00) (20.04-21.82)
Lexical decision	*36*	32.80 (2.40) (32.15-33.44)	34.47 (1.62) (33.99-34.96)
Reading fluency: Reading Test Alpha	*279*	63.16 (26.14) (56.16-70.16)	91.86 (24.73) (84.52-99.21)
		DDLD group (*n* = 10) *M* (*SD*)	**TD group (*n* = 37)** *M* (*SD*)
Syllable reading: Test of DIRD	*24*	20.50 (4.45) (17.31-23.68)	22.32 (2.33) (21.54-23.10)
Nonword reading: Test of DIRD	*24*	17.50 (6.02) (13.19-21.80)	21.18 (2.41) (20.38-21.99)

Notes: Reading Test Alpha; Test of DIRD, Test of Detection and Investigation of Reading Difficulties.

Moreover, Figure 3.5. shows mean performance on literacy tasks in the DDLD group in *z* scores. From this Figure, it becomes apparent that the DDLD group performed within -1 *SD* of the TD group's mean on l'Alouette words read correctly, spelling and syllable reading and below -1 *SD* of the TD group's mean on reading accuracy, reading fluency and nonword reading, with performance on reading accuracy for older children and on nonword reading for younger children being the lowest (respectively, -1.34 and -1.52).

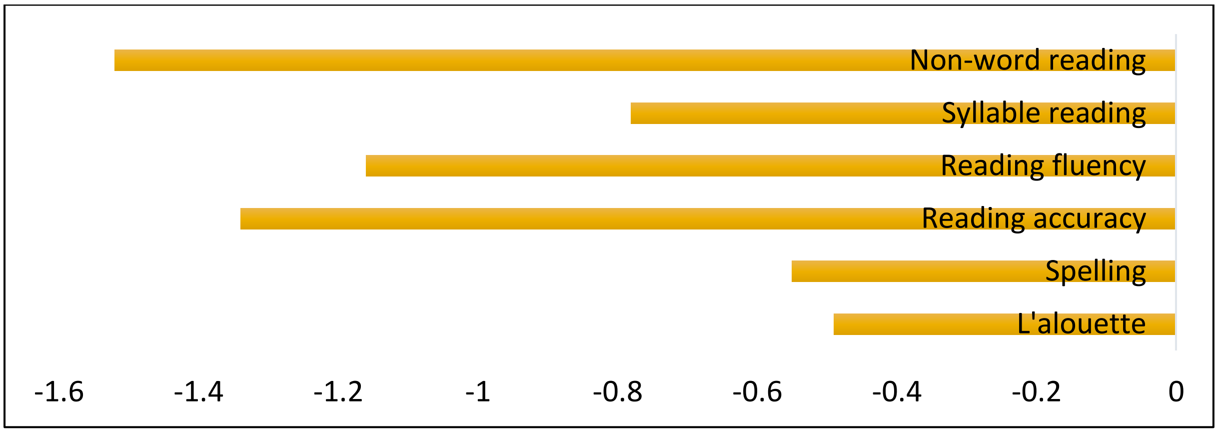


Figure 3.5. Mean performance on literacy tasks in the DDLD group in *z* scores

As shown in Table 3.6., regression analyses also revealed that age in months was a significant predictor of three literacy tasks, namely, l'Alouette, spelling and reading fluency. NVIQ was a significant predictor of almost all literacy tasks except for syllable reading. Group was also a significant predictor of almost all literacy tasks except for syllable reading. It should be noted that in the Table below, l'Alouette and spelling were the two tasks used in the overall sample. Reading accuracy and reading fluency were used in a sample of 102 children (i.e. 3rd Graders onwards), and syllable reading and nonword reading were used in a sample of 47 children (i.e. 1st and 2nd Graders).

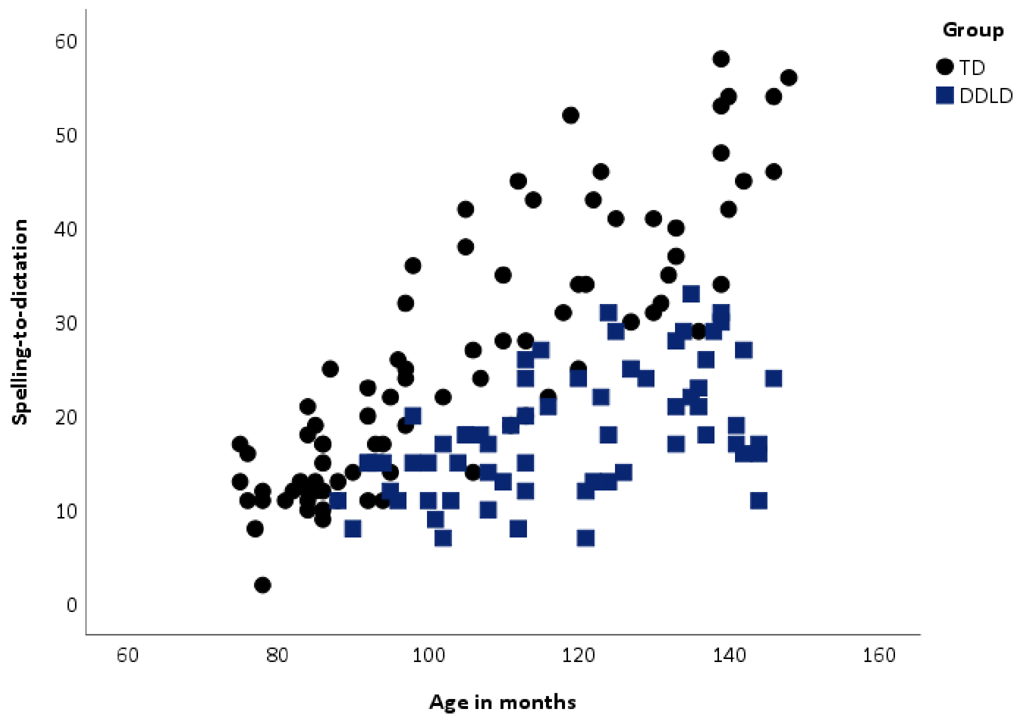
Table 3.6. Percentage of variance (R^2) on literacy tasks explained by age in months and NVIQ entered in Step 1 and proportion of variance (ΔR^2) explained by group entered in Step 2 in the overall sample

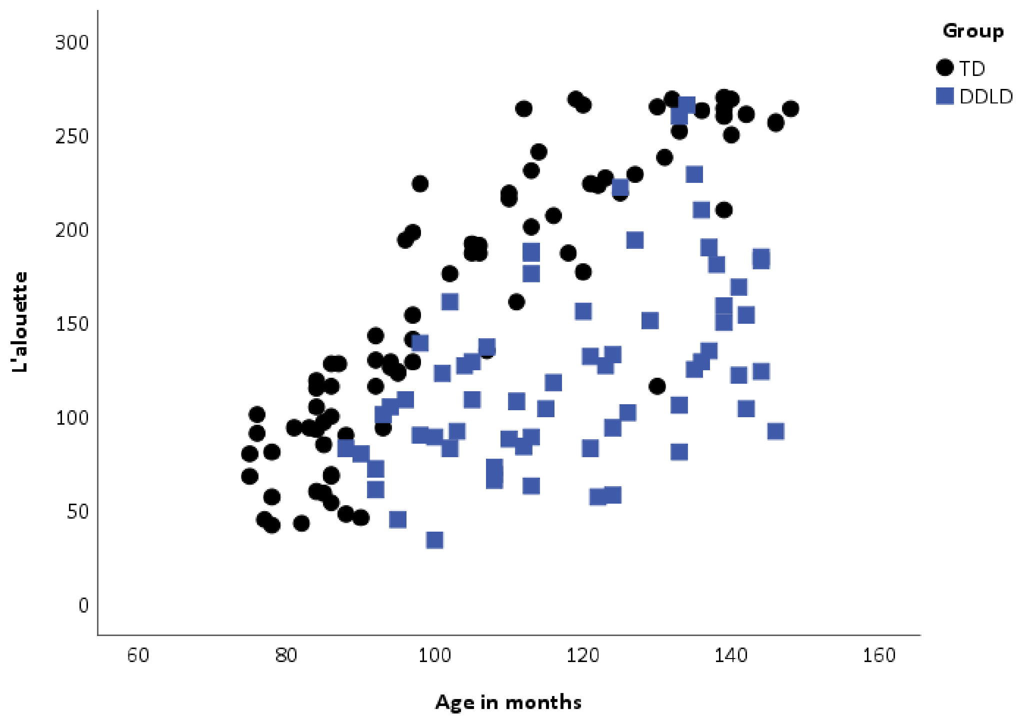
Explanatory variables	*N*	R^2	ΔR^2	*β* age	*β* NVIQ	*β* group
L'Alouette	149	39.5^{***}	$.235^{***}$	$.613^{***}$	$.132^{*}$	$-.537^{***}$
Spelling	149	38.1^{***}	$.258^{***}$	$.583^{***}$	$.198^{**}$	$-.562^{***}$
Reading accuracy	102	10.3^{**}	$.128^{***}$	.039	$.317^{**}$	$-.379^{***}$
Reading fluency	102	16.9^{***}	$.203^{***}$	$.305^{**}$	$.267^{**}$	$-.478^{***}$
Syllable reading	47	.05	.055	-.060	.230	-.294
Nonword reading	47	11.1	$.112^{*}$	-.169	$.300^{*}$	$-.420^{*}$

Notes: For each regression, age in months and NVIQ were entered in Step 1 and group was entered in Step 2. For Step 1, information is provided on the amount of variance accounted for by age in months and NVIQ (R^2), and for Step 2, information is provided on the proportion of variance accounted for by group (ΔR^2). The *β*-values for the three predictor variables are presented in Columns 4, 5 and 6. Significance values are given where they are relevant; $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Checks indicated that Mahalanobis distances were less than 15 for all the children for all six literacy tasks, therefore all cases were included in regressions (Field, 2013). The regression model was significant for l'Alouette, $F_{(3, 145)} = 82.321$, $p < .001$, accounting for 63% of the variance, for spelling, $F_{(3, 145)} = 85.354$, $p < .001$, accounting for 63.8% of the variance, for reading accuracy, $F_{(3, 98)} = 11.881$, $p < .001$, accounting for 26.7% of the variance, for reading fluency, $F_{(3, 98)} = 32.904$, $p < .001$, accounting for 50.2% of the variance, and for nonword reading, $F_{(3, 43)} = 4.106$, $p = .012$, accounting for 22.3% of the variance. The model was nonsignificant for syllable reading, $F_{(3, 43)} = 1.449$, $p = .242$.

It is evident from the analyses that of the literacy tasks, the largest amount of variance explained by group was for spelling and l'Alouette (respectively, 25.8 and 23.5%). Figures 3.6. and 3.7. show performance in spelling and l'Alouette in raw scores in the TD and DDLD groups, plotted against age in months. The spelling task had many difficult items, and l'Alouette task comprised a text with no meaning and misleading contexts, and therefore both tasks were difficult for children with DDLD.





Figures 3.6. and 3.7. Scatterplots showing performance in spelling and l'Alouette in raw scores in the TD and DDLD groups, plotted against age in months

3.5.4. Phonological tasks. Table 3.7. presents means (*SD*s) and 95% *CI*s (lower line) on phonological measures for the DDLD and the TD group. The results showed that the TD group outperformed the DDLD group on every phonological measure.

Table 3.7. Means (*SD*s) and 95% *CI*s (lower line) on phonological measures for the DDLD and the TD group

Phonological measures	*Max score*	DDLD group *M* (*SD*)	TD group *M* (*SD*)
Phoneme deletion measures: EVALEC			
Accuracy of CVCVCV items	*10*	6.84 (2.45)	8.07 (2.01)
		(6.23-7.45)	(7.63-8.51)
Speed of CVCVCV items (in sec)		59.19 (22.87)	55.99 (27.95)
		(53.56-64.81)	(49.81-62.17)
Accuracy of CVC items	*12*	11.01 (1.29)	10.96 (1.37)
		(10.69-11.33)	(10.66-11.26)
Speed of CVC items (in sec)		43.67 (20.70)	47.03 (29.17)
		(38.58-48.76)	(40.58-53.48)
Accuracy of CCV items	*12*	9.13 (2.81)	10.30 (2.04)
		(8.44-9.83)	(9.85-10.74)
Speed of CCV items (in sec)		59.51 (21.46)	56.34 (29.31)
		(54.23-64.79)	(49.86-62.82)
NWR measures: EVALEC			
Accuracy	*24*	14.08 (3.88)	17.66 (4.13)
		(13.08-15.07)	(16.75-18.56)
Speed (in sec)		92.78 (12.90)	91.28 (12.39)
		(89.48-96.09)	(88.57-93.99)
RAN measure: PhAB			
Composite score (in sec)		140.50 (43.76)	118.72 (42.02)
		(129.74-151.25)	(109.54-127.89)

Notes: EVALEC, Evaluation de la Lecture; PhAB, Phonological Assessment Battery.

Figure 3.8. shows mean performance on phonological measures in the DDLD group in *z* scores. From this Figure, it becomes apparent that the DDLD group performed within -1 *SD* of the mean on all phonological measures. Specifically, the DDLD group performed within -1 *SD* of the mean on a composite accuracy score of phoneme deletion tasks of CVCVCV, CVC and CCV items (phonology accuracy variable), a composite speed score of phoneme deletion tasks of CVCVCV, CVC and CCV items (phonology speed variable), RAN, NWR accuracy and NWR speed. Performance on NWR accuracy was -.82, performance on RAN was-.51 and performance on phonology accuracy was -.48.

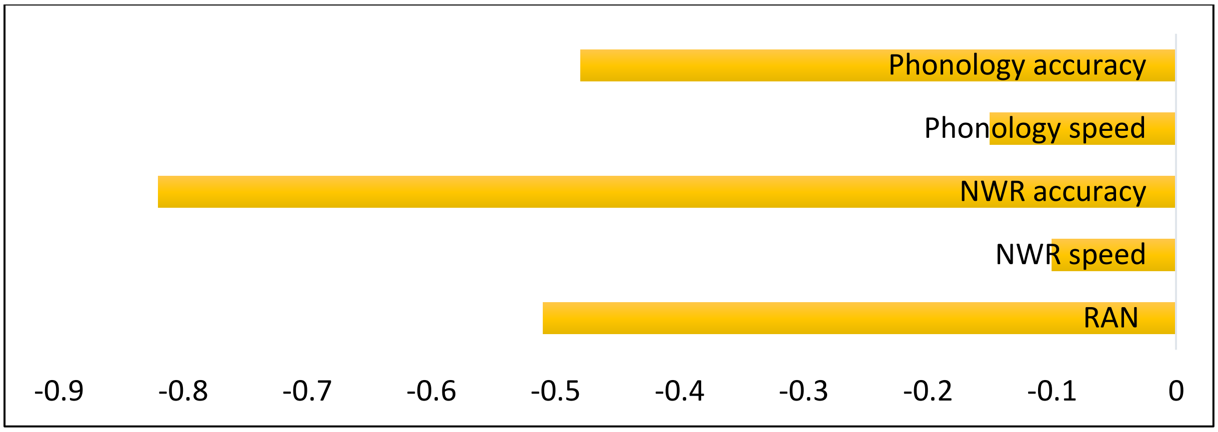


Figure 3.8. Mean performance on phonological measures in the DDLD group in *z* scores

Checks indicated that Mahalanobis distances were less than 15 for all the children for all phonological measures, therefore all cases were included in regressions (Field, 2013). The regression model was significant for accuracy of CVCVCV items, $F_{(3, 144)} = 15.158$, $p < .001$, accounting for 24% of the variance, for accuracy of CCV items, $F_{(3, 145)} = 10.606$, $p < .001$, accounting for 18% of the variance, for NWR accuracy, $F_{(3, 142)} = 21.424$, $p < .001$, accounting for 31.2% of the variance. For the speeded measures, the model was significant for speed of CVCVCV items, $F_{(3, 144)} = 20.000$, $p < .001$, accounting for 29.4% of the variance, for speed of CVC items, $F_{(3, 145)} = 18.261$, $p < .001$, accounting for 27.4% of the variance, for speed of CCV items, $F_{(3, 144)} = 19.439$, $p < .001$, accounting for 28.8% of the variance, for NWR speed, $F_{(3, 142)} = 5.173$, $p = .002$, accounting for 9.9% of the variance, and for RAN speed composite score, $F_{(3, 145)} = 39.000$, $p < .001$, accounting for 44.7% of the variance. The model was nonsignificant for accuracy of CVC items, $F_{(3, 145)} = 1.511$, $p = .214$.

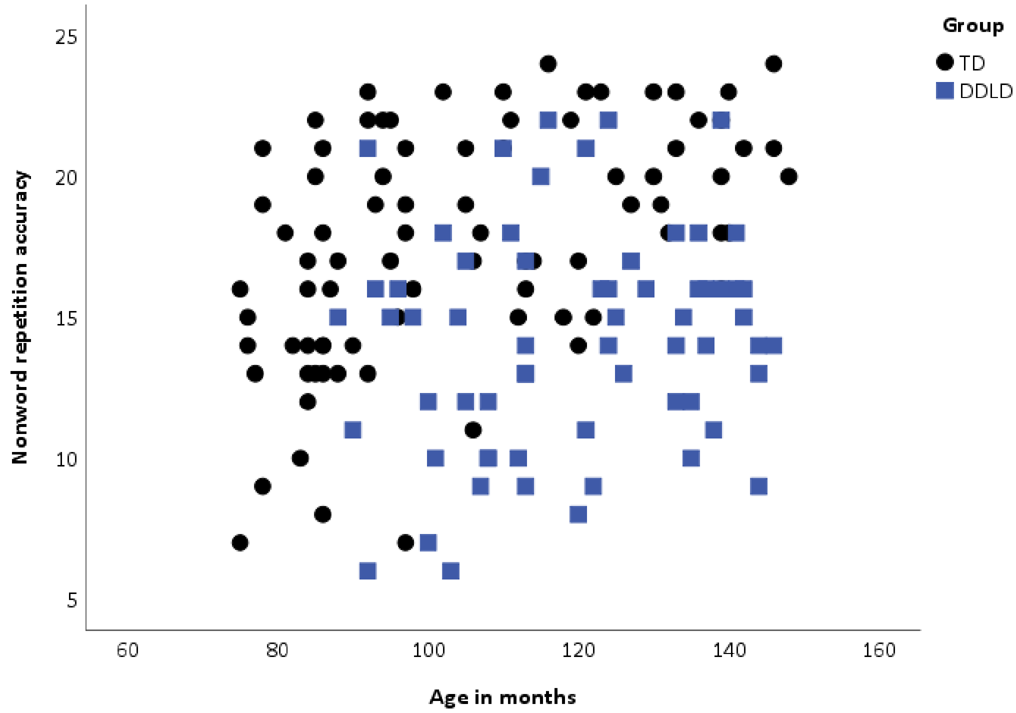
As shown in Table 3.8., age in months was a significant predictor of almost all phonological measures except for accuracy of CVC items. NVIQ was a significant predictor of almost all phonological measures except for accuracy of CVC items, accuracy and speed of CCV items, and NWR speed. Group was also a significant predictor of almost all phonological measures except for accuracy and speed of CVC items and NWR speed.

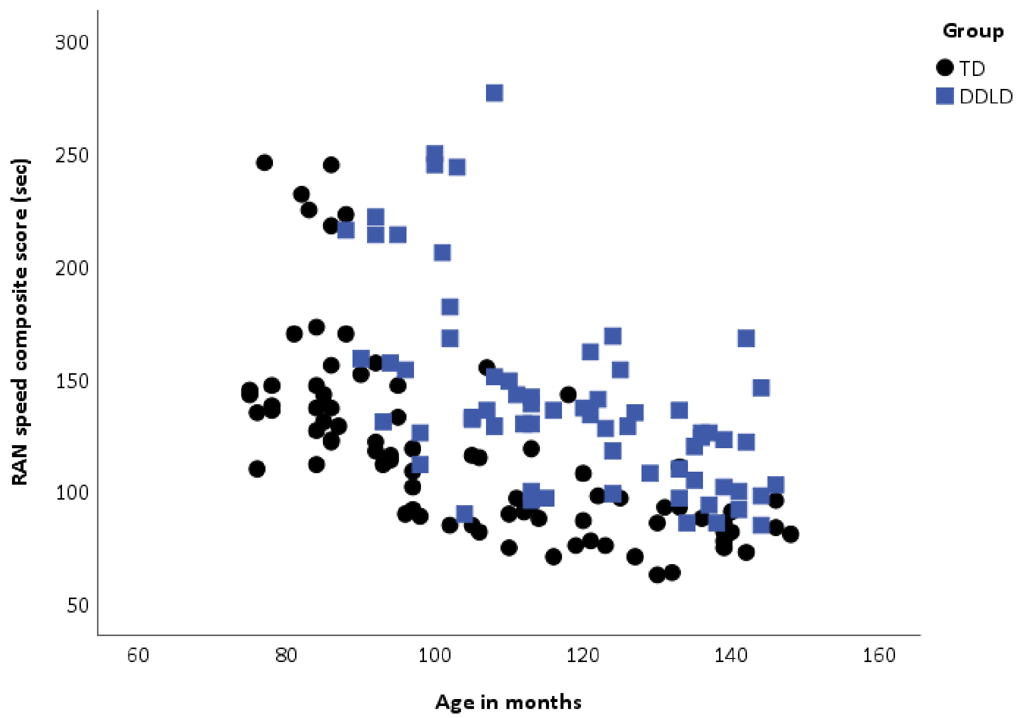
Table 3.8. Percentage of variance (R^2) on phonological measures explained by age in months and NVIQ entered in Step 1 and proportion of variance (ΔR^2) explained by group entered in Step 2 in the overall sample

Explanatory variables	R^2	ΔR^2	β age	β NVIQ	β group
Accuracy of CVCVCV items	15.1***	.091***	.256**	.291***	-.335***
Speed of CVCVCV items	25.6***	.041**	-.478***	-.165*	.224**
Accuracy of CVC items	3.0	.000	.154	.078	-.012
Speed of CVC items	27.2***	.002	-.478***	-.203**	.054
Accuracy of CCV items	8.5**	.095***	.252**	.144	-.341***
Speed of CCV items	23.5***	.055**	-.481***	-.063	.260**
NWR accuracy	16.4***	.160***	.189*	.359***	-.443***
NWR speed	8.0**	.019	-.274**	-.069	.155
RAN speed composite score	29.0***	.157***	-.502***	-.190**	.439***

Notes: For each regression, age in months and NVIQ were entered in Step 1 and group was entered in Step 2. For Step 1, information is provided on the amount of variance accounted for by age in months and NVIQ (R^2), and for Step 2, information is provided on the proportion of variance accounted for by group (ΔR^2). The β-values for the three predictor variables are presented in Columns 4, 5 and 6. Significance values are given where they are relevant; $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

It is evident from the analyses that of the phonological tasks, the largest amount of variance explained by group was for NWR accuracy and RAN speed composite score (respectively, 16 and 15.7%). Figures 3.9. and 3.10. show accuracy performance in NWR and RAN speed composite score in raw scores in the TD and DDLD groups, plotted against age in months.





Figures 3.9. and 3.10. Scatterplots showing accuracy performance in NWR and RAN in raw scores in the TD and DDLD groups, plotted against age in months

3.5.5. Summary of descriptives of language, literacy and phonological tasks

3.5.5.a. Group differences on language, literacy and phonological tasks. Regression analyses with age in months and NVIQ entered in Step 1 and group entered in Step 2 revealed that group was a significant predictor of almost all language tasks except for syntax comprehension, a significant predictor of almost all literacy tasks except for syllable reading, and a significant predictor of almost all phonological measures except for accuracy and speed of CVC items and NWR speed. Table 3.9. shows the proportion of variance accounted for by group for language, literacy and phonological measures and *p*-values of significance.

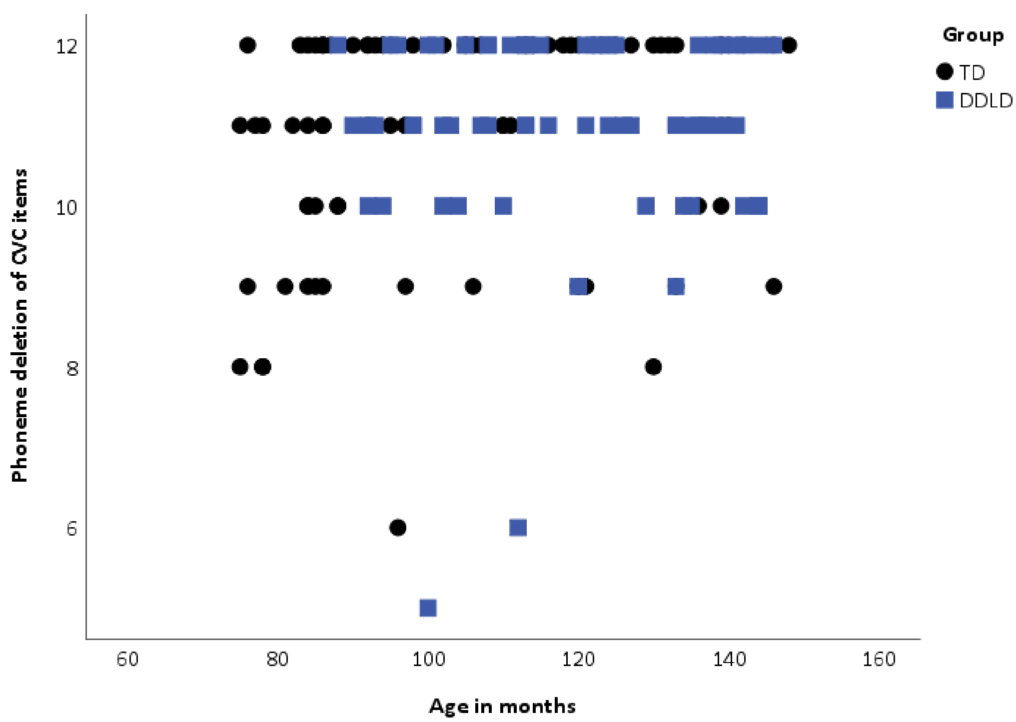
Table 3.9. Proportion of variance accounted for by group for language, literacy and phonological measures and *p*-values of significance after controlling for age in months and NVIQ

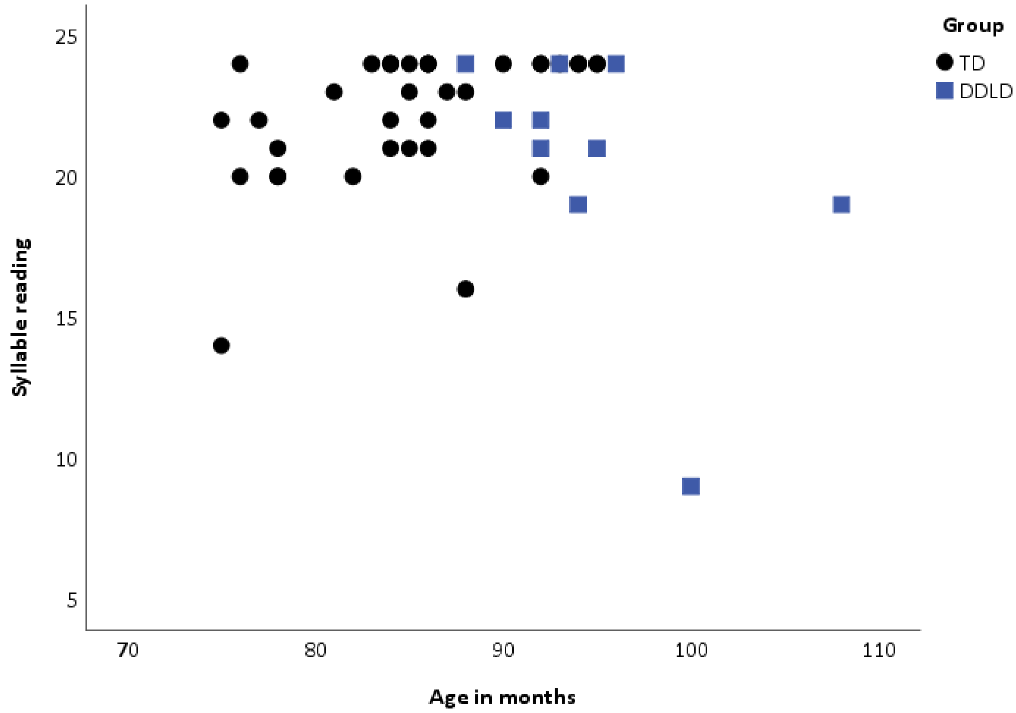
Language measures	ΔR^2	*p*
Receptive vocabulary: PPVT-R	.050	< .001
Syntax comprehension: DVIQ Test	.008	n.s.
Sentence repetition: DVIQ Test	.118	< .001
Verbal comprehension: WISC Similarities	.050	< .001
Verbal comprehension: WISC Vocabulary	.117	< .001
Literacy measures		
Text-reading fluency: L'Alouette	.235	< .001
Spelling ability: Spelling-to-dictation	.258	< .001
Reading accuracy: Reading Test Alpha	.128	< .001
Reading fluency: Reading Test Alpha	.203	< .001
Syllable reading: Test of DIRD	.055	n.s.
Nonword reading: Test of DIRD	.112	.012
Phonological measures		
CVCVCV items accuracy: EVALEC	.091	< .001
CVCVCV items speed: EVALEC	.041	.004
CVC items accuracy: EVALEC	.000	n.s.
CVC items speed: EVALEC	.002	n.s.
CCV items accuracy: EVALEC	.095	< .001
CCV items speed: EVALEC	.055	.001
NWR accuracy: EVALEC	.160	< .001
NWR speed: EVALEC	.019	n.s.
RAN speed composite score: PhAB	.157	< .001

3.5.5.b. Task limitations. The aim of this section is to explore the tasks that did not show significant group differences. Phoneme deletion of CVC items and syllable reading showed limited ability to discriminate between participants at the higher range of the distribution because the tasks suffer from ceiling effects. Further below are presented these analyses in detail.

Phoneme deletion of CVC items. The task did not correlate with age in months in the overall sample, $r_{(149)} = .15$, $p = .05$, in the DDLD group, $r_{(66)} = .05$, $p = .64$, and in the TD group, $r_{(83)} = .21$, $p = .05$. Further, a large proportion of children obtained maximum scores in this task. This was an easy task, so both groups scored high and there was no group difference. An extensive overlap was also found between the two groups, as Figure 3.11. shows.

Syllable reading. The task did not correlate with age in months in the overall sample, $r_{(47)} = -.04$, $p = .78$, because of ceiling effects (see Figure 3.12.). Analyses by subgroup revealed a nonsignificant correlation between the task and age in months in the DDLD group, $r_{(10)} = -.51$, $p = .13$, but a moderate respective association in the TD group, $r_{(37)} = .44$, $p = .006$. This was also an easy task, which is why there were no group differences.





Figures 3.11. and 3.12. Scatterplots showing performance in phoneme deletion of CVC items and syllable reading in raw scores in the TD and DDLD groups, plotted against age in months

3.6. Experimental measures

Verbal fluency categories, namely, semantic and phonological, a design fluency task, three phoneme deletion tasks, a NWR, a RAN task and a spelling-to-dictation task were used as experimental measures. It should be noted that the three phoneme deletion tasks, the NWR task, the RAN task, and the spelling-to-dictation task were also used as descriptive measures to assess children's phonology and spelling performance. More fine-grained analyses derived from the data of all those tasks were carried out when they have been used as experimental measures.

3.6.1. Verbal fluency tasks

3.6.1.a. Semantic fluency tasks. Semantic fluency used the semantic categories 'ζώα' ('animals'), 'τρόφιμα' ('foods') and 'πράγματα του σπιτιού' ('objects from around the house'; hereafter 'objects'), in that order. Children were instructed to produce as many different words belonging to the target category as possible, allowing 60 s for each category. No examples were given, but 'χώρες' ('countries') was used as a practice category. The first two categories were chosen because they are the most widely used categories in the spoken language literature (e.g. Kosmidis et al., 2004; Nash & Snowling, 2008). The last category was chosen on the basis that it would be an easy category, when it comes to word productivity, for most children. Moreover, all three categories are natural, that is, well defined categories with clear boundaries separating category members from non-members, allowing category members to be classified as such unambiguously. Specifically, Martin (2007) argues that there is a distinction between living (animals and plants) and non-living (artifacts) objects. The number of correct responses retrieved for the three semantic categories was combined to create a composite score of semantic fluency. Gender differences have been noted for some categories, as it has been reported, for example, by Woods et al. (2016a) for the categories of 'cars' and 'tools'. Woods et al. also reported that females are generally better than males in the semantic category of 'fruits', while for the semantic category of 'animals', a nonsignificant difference between males and females has been reported. It is expected, therefore, that the semantic categories used in the current study will not reveal gender differences.

3.6.1.b. Phonological fluency tasks. Phonological fluency used the letters 'χ' (chi), 'σ' (sigma) and 'α' (alpha) of the Greek alphabet, in that order. Children were instructed to produce as many different words beginning with the target letter as possible, allowing 60 s for each letter. The letter 'τ' (tau) was provided as an example. The particular letters were chosen because they have been previously used in a large study of verbal fluency in Greek adults (Kosmidis et al., 2004). Kosmidis et al. (2004) reported that the particular letters were chosen based on the ratio of words in Greek beginning with 'χ', 'σ' and 'α' relative to the total number of words in a Greek dictionary, which corresponds to the ratio of words in English beginning with the letters 'f', 'a' and 's' relative to the total number of words in an English dictionary. The letters 'f', 'a' and 's' are the set of letters most commonly used in English. The number of correct responses retrieved for the three phonological categories was combined to create a composite score of phonological fluency.

3.6.1.c. Coding of verbal fluency responses

Responses in semantic and phonological fluency tasks were audio-recorded using Audacity for Windows 7, and they were then entered, after transcription, into an Excel database, timed (i.e. it was calculated how many words produced in the first 15 s and the subsequent 45 s of the minute), and coded as correct or incorrect. Correct words produced after the 60 s test period were not included in the number of correct responses.

There were 4 types of incorrect responses scoring zero points:

- repeated responses (see below which lexical items were considered as repeated);
- made-up responses (e.g. 'ανοποιός');
- out-of-category responses (i.e. real but irrelevant words to the target category, such as producing 'ναύτης' [naftis]; i.e. <sailor> in the category of the letter alpha); and
- unintelligible responses (i.e. made-up words or words which could not be transcribed).

All other responses were judged as correct and scored one point each. Correct responses in phonological categories, beginning with the target letter or with a letter having the same sound in Greek with the target letter as specified below, were also:

- expressions of two words functioning as an adverb (e.g. 'σίμα-σίμα' <side-by-side>);
- two words produced together functioning as a noun (e.g. 'χιονοδρομικό κέντρο' <ski center>);
- foreign words used in Greek (e.g. 'Samsung', 'CD', 'snowboard');
- two words connected with an apostrophe in written language produced as such (e.g. 'άσ'το' originating from 'άσε το' <let it be>, or 'σ'αγαπώ' originating from 'σε αγαπώ' <I love you>);

- two words produced together with the first functioning as a preposition that complements the meaning of verbs, adjectives, or nouns (e.g. 'σε ξέρω' <I know you>, 'σε σένα' <to you>);
- words beginning with the target letter in written language, but this letter is part of a digraph representing a different sound in oral language (e.g. 'αίνιγμα' <enigma> is pronounced 'enigma', or 'αίθουσα' and 'Αιτωλοακαρνανία' where again their first digraph is orally pronounced as /e/);
- idiomatic words (e.g. 'χαψί' known as 'ψάρι' <fish> in common Greek); or
- relevant to semantic categories, subcategory names (e.g. 'mammal', 'fish', 'fruits') even when followed words in the subcategory (e.g. 'cow', 'trout', 'apple').

Aside from those words repeated exactly as before, all regular inflections were counted as correct responses: (*a*) different verb forms (e.g. 'αγοράζω' followed by 'αγόρασα' <I buy, I bought>); (*b*) different noun forms (e.g. 'μήλο' <apple>, a singular noun, followed by 'μήλα' <apples>, a plural noun); (*c*) different adjective forms for the positive, the comparative and the superlative (e.g. 'ευγενικός', 'ευγενέστερος', 'ευγενέστατος' <polite, more polite, most polite>); and (*d*) different pronoun forms (e.g. 'άλλος' <another>, a pronoun associated with a masculine gender, followed by 'άλλη', a pronoun associated with a feminine gender). Thus, two or more words in the sequence that comprise variant forms of the same lexeme were counted as correct responses, unlike previous coding reported in the literature considering these as repeated responses and thus scoring zero points (e.g. Kosmidis et al., 2004). The rationale for this was that children were asked to try to avoid producing the same word, but they were not instructed to avoid producing different forms of a word (e.g. a singular form of a word followed by a plural form of the same word in a row). Although repeated, made-up, out-of-category and unintelligible responses were excluded from the number of correct responses, they were included in the number of total responses, indicating a child's word productivity. Repeated, made-up and out-of-category responses were included in any cluster analysis, described in the next paragraph, as they might have aided children to retrieve words classified into clusters. For example, in the phonological fluency condition, a child might have retrieved the word ‘άλλος’ which was counted by the experimenter as a repeated word; however, the word ‘άλλος’ was followed by the word ‘Αλίκη’, and together the two words are a phonological cluster. It is considered therefore that an incorrect word (repeated in the example given above) might have aided the child to retrieve another phonologically-related word immediately afterwards. Wrongly articulated responses were also counted as correct responses, since there was sufficiently unambiguous evidence that a correct word has been retrieved (e.g. areoplano). Correct and incorrect responses were therefore coded according to semantic or phonological clusters using the coding instructions described thereafter.

3.6.1.d. Semantic clusters. The number of semantic clusters was computed for each semantic category, where a semantic cluster was considered as two or more successive responses belonging to a conventional subcategory. For example, in the category of ‘animals’, PIG followed by CHICKEN followed by GOAT belonged to the subcategory of ‘farm animals’. Cluster size was also computed for each semantic category. In the following sequence, for example, PIG, CHICKEN, GOAT, SHEEP, WHALE, OCTOPUS, PARROT, SPARROW and SWALLOW, three semantic clusters can be identified: ‘farm animals’, ‘water animals’ and ‘birds’, in that order, with cluster sizes of 4, 2 and 3, respectively, and an average cluster size of 3 items (as measured by taking the total number of items in a cluster and dividing it by the total number of clusters identified).

In determining semantic clusters, the guidelines provided in the studies by Kosmidis et al. (2004), Marshall et al. (2013) and Troyer et al. (1997) were followed. For example, ‘foods’ subcategories included, amongst others: ‘fruits’, ‘vegetables’, ‘meat’, ‘carbohydrates’, ‘sweets’, ‘breakfast foods’, ‘dairy’, ‘legumes’, ‘nuts’, ‘fish’ and ‘liquid food’. However, unlike a priori subcategory methods followed by Troyer et al. (1997), subcategories were driven by the data and were non-exclusive. This means that for semantic clusters, as both thematic (e.g. pets, farm animals) and taxonomic (e.g.

birds, fruits, vegetables) clusters identified from children's responses, an item could fall into one subcategory on one occasion, but to a different subcategory on another; that is, the subcategory was chosen based on the responses that immediately preceded and followed each item.

3.6.1.e. Computational modelling. Although clustering patterns shed light on lexical organization, do not allow researchers to investigate individuals' semantic network. As Kenett et al. (2013) argue this is due to the random nature of both the retrieval exhaustion of subcategories, and the switching process between subcategories. To this end, computational modelling was also used to investigate any group differences in children's semantic network using solely the semantic category of animals (Davelaar, Marshall, & Mengisidou, in preparation). Recent work used the knowledge originating from the computational methods and developed a new metric to capture the likelihood of a speaker producing a particular sequence in the fluency task in order to investigate the underlying structure of semantic memory as measured using semantic fluency tasks (Davelaar et al., 2015). This offers the opportunity to compare networks across different groups in an alternative way. This new metric was used in this study too. Analysis proceeded as follows. Sequences were combined into a single network by overlaying individual paths. In order to analyse the animal fluency data, we converted the responses for all participating children into an Excel database. This database was constructed such that the first column contains each child's code, with all responses of a single child having the same code in the column (e.g. 1001), as Table 3.10. shows. In the second column (identified with number 1 in Table 3.10.), the sequence of responses produced by a child is given. In the third column (identified with number 2 in Table 3.10.), the sequence of responses produced by a child is again given, however, in this column, the first response of each child is omitted. That is, each row of the Table presents two successive responses in the task.

Table 3.10. An example for animal fluency data preparation of the first participating child

Code	1	2
1001	Elephant	Pig
1001	Pig	Hen
1001	Hen	Goat
1001	Goat	Sheep
1001	Sheep	Cat
1001	Cat	Dog
1001	Dog	Mouse
1001	Mouse	Tiger
1001	Tiger	Cheetah
1001	Cheetah	Flamingo
1001	Flamingo	Frog
1001	Frog	Whale
1001	Whale	Octopus
1001	Octopus	Parrot
1001	Parrot	Sparrow
1001	Sparrow	Swallow
1001	Swallow	Flying fish

The resulting network (i.e. reference network) was directed and weighted. The following could then be computed: (i) a maximum likely sequence through this network; and (ii) for each individual's sequence, the likelihood of being produced by this network. As such, a ratio score for each individual is the metric of interest, namely, own sequence likelihood divided by maximum likely sequence. The closer this number is to 1, the more typical the individual's sequence is. The ratio was calculated (or controlled) for sequence length and the reference network was conditioned on the time allocated for producing the sequences. Moreover, the ratio was based on the log of the likelihoods. This leads to very unlikely sequences (in the case of people with dementia as reported by Davelaar et al., 2015) to have ratios larger than 1. Thus, the ratio score goes from 1 to some large number.

3.6.1.f. Phonological clusters. The number of phonological clusters was computed for each phonological category, where a phonological cluster was considered to be two or more successive responses that could be classified into the following types of cluster:

- words that shared the same first syllable (e.g. σέλα-σελήνη), or words that shared the same first two or more letters (e.g. σκάω-σκεπή) or sounds (e.g. χελώνα-χαίτη-χέλι) irrespective of spelling;
- words that differed only in a single vowel sound irrespective of spelling, e.g. Σίσυ-σούσι, σήμα-σώμα, σήμα-σύρμα, ανοίγω-ανήκω, Χίος-χάος;
- words that differed only in a single consonant sound, αδύναμος-αδύνατος, e.g. σόμπα-σόλα, στήνω-στίβω, χήρα-χήνα;
- words that shared exactly the same phonemes but spelled and pronounced slightly differently because of a different graph used for a vowel sound and a different syllable stress, e.g. χώρος-χορός;
- words that were homographs spelled the same but differed in syllable stress, and therefore pronounced differently, e.g. Σταύρος-σταυρός <Stavros, given name, and cross>;
- words that were homophones pronounced the same but spelled differently, e.g. αυτή-αυτί <she-ear>, χοίρος-χήρος <widowed-pig>; and
- words that were homonyms both spelled and pronounced the same, but with different meanings, and they were often identified as such by children when produced, e.g. by saying "Αγγελική, the given name, and αγγελική, the plant".

In classifying words into phonological clusters, the experimenter tried to be as inclusive as possible (Mielnik et al., 2015). For example, the following sequence of words, 'χαριτωμένο-χαρτί-χαρτοπετσέτα' (cute-paper-napkin), was identified as a phonological cluster of three words, even though there is more phonological overlap between the second and the third word (χαρτ) than between the last two words with the first word (χαρ). Another example is the following, 'άλογο-άλμα-αλφαβήτα-άλφα’ <horse-jump-ABC-alpha>, in which all four words were counted as belonging to one single phonological cluster sharing the first two phonemes, even though there is more overlap between the two items (‘αλφαβήτα-άλφα’) sharing the first four phonemes than between the other two items in the sequence (‘άλογο-άλμα’) sharing the first two phonemes. Repeated responses were counted in computing the number and the size of semantic and phonological clusters. The rationale is that even repeated responses might have aided children’s semantic and phonological clustering. Given that neither intrusions nor unintelligible responses could contribute to a cluster, they were not relevant for computing the number and the size of clusters.

For semantic and phonological fluency tasks, the following composite scores based on all three semantic categories and all three phonological categories were computed:

- Total number of responses inclusive of errors
- Total number of responses in the first 15 s of the test period
- Total number of responses in the subsequent 45 s of the test period
- Number of correct responses
- Number of incorrect responses, namely,
    i. repeated
    ii. made-up
    iii. out-of-category
    iv. unintelligible
- Number of switches, where switches were counted as the number of transitions between the semantic or phonological clusters but also between non-clustered responses
- Number of clusters
- Average cluster size. Cluster size was counted beginning with the first word in a cluster (i.e. a two-word cluster was given a size of 2, a three-word cluster a size of 3, etc.). In order to

compute average cluster size, a mean cluster size for each of the three categories was computed. That is, for each semantic and phonological category, the number of responses belonging to clusters was divided by the number of clusters, e.g. 4 responses belonging to clusters divided by 2 clusters produced, gives a mean cluster size equal to 2). Then, the three mean cluster sizes of the three categories were summed and that number was divided by the number of categories, namely, 3, as shown in the parenthesis (e.g. mean cluster size of the first category = 2 + mean cluster size of the second category = 2.33 + mean cluster size of the third category = 2 divided by the number of categories, namely, 6.33/3 = an average cluster size of 2.11).

Two examples of coding of verbal fluency responses can be found in the Appendix (D). Moreover, with respect to the number of incorrect responses, a measure of proportion of errors will be computed in the current study. Henry et al. (2015) criticised Weckerly et al. (2001) for reporting that children with DLD showed no differences in the numbers of overall errors compared to controls. Henry et al. argued that Weckerly et al. did not take into account overall level of performance, and that a measure of proportion of errors is more appropriate. It should be noted that the first 15 s and the subsequent 45 s of the test period are numbers of total responses including incorrect responses. Even though this increases somehow the number of responses in the first 15 s and the subsequent 45 s of the test period, considering that the number of incorrect responses is very low, it is argued that the results would not look different if such responses were excluded.

3.6.2. The design fluency task

The English version of the NEuroPSYchological Assessment (NEPSY-II) design fluency subtask (Korkman, Kirk, & Kemp, 1998) designed for children aged 6-11 years was used. The NEPSY-II design fluency subtask is a nonverbal fluency test that measures EFs. The test consists of two response booklets. Each booklet contains 35 five-dot designs arranged in five columns and five rows. Each booklet has a different stimulus design, with the five-dot design in structured or random arrays. Children were given 60 s for each page to create as many different designs as fast as they can by connecting two or more dots in each square. Timing started after the child started drawing the first design. The experimenter explicitly emphasized to children that they can connect only two, only three, only four, or all five dots to create a design by drawing on his own the respective designs. Children were required to include at least 1-line design in each box, with a possibility of 10 correct (unique) 1-line designs in each array, and therefore more designs would be expected on the NEPSY design fluency subtask, where 1-line designs are permitted, than on 4-line design fluency tasks. Woods, Wyma, Herron, and Yund (2016b) found that each line adds approximately 1 s to design completion time using a computerized test of design fluency in adults. This means that a 5-line design would require roughly 4 s more than a 1-line design.

The children were instructed as follows: “In every box, connect two or more dots with straight lines. Work as quickly as you can, and make every design different. Start here (the experimenter points to the upper left box relative to the child) and go this way (indicating left to right). When you will have finished with this row, go to the next one (pointing to the next line). Remember to make all designs different, make your lines straight and connect the dots. Ready? Begin”. Pointing to the upper left box and left to right direction for each row in the response booklet was important in order to determine if repeated designs had occurred. The experimenter kept notes after the first design not following left to right direction in order to compute the right number of repeated designs. As previous designs performed remain visible, they can be used as cues for performing subsequent designs and avoiding repeated designs. For each array, the experimenter performed two designs as examples of correct designs in the practice booklets, and then the child performed two designs as practice trials. After each practice trial, feedback was provided indicating the type of error that occurred, if any. Children were allowed to reproduce the designs that were used for practice purposes. Each correct design scored one point, and repeated designs, trials with curved lines, trials in which lines did not connect dots and trials with non-continuous lines scored zero points. In the current study, performance on the NEPSY-II design fluency subtask is expressed as the number of total, correct, incorrect and repeated designs of both booklets. The relationship between the total number of correct designs and the total number of incorrect designs (incorrect and repeated) is sometimes expressed as an error ratio: the total number of incorrect designs is divided by the total number of correct designs. This error ratio is reported in the current study. The NEPSY-II design fluency subtask is based on Regard et al. (1982). The mean test-retest reliability of the task is .59. Examples of scoring of design fluency responses can be found in the Appendix (E).

3.6.3. Phonological ability tasks

Two types of phonological tasks—namely, phoneme deletion and RAN, which are widely used to test the phonological processing skills of children with dyslexia and DLD, were used to potentially tease apart the two phonological hypotheses of dyslexia and DLD. To this end, the third subtask of the deletion tasks described below was used as a phoneme deletion rather than as a syllable deletion task as it was initially designed for, and in the RAN task, the number of phonological errors found was computed.

3.6.3.a. Phoneme deletion tasks. Three phoneme deletion subtasks of the computerised battery Evaluation de la Lecture (EVALEC; Sprenger-Charolles et al., 2005) adapted into Greek by Talli (2010) were used. The first two subtasks contain 12 nonwords of one syllable each, the one with items with a simple CVC syllable structure and the other with items with a complex CCV syllable structure. The third subtask contains 10 nonwords of three syllables with items with a simple CVCVCV syllable structure. Children had to produce the word without the initial consonant, with three examples given for each subtask. Responses were scored as correct (one point) or incorrect (zero points), and the total time to complete each subtask was also measured. EVALEC computerised battery estimates precisely how many millisec children spend on each sub-lexical stimulus presented, and it also provides the total time (in sec) children spend on each of the three phoneme deletion tasks as a whole. There will be no concerns therefore about the reliability of the manipulation time data considered in this study of these sub-lexical stimuli, and also of the time data of the NWR task of the same battery presented below. A child's accuracy score was the total number of correct responses in each subtask, and a child's speed score was the total time (in sec) to complete each subtask.

3.6.3.b. Nonword repetition (NWR) task. The NWR subtask of the computerised battery EVALEC (Sprenger-Charolles et al., 2005) adapted into Greek by Talli (2010) was used to assess phonological STM. The subtask contains 24 items with equal numbers (six each) of nonwords with three, four, five and six syllables. Three nonwords of each length used syllables with a CVC syllable structure, in addition to syllables with a CV structure, while the remaining three used a CV syllable structure only. Nonwords are presented according to their length, that is, three-syllable nonwords are presented first, followed by four-, five-, and six-syllable nonwords. Resemblance to real words was avoided by not including grammatical morphemes. Children were instructed as follows: "Please listen carefully to some made-up words and repeat them as better as you can. Now you will hear the first word. Ready?". Responses were scored as correct (one point) or incorrect (zero points), and the total time to complete the subtask was measured. The number of nonwords repeated correctly was the child's score. Verbal memory is a critical skill for a successful completion of phoneme deletion tasks in that Diamanti et al. (2018) reported no lexicality effect (or interaction) in phoneme deletion tasks in Greek and English children with dyslexia.

3.6.3.c. Rapid automatic naming (RAN) task. The picture naming subtask of the Phonological Assessment Battery (PhAB; Frederickson, Frith, & Reason, 1997) was used. Each of the two cards contains five pictures (table, door, ball, hat and box) repeated ten times on each card. RAN was designed to assess children's speed of phonological production, involving retrieval of phonological coding at the whole word level. Children were instructed to name the pictures as fast as possible while trying not to make any mistakes. The score was determined by taking the mean naming time of the two cards. Any semantic (e.g. 'desk' instead of table) or phonological errors (e.g. 'πότρα' instead of 'πόρτα' [door]) and any omissions were reported 'offline'. Concomitantly, accurate picture naming was considered to be that which was semantically and phonologically correct. However, of greater interest was the time children spent to name the pictures in the two cards, a variable which has been reported to be significantly related to children's reading fluency performance in transparent orthographies (e.g. Wimmer, Mayringer, & Landerl, 1998). RAN was assessed using pictures, because the ability to name pictures (or colours) is assumed to be less dependent on reading level than the ability to name letters or digits (Parrila, Kirby, & McQuarrie, 2004). Children's scores were the average naming time (in sec) taken for the two cards, the raw numbers of semantic and phonological errors, and the raw number of omissions.

3.6.4. The spelling-to-dictation task

The spelling-to-dictation task developed for Greek students in Grades 2-6 was used (Mouzaki & Protopapas, 2010; Mouzaki, Sideridis, Protopapas, & Simos, 2007; Sideridis et al., 2008). The task comprises 60 words presented orally in the context of a short sentence. Words included in the task are nouns, verbs, adjectives, pronouns, conjunctions, adverbs, prepositions and participles. First the word is read aloud, then the sentence including the target word, and then again the word, and therefore the children write each word after it has been read three times by the experimenter. Any word with correct spelling scored one point and the task was discontinued after six consecutive errors. The task has very good psychometric characteristics (internal consistency in the overall sample: Cronbach's α = .95; test-retest reliability one year later: r = .91; internal consistency for Grade 2 onwards: respectively, α = .89, α = .93, α = .94, α = .95, and α = .94). The order of words on the basis of their difficulty strongly correlated with the order the words are administered (in the overall sample, ρ = .94), and this was confirmed by the use of modern techniques of analysis (Rasch model), as reported by Mouzaki et al. (2007). This confirmation is very important as the task is administered with a discontinue rule.

According to Mouzaki et al. (2007), single words were chosen from primary school reading primers, and they included a wide range of morpho-syntactic rules. Words chosen were prone to phonological, grammatical and orthographic spelling errors. Even the simplest word, first in the list (από), can be misspelled, preserving, however, its phonological structure. With respect to phonological errors, considered to be those that change the phonological structure of a word, there are opportunities, using this spelling task, for omission or substitution of graphemes, and for simplification or inversion of digraphs (i.e. letter combinations such as 'μπ', 'ου' and 'αι' used to represent phonemes). Moreover, on the basis that the alternative spellings for the vowels are governed by morpho-syntactic rules (e.g. the first person of verbs ends with the vowel grapheme <-ω> /o/, while nouns end with <-ο> /o/), chosen words are also prone to grammatical spelling errors, in the case that the child does not know the appropriate spelling of the vowel grapheme. Given that many sounds, and in particular vowels, can be written in different ways, and given that the different ways that many sounds can be written do not always depend on grammatical knowledge, chosen words are also prone to orthographic spelling errors (e.g. the first 'ο' in 'κόπος' cannot be predicted either by the knowledge of the part of speech of this word or the word's morphological type, and there is no way for one who does not know this word and its correct spelling to guess how this word can be spelled correctly). Mouzaki et al. (2007) also reported that because of their large number, stress assignment errors or omissions have not been counted for this task.

3.6.4.a. Error classification

In this study, following the classification system of spelling errors proposed by Protopapas et al. (2013), spelling errors were classified into three broad categories of errors: phonological (grapho-phoneme mappings), grammatical (inflectional suffixes) and orthographic (word stems). Stress assignment errors were counted in this study (e.g. *ταμείο* /tami' o/ spelled τά̲μειο) concerning the stress diacritic, which obligatorily marks the vowel of the stressed syllable in every Greek word with two or more syllables. As such, a stress assignment error alters the word's pronunciation by putting the stress diacritic in a vowel sound which is not that of the stressed syllable in a particular word. However, stress omissions were not counted because of their large number and also because the author considers that children tended to omit the stress not because they did not know which part of a word is stressed - they usually stressed correctly when they were asked to - but because they seemed not to consider it so important in their spelling. Indeed, it has been reported that many children, both TD children and children with dyslexia, omit the stress diacritic in their spelling (Anastasiou & Protopapas, 2015; Protopapas et al., 2013). Protopapas (2017) argues that this might be because the stress diacritic does not seem to facilitate lexical access and disambiguation. Unclassifiable errors defined as miscellaneous infrequent errors, such as mirrored letters (e.g. *έτσι* /etsi/ spelled '3τσι) were also not considered further as they were very rare in this sample.

Protopapas and his colleagues identified for each of the 60 words in the task which graphemes or parts of words (e.g. derivational morphemes or inflectional suffixes) could be possibly misspelled, and into which category of errors each possible misspelled grapheme or part of word could be classified. This was explicit for the categories of grammatical and orthographic errors (e.g. a grammatical error in a noun, an adjective or a verb is always in its inflectional suffix, or an orthographic error in a noun, an adjective or a verb is always in its stem) but not for the category of phonological errors since phonological errors could occur on any grapheme or digraph and on any syllable in a word, as discussed further below.

3.6.4.b. Phonological errors. The first error category concerned phonological errors, or errors that were phonologically implausible. A phonological error occurs where the orthographic representation does not map onto the word's phonological representation. Specifically, each sound can be written in one or more specific ways. For example, sound [ð] is always written with letter 'δ'. In contrast, sound [v] can be written with letter 'β' (καράβι), double letter 'ββ' (Σάββατο), letter 'υ' (αυγή), or with letter combination 'υβ' (ευβοϊκός). Phonological errors alter the word's phonological form, so that the written word is pronounced differently from the one intended. Καράφι (KARAFI instead of KARAVI <ship>) and αυχή (AVCHI instead of AVGI <dawn>) are phonologically incorrect. Thus, if one reads the written word and it sounds like the correct word, then the written word is considered to be phonologically correct, and the misspelling is *not* counted as a phonological error. For example, αβγή written with the letter 'β' instead of the correct letter 'υ', still sounds like AVGI <dawn>; Σά_βατο written with the letter 'β' instead of the correct double letter 'ββ', still sounds like SAVATO <Saturday>. However, those spelling errors that affect the pronunciation of the word, altering its phonological structure by misusing the phoneme-to-grapheme mappings of Greek, are considered phonological errors. Only phonologically implausible errors belong to this category, and not errors that are phonologically plausible, considered to be those containing existing phoneme-to-grapheme mappings in Greek (e.g. είναι spelled είνε, or δίχτυ spelled δύχτι, where the sound /i/ is used but with an alternative grapheme). The major phonological category consists of 8 minor subcategories: substitution of syllables or phonemes, insertion of syllables or phonemes, omission of syllables or phonemes, and inversion of syllables or phonemes. In the list of words used for this study, even the simplest words, first in the list (από, έλα, και), can be misspelled, preserving, however, their phonological structure (*απώ* instead of από, *αίλα* instead of έλα, *κε* instead of και). Phonological errors shed light on children's sub-lexical, or phono-graphemic skills. Phonological errors are relatively rare in children with dyslexia as reported by studies in the Greek orthography

(Nikolopoulos et al., 2003; Niolaki & Masterson, 2013; Protopapas et al., 2013). Table 3.11. shows examples of minor subcategories of phonological errors.

Table 3.11. Examples of minor subcategories of phonological errors using the word ‘άλογο’ <horse>

Word	Substitution		Insertion		Omission		Inversion	
	Syllable	Phoneme	Syllable	Phoneme	Syllable	Phoneme	Syllable	Phoneme
Άλογο	A**ka**go	Al**a**go	A**ka**logo	Alogo**s**	A_go	A_ogo	Ago**lo**	Ao**l**go

Following the classification system proposed by Protopapas et al. (2013), a word could be either phonologically correct or phonologically incorrect; thus, phonological errors, also referred to as phonologically implausible errors in the literature, were considered in the current study. There were many different words that resulted in the production of phonological errors with the children altering the phonological structure of those words in many different ways. In this point, a clarification should be made. In the previous paragraph, the author gave as an example of phonologically correct spellings the words ‘αβγή’ and ‘Σά_βατο’. Another example can be the word ‘αυτός’ spelled as ‘αφτός’. Indeed, the three words sound like the correct words and therefore can be considered as phonologically correct. One might consider though that those are types of phonologically plausible spelling errors, considered those spelling errors which use an alternative grapheme for the same phoneme without changing the phonological structure of the word; however, they are not since those specific graphemes consist parts of orthographic knowledge and therefore spelling errors in those parts should be classified into the category of orthographic errors discussed below. Following the classification system proposed by Protopapas et al. (2013) therefore and using the specific spelling task, the experimenter could not find errors referred to in the literature as phonologically plausible errors. Table 3.12. shows examples of phonological errors found.

Table 3.12. Examples of phonological errors where the phonological errors are shown in bold and omitted grapheme is shown with _ in each word

‘Αυ**α**τός’ instead of ‘αυτός’
‘**Τ**οτίζο’ instead of ‘ποτίζω’
‘Δί**κ**_ι’ instead of ‘δίχτυ’

3.6.4.c. Grammatical errors. The second error category related to grammar. A grammatical error alters the word’s written representation by substituting alternative graphemes for the same phonemes, and therefore grammatical spelling errors concern alternative, phonologically equivalent (i.e. they maintain the word’s correct pronunciation), spellings of inflectional suffixes. Specifically, many times the correct spelling of a word depends on the part of speech the word belongs to or the grammatical type of the word in a specific context. Grammatical errors are considered those errors where the written word does not depict correctly its grammatical type (part of speech and inflection). For example, κόπος (kopos; toil) is a noun and ‘ος’ has to be written with an ‘ο’ (omicron) not an ‘ω’ (omega), and κάπως (kapos; somehow) is an adverb and ‘ως’ has to be written with an ‘ω’ not an ‘ο’. Κόπως and κάπος are therefore phonologically correct but grammatically incorrect. Thus, grammatical errors maintain the word’s correct pronunciation but alter its written representation by substituting alternative graphemes for the same phonemes in that part of the word which depicts a word’s grammatical type (as in κόπος or κάπως).

As shown in Table 3.13., the major grammatical category consists of 4 minor subcategories: error in a noun, error in an adjective/pronoun and error in a verb, with all three belonging to the category of inflectional suffixes, and error in an uninflected suffix. An error type in an inflectional suffix concerns graphemes with multiple phonologically equivalent spellings in inflectional suffixes of inflected parts of speech (articles, verbs, nouns, adjectives, pronouns). An error type in an uninflected suffix concerns graphemes with multiple phonologically equivalent spellings in inflectional suffixes of uninflected parts of speech, such as adverbs and gerunds. Grammatical errors shed light on children's mastery of inflectional morphology. Grammatical errors are more frequent than phonological errors in Greek children with dyslexia, as reported by Protopapas et al. (2013).

Table 3.13. Examples of minor subcategories of grammatical errors using the words 'άλογο' <horse>, 'αυτός' <he>, 'είναι' <is> and 'ξεφυλλίζοντας' <flipping through>

Words	Inflected suffixes			Uninflected suffix
	Noun	Pronoun	Verb	Participle
Άλογ**ο**	**ο**			
Αυτ**ός**		**ός**		
Είν**αι**			**αι**	
Ξεφυλλίζ**ο**ντας				**ο**

3.6.4.d. Orthographic errors. The third and final category of errors related to orthographic errors. An orthographic error alters the word's written representation by substituting alternative graphemes for the same phonemes, and therefore orthographic spelling errors concern alternative, phonologically equivalent (i.e. they maintain the word's correct pronunciation), spellings of word stems, including roots and any derivational morphemes preceding the obligatory inflectional suffix. For example, the first /o/ sound in κόπος (kopos) has to be written with an 'ο' (omicron) not an 'ω' (omega). Thus, these errors are phonologically correct (e.g. κώπος instead of κόπος (toil); δύχτυ for δίχτυ (net); or ποτήζω rather than ποτίζω (water)) but orthographically incorrect. This orthographic spelling is not dependent on grammatical knowledge - it cannot be predicted by the morphological type of the word and it cannot be justified in any way. It is applied, however, to all words belonging to the same family with a common origin; e.g. κοπιαστικός (adjective; tiring), κοπιάζω (verb; toil), άκοπα (adverb; effortlessly).

As Table 3.14. shows, the major orthographic category consists of 4 minor subcategories: error of a thematic/rule, error of a thematic/exception, error of an etymological vowel and error of an etymological consonant. Thematic/rule refers to graphemes with multiple phonologically equivalent spellings in derivational morphemes taught in school as rules (e.g. 'ποτίζω' is a verb in which 'ίζω' is a derivational morpheme written always with 'ί' as in almost all other verbs with the same derivational morpheme; thus, ποτήζω is considered as a misspelling of a thematic/rule). Thematic/exception refers to graphemes with multiple phonologically equivalent spellings in derivational morphemes violating school rules, or taught exceptions (e.g. 'δίχτυ' is a noun in which 'υ' is a derivational morpheme written with 'υ' and not with 'ί' as one might expect given that nouns usually end with 'ί'; thus, 'δίχτι' is considered as a misspelling of a thematic/exception). Error of an etymological vowel refers to an error in a root vowel grapheme with multiple phonologically equivalent spellings, and error of an etymological consonant refers to an error in double consonant letters, as shown in Table 3.14. Orthographic errors shed light on children's knowledge of word-specific (or root-specific) knowledge. Orthographic errors are also more frequent than phonological errors in Greek children with dyslexia, as reported by Protopapas et al. (2013).

Table 3.14. Examples of minor subcategories of orthographic errors using the words 'ποτίζω' <water>, 'δίχτυ' <net>, δανείζω <loan>, αυτός <he>, 'χείμαρρος' <torrent> and φωτισμένος <lit>

Words	Thematic/rule	Thematic/exception	Etymological Vowel	Etymological Consonant
Ποτ**ί**ζω	**ί**		**ο**	
Δ**ί**χτ**υ**		**υ**	**ί**	
Δαν**εί**ζω		**εί**		
Α**υ**τός				**υ** (sounds f)
Χ**εί**μα**ρρ**ος			**εί**	**ρρ**
Φ**ωτισ**μ**έ**νος		**ι, ε**	**ω**	**σ**

Thus, an error in a word can be classified according to the above-presented categories of errors. Moreover, misspelled words can have more than one error each and therefore a word can be both phonologically and grammatically incorrect, or both phonologically and orthographically incorrect or both grammatically and orthographically incorrect, or even phonologically, grammatically and orthographically incorrect. Concomitantly, a word spelled correctly is a word which is phonologically, grammatically and orthographically correct. Three composite scores were computed: one for phonological errors, one for grammatical, and one for orthographic errors. Composite scores of phonological, grammatical and orthographic errors were computed by summing the number of errors in all minor subcategories for each of the three major categories. Initially therefore, errors were classified according to those minor subcategories and then a composite score of phonological, grammatical and orthographic errors was computed. Given that not all the children attempted all the words of the list, group comparisons on composite scores of phonological, grammatical and orthographic errors were conducted using proportional numbers of composites scores, proportionally to the number of total words spelled (e.g. for each child, the composite score of phonological errors was divided by the number of total words spelled and then multiplied by 100). Specifically, the raw numbers for each of the major categories of errors was estimated. Next, each of the three raw numbers was divided by the total number of words spelled, with this leading to the three proportional scores of phonological, grammatical and orthographic errors.

Moreover, comparisons were conducted between groups only and not within groups, namely, the DDLD group’s proportional numbers of the three categories of errors was compared to TD group’s respective numbers; however, within group comparisons of the three categories of errors were considered inappropriate since they were not independent. This is because there were children who produced phonological errors, for example, and the same children produced grammatical and orthographic errors, and as such, the means of the three composite scores were not independent of each other. Another relevant issue is that if there were parts in a word that are more difficult than other parts, this cannot be ruled out in the analyses. For example, there might be several vowel phonemes in a stem in a word that can be misspelled but only one vowel phoneme in the inflectional suffix that can be misspelled. This makes comparisons among the different categories of errors a difficult task. As Table 3.15. shows, indeed there were more opportunities for orthographic than grammatical errors, and therefore it would not be surprising to find in the sample of the current study, more orthographic than grammatical errors. Moreover, as Table 3.15. shows, as children attempt more and therefore more difficult words in the list, they are more likely to produce more grammatical and orthographic errors since more difficult words offer more error opportunities for grammatical and orthographic errors (see in Table 3.15. the total number of error opportunities in the first 30 words and the subsequent 30 words of the list). Diamanti et al. (2014) found that misspelled inflections in verbs are more frequent than misspelled inflections in nouns as the former are less consistent than the latter.

Specifically, in Table 3.15. below, for the category of phonological errors, the first 30 words offer 160 opportunities for a substitution or for an omission of a phoneme as counted based on the number of graphemes or digraphs in each word. Thus, error opportunities for the category of phonological errors concern only the two minor subcategories of errors mentioned above. This offers just an estimation for the reader of the number of phonological errors just for the two minor subcategories, and shows that the subsequent 30 words offer more opportunities for phonological errors than the first 30 words, and concomitantly that the same applies for the remaining minor subcategories of phonological errors. Moreover, the first 30 words offer 20 opportunities for grammatical errors and 42 opportunities for orthographic errors. The subsequent 30 words offer 302 opportunities for a substitution or for an omission of a phoneme, 28 for grammatical errors and 93 opportunities for orthographic errors. Thus, the list of 60 words offers 462 opportunities for a substitution or for an omission of a phoneme, 48 opportunities for grammatical errors and 135 opportunities for orthographic errors.

Table 3.15. The total number of error opportunities for the three major categories of errors in the first 30 words, in the subsequent 30 words and in total 60 words of the spelling-to-dictation task

	Phonological[1]	Grammatical	Orthographic
First 30 words	160	20	42
Subsequent 30 words	302	28	93
Total 60 words	462	48	135

Note: [1]the composite score of phonological errors concerned only two minor subcategories of errors, substitution and omission of phonemes, as the number of opportunities was counted based on the number of graphemes and digraphs in each word.

3.7. Assessment procedure

Following the programme of a school-day in primary education, children were tested in one session (of 90 min) in a school classroom between October 2015 and June 2016. Some children with DDLD were also assessed in July 2016. Certain children with DDLD were tested in the referral centre where they were receiving speech and language therapy, but the majority were assessed in a school classroom as they were identified in schools and not in a referral centre. Responses were recorded when needed, using Audacity for Windows 7 and a microphone for later transcriptions.

3.8. Calculation of statistical power

With respect to sample size for each group of children, power analysis can be used to calculate the minimum sample size required so that an effect of a given size can be detected, if the effect actually exists. Put another way, the concept of statistical power refers to the probability of rejecting the null hypothesis when it is not correct, with a high statistical power indicating that the probability of making a Type II error, or concluding there is no effect when there is an effect, goes down. Power analysis can either be used before (a priori power analysis) or after (*post hoc* power analysis) data are collected. A priori power analysis is conducted prior to the research study, and is used in estimating sufficient sample sizes to achieve adequate power. *Post-hoc* power analysis is conducted after a study has been completed, and uses the obtained sample size and effect size to determine what the power was in the study, assuming the effect size in the sample is equal to the effect size in the population. There are no formal standards for power (π), even though most researchers assess the power of their tests using $\pi = 0.80$ as a standard for adequacy. Cohen (1988) has chosen a β error of 0.2 for statistical power, which is an arbitrary level but it has been chosen by other researchers for decades, and the same level is used for this study.

Sullivan and Feinn (2012) argued that: “statistical power must be calculated prior to starting the study as post-hoc calculations, sometimes reported when prior calculations are omitted, have limited value due to the incorrect assumption that the sample effect size represents the population effect size" (p. 281). Moreover, Hoenig and Heisey (2001) argued that observed power is determined by the observed significance level of a test statistic. In their own words: "for any test the observed power is a 1:1 function of the *p* value" (p. 2).

A broad age range of children were included in the current study, and therefore appropriate power analyses should be carried out to ensure significant results will be detected when they really exist. Calculation of sample size used G*Power 3.1 software (retrieved from http://www.gpower.hhu.de/en.html; Faul, Erdfelder, Lang, & Buchner, 2007; Faul, Erdfelder, Buchner, & Lang, 2009). Table 3.17. shows a priori power analyses computing required overall sample size for a power of at least 0.80 to detect a given effect size of the analyses with 2 participant groups, 2 covariates (age in months and NVIQ) and $\alpha = 0.05$. For regression analyses, 3 predictors were considered corresponding to age in months, NVIQ and group for part of the analyses and also to age in months, NVIQ and the Language and Literacy variable for some other analyses. The smaller an effect size is, the bigger the overall sample size needs to be. Thus, if a small effect size can be detected with the overall sample size of 149 children of this study, a medium and a large effect, with a smaller number of participants, would also be detected. As presented in Table 3.16., this is the case for MANOVAs. For correlation, independent samples *t*-test, Wilcoxon Mann-Whitney and regression analyses only medium and large effect sizes would be detected with the overall sample size of the current study (i.e. $n = 149$), but not a small effect size. Further, Chuard, Vrtílek, Head, and Jennions (2019) argued that earlier studies often used univariate tests (for example, *t* tests or Mann-Whitney U tests) that ignored major confounding variables (i.e. implicitly assumed that they did not differ between a control group and a clinical group, for example). In contrast, recent studies usually add potential confounding variables as covariates that are 'corrected for' before examining the effect of a variable. The current study is such a study in that confounding variables (i.e. age and NVIQ) were added as covariates that were 'corrected for' before examining the effect of group on semantic and phonological fluency variables, which are the main variables of interest in the current study. According to Chuard et al. (2019), this statistical approach renders unlikely the interpretation of the current study to be affected by group differences in confounding variables.

Table 3.16. A priori power analyses computing required overall sample sizes

Test statistics	Effect sizes	Overall sample sizes
Correlation	.30 (medium)	82
Independent samples *t*-test	.50 (medium)	128
Wilcoxon Mann-Whitney	.50 (medium)	134
MANOVA (repeated measures, within factors)	.25 (small)	24
MANOVA (repeated measures, between factors)	.25 (small)	96
Linear multiple regression (R^2 deviation from zero)	.15 (medium)	77
Linear multiple regression (R^2 increase)	.15 (medium)	77

3.9. Threshold of statistical significance

Benjamin et al. (2018) reported that the threshold for defining statistical significance should be redefined and proposed a change from $p < .05$ to $p < .005$. They argued that one of the main reasons for a lack of reproducibility of scientific studies is that a p value lower than .05 is too high to be deemed as an evidence of statistical significance as it results in a high rate of false positives *even in the absence of other experimental, procedural and reporting problems*. Put another way, they considered that findings being significant at the $p < .05$ level of significance should not be associated with statistical significance. They proposed that this change would improve the reproducibility of scientific studies. According to the researchers, results meeting the new threshold, i.e. $p < .005$. should be called significant and results meeting the old threshold (i.e. $p < .05$) should be called "suggestive". The new criterion means that there is a .05% probability (instead of 5%) for accepting the alternative hypothesis when it is false. In the current study, statistically significant results are considered to be those meeting the old threshold. However, because many studies have failed to replicate previous findings and scientists argue that there is indeed a reproducibility crisis in science

(Nosek et al., 2015), in the present thesis, (significant) results meeting the proposed threshold by Benjamin et al. and (suggestive) results meeting the old threshold will be presented in the Results. The aim is to inform the reader about findings which are likely to be replicated and about those which are less likely to be replicated.

Chapter 4. Results

The results are presented in six parts. The first part (4.1.) compares the groups on differences on semantic fluency tasks including patterns of lexical retrieval (clustering, switching, and average cluster size), incorrect responses, and responses in the first 15 s and in the subsequent 45 s of the test period. In the second part (4.2.), the groups' performance and group differences on phonological fluency tasks are presented. The third part (4.3.) presents group comparisons on phoneme deletion, NWR and RAN tasks followed by the fourth part (4.4.) which presents group comparisons on types of spelling errors. In the fifth part (4.5.), the relationship between semantic and phonological fluency and children's language and literacy skills is investigated, in addition to the relationship between automatic and controlled processing and children's language and literacy skills. The last part (4.6.) compares the groups on the design fluency task. Prior to the Discussion chapter, two more sections are presented at the end of the Results chapter, namely, a section (4.7.) presenting results meeting the old and the new threshold of statistical significance followed by a summary of the results (4.8.).

4.1. Groups' Performance and Group Differences on Semantic Fluency Tasks

Semantic clusters in semantic fluency tasks were used to answer the following research question: Which model better characterises lexical difficulties in semantic categories in dyslexia and DLD: the Poor Lexical-Semantic Structure Model (which attributes lexical difficulties to children's impaired semantic structure) or the Slow-Retrieval Model (which attributes lexical difficulties to slow retrieval processes while children's semantic structure is intact)? Computational modelling was also used, offering the opportunity to compare semantic networks across different groups in an alternative way, namely, independently of the experimenter's coding of semantic clusters. A further research question was: Do cluster number and/or cluster size drive productivity in semantic fluency tasks in TD children and children with DDLD?

4.1.1. Statistical methods. Group differences in semantic fluency variables were assessed with regression analyses, as proposed by Henry et al. (2015) who argue that regression techniques are more robust than univariate analyses of variance. Linear regression analyses were carried out on the number of total, correct and incorrect responses, the number of total responses in the first 15 s and the subsequent 45 s of the test period, the number of switches, the number of clusters, and the average size of clusters in semantic fluency tasks. For each semantic fluency variable, the linear regression controlled for age in months and NVIQ in Step 1, and participant group was entered in Step 2 to investigate how much of the variance in each semantic fluency variable was accounted for by group after controlling for age in months and NVIQ. Linear regression analysis was also carried out to investigate whether the ratio score derived from computational modelling is different in the two groups. Pearson correlations were used for the associations between word productivity with the number of switches, number of clusters and average cluster size. Two independent samples Mann-Whitney *U* tests were used to compare the two groups on the proportional scores of types of incorrect responses and the error ratio. A non-parametric test was used for these analyses since the explanatory variables of interest were continuous but not normally distributed.

4.1.2. Regression analyses on the semantic fluency variables

Significant and nonsignificant effects of group. Group comparisons on semantic fluency variables are presented in Table 4.1.1.

Table 4.1.1. Linear regression analyses conducted on the semantic fluency variables

Semantic fluency variables	DDLD group *M* (*SD*)	TD group *M* (*SD*)	R^2	ΔR^2	β age	β NVIQ	β group
Total responses (TR)	43.71 (12.10)	44.87 (13.85)	29.8^{***}	$.020^{*}$	$.488^{***}$	$.241^{**}$	$-.158^{*}$
Correct responses	41.46 (11.25)	43.34 (13.59)	32.1^{***}	$.034^{**}$	$.508^{***}$	$.246^{***}$	$-.204^{**}$
Incorrect responses	2.25 (2.24)	1.53 (1.57)	0.0	$.043^{*}$	-.021	-.030	$.229^{*}$
TR in the first 15 s	18.75 (4.27)	20.29 (5.33)	27.6^{***}	$.074^{***}$	$.475^{***}$	$.221^{**}$	$-.301^{***}$
TR in the subsequent 45 s	24.96 (8.57)	24.58 (9.41)	21.8^{***}	.013	$.437^{***}$	$.159^{*}$	-.125
Number of switches	24.06 (6.18)	22.50 (6.59)	18.3^{***}	.000	$.409^{***}$	.119	.024
Number of clusters	9.46 (3.62)	10.73 (4.55)	17.2^{***}	$.065^{**}$	$.380^{***}$	$.161^{*}$	$-.282^{**}$
Average cluster size	2.81 (.42)	2.95 (.72)	0.0	.004	-.017	$.198^{*}$	-.066

Notes: For each regression, age in months and NVIQ were entered in Step 1 and participant group was entered in Step 2. For Step 1, information is provided on the variance accounted for by age in months and NVIQ (R^2), and for Step 2, information is provided on the proportion of variance accounted for by group (ΔR^2). The β-values for the three predictor variables are given. Significance values are given where they are relevant. Means and *SD*s for the two participant groups are also presented; $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$; Table presents total scores across 3 semantic categories (excluding average cluster size).

Checks indicated that Mahalanobis distances were less than 15 for all the children for all variables, therefore all cases were included in regressions (Field, 2013). The regression model was significant for total responses, $F_{(3, 145)} = 22.571$, $p < .001$, accounting for 31.8% of the variance, for correct responses, $F_{(3, 145)} = 26.616$, $p < .001$, accounting for 35.5% of the variance, for total responses in the first 15 s, $F_{(3, 145)} = 25.963$, $p < .001$, accounting for 34.9% of the variance, for total responses in the subsequent 45 s, $F_{(3, 145)} = 14.465$, $p < .001$, accounting for 23% of the variance, for the number of switches, $F_{(3, 145)} = 10.840$, $p < .001$, accounting for 18.3% of the variance, and for the number of clusters, $F_{(3, 145)} = 14.979$, $p < .001$, accounting for 23.7% of the variance. The model was nonsignificant for the number of incorrect responses, $F_{(3, 145)} = 2.231$, $p = .085$, and for average cluster size, $F_{(3, 145)} = 2.177$, $p = .093$.

Group was a significant predictor of the variance on total, correct and incorrect responses produced, on the number of total responses in the first 15 s of the test period, and on the number of clusters. Group accounted for 2, 3.4 and 4.3% of the variance in the number of total, correct and incorrect responses produced, respectively, 7.4 and 6.5% of the variance in the number of total responses in the first 15 s and the number of clusters, respectively. Group was a nonsignificant predictor of the number of total responses in the subsequent 45 s, the number of switches and average cluster size. However, as Figure 4.1.1. shows, there was a lot of overlap between the two groups with respect to the number of correct responses, even though group was a significant predictor of the number of correct responses.

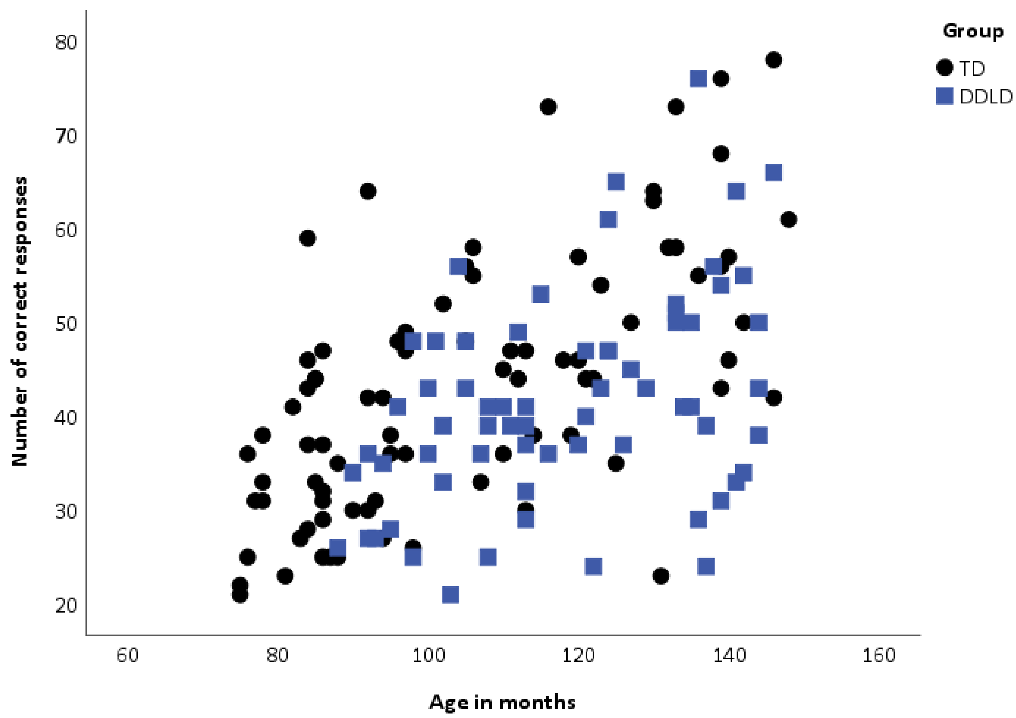


Figure 4.1.1. Scatterplot showing the number of correct responses in semantic categories in the DDLD group and the TD group, plotted against age in months

Summary: Significant and nonsignificant effects of group in semantic fluency variables

Children with DDLD produced significantly fewer total and correct responses and significantly more incorrect responses than TD children after controlling for age in months and NVIQ in the regression models. Children with DDLD also produced fewer total responses in the first 15 s of the test period, and fewer clusters than TD children. However, the two groups did not differ on the number of total responses in the subsequent 45 s of the test period, on the number of switches, or on average cluster size.

4.1.3. Computational analysis for the semantic category of animals

Computational modelling was also used to investigate any group differences in children's semantic network using the category of animals (Davelaar et al., in preparation). Analysing the data originating from this category, it was found that the most likely sequence for the category of animals was the following: *cat*, *dog*, *cow*, *sheep*, *goat*, *horse*, *lion*, *tiger*, *elephant*, *giraffe*, *zebra*, *monkey*, *gorilla*, *wolf*, *fox*, *bear*, *deer*, *mouse*, *rat*, *cheetah*, *snake*, *crocodile*, *shark*, *fish*, *bird*, *eagle*, *hare*, *tortoise*, *dolphin*, *seal*, *whale*, *rhino*. Mean ratio score (*SD*) for children with DDLD was 1.27 (.08) and for TD children was 1.26 (.08). As a reminder to the reader, the ratio score is a measure of how far a child's score deviates from the most likely sequence for the category of animals presented above. In the regression analysis, age in months and NVIQ entered in Step 1 and group entered in Step 2. Analysis revealed that age in months and group were nonsignificant predictors of the variance in ratio score, but NVIQ was a marginally significant predictor of the variance in ratio score; age: *Beta* = -.099, t = -1.221, p = .224; NVIQ: *Beta* = -.174, t = -2.148, p = .033; group: *Beta* = .019, t = .210, p = .834, and the model with the three predictors was nonsignificant, $F_{(3, 145)}$ = 2.053, p = .109. This finding is therefore consistent with the finding that the two groups did not differ on average cluster size: both findings suggest that the semantic network of children with DDLD is not significantly different from that of TD children.

4.1.4. Associations between the number of correct responses with the number of switches, the number of clusters and average cluster size in the TD and DDLD groups

In order to understand whether semantic fluency performance in each group was related to the production of a greater number of clusters or to the production of bigger clusters, correlations between the number of correct responses and the number of clusters, the number of switches, and average cluster size were investigated. In the TD group, the number of correct responses correlated strongly with the number of clusters, $r_{(80)} = .64$, $p < .001$, and the number of switches, $r_{(80)} = .72$, $p < .001$, but not with average cluster size, $r_{(80)} = .10$, $p = .336$. Likewise, in the DDLD group, the number of correct responses correlated with cluster number, $r_{(63)} = .77$, $p < .001$, and the number of switches, $r_{(63)} = .76$, $p < .001$, but not with average cluster size, $r_{(63)} = .00$, $p = .964$. Thus, in both groups, productivity in semantic fluency tasks is driven by the production of more clusters and more switches, but not by bigger clusters.

4.1.5. Proportional scores of types of incorrect responses in semantic fluency categories

Table 4.1.2. shows types of incorrect responses in proportional scores in the semantic condition in the TD and DDLD groups. Two independent samples Mann-Whitney *U* tests revealed that there were nonsignificant differences between the two groups with respect to the proportional scores of repeated, $U = 2337.50$, $p = .08$, made-up, $U = 2697.50$, $p = .26$, out-of-category, $U = 2338.50$, $p = .10$, and unintelligible responses, $U = 2581.00$, $p = .23$.

Table 4.1.2. Types of incorrect responses in proportional scores in the semantic condition in the TD and DDLD groups

Types of incorrect responses	TD group *M* (*SD*)	DDLD group *M* (*SD*)
Repeated	1.23 (1.98)	1.90 (2.52)
Made-up	.12 (.55)	.04 (.17)
Out-of-category	1.73 (2.50)	2.25 (2.60)
Unintelligible	.29 (.85)	.49 (2.98)

4.1.6. Error ratio in semantic fluency categories

The error ratio in semantic categories was .03 (*SD* = .04) in the TD group and .05 (*SD* = .04) in the DDLD group. Two independent samples Mann-Whitney *U* tests revealed that there was a marginally significant difference in the error ratio in semantic fluency categories between the two groups, U = 2194.00, Z = -2.10, p = .012.

Further analyses in the two groups separately revealed that the number of correct responses correlated strongly with age in the TD group, $r_{(83)} = .63$, $p < .001$, and moderately in the DDLD group, $r_{(66)} = .42$, $p < .001$. NVIQ performance correlated weakly with semantic fluency in the TD group, $r_{(80)} = .23$, $p = .032$, but did not correlate in the DDLD group, $r_{(63)} = .18$, $p = .132$.

4.1.7. Summary

It was predicted that using analysis of clustering behaviour in the semantic condition, if the Poor Lexical-Semantic Structure Model holds true, the DDLD group would produce significantly fewer correct responses and a significantly smaller average cluster size than the TD group. This is explained by poor semantic structure in the DDLD group. In contrast, if the Slow-Retrieval Model holds true, the DDLD group would produce significantly fewer correct responses than the TD group but the two groups would not differ on average cluster size. This is explained by slow retrieval processes while lexical-semantic representations are intact in the DDLD group. Moreover, using computational modelling analysis, the Poor Lexical-Semantic Structure Model predicts that the DDLD group would produce a significantly larger ratio score than the TD group, suggesting structural differences in the semantic network of children with DDLD. In contrast, the Slow-Retrieval Model predicts that the two groups would not differ on the computed ratio score, suggesting an adequate semantic network but difficulties in accessing semantic information quickly and efficiently, that is, suggesting retrieval differences between the two groups. It was also predicted that in both groups, semantic fluency performance would be driven by the production of more switches and more clusters rather than by the production of bigger clusters.

Results showed that children with DDLD produced significantly fewer total and correct responses and significantly more incorrect responses than TD children after controlling for age in months and NVIQ in the regression models. Children with DDLD also produced fewer total responses in the first 15 s of the test period, and fewer clusters than TD children. However, the two groups did not differ on the number of total responses in the subsequent 45 s, on the number of switches, or on average cluster size. Group was a nonsignificant predictor of the variance in ratio score which was computed based on computational analysis techniques after controlling for age in months and NVIQ. This finding is consistent with the finding that group was a nonsignificant predictor of the variance in average cluster size in semantic categories. Together the two findings suggest that the semantic network of children with DDLD is not significantly different from that of TD children. In both groups, productivity in semantic fluency tasks was driven by the production of more clusters and more switches, but not by bigger clusters. The two groups did not differ on any type of incorrect responses using proportional scores in the analyses. A significantly higher error ratio was observed, however, in the DDLD group compared to the TD group, suggesting difficulties with EFs in the DDLD group.

4.2. Groups' Performance and Group Differences on Phonological Fluency Tasks

Phonological fluency tasks were used to answer the following research questions. Which of the two prominent phonological hypotheses better characterises the locus of the phonological deficit in children with dyslexia and DLD: the Degraded Phonological Representations Hypothesis (which claims that children with DDLD will show impaired phonological representations) or the Deficient Phonological Access Hypothesis (which claims that children with DDLD will show impaired explicit access but intact implicit access to phonological representations)? Do cluster number and/or cluster size drive productivity in phonological fluency tasks in TD children and children with DDLD?

4.2.1. Statistical methods. Group differences on phonological fluency variables were assessed with regression analyses, as presented above in the section of group differences on semantic fluency variables. The remaining analyses presented in this section were also carried out using the same statistical tests as presented in section 4.1.1.

4.2.2. Regression analyses on the phonological fluency variables

Significant and nonsignificant effects of group. Group comparisons on phonological fluency variables are presented in Table 4.2.1.

Table 4.2.1. Linear regression analyses conducted on the phonological fluency variables

Phonological fluency variables	DDLD group *M* (*SD*)	TD group *M* (*SD*)	R^2	ΔR^2	*β* age	*β* NVIQ	*β* group
Total responses (TR)	21.64 (8.07)	22.64 (9.59)	26.5^{***}	.017	$.430^{***}$	$.280^{***}$	-.142
Correct responses	20.24 (8.02)	21.68 (9.59)	28.5^{***}	$.027^{*}$	$.450^{***}$	$.284^{***}$	$-.183^{*}$
Incorrect responses	1.40 (1.86)	.96 (1.29)	0.0	$.034^{*}$	-.118	.000	$.204^{*}$
TR in the first 15 s	9.85 (4.06)	9.98 (4.06)	21.2^{***}	.012	$.416^{***}$	$.193^{*}$	-.120
TR in the subsequent 45 s	12.79 (4.94)	12.66 (6.08)	17.6^{***}	$.023^{*}$	$.379^{***}$	$.178^{*}$	$-.169^{*}$
Number of switches	10.71 (5.25)	11.84 (6.43)	24.9^{***}	$.025^{*}$	$.399^{***}$	$.295^{***}$	$-.176^{*}$
Number of clusters	4.74 (2.73)	4.71 (2.68)	14.4^{***}	.002	$.319^{***}$	$.202^{**}$	-.052
Average cluster size	2.79 (.80)	2.63 (.64)	0.2	.000	.149	-.040	-.015

Notes: For each regression, age in months and NVIQ were entered in Step 1 and participant group was entered in Step 2. For Step 1, information is provided on the variance accounted for by age in months and NVIQ (R^2), and for Step 2, information is provided on the proportion of variance accounted for by group (ΔR^2). The *β*-values for the three predictor variables are given. Significance values are given where they are relevant. Means and *SD*s for the two participant groups are also presented; $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$. Table presents total scores across 3 phonological categories (excluding average cluster size).

Checks indicated that Mahalanobis distances were less than 15 for all the children for all variables, therefore all cases were included in regressions (Field, 2013). The regression model was significant for the number of total responses, $F_{(3, 145)} = 18.984$, $p < .001$, accounting for 28.2% of the variance, for the number of correct responses, $F_{(3, 145)} = 21.942$, $p < .001$, accounting for 31.2% of the variance, for the number of total responses in the first 15 s, $F_{(3, 145)} = 13.922$, $p < .001$, accounting for 22.4% of the variance, for the number of total responses in the subsequent 45 s, $F_{(3, 145)} = 12.043$, $p <$

.001, accounting for 19.9% of the variance, for the number of switches, $F_{(3, 145)} = 18.218$, $p < .001$, accounting for 27.4% of the variance, and for the number of clusters, $F_{(3, 145)} = 8.248$, $p < .001$, accounting for 14.6% of the variance. The model was nonsignificant for the number of incorrect responses, $F_{(3, 145)} = 2.444$, $p = .066$, and for average cluster size, $F_{(3, 145)} = 1.189$, $p = .316$.

Group was a significant predictor of the variance on the number of correct and incorrect responses produced, on the number of total responses in the subsequent 45 s of the test period, and on the number of switches. Group accounted for 2.7, 3.4 and 2.3% of the variance in the number of correct responses, incorrect responses, and total responses in the subsequent 45 s, respectively, and 2.5% of the variance in the number of switches. Group was a nonsignificant predictor of the number of total responses, total responses in the first 15 s, number of clusters and average cluster size. However, as Figure 4.2.1. shows, there was a lot of overlap between the two groups with respect to the number of correct responses, even though group was a significant predictor of the number of correct responses.

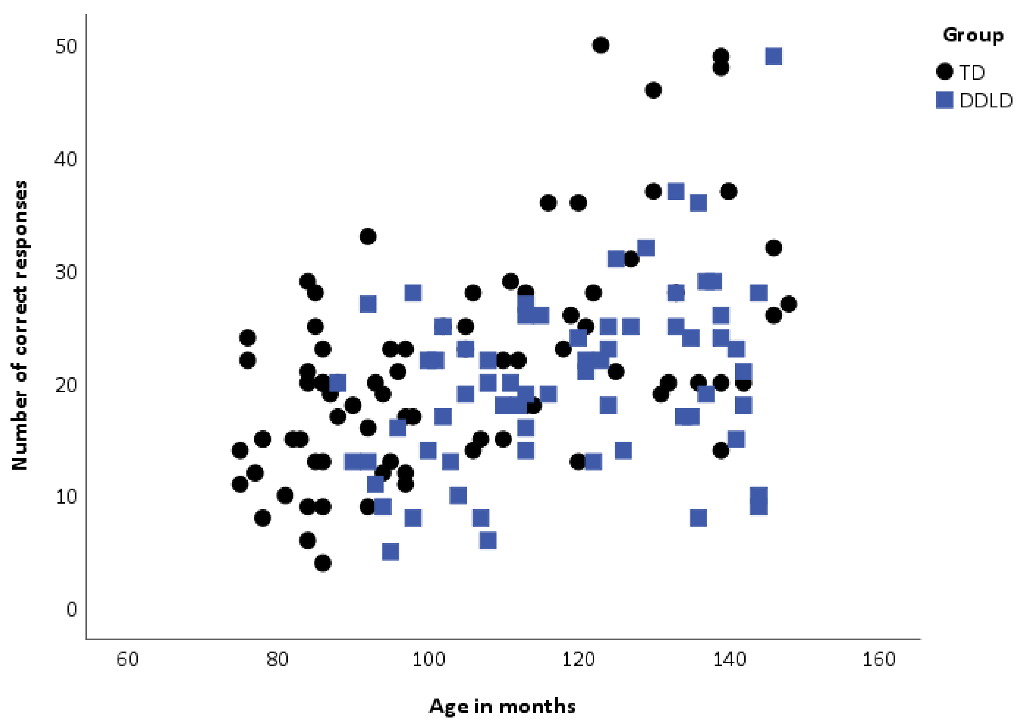


Figure 4.2.1. Scatterplot showing the number of correct responses in phonological categories in the DDLD group and the TD group, plotted against age in months

Summary: Significant and nonsignificant effects of group in phonological fluency variables

Children with DDLD produced significantly fewer correct responses and more incorrect responses than TD children after controlling for age in months and NVIQ. Children with DDLD also produced fewer total responses in the subsequent 45 s of the test period and fewer switches than TD children. However, the two groups did not differ on the number of total responses, total responses in the first 15 s of the test period, number of clusters and average cluster size.

4.2.3. Associations between the number of correct responses with the number of switches, the number of clusters and average cluster size in the TD and DDLD groups

In order to understand whether phonological fluency performance in each group was related to the production of a greater number of clusters or to the production of more items within a cluster, partial Pearson correlations (controlling for age) were used between the number of correct responses and the number of clusters, the number of switches, and average cluster size. In the TD group, the number of correct responses correlated strongly with the number of clusters, $r_{(80)} = .74$, $p < .001$, and the number of switches, $r_{(80)} = .83$, $p < .001$, but not with average cluster size, $r_{(80)} = .10$, $p = .340$. Likewise, in the DDLD group, the number of correct responses correlated with cluster number, $r_{(63)} = .73$, $p < .001$, and the number of switches, $r_{(63)} = .78$, $p < .001$, but again not with average cluster size, $r_{(63)} = .21$, $p = .083$. Thus, in both groups, the production of more clusters and more switches drives word productivity, and not the production of more items within a cluster (i.e. bigger clusters).

4.2.4. Proportional scores of types of incorrect responses in phonological fluency categories

Table 4.2.2. shows types of incorrect responses in proportional scores in the phonological condition in the TD and DDLD groups. Two independent samples Mann-Whitney *U* tests revealed that there were nonsignificant differences between the two groups with respect to the proportional

scores of repeated, *U* = 2565.50, *p* = .33, made-up, *U* = 2477.50, *p* = .18, out-of-category, *U* = 2572.50, *p* = .32, and unintelligible responses, *U* = 2582.50, *p* = .18.

Table 4.2.2. Types of incorrect responses in proportional scores in the phonological condition in the TD and DDLD groups

Types of incorrect responses	TD group *M* (*SD*)	DDLD group *M* (*SD*)
Repeated	.95 (2.56)	1.23 (2.82)
Made-up	1.78 (4.40)	2.54 (5.50)
Out-of-category	1.33 (4.57)	1.70 (4.64)
Unintelligible	.35 (1.84)	.71 (2.54)

4.2.5. Error ratio in phonological fluency categories

The error ratio in phonological categories was .06 (*SD* = .15) in the TD group and .09 (*SD* = .17) in the DDLD group. Two independent samples Mann-Whitney *U* tests revealed that this difference was not significant, *U* = 2348.00, *Z* = -1.56, *p* = .11.

Further analyses in the two groups separately showed that the number of correct responses correlated strongly with age in the TD group, $r_{(83)}$ = .57, *p* < .001, and moderately in the DDLD group, $r_{(66)}$ = .37, *p* = .002. NVIQ performance correlated moderately with the number of correct responses in the TD group, $r_{(83)}$ = .40, *p* < .001, but did not correlate in the DDLD group, $r_{(66)}$ = .12, *p* = .315.

4.2.6. Summary

It was predicted that if the Degraded Phonological Representations Hypothesis holds true, the DDLD group would produce significantly fewer correct responses and a significantly smaller average cluster size than the TD group. This is explained by impaired phonological representations in the DDLD group. In contrast, if the Deficient Phonological Access Hypothesis holds true, the DDLD group would produce significantly fewer correct responses than the TD group but the two groups would not differ on average cluster size. This is explained by impaired explicit access but intact implicit access to phonological representations in the DDLD group. It was also predicted that if the Deficient Phonological Access Hypothesis holds true, in both groups, phonological fluency performance would be driven by the production of more switches and more clusters rather than by the production of bigger phonological clusters.

Results showed that children with DDLD produced significantly fewer correct responses and more incorrect responses than TD children after controlling for age in months and NVIQ. Children with DDLD also produced fewer total responses in the subsequent 45 s of the test period and fewer switches than TD children. However, the two groups did not differ on the number of total responses, total responses in the first 15 s of the test period, number of clusters and average cluster size. In both groups, productivity in phonological fluency tasks was driven by the production of more clusters and more switches, but not by bigger clusters. The two groups did not differ on any type of incorrect responses using proportional scores in the analyses, or on error ratio.

4.3. Groups' Performance and Group Differences on Phonological Tasks

Which hypothesis better characterises the locus of the phonological deficit in children with DDLD in phonological tasks—namely, phoneme deletion, NWR and RAN tasks: The Degraded Phonological Representations Hypothesis or the Deficient Phonological Access Hypothesis?

4.3.1. Statistical methods. Group differences on phoneme deletion and NWR tasks were analysed further using multivariate analysis of variance (MANOVA). Wilks' Lambda was reported as it is the most widely used test statistic in MANOVA. Conventions for interpreting partial eta squared originating from analyses of group differences on phoneme deletion and NWR tasks, as proposed by Cohen (1988) were used in this study.

4.3.2. Phoneme deletion tasks

Prior to statistical analyses, *z* scores are presented as illustrative measures which profile the DDLD group. As shown in Figure 4.3.1. presenting accuracy performance in *z* scores in phoneme deletion tasks in the two groups, the DDLD group performed within -1 *SD* of the TD group's mean on accuracy on phoneme deletion of CVCVCV, CVC and CCV items. Deleting phonemes of monosyllable nonwords with a simple CVC syllable structure was not challenging for the DDLD group (*z* score = -.02). The DDLD group were less accurate, however, at deleting phonemes of trisyllable nonwords with a simple CVCVCV syllable structure (*z* score = -.77) and phonemes of monosyllable nonwords with a complex CCV syllable structure (*z* score = -.73).

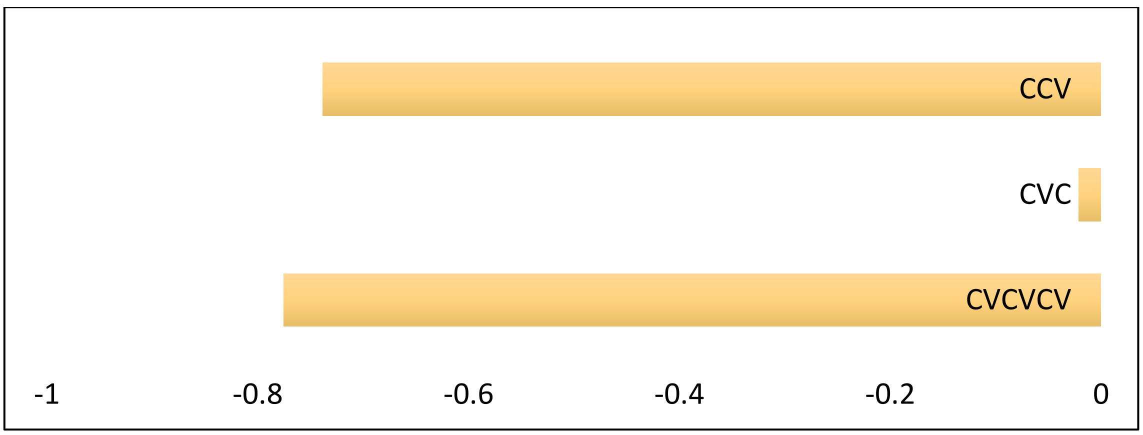


Figure 4.3.1. Accuracy performance in *z* scores in phoneme deletion tasks in the DDLD group

Moreover, as shown in Figure 4.3.2. presenting time performance in *z* scores in phoneme deletion tasks in the two groups, the DDLD group performed within -1 *SD* of the TD group's mean on speed on phoneme deletion of CVCVCV, CVC and CCV items. The DDLD group were slower at deleting phonemes of trisyllable nonwords with a simple CVCVCV syllable structure (*z* score = -.32) and phonemes of monosyllable nonwords with a complex CCV syllable structure (*z* score = -.31), than in deleting phonemes of monosyllable nonwords with a simple CVC syllable structure (*z* score = -.07).

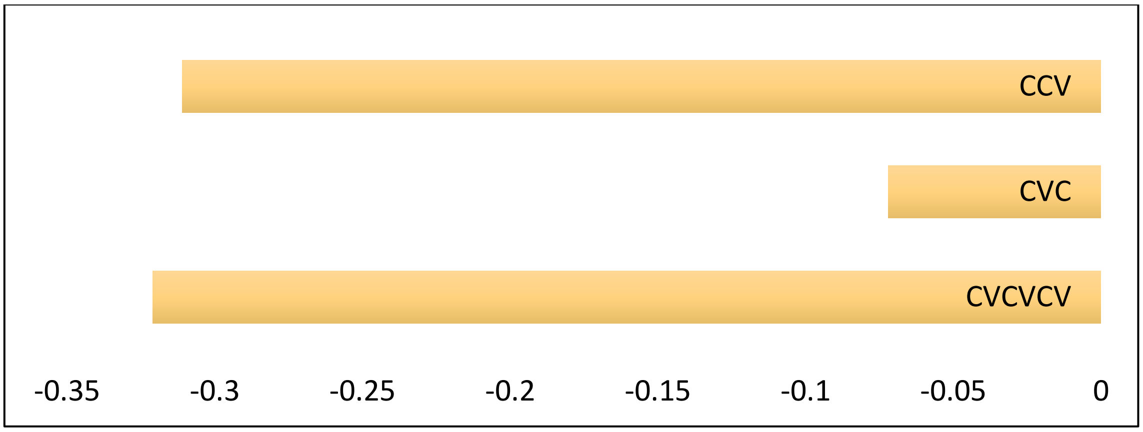


Figure 4.3.2. Time performance in *z* scores in phoneme deletion tasks in the DDLD group

Phoneme deletion variables were analysed further using a statistical test. For these analyses, first Graders were excluded given that using age as a covariate variable is not justified in order to control for the confounding effect of age. After excluding first Graders, the two groups did not differ significantly on age, $t_{(127)} = -1.92$, $p = .05$. However, the two groups differed on NVIQ, $t_{(127)} = 3.92$, $p < .001$, and therefore NVIQ was used as a covariate variable in the analyses presented next. Two 3 (phoneme deletion of CVCVCV, CVC and CCV items) by 2 (group) MANOVAs, one for accuracy and one for speed, were conducted based on *z* scores computed for the phoneme deletion tasks as there was not an equal number of nonwords in all three tasks. Analyses for accuracy revealed a nonsignificant effect of task, Wilks' Lambda = .950, $F_{(3, 123)} = 2.14$, $p = .09$, multivariate $\eta_p^2 = .05$, a nonsignificant effect of NVIQ, Wilks' Lambda = .960, $F_{(3, 123)} = 1.66$, $p = .17$, multivariate $\eta_p^2 = .03$, but a significant effect of group, Wilks' Lambda = .847, $F_{(3, 123)} = 7.39$, $p < .001$, multivariate $\eta_p^2 = .15$. Children with DDLD performed significantly less accurately on deleting the initial phoneme of items with CVCVCV syllable structure, $F_{(1, 125)} = 16.28$, $p < .001$, $\eta_p^2 = .11$, and items with CCV syllable structure, $F_{(1, 125)} = 14.69$, $p < .001$, $\eta_p^2 = .10$, but the two groups did not differ on how accurately they deleted the initial phoneme of items with CVC syllable structure, $F_{(1, 125)} = .33$, $p = .56$, $\eta_p^2 = .00$.

Analyses for speed revealed a nonsignificant effect of task, Wilks' Lambda = .96, $F_{(3, 122)} = 1.67$, $p = .17$, multivariate $\eta_p^2 = .04$, a nonsignificant effect of NVIQ, Wilks' Lambda = .94, $F_{(3, 122)} = 2.37$, $p =$

.07, multivariate η_p^2 = .05, but a significant effect of group, Wilks' Lambda = .91, $F_{(3, 122)}$ = 3.83, p = .012, multivariate η_p^2 = .08. Children with DDLD performed significantly slower on deleting the initial phoneme of items with CVCVCV syllable structure, $F_{(1, 124)}$ = 6.14, p = .015, η_p^2 = .04, and items with CCV syllable structure, $F_{(1, 124)}$ = 6.46, p = .012, η_p^2 = .05, but the two groups did not differ on how fast they deleted the initial phoneme of items with CVC syllable structure, $F_{(1, 124)}$ = .23, p = .62, η_p^2 = .00. In sum, children with DDLD performed just as accurately and fast as TD children on phoneme deletion tasks with monosyllable CVC stimuli, but were significantly less accurate and slower on longer stimuli (CVCVCV) and stimuli with a more complex syllable structure (CCV).

4.3.3. Nonword repetition (NWR) task

The NWR data were analysed further. The NWR data in Table 4.3.2. represents the number of correctly repeated nonwords in terms of nonword length in the two groups. These analyses were performed using 4 (3-, 4-, 5- & 6-syllable nonwords) by 2 (group) MANOVA. For these analyses, first Graders were excluded given that using age as a covariate variable is not justified in order to control for the confounding effect of age. After excluding first Graders, the two groups did not differ significantly on age, $t_{(127)}$ = -1.92, p = .05. However, the two groups differed on NVIQ, $t_{(127)}$ = 3.92, p < .001, and therefore NVIQ was used as a covariate variable in the analyses presented in Table 4.3.2. As presented in Table 4.3.2., the two groups differed significantly on the number of 3-, 4-, 5- and 6-syllable nonwords repeated correctly, with the TD group outperforming the DDLD group.

Table 4.3.2. Means (*SD*s), *CI*s and group comparisons of the number of correctly repeated nonwords in terms of the number of syllables in the DDLD group and the TD group

	DDLD group (n = 62)	TD group (n = 63)	Group comparisons		
Number of syllables	*M* (*SD*)	*M* (*SD*)	*F*	*p*	η_p^2
3	5.14 (1.49) (4.86-5.43)	5.74 (.59) (5.46-6.02)	4.22	.042	.03
4	4.16 (1.59) (3.83-4.48)	5.26 (.91) (4.94-5.59)	15.62	< .001	.11
5	3.17 (1.64) (2.77-3.57)	4.57 (1.52) (4.17-4.96)	15.44	< .001	.11
6	1.20 (1.55) (.76-1.65)	2.95 (1.97) (2.50-3.39)	20.68	< .001	.14

As Table 4.3.2. shows, a medium effect of group was found for accuracy scores on 3-syllable nonwords, and large effects of group were found for accuracy scores on 4-, 5- and 6-syllable nonwords. This finding implies that the longer the nonword is, the bigger the gap between the two groups, with the DDLD group's performance falling more sharply than the TD group's performance as nonword length increases.

4.3.4. Rapid automatic naming (RAN) task

Children with DDLD took on average 140.50 s (*SD* = 43.76) to name all the pictures in the RAN task and TD children took on average 118.72 s (*SD* = 42.02). The numbers of phonological errors, semantic errors and omissions were negligible in both groups. Table 4.3.1. shows means (*SD*s) and *CI*s (lower line) of RAN measures in the DDLD group and the TD group.

Table 4.3.2. Means (*SD*s) and *CI*s (lower line) of RAN measures in the DDLD group and the TD group

RAN measures	DDLD group *M* (*SD*)	TD group *M* (*SD*)
Composite score (in sec)	140.50 (43.76)	118.72 (42.02)
	(129.74-151.25)	(109.54-127.89)
Phonological errors	.06 (.38)	.04 (.43)
	(-.03-.15)	(-.04-.14)
Semantic errors	1.27 (1.61)	.62 (1.08)
	(.87-1.66)	(.38-.86)
Omissions	.42 (1.39)	.13 (.53)
	(.08-.76)	(.01-.24)

Table 4.3.3. Percentage of variance (R^2) on RAN measures explained by age in months and NVIQ entered in Step 1 and proportion of variance (ΔR^2) explained by group entered in Step 2 in the overall sample

Explanatory variables	R^2	ΔR^2	β age	β NVIQ	β group
RAN phonological errors	0.0	.000	-.100	$-.173^{*}$	.000
RAN semantic errors	0.0	$.052^{**}$	.030	-.064	$.252^{**}$
RAN omissions	0.0	.014	.076	-.038	.132

Notes: For each regression, age in months and NVIQ were entered in Step 1 and group was entered in Step 2. For Step 1, information is provided on the amount of variance accounted for by age in months and NVIQ (R^2), and for Step 2, information is provided on the proportion of variance accounted for by group (ΔR^2). The β-values for the three predictor variables are presented in Columns 4, 5 and 6. Significance values are given where they are relevant; $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Checks indicated that Mahalanobis distances were less than 15 for all the children for all variables, therefore all cases were included in regressions (Field, 2013). The regression model (Table 4.3.3.) was nonsignificant for the number of phonological errors in RAN, $F_{(3, 145)} = 2.025$, $p = .113$, and for the number of omissions in RAN, $F_{(3, 145)} = 1.057$, $p = .369$. The model was significant for the number of semantic errors in RAN, $F_{(3, 145)} = 2.920$, $p = .036$, accounting for 5.7% of the variance in the number of semantic errors. Regression analyses showed nonsignificant differences between the two groups on the number of phonological errors and the number of omissions. However, significantly more semantic errors were found in the DDLD group than in the TD group.

4.3.5. Summary

It was predicted that if the Degraded Phonological Representations Hypothesis holds true, the TD group would outperform the DDLD group in accuracy and speed performance in all three phoneme deletion tasks. This is explained by impaired phonological representations in the DDLD group. In contrast, if the Deficient Phonological Access Hypothesis holds true, accuracy and speed performance in the phoneme deletion task of monosyllable items with simple CVC syllable structure would be equivalent for the two groups. This is explained by intact explicit access to phonological representations in the phoneme deletion task requiring metacognitive access to phonological representations without loading on phonological STM skills since nonwords were short and structurally simple. The Deficient Phonological Access Hypothesis also predicts that accuracy and speed performance in the phoneme deletion task of trisyllable items with simple CVCVCV syllable structure and in the phoneme deletion task of monosyllable items with complex CCV syllable structure will be poorer for the DDLD group than the TD group. This is explained by impaired explicit access to phonological representations in the phoneme deletion tasks requiring metacognitive access to phonological representations whilst loading on phonological STM skills since nonwords were long or had a complex syllable structure. With respect to phoneme deletion tasks, results showed that children with DDLD performed just as accurately and fast as TD children on phoneme deletion tasks with monosyllable CVC stimuli, but were less accurate and slower on longer stimuli (phoneme deletion of CVCVCV items) and stimuli with a more complex (CCV) syllable structure. With respect to the NWR task, results showed a large effect of nonword length, and the effect size increased with increasing nonword length, implying that the longer the nonword was, the bigger the gap between the two groups.

Using the RAN task, it was predicted that if the Degraded Phonological Representations Hypothesis holds true, the DDLD group would make phonological errors in their picture naming. Phonologically inaccurate performance is explained by inaccurate phonological representations. If the Deficient Phonological Access Hypothesis, however, holds true, the DDLD group would name pictures significantly slower than the TD group but the two groups would not differ on phonological accuracy. Slower naming performance is explained by the fact that the phonological access deficit in the DDLD group renders performance on tasks requiring speeded access to phonological representations particularly slow. Phonologically accurate performance is explained by intact access to phonological representations in the RAN task not requiring metalinguistic manipulation. The DDLD group named items in the RAN task significantly slower than the TD group; however, the two groups did not differ on the number of phonological errors observed in the RAN task.

4.4. Groups' Performance and Group Differences on Types of Spelling Errors

Which hypothesis better characterises the locus of the phonological deficit in children with DDLD: the Degraded Phonological Representations Hypothesis or the Deficient Phonological Access Hypothesis? Spelling errors in dictation were assigned in three types of errors, namely, phonological, grammatical and orthographic in order to answer this research question, shedding light on the spelling processes in which Greek children with DDLD show difficulty relative to TD children.

4.4.1. Statistical methods. With respect to the spelling task, two independent samples Mann-Whitney *U* test, a non-parametric test, was used to compare the groups on the proportional numbers of phonological and grammatical errors since the explanatory variables of interest were continuous but not normally distributed. However, an independent samples *t*-test, a parametric test, was used to compare the two groups on the proportional number of orthographic errors since the distribution of the proportional number of orthographic errors was normal. For group comparisons on the proportional numbers of composite scores of phonological, grammatical and orthographic errors, first Graders were excluded, and the two groups did not differ significantly on age in months. For parametric data, Pearson correlation was used, and for non-parametric data, Spearman correlation was used.

4.4.2. Associations between the number of total words spelled with phonological, grammatical and orthographic errors

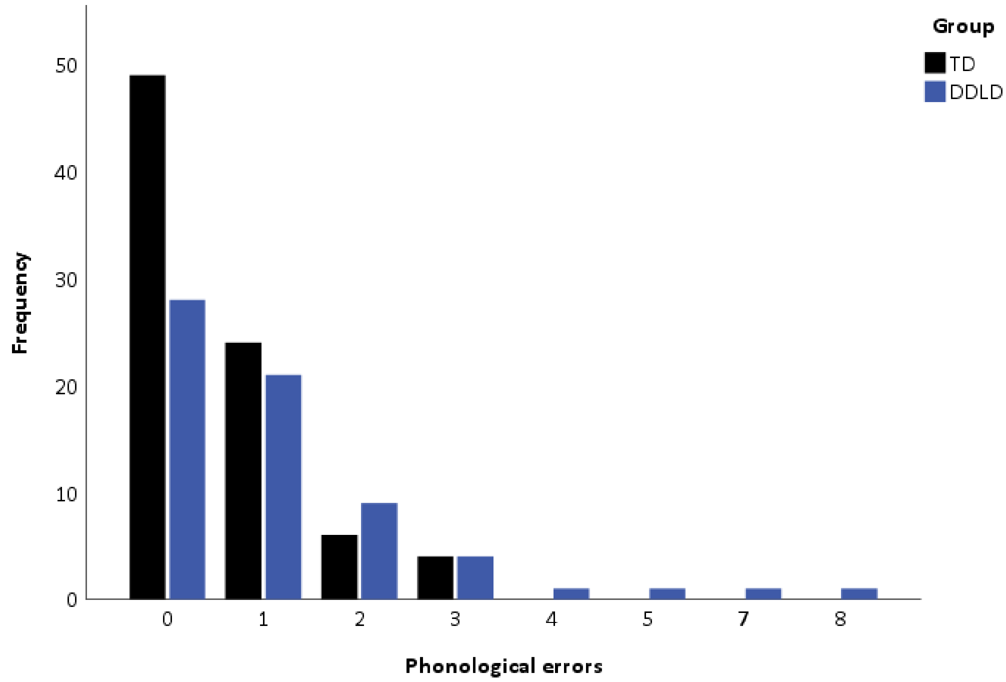
The total number of words spelled correlated weakly with the composite score of phonological errors, $r_{s(149)} = -.21$, $p = .009$. This result shows that children who spelled more words in the task produced more phonological errors. The number of total words spelled showed a negative moderate correlation with the composite score of grammatical errors, $r_{s(149)} = -.38$, $p < .001$, and a strong positive correlation with the composite score of orthographic errors, $r_{(149)} = .63$, $p < .001$. These results show that children who spelled more words in the task produced fewer grammatical errors but more orthographic errors. It seems therefore that the more the words spelled the more the orthographic errors produced, indicating that good spellers produced more orthographic errors. The positive correlations found between the number of total words spelled with phonological and orthographic errors can be accounted for as follows: children who spelled more words in the task were more likely to produce more phonological and orthographic errors in that they were more likely to meet more challenging words, and misspell them.

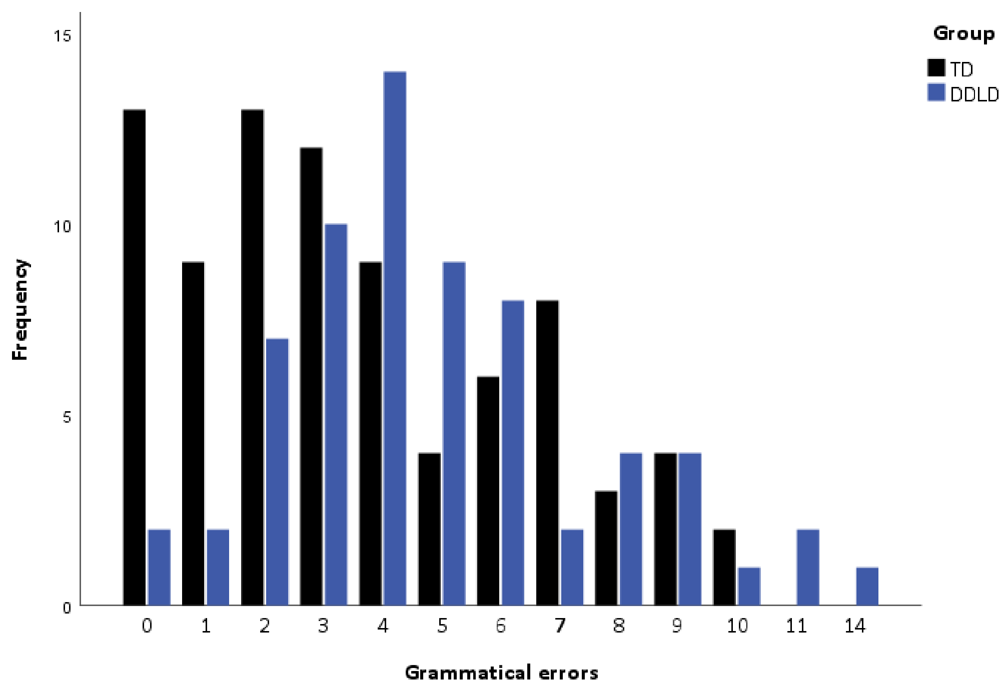
4.4.3. Group comparisons of phonological, grammatical and orthographic errors

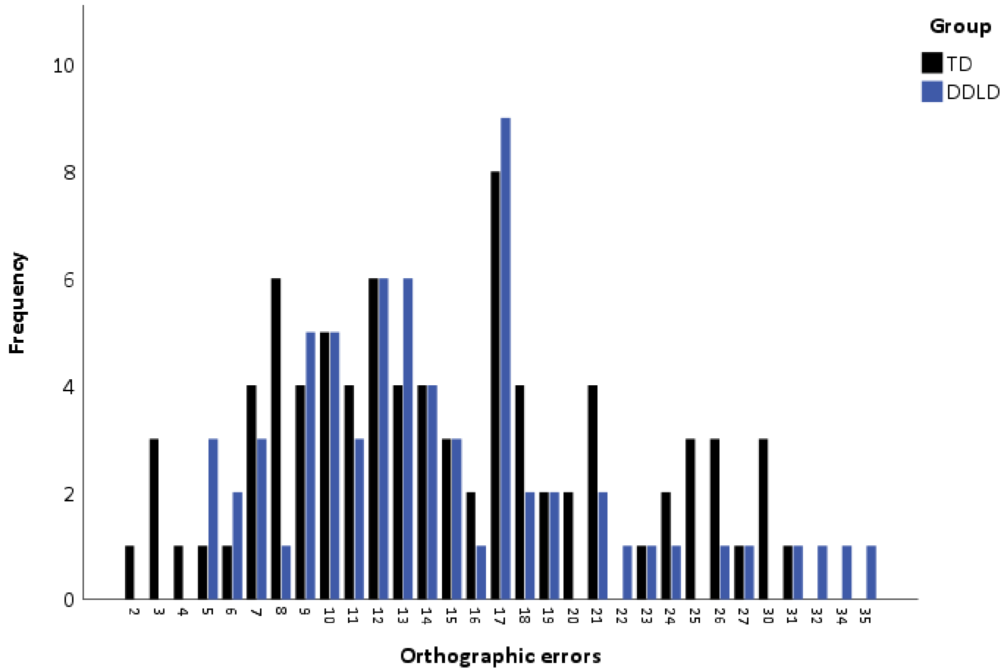
In order to compare the two groups statistically, first Graders were excluded ($n = 20$), and proportions of phonological, grammatical and orthographic errors were computed based on the composite scores of all three categories of errors divided by the number of total words spelled (e.g.

the composite score of phonological errors was divided by the number of total words spelled and then multiplied by 100). As discussed in the Methods chapter, it would not be appropriate to conduct group comparisons based on raw scores since a stop rule was applied in the task - testing was discontinued after six successively incorrectly spelled words - and therefore not all the children attempted to spell the same words. An independent samples *t*-test revealed that after excluding first Graders, the two groups did not differ significantly on age in months, $t_{(127)} = -1.92$, $p = .05$, and therefore age was not controlled in the analyses presented below.

The distributions of the proportional numbers of phonological and grammatical errors were strongly skewed to the right. This was because there were children who did not produce any phonological or grammatical errors. Thus, a non-parametric test, two independent samples Mann-Whitney *U* test, was used to compare statistically the two groups on the proportional numbers of phonological and grammatical errors since the explanatory variables of interest were continuous but not normally distributed. The distribution of the proportional number of orthographic errors was normal. Thus, a parametric test, an independent samples *t*-test, was used to compare statistically the two groups on the proportional number of orthographic errors. Prior to presenting the analyses of group comparisons, Figures 4.4.1., 4.4.2. and 4.4.3. show the distribution of phonological, grammatical and orthographic errors in the TD group and in the DDLD group.


Group
TD
DDLD
Frequency
50
40
30
20
10
0
0
1
2
3
4
5
7
8
Phonological errors


Group
TD
DDLD
Frequency
15
10
5
0
0
1
2
3
4
5
6
7
8
9
10
11
14
Grammatical errors



Figures 4.4.1., 4.4.2. and 4.4.3. The distribution of phonological, grammatical and orthographic errors in the TD group and in the DDLD group

The pattern of results for the number of errors in the three categories is noteworthy, in particular for the category of phonological errors, which is relevant to a research question of the thesis, namely, whether the phonological deficit in children with DDLD can be better accounted for by the Degraded Phonological Representations Hypothesis or the Deficient Phonological Access Hypothesis. Table 4.4.1. presents proportions of phonological, grammatical and orthographic errors in the DDLD group and the TD group and group comparisons. The pattern of results based on proportional scores of phonological, grammatical and orthographic errors found was same in both groups, that is, phonological < grammatical < orthographic errors. Statistical group comparisons revealed that the DDLD group produced significantly more phonological, grammatical and orthographic errors compared to the TD group, as Table 4.4.1. shows. However, even though children with DDLD produced significantly more phonological errors than TD children, the majority of errors (96%) made by the DDLD group did not change the phonology of the word.

Table 4.4.1. Mean (*SD*s) proportions of phonological, grammatical and orthographic errors and 95% *CI*s (lower line) in the DDLD group and the TD group

Composite scores	DDLD group (*n* = 66)	TD group (*n* = 63)	Group comparisons		
	M (*SD*)	*M* (*SD*)	*U*	*Z*	*p*
Phonological errors	3.61 (5.20) (2.33-4.89)	1.22 (2.24) (.65-1.78)	1408.50	-3.43	< .01
Grammatical errors	15.30 (8.74) (13.15-17.45)	7.51 (8.31) (5.42-9.61)	940.00	-5.37	< .001
			t	*p*	η_p^2
Orthographic errors	41.32 (9.99) (38.86-43.78)	36.52 (12.63) (33.33-39.70)	-2.386	< .05	.03

Group comparisons based on proportional scores should be treated with caution. Considering that not all the children attempted to spell all the words in the list, an issue was that children who stopped earlier, attempted to spell only the easier words since the list begins with easier words and continues with more difficult words. As reported by Mouzaki et al. (2007), Rasch model analysis confirmed that the order of words on the basis of their difficulty strongly correlated with the order the words are administered (in the overall sample, $\rho = .94$). Concomitantly, this implies that children who attempted to spell more and therefore more difficult words in the list, were also more likely to produce more phonological, grammatical and orthographic errors. This issue justifies why raw scores of the three types of spelling errors were not considered appropriate in the analyses, although it should be noted that even proportional scores do not correct this issue.

The findings showed that for both groups, phonological errors were rare in the spelling task. With respect to the category of phonological errors, it is likely that children who produced words early in the task might be more likely to spell words phonologically correct, and concomitantly, that children who spelled words later in the task might be likely to spell more words phonologically incorrect, as it is the case that, it terms of spelling, words might be more difficult later in the task. Words later in the task can be more difficult since they can be less frequent, multisyllable and/or with a complex syllable structure. This explanation is illustrated above where analyses revealed that the number of total words spelled showed a positive association with the number of phonological and orthographic errors. This issue raises a question which concerns children's phonological processing skills and the experimenter's ability to interpret correctly children's phonological processing skills based on a spelling task like the one used in this study. It could be argued that more difficult words involve phonological processing skills, in addition to phonological representations, to a greater extent than easy words. Therefore, it might be difficult for the experimenter to tease apart using a spelling task whether children's phonological errors could be attributed to degraded phonological representations, or alternatively that children's phonological representations are intact, and that rather phonological processing skills are less efficient.

Another methodological limitation of the study using the spelling task with a discontinue rule is that in the case where there were parts in a word that were more difficult than other parts, this could not be ruled out in the analyses. Diamanti et al. (2014), for example, found that in Greek, misspelled inflections in verbs are more frequent than misspelled inflections in nouns as the former inflections are less consistent than the latter since verb inflections are more complex than noun inflections. In the word list used, there were several vowel phonemes in a stem in a word that could be misspelled but only one vowel phoneme in the inflectional suffix that could be misspelled. This indicates that using the word list, there were more opportunities for orthographic than grammatical errors, and therefore it would not be surprising to find more orthographic than grammatical errors. Having said that, according to Diamanti et al. (2014), derivational spellings are more difficult than inflectional spellings because since inflections are more common than derivations, children are more familiar with the former than the latter. Ultimately, it might be that children with different profiles (e.g. two types of dyslexia described in the literature, *surface* and *phonological* dyslexia (Castles & Coltheart, 1993), produce different types of spelling errors. Douklias, Masterson, and Hanley (2010) found that children who match a surface dyslexia reading profile made a lot of orthographic spelling errors in irregular words whereas those who match a phonological dyslexia reading profile made phonological spelling errors in nonwords.

4.4.6. Summary

It was predicted that if the Degraded Phonological Representations holds true, qualitative analysis of spelling errors would reveal that the DDLD group will produce a higher proportion of phonological spelling errors than the TD group. This is explained by inaccurate phonological representations in the DDLD group. If the Deficient Phonological Access Hypothesis, however, holds true, qualitative analysis of spelling errors will reveal a similar proportion of phonological spelling errors in the two groups. This is explained by accurate phonological representations but inappropriate orthographic encoding of words using phoneme-to-grapheme mappings that are inappropriate for a particular context in the DDLD group. It was found that children with DDLD produced significantly more phonological errors than TD children, supporting the Degraded Phonological Representations Hypothesis. Importantly, however, in the Greek orthography, phonologically correct spelling was not a challenging task for children with DDLD since the majority of spelling errors (96%) found were phonologically correct, supporting the Deficient Phonological Access Hypothesis.

4.5. Relation of language and literacy skills with semantic and phonological fluency

As presented in Chapter 1b in the section presenting the role of semantic skills in reading ability, another issue concerns whether children’s language and literacy skills have an impact on how efficiently they retrieve lexical items from the mental lexicon. In the current study, the impact of language and literacy skills on word productivity in semantic and phonological fluency tasks is considered. If children’s word productivity in semantic and phonological fluency tasks is indeed predicted by language and literacy skills, poorer semantic and phonological fluency performance in the disordered children will be partly attributed to their inferior language and literacy skills. To this end, the associations between a range of language and literacy tasks to semantic and phonological fluency performance will be first investigated, and next, how much of the variance in semantic and phonological fluency is explained by children's language and literacy skills will be explored in regression models.

The rationale for this investigation is as follows. Nation (2017) proposed the Lexical Legacy Hypothesis which argues that reading experience allows a reader to read words in different semantic contexts, and this leads to a rich and nuanced database about a word and its connections to other words. This hypothesis therefore states that word knowledge is based on lexical co-occurrence in the sense that a word is known as it is related in meaning with other words. In a similar vein, the Lexical Quality Hypothesis claims that "a lexical representation has high quality to the extent that it has a fully specified orthographic representation (a spelling) and redundant phonological representations (one from spoken language and one recoverable from orthographic-to-phonological mappings)" (Perfetti & Hart, 2001, p. 68). Lexical quality therefore concerns the knowledge of the form and the meaning of the word and leads to rapid processing (Perfetti & Hart, 2001). The origin of high-quality representations may therefore be sought in the amount of experience with both oral and written language. Concomitantly, this suggests a relation between children's lexical-semantic representations and their language and literacy skills, and a valid index of lexical quality is performance on semantic and phonological fluency tasks. In support of the Lexical Quality Hypothesis, Dyson et al.'s (2017) intervention study suggested that it is through access to the meaning of the word after pronouncing the word correctly that improves children's ability to learn to read. Furthermore, is automatic (first 15 s of the test period) and controlled (45 s of the test period) processing in the two verbal fluency categories related to children's language and literacy skills? It is hypothesised that if automatic processing reflects the lexico-semantic structure and controlled processing reflects EFs, children's language and literacy skills will have a greater effect on automatic than controlled processing in the regression models.

The TD group and the DDLD group of children were included in the PCA with oblique rotation and in the regression analyses presented in this chapter to investigate the amount of variance in semantic and phonological fluency accounted for by language and literacy skills. The PCA technique attempts to produce a smaller number of linear combinations of the original variables in a way that captures (or accounts for) most of the variability in the pattern of correlations. Tabachnick and Fidell (2007) concluded that: "If you are interested in a theoretical solution uncontaminated by unique and error variability … factor analysis is your choice. If, on the other hand, you simply want an empirical summary of the data set, PCA is the better choice" (p. 635). The PCA will be used to extract the number of factors that could be entered in the linear regression models. Linear regression analyses will be used to further identify factors with independent influences on semantic and phonological fluency performance.

4.5.1. Statistical methods. All the associations were calculated as Pearson's correlation coefficient. Partial associations (controlling for age) among all language measures and among all literacy measures in the overall sample are presented first. Partial associations between language and literacy measures with semantic and phonological fluency in the overall sample and by subgroup are then presented. As described below, raw scores of those language and literacy measures which were correlated significantly with semantic and phonological fluency were converted to *z* scores. *Z* scores of language and literacy measures associated with semantic and phonological fluency were entered into the PCA. The PCA revealed that language and literacy measures loaded on one single factor, named the Language and Literacy variable (hereafter LangLit variable), and then the LangLit variable was entered in the regression analyses. For semantic and phonological fluency, the linear regression controlled for age in months and NVIQ in Step 1, and the LangLit variable was entered in Step 2 to investigate how much of the variance in semantic and phonological fluency was accounted for by the LangLit variable after controlling for age and NVIQ in Step 1. Regression analyses were first conducted in the overall sample and then by subgroup.

4.5.2. Associations between language and literacy measures with semantic and phonological fluency in the overall sample

Moderate and strong significant partial correlations (controlling for age) were found in the overall sample among the language and among the literacy measures (respectively, *r*s from .31 to .67; *r*s from .38 to .77), as shown in Tables 4.5.1. and 4.5.2.

Table 4.5.1 Partial correlations (controlling for age) among all language measures in the overall sample

Variables		1	2	3	4	5
1	Receptive vocabulary		.46***	.33***	.44***	.52***
2	Syntax comprehension			.32***	.31***	.34***
3	Sentence repetition				.38***	.47***
4	WISC Similarities					.67***
5	WISC Vocabulary					

Note: ****p* < .001.

Table 4.5.2. Partial correlations (controlling for age) among literacy measures in the overall sample

Variables		1	2	3	4
1	L'Alouette		.77***	.50***	.62***
2	Spelling			.38***	.44***
3	Reading accuracy				.68***
4	Reading fluency				

Note: ****p* < .001.

Having presented the partial associations among all language measures and among all literacy measures in the overall sample, Table 4.5.3. presents partial correlations (controlling for age) between language and literacy measures with semantic fluency in the overall sample, in the DDLD group and in the TD group. Partial correlations in the overall sample revealed that semantic fluency correlated significantly with all five language measures, namely, receptive vocabulary, syntax comprehension, sentence repetition, WISC Similarities and WISC Vocabulary. Semantic fluency also correlated significantly with two literacy measures, namely, l'Alouette and spelling tasks. Given that nonsignificant correlations were found between semantic fluency with reading accuracy, reading fluency, syllable reading and nonword reading, the four measures are not considered further.

Table 4.5.3. Partial correlations (controlling for age) between semantic fluency (number of correct responses) and language and literacy tasks in the overall sample, in the DDLD group and in the TD group

	Overall sample		DDLD group		TD group	
	r	*p*	*r*	*p*	*r*	*p*
Language Skills						
Verbal comprehension: WISC Similarities	.313***	< .001	.336**	.006	.075	.503
Verbal comprehension: WISC Vocabulary	.323***	< .001	.254*	.041	.124	.266
Syntax comprehension: DVIQ Test	.248**	.002	.238	.056	.142	.202
Sentence repetition: DVIQ Test	.248**	.002	.255*	.040	-.008	.942
Receptive vocabulary: PPVT-R	.280**	.001	.162	.198	.200	.072
Literacy Skills						
Text-reading fluency: L'Alouette task	.263**	.001	.092	.468	.069	.538
Text-reading fluency: Reading Test Alpha	.183	.067	.107	.438	-.103	.501
Reading accuracy: Reading Test Alpha	.097	.335	.007	.960	-.101	.509
Syllable reading: Test of DIRD	.015	.922	-.522	.149	.002	.990
Nonword reading: Test of DIRD	-.071	.638	-.731*	.025	-.031	.857
Spelling ability: Spelling-to-dictation task	.268**	.001	.191	.128	-.005	.962

Notes: WISC, Wechsler Intelligence Scale for Children; DVIQ Test, Diagnostic Verbal Intelligence Test; PPVT-R, Peabody Picture Vocabulary Test-Revised; Test of DIRD, Test of Detection and Investigation of Reading Difficulties; $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Having presented the partial correlations (controlling for age) between phonological fluency (number of correct responses) and language and literacy tasks in the overall sample, in the DDLD group and in the TD group, Table 4.5.4. presents partial correlations (controlling for age) between language and literacy measures with phonological fluency in the overall sample, in the DDLD group and in the TD group. Partial correlations in the overall sample revealed that phonological fluency correlated significantly with all five language measures, namely, receptive vocabulary, syntax comprehension, sentence repetition, WISC Similarities and WISC Vocabulary. Phonological fluency also correlated significantly with two literacy measures, namely, l'Alouette and spelling tasks. Given that nonsignificant correlations were found between phonological fluency with reading accuracy, reading fluency, syllable reading and nonword reading, the four measures are not considered further.

Table 4.5.4. Partial correlations (controlling for age) between phonological fluency (number of correct responses) and language and literacy tasks in the overall sample, in the DDLD group and in the TD group

	Overall sample		DDLD group		TD group	
	r	*p*	*r*	*p*	*r*	*p*
Language Skills						
Verbal comprehension: WISC Similarities	.378***	< .001	389**	.001	.208	.060
Verbal comprehension: WISC Vocabulary	.422***	< .001	.451***	< .001	.222*	.045
Syntax comprehension: DVIQ Test	.288***	< .001	.355**	.004	.119	.289
Sentence repetition: DVIQ Test	.334***	< .001	.390**	.001	.084	.455
Receptive vocabulary: PPVT-R	.291***	< .001	.153	.224	.246*	.026
Literacy Skills						
Text-reading fluency: L'Alouette task	.215**	.009	.027	.832	.035	.754
Text-reading fluency: Reading Test Alpha	.170	.089	.081	.558	-.047	.760
Reading accuracy: Reading Test Alpha	.112	.267	-.013	.923	.039	.799
Syllable reading: Test of DIRD	.105	.488	-.385	.306	.117	.496
Nonword reading: Test of DIRD	-.018	.903	-.447	.228	.048	.783
Spelling ability: Spelling-to-dictation task	.363***	< .001	.242	.052	.240	.030

Notes: WISC, Wechsler Intelligence Scale for Children; DVIQ Test, Diagnostic Verbal Intelligence Test; PPVT-R, Peabody Picture Vocabulary Test-Revised; Test of DIRD, Test of Detection and Investigation of Reading Difficulties; $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

4.5.5. Explanatory factor analysis

In order to investigate further the relationship between semantic fluency and language and literacy measures, raw scores of all five language measures and the two literacy measures correlated significantly with semantic fluency were converted to *z* scores. *Z* scores were computed relative to the TD group's mean and standard deviation for each task, with the mean *z*-score being equal to 0 and *SD* equal to 1 for all tasks. *Z* scores of all seven tasks associated significantly with semantic fluency were entered into the PCA. The PCA was used to extract the factors followed by oblique rotation of factors using Oblimin rotation (delta = 0). The number of factors to be retained was guided by two decision rules: Kaiser's criterion (eigenvalues above 1) and by inspection of the Scree plot. The PCA revealed that all five language measures and the two literacy measures loaded on a single factor, revealing just one component rather than two separate language and literacy components, as Figure 4.5.1. presents. The component identified from the PCA was named Language and Literacy variable (hereafter LangLit variable). The mean (*SD*) for the LangLit variable was .00 (.85) for the TD group and -.43 (.52) for the DDLD group. Prior to presenting the results of the regression analyses, some details on the PCA are provided.

The sample was first assessed for its suitability for factor analysis. The Correlation Matrix table showed that all correlation coefficients were of .3 and above. This verified that the data set is suitable for factor analysis. Bartlett's Test of Sphericity was highly significant ($p < .001$; Bartlett, 1954) and the Kaiser-Meyer-Olkin measure of sampling adequacy value of .871 supported the factorability of the matrix. The PCA revealed one eigenvalue exceeding 1, explaining 63.27% of the variance. Inspection of the Scree plot (see Figure 4.5.1.) also supported a one factor solution as a clear break was found between the first and the second component, indicating that component one explains much more of the variance than the remaining components. These analyses were based on Pallant (2010).

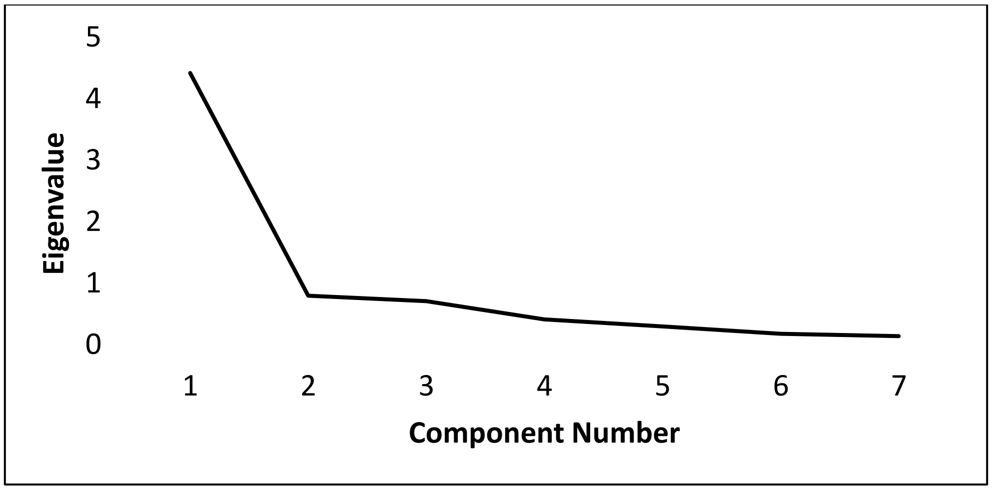


Figure 4.5.1. Scree plot showing the number of components generated by the Principal Component Analysis (PCA) with oblique rotation in the overall sample and each component's eigenvalue

Results of the PCA revealed that all variables loaded quite strongly (above .4) on the first component. All steps in procedure 1 were repeated but selecting one factor to extract of SPSS output. This procedure showed that 62.39% of the variance was explained, for the one-factor solution. Inspection of Communalities (an index of how much of the variance is explained by any variable) indicated, however, that sentence repetition did not fit well with the other variables in its component, having the lowest communality value (.279) and the lowest loading (.528). In order to improve the factor, this information was used to remove sentence repetition. The PCA with direct oblimin (delta = 0) was repeated without sentence repetition, and the LangLit variable presented below excludes therefore sentence repetition. The repeated PCA showed that all variables loaded above .67 on the one-single factor, and Component 1 explained 69.97% of the variance. The LangLit variable correlated moderately with semantic and phonological fluency, respectively, $r = .38$, $p < .001$; $r = .43$, $p < .001$. Figure 4.5.2. shows the LangLit variable plotted against age in months in the two groups and that there is a clear split between children with and without DDLD.

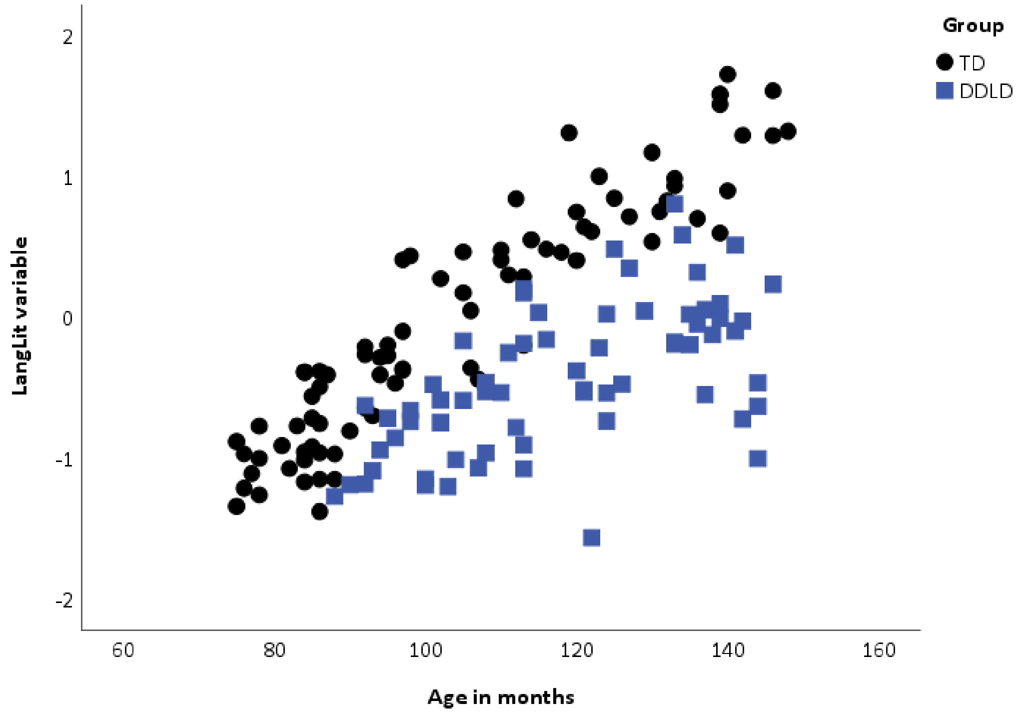


Figure 4.5.2. The LangLit variable in *z* scores, plotted against age in months, in the DDLD group and the TD group

4.5.6. Regression analyses in the overall sample

As table 4.5.5. shows, in the overall sample, linear regression with age in months and NVIQ as factors accounted for 32.1% of the variance in semantic fluency and 28.5% of the variance in phonological fluency. The LangLit variable entered in Step 2 in the linear regression model accounted for 6 and 8.8% of the variance in semantic and phonological fluency, respectively, after controlling for age in months and NVIQ in Step 1. The contribution of the LangLit variable was highly significant in both conditions. The overall percentage variance in the number of correct responses produced in semantic and phonological fluency being accounted for by the model (age in months, NVIQ and the LangLit variable) was, respectively, 38.1 and 37.3%. More than half of the variance therefore was not accounted for by the model, implying that unmeasured variables and measurement error are relevant.

Table 4.5.5. Percentage of variance in semantic and phonological fluency explained by age in months and NVIQ entered in Step 1 and proportion of variance explained by the LangLit variable entered in Step 2 in the overall sample

Explanatory variables	R^2	ΔR^2	β age	β NVIQ	β LangLit
Semantic fluency	32.10***	.060***	.508***	.246***	.363***
Phonological fluency	18.50***	.088***	.450***	.284***	.440***

Notes: For each regression, age in months and NVIQ entered in Step 1 and the LangLit variable was entered in Step 2. For Step 1, information is provided on the amount of variance accounted for by age in months and NVIQ (R^2), and for Step 2, information is provided on the amount of variance accounted for by the LangLit variable (ΔR^2). The β-values for the three predictor variables are presented in Columns 4, 5 and 6. Significance values are given where they are relevant; $^{***}p < .001$.

Checks indicated that Mahalanobis distances were less than 15 for all the children for the number of correct responses both in semantic and phonological categories, therefore all cases were included in regressions (Field, 2013). The regression model was significant both for semantic fluency, $F_{(3, 145)} = 29.790$, $p < .001$, accounting for 38.1% of the variance, and for phonological fluency, $F_{(3, 145)} = 28.803$, $p < .001$, accounting for 37.3% of the variance.

4.5.7. Regression analyses by subgroup

Regression analyses by subgroup were also conducted as a large amount of the variance could not be accounted for by the model in the overall sample. In the DDLD group, the LangLit variable predicted 7.9 and 15.3% of the variance in semantic and phonological fluency, respectively, after controlling for age in months and NVIQ. The LangLit variable was a significant predictor of both conditions. Age in months and NVIQ accounted for 21.3 and 15.9% of the variance in semantic and phonological fluency, respectively. In the TD group, the LangLit variable did not predict any of the variance in semantic and phonological fluency, after controlling for age in months and NVIQ. Age in months and NVIQ accounted for 43.8 and 42.1% of the variance in semantic and phonological fluency, respectively. Table 4.5.6. presents the percentage of variance in semantic and phonological fluency explained by age in months and NVIQ entered in Step 1 and the proportion of variance in semantic and phonological fluency explained by LangLit variable entered in Step 2 in the DDLD group and the TD group.

Table 4.5.6. Percentage of variance in semantic and phonological fluency explained by age in months and NVIQ entered in Step 1 and proportion of variance explained by the LangLit variable entered in Step 2 in the DDLD group and in the TD group

		DDLD group			
Explanatory variables	R^2	ΔR^2	*β* age	*β* NVIQ	*β* LangLit
Semantic fluency	21.30**	.079*	.427***	.172	.362*
Phonological fluency	15.90**	.153***	.378**	.123	.504***
		TD group			
Explanatory variables	R^2	ΔR^2	*β* age	*β* NVIQ	*β* LangLit
Semantic fluency	43.80***	.001	.601***	.187*	.086
Phonological fluency	42.10***	.014	.518***	.305**	.367

Notes: For each regression, age in months and NVIQ were entered in Step 1 and the LangLit variable was entered in Step 2. For Step 1, information is provided on the amount of variance accounted for by age in months and NVIQ (R^2), and for Step 2, information is provided on the amount of variance accounted for by the LangLit variable (ΔR^2). The *β*-values for the three predictor variables are presented in Columns 4, 5 and 6. Significance values are given where they are relevant; $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Checks by subgroup indicated that Mahalanobis distances were less than 15 for all the children for the number of correct responses both in semantic and phonological categories, therefore all cases were included in regressions (Field, 2013). Further, the assumptions for the linear regression by subgroup, as reported above, were all met. In the TD group, the regression model was significant both for semantic fluency, $F_{(3, 79)} = 29.569$, $p < .001$, accounting for 43.9% of the variance, and for phonological fluency, $F_{(3, 79)} = 20.266$, $p < .001$, accounting for 43.5% of the variance. In the DDLD group, the regression model was significant both for semantic fluency, $F_{(3, 62)} = 8.513$, $p < .001$, accounting for 29.2% of the variance, and for phonological fluency, $F_{(3, 62)} = 9.345$, $p < .001$, accounting for 31.1% of the variance.

4.5.8. Relation of language and literacy skills with automatic and controlled processing in semantic and phonological fluency categories

As table 4.5.7. shows, in the overall sample, the LangLit variable entered in Step 2 in the linear regression model accounted for word productivity throughout the one-minute test period in both verbal fluency categories, after controlling for age in months and NVIQ in Step 1. The contribution of the LangLit variable was highly significant in the first 15 s of the semantic condition, accounting for 8.2% of the variance in semantic fluency performance in the first 15 s. The contribution of the LangLit variable was also significant in the first 15 s and the subsequent 45 s of the phonological condition, accounting for 6% of the variance in phonological fluency performance both in the first 15 s and the subsequent 45 s.

Table 4.5.7. Percentage of variance in the first 15 s and the subsequent 45 s in semantic and phonological categories explained by age in months and NVIQ entered in Step 1 and proportion of variance explained by the LangLit variable entered in Step 2 in the overall sample

Explanatory variables	R^2	ΔR^2	β age	β NVIQ	β LangLit
Semantic fluency first 15 s	27.60^{***}	$.082^{***}$	$.475^{***}$	$.221^{**}$	$.424^{***}$
Semantic fluency subsequent 45 s	21.80^{***}	$.033^{*}$	$.437^{***}$	$.159^{*}$	$.270^{*}$
Phonological fluency first 15 s	21.20^{***}	$.060^{**}$	$.416^{***}$	$.193^{*}$	$.363^{**}$
Phonological fluency subsequent 45 s	17.60^{**}	$.060^{**}$	$.379^{***}$	$.178^{*}$	$.363^{**}$

Notes: For each regression, age in months and NVIQ entered in Step 1 and the LangLit variable was entered in Step 2. For Step 1, information is provided on the amount of variance accounted for by age in months and NVIQ (R^2), and for Step 2, information is provided on the amount of variance accounted for by the LangLit variable (ΔR^2). The β-values for the three predictor variables are presented in Columns 4, 5 and 6. Significance values are given where they are relevant; $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Checks indicated that Mahalanobis distances were less than 15 for all the children for the number of total responses in the first 15 s and the subsequent 45 s both in semantic and phonological categories, therefore all cases were included in regressions (Field, 2013). For semantic fluency, the regression model was significant for the first 15 s, $F_{(3, 145)} = 26.921$, $p < .001$, accounting for 35.8% of the variance, and the subsequent 45 s, $F_{(3, 145)} = 16.190$, $p < .001$, accounting for 25.1% of the variance. For phonological fluency, the regression model was significant for the first 15 s, $F_{(3, 145)} = 18.076$, $p < .001$, accounting for 27.2% of the variance, and the subsequent 45 s, $F_{(3, 145)} = 14.968$, $p < .001$, accounting for 23.6% of the variance.

4.5.9. Summary

It was predicted that some of the variance in semantic and phonological fluency performance would be accounted for by children's language and literacy measures after controlling for age in months and NVIQ in the analyses. The PCA technique revealed a single language and literacy component which was defined by tasks of verbal comprehension, syntax comprehension, receptive vocabulary, text-reading rate, and spelling (sentence repetition was not included as it did not fit well with the other variables in its component). The component identified by the PCA was named Language and Literacy variable (LangLit variable). Findings showed that in the overall sample, after controlling for age in months and NVIQ, 6 and 8.8%, respectively, of the variance in semantic and phonological fluency was accounted for by the LangLit variable. Analyses by subgroup showed that in the DDLD group, after controlling for age in months and NVIQ, 7.9 and 15.3%, respectively, of the variance in semantic and phonological fluency was accounted for by the LangLit variable. In the TD group, after controlling for age in months and NVIQ, the LangLit variable was a nonsignificant predictor both of semantic and phonological fluency. The results demonstrated that children's language and literacy skills uniquely and significantly predicted semantic and phonological fluency performance in the overall sample, but that the pattern was different when analyses by subgroup were carried out.

Analyses by subgroup revealed that children's language and literacy skills uniquely and significantly predicted semantic and phonological fluency performance in the DDLD group, but not in the TD group. Language and literacy skills related both to automatic and controlled processing in semantic and phonological fluency tasks.

4.6. Groups' Performance and Group Differences on the Design Fluency Task

How specific is the verbal fluency deficit in children with DDLD: Does it extend to a nonverbal task (design fluency)?

4.6.1. Statistical methods. The associations were calculated as Pearson's correlation coefficient. A paired samples *t*-test was initially used to test the difference in the number of correct designs in the two arrays, structured and random, of the design fluency task. Group differences on the number of total, correct, incorrect and repeated designs produced were tested using regression analysis. For each regression, age in months and NVIQ were entered in Step 1 and participant group was entered in Step 2. For Step 1, information is provided on the amount of variance accounted for by age in months and NVIQ (R^2), and for Step 2, information is provided on the proportion of variance accounted for by group (ΔR^2). The β-values for the three predictor variables are also given.

4.6.2. Regression analyses on the design fluency variables

Prior to presenting the results of regression analyses on the design fluency analyses, the rationale for entering in the analyses composite scores based on the number of total, correct, incorrect and repeated designs in the two arrays is presented. The number of correct designs in the structured array was strongly associated with the number of correct designs in the random array both in the TD group, $r_{(80)} = .58$, $p < .001$, and in the DDLD group, $r_{(63)} = .55$, $p < .001$. Further, in both groups, a paired samples *t*-test showed that the number of correct designs did not significantly differ in the two arrays; for the TD group: $t_{(82)} = -1.12$, $p = .26$; for the DDLD group: $t_{(65)} = .63$, $p = .52$. Thus, composite scores based on the number of total, correct, incorrect and repeated designs in the two arrays were computed, as Table 4.6.1. shows. In both groups, design fluency correlated strongly with age; for the TD group: $r_{(83)} = .60$, $p < .001$; for the DDLD group: $r_{(66)} = .52$, $p < .001$. Nonsignificant correlations were found between design fluency and NVIQ in the TD group, $r_{(80)} = .08$, $p = .47$, and the DDLD group, $r_{(63)} = .05$, $p = .68$.

Table 4.6.1. Linear regression analyses conducted on the design fluency variables

Design fluency variables	DDLD group *M* (*SD*)	TD group *M* (*SD*)	R^2	ΔR^2	*β* age	*β* NVIQ	*β* group
Total designs	25.78 (9.08)	24.04 (7.24)	19.3^{***}	.001	$.405^{***}$	$.167^{*}$	.025
Correct designs¥	21.72 (7.55)	20.36 (6.23)	39.5^{***}	.000	$.556^{***}$	$.288^{***}$	-.003
Incorrect designs	1.57 (3.27)	1.02 (1.47)	0.0	.016	.002	.010	.140
Repeated designs	2.48 (3.26)	2.66 (3.36)	11.9^{***}	.000	$-.266^{**}$	$-.217^{**}$	-.019

Notes: For each regression, age in months and NVIQ were entered in Step 1 and participant group was entered in Step 2. For Step 1, information is provided on the variance accounted for by age in months (R^2), and for Step 2, information is provided on the proportion of variance accounted for by group (ΔR^2). The *β*-values for the three predictor variables are given. Significance values are given where they are relevant. Means and *SD*s for the two participant groups are also presented; $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$; ¥max correct score is 35; Table presents total scores in the two design fluency conditions (i.e. structured and random).

Checks indicated that Mahalanobis distances were less than 15 for all the children for the number of total, correct, incorrect and repeated designs, therefore all cases were included in regressions (Field, 2013). The regression model was significant for the number of total designs, $F_{(3, 145)} = 11.619$, $p < .001$, accounting for 19.4% of the variance, for the number of correct designs, $F_{(3, 145)} = 31.529$, $p < .001$, accounting for 39.5% of the variance, and for the number of repeated designs, $F_{(3, 145)} = 3.320$, $p = .022$, accounting for 6.4% of the variance. The model was nonsignificant for the number of incorrect designs, $F_{(3, 145)} = .795$, $p = .499$.

Significant and nonsignificant effects of age and NVIQ. Age and NVIQ accounted for 19.3 and 39.5% of the variance in total and correct designs, respectively, being a significant predictor of both measures. Age and NVIQ were nonsignificant predictors of the number of incorrect designs, but both were significant predictors of the number of repeated designs accounting for 11.9% of the variance on repeated designs.

Nonsignificant effects of group. Group was a nonsignificant predictor of the number of total, correct, incorrect and repeated designs. Thus, design fluency cannot differentiate well between children with and without DDLD. Further, as Figure 4.6.1. shows, there was a lot of overlap between the two groups with respect to the number of correct designs.

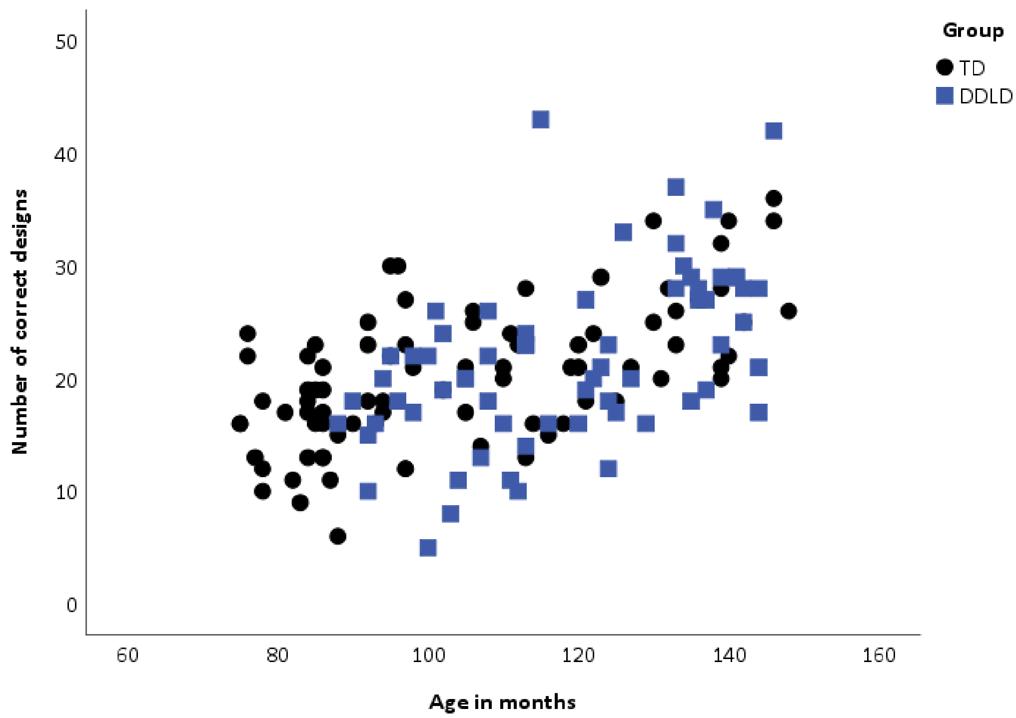


Figure 4.6.1. Scatterplot showing the number of correct designs produced in the DDLD group and the TD group, plotted against age in months

Furthermore, children with DDLD showed poorer semantic and phonological fluency performance relative to their TD peers even after design fluency performance was controlled, demonstrating the specificity of their verbal fluency deficit. Specifically, in the overall sample, a partial (controlling for age) correlation revealed that the number of correct responses produced in semantic fluency tasks was weakly correlated with the number of correct designs generated in the design fluency task, $r_{(146)} = .188$, $p = .022$. Therefore, in order to assess the specificity of the semantic fluency deficit in children with DDLD, an ANCOVA was carried out, with the number of correct responses in semantic fluency tasks as a dependent variable, group as a fixed factor, and age in months and the number of correct designs generated in the design fluency task as covariate variables. ANCOVA revealed that there were group differences for the mean number of correct responses produced in semantic fluency tasks, $F_{(1, 145)} = 11.520$, $p = .001$, $\eta p^2 = .074$. Likewise, in the overall sample, a partial (controlling for age) correlation revealed that the number of correct responses produced in phonological fluency tasks was weakly correlated with the number of correct designs generated in the design fluency task, $r_{(146)} = .268$, $p = .001$. In order to assess the specificity of the phonological fluency deficit in children with DDLD, an ANCOVA was carried out, with the number of correct responses in phonological fluency tasks as a dependent variable, group as a fixed factor, and age in months and the number of correct designs generated in the design fluency task as covariate variables. ANCOVA revealed that there were group differences for the mean number of correct responses produced in phonological fluency tasks, $F_{(1, 145)} = 9.687$, $p = .002$, $\eta p^2 = .063$. Together the results demonstrate that after the effects of age and design fluency performance were controlled, children with DDLD still show lexical retrieval difficulties in semantic and phonological fluency tasks, arguing for the specificity of the verbal fluency deficit in children with DDLD.

Summary: Significant and nonsignificant effects of age, NVIQ and group in the design condition, and the specificity of the verbal fluency deficit

Age and NVIQ were significant predictors of the number of total, correct and repeated designs, and nonsignificant predictors of the number of incorrect designs. Group was a nonsignificant predictor of all four design fluency variables, namely, the number of total, correct, incorrect and repeated designs. The specificity of the verbal fluency deficit is supported by evidence showing that after the effects of age and design fluency performance were controlled, children with DDLD still showed lexical retrieval difficulties in semantic and phonological fluency tasks.

4.6.3. Summary

It was predicted that in accordance with the two phonological hypotheses and the two lexical-semantic models, the DDLD group will not differ on design fluency performance from the TD group. It was hypothesized that if there is a slower processing speed in children with DDLD accounting for lower semantic and phonological fluency performance, lower design fluency performance would be also found in the DDLD group; however, if only verbal processing difficulties were to underlie poorer semantic and phonological fluency performance in children with DDLD, the two groups would show similar design fluency performance. The findings support the two phonological hypotheses and the two lexical-semantic models considered in that children with DDLD did not differ from TD children on the number of correct designs generated in the design fluency task, implying that children with DDLD perform age-appropriately. The specificity of the verbal fluency deficit is supported by evidence showing that after the effects of age and design fluency performance were controlled, children with DDLD still showed lexical retrieval difficulties in semantic and phonological fluency tasks.

4.7. Results meeting the old and the new threshold of statistical significance

Prior to presenting a summary of the Results, results meeting the old and the new threshold of statistical significance are presented in this section. Benjamin et al. (2018) reported that the threshold for defining statistical significance should be redefined, proposing a change from $p < .05$ to $p < .005$. They proposed that results meeting the new threshold should be called significant and results meeting the old threshold (i.e. $p < .05$) should be called “suggestive”. In the current study, statistically significant results are considered to be those meeting the old threshold. However, the aim of this section is to inform the reader for findings which would be likely to be replicated and for those which would be less likely to be replicated based on the threshold of statistical significance the results reached, as discussed in the Methods chapter.

4.7.1. Results meeting the new threshold of statistical significance

- Switching and clustering behaviour in semantic and phonological fluency tasks is associated with semantic and phonological fluency performance in TD children and children with DDLD
- Semantic and phonological fluency are strongly correlated
- The DDLD group has poorer semantic fluency performance in the first 15 s of the test period compared to the TD group
- The DDLD group produces a smaller number of semantic clusters compared to the TD group.
- Language and literacy skills predict semantic and phonological fluency performance in the overall sample
- DDLD children's language and literacy skills predict their phonological fluency performance
- Language and literacy skills predict semantic fluency performance in the first 15 s of the test period in the overall sample.

4.7.2. Results meeting the old threshold of statistical significance

- Compared to the TD group, the DDLD group produced fewer total, correct and incorrect responses in semantic fluency tasks, in addition to fewer total responses in the subsequent 45 s of the test period
- Compared to the TD group, the DDLD group produced fewer total, correct and incorrect responses in phonological fluency tasks, in addition to fewer total responses in the first 15 s and in subsequent 45 s of the test period and a smaller number of switches.

4.8. Summary of the Results

Research Question 1. What is the structure of the lexicon in children with DDLD compared to TD children? Is poorer semantic fluency performance in children with DDLD better explained by impoverished semantic structure or slower retrieval processes?

The DDLD group showed significantly poorer semantic fluency performance and produced significantly fewer clusters compared to the TD group; however, the two groups did not differ on the number of switches and average cluster size, as Figure 4.8.1. illustrates. Given that a similarly-sized average cluster was found in the two groups, the findings indicate that lexical-semantic structure does not differ in the two groups, and that the poorer semantic fluency performance in children with DDLD relative to TD children is driven by slower lexical retrieval process. It is concluded that the lexical retrieval difficulties experienced by children with DDLD in semantic fluency tasks are better explained by the Slow-Retrieval Model than by the Poor Lexical-Semantic Structure Model.

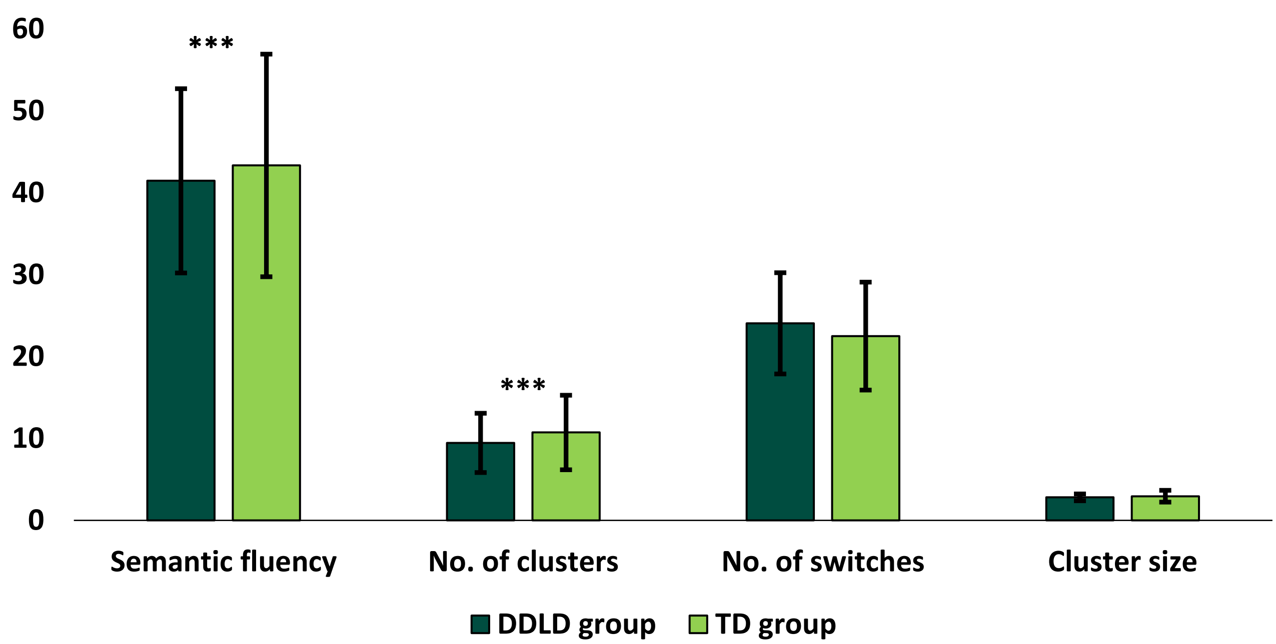


Figure 4.8.1. Raw scores (bars represent *SD*s) of the number of correct responses (semantic fluency), the number of switches, the number of clusters and average cluster size in semantic fluency categories in the two groups

Notes: $^{***}p < .001$; statistical significance is based on regression analyses.

Further, there were nonsignificant differences based on proportional scores for all types of incorrect responses. The DDLD group produced, however, a significantly higher error ratio in the semantic condition than the TD group, suggesting that monitoring of word search, retrieval processes and executive control was somehow harder for the DDLD group than the TD group.

Research Question 2a. Where does the phonological deficit in children with DDLD lie? Is poorer phonological fluency performance in children with DDLD better explained by degraded phonological representations or by deficient explicit access to (intact) phonological representations?

The DDLD group showed significantly poorer phonological fluency performance and produced significantly fewer switches compared to the TD group; however, the two groups did not differ on the number of clusters and average cluster size, as Figure 4.8.2. illustrates. A similarly-sized average cluster, considered to be an implicit phonological measure of the quality of phonological representations, suggested that in children with DDLD phonological representations were as robust and distinct as those of TD children. This is consistent with the Deficient Phonological Access Hypothesis. Further, there were nonsignificant differences based on proportional scores for all types of incorrect responses, and a nonsignificant difference on error ratio, suggesting no difficulty with monitoring of word search, retrieval processes and executive control for the DDLD group relative to the TD group.

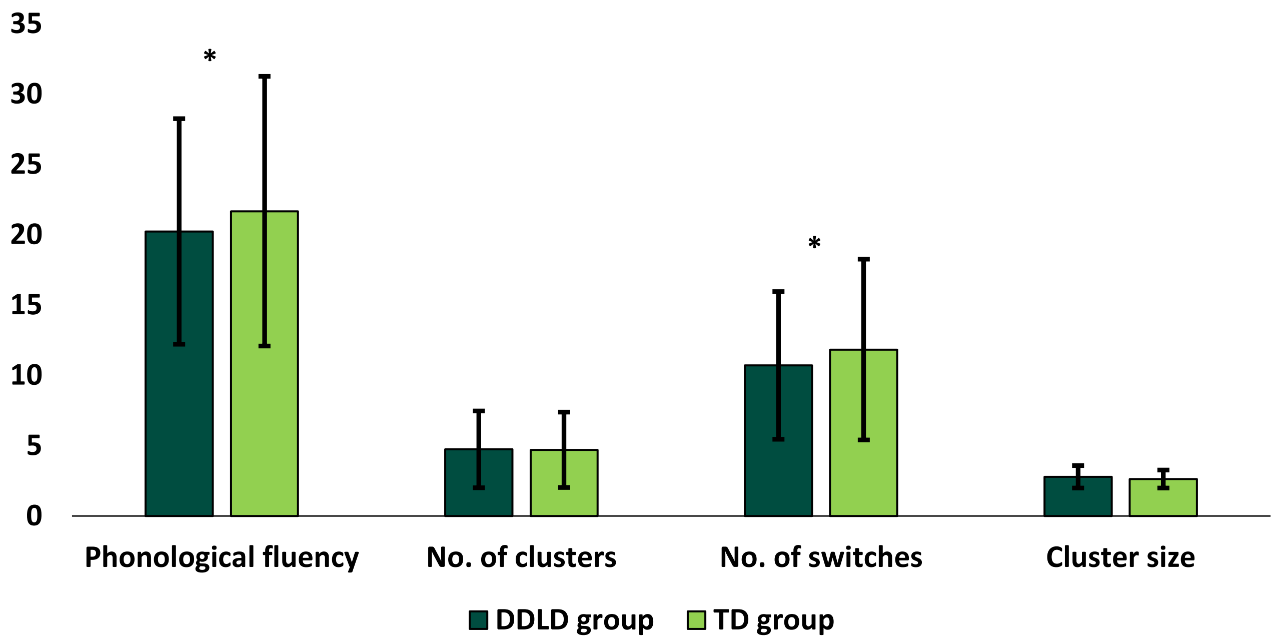


Figure 4.8.2. Raw scores (bars represent *SD*s) of the number of correct responses (phonological fluency), the number of switches, the number of clusters and average cluster size in phonological fluency categories in the two groups

Notes: $^{*}p < .05$; statistical significance is based on regression analyses.

Research Question 2b. Is the phonological deficit in children with DDLD better explained by degraded phonological representations or by deficient explicit access to (intact) phonological representations using a range of phonological tasks and a spelling-to-dictation task?

With respect to phoneme deletion tasks, findings showed that nonword-initial phonemes can be deleted similarly accurately and fast when the target nonword was monosyllabic and had a simple syllable structure (CVC) in the two groups, but when the target nonword was either trisyllabic with a simple syllable structure (CVCVCV) or monosyllabic with a complex syllable structure (CCV), children with DDLD showed significantly poorer performance relative to TD children (Figure 4.8.3.). These findings in turn imply that whenever high phonological STM demands are placed upon DDLD children's phonological representations, as evidenced in the phoneme deletion tasks of CVCVCV and CCV items, the phonological deficit in children with DDLD becomes evident. Age-appropriate performance was, however, observed in children with DDLD when phonological STM demands are not placed upon DDLD children's phonological representations, as evidenced in the phoneme deletion task of CVC items. The findings are limited, however, by the ceiling effects observed in the phoneme deletion task of CVC items, implying that if more a more sensitive task of phoneme deletion of CVC items was used, group differences might possibly be observed.

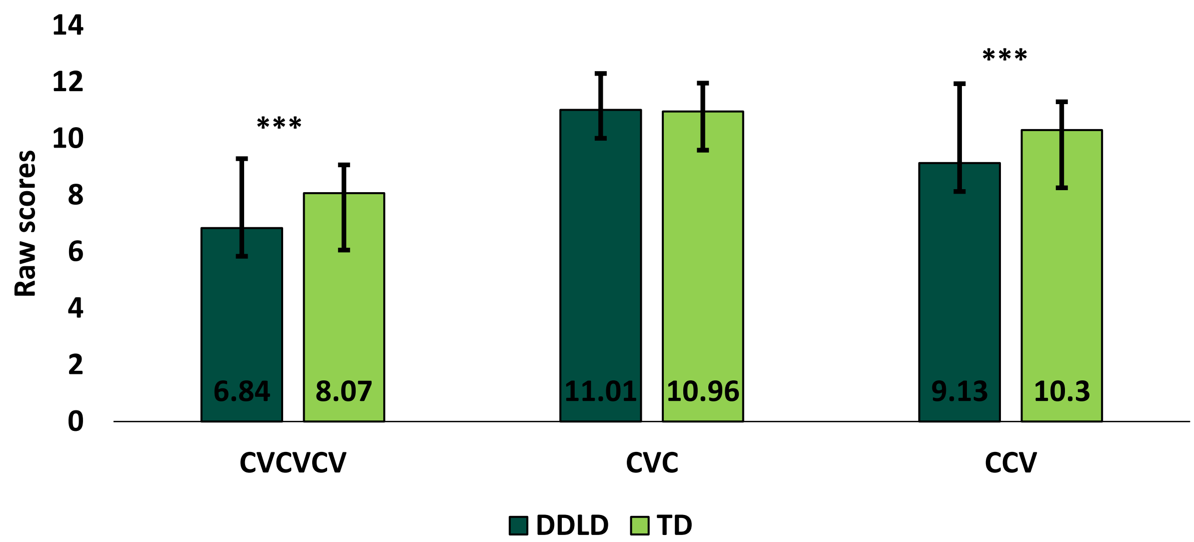


Figure 4.8.3. Raw scores of accuracy (bars represent *SD*s) for phoneme deletion tasks of CVCVCV items, CVC items and CCV items

Notes: $^{***}p < .001$; statistical significance is based on regression analyses.

A NWR task was used to investigate phonological STM deficits in children with DDLD in order to support the above-mentioned argument that phonological tasks are harder when phonological STM demands are high. The DDLD group's performance was found to fall more sharply than the TD group's performance as length increased (Figure 4.8.4.).

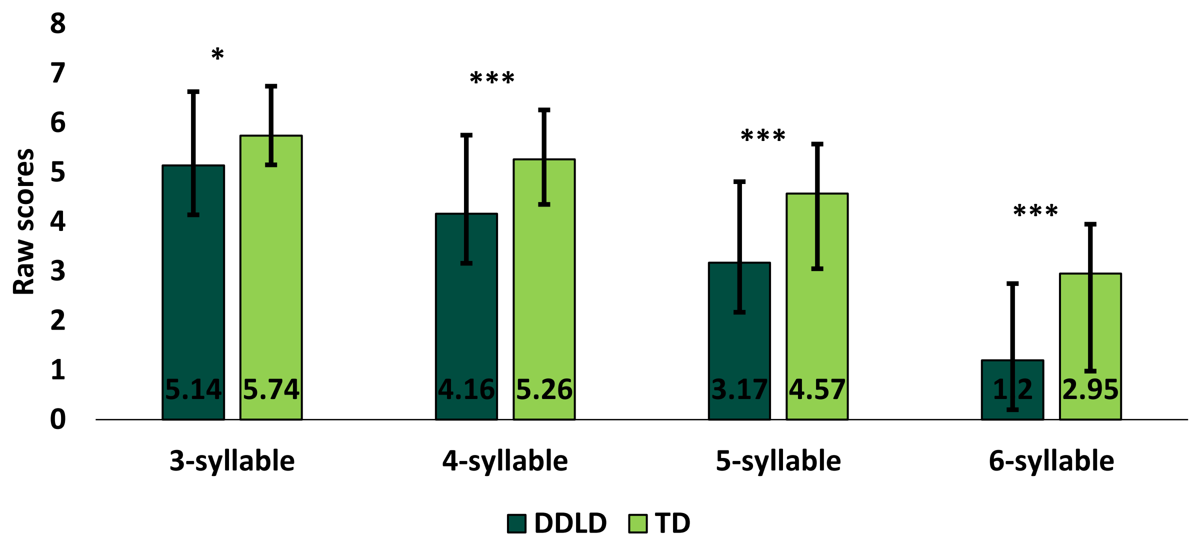


Figure 4.8.4. Raw scores of accuracy (bars represent *SD*s) for 3-, 4-, 5- and 6-syllable items of the NWR task in the two groups

Notes: $^{*}p < .05$, $^{***}p < .001$; statistical significance is based on MANOVAs.

With respect to the RAN task, findings showed that the DDLD group differed from the TD group in the time they spent naming all pictures (Figure 4.8.5.); however, the two groups did not differ on the number of phonological errors observed in the RAN task (Figure 4.8.6.). The findings in turn imply that children with DDLD have difficulty in accessing phonological representations of the items in the RAN task, with their phonological representations being accurate as evidenced by the low number of phonological errors observed in the DDLD group. The number of semantic errors in the RAN task were also very low; however, considering that the DDLD group found to show significantly more semantic errors than the TD group, it can be assumed that their difficulty in accessing the correct names of pictures might have resulted in semantic errors (e.g. naming the picture of 'table' as 'desk').

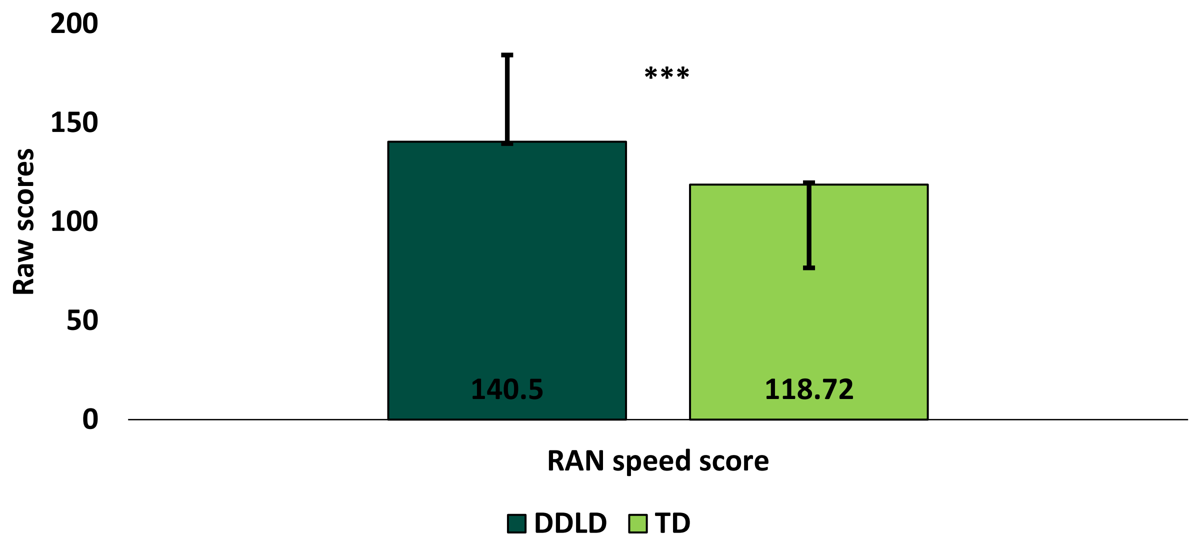


Figure 4.8.5. Raw scores of speed in sec (bars represent *SD*s) for the RAN task in the two groups

Notes: $^{***}p < .001$; statistical significance is based on regression analysis.

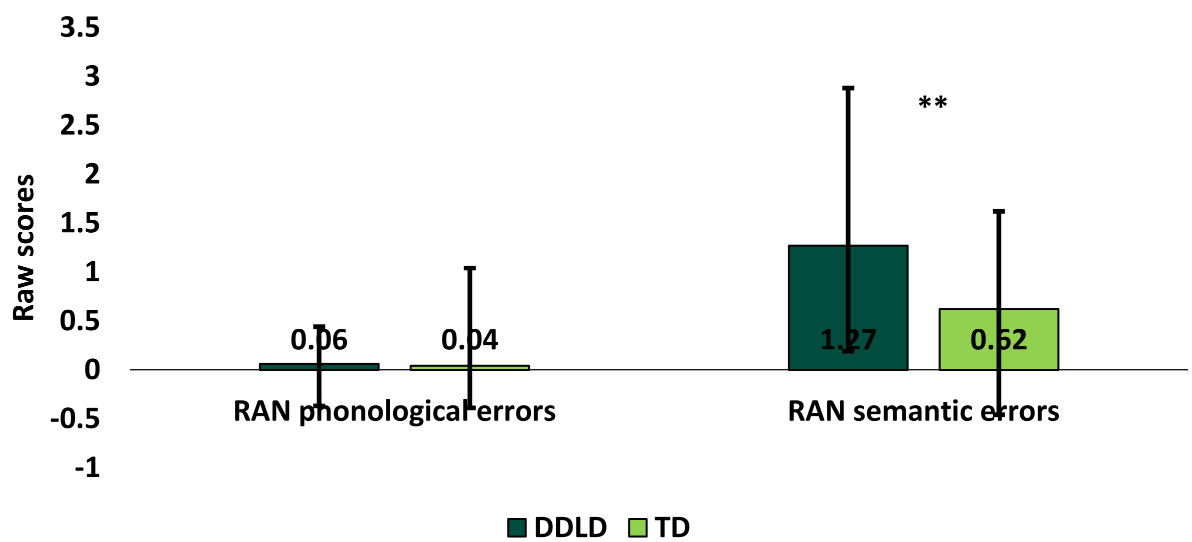


Figure 4.8.6. Raw scores (bars represent *SD*s) of the number of phonological and semantic errors in the RAN task in the two groups

Notes: $^{**}p < .01$; statistical significance is based on regression analyses.

With respect to the spelling task, the TD group produced significantly smaller proportional numbers of phonological, grammatical and orthographic errors than the DDLD group (Figure 4.8.7.). Importantly, however, in the Greek orthography, phonologically correct spelling was not a challenging task for children with DDLD since the majority of spelling errors (96%) were nonphonological in nature. The low proportional number of phonological spelling errors found in the DDLD group is consistent with the Deficient Phonological Access Hypothesis in that DDLD children's phonological representations appear to be as robust as those of TD children.

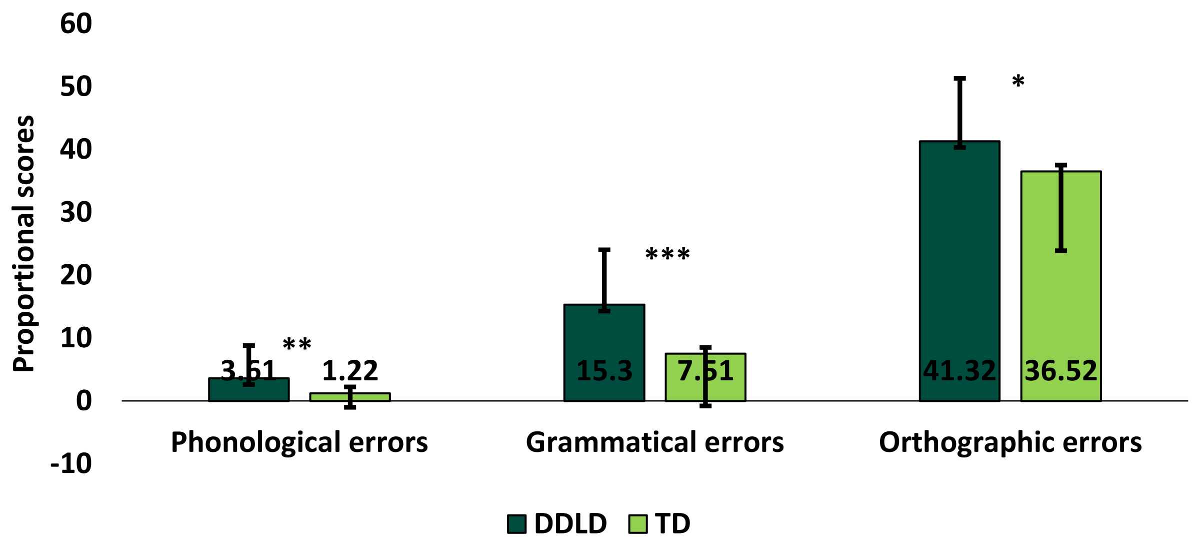


Figure 4.8.7. Proportional scores (bars represent *SD*s) of the number of phonological, grammatical and orthographic errors in the two groups

Notes: $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$; statistical significance is based on Wilcoxon Mann-Whitney for the number of phonological and grammatical errors, and independent-samples *t*-test for the number of orthographic errors.

Research Question 3. Does semantic and phonological fluency performance relate to children's language and literacy skills?

The LangLit variable is the variable generated by the PCA based on *z*-scores of all the language and literacy tasks that correlated significantly to semantic and phonological fluency tasks. *Z*-scores were computed relative to the control group's mean and standard deviation for each task. The PCA demonstrated that all five language measures used in the study and two literacy measures, namely, l'aloutte and spelling, loaded on a single factor. In the overall sample, variations in verbal fluency performance were related to the LangLit variable, with the LangLit variable accounting for 6% of the variance in semantic fluency and 8.8% of the variance in phonological fluency after controlling for age in months and NVIQ. Analyses by subgroup showed that the LangLit variable in the DDLD group accounting for 7.9% and 15.3% of the variance in semantic and phonological fluency respectively after controlling for age in months and NVIQ. In the TD group, after controlling for age in months and NVIQ, the LangLit variable was a nonsignificant predictor of semantic and phonological fluency. The findings suggest that poorer semantic and phonological fluency performance is partly attributed to DDLD children's poor language and literacy skills.

Further, regression analyses showed that in the overall sample, the LangLit variable accounted for word productivity throughout the one-minute test period in both verbal fluency categories, after controlling for age in months and NVIQ in Step 1. The contribution of the LangLit variable was highly significant in the first 15 s of the semantic condition, accounting for 8.2% of the variance in semantic fluency performance in the first 15 s and 3.3% of the variance in semantic fluency performance in the subsequent 45 s. The contribution of the LangLit variable was also significant in the first 15 s and the subsequent 45 s of the phonological condition, accounting for 6% of the variance in phonological fluency performance both in the first 15 s and the subsequent 45 s. The findings imply that children's language and literacy skills are relevant throughout the one-minute test period in both verbal fluency conditions.

Research Question 4. Do cluster number and/or cluster size drive productivity in semantic and phonological fluency tasks in TD children and children with DDLD?

Both in the TD and DDLD groups, variations in both verbal fluency conditions were related to the number of clusters and the number of switches, but not to cluster size, as Tables 4.8.1. and 4.8.2. show, revealing that the semantic lexicon is organized in a similar way in both groups.

Tables 4.8.1. and 4.8.2. Partial correlations (controlling for age) between word productivity with the number of switches, the number of clusters and average cluster size in both verbal fluency categories in the TD and DDLD groups

	TD group		
	No. of Switches	No. of Clusters	Cluster size
Semantic fluency	$r_{(80)}$=.72, p<.001	$r_{(80)}$=.64, p<.001	$r_{(80)}$=.10, p=.336
	DDLD group		
	No. of Switches	No. of Clusters	Cluster size
Semantic fluency	$r_{(63)}$=.76, p<.001	$r_{(63)}$=.77, p<.001	$r_{(63)}$=.00, p=.964

	TD group		
	No. of Switches	No. of Clusters	Cluster size
Phonological fluency	$r_{(80)}$=.83, p<.001	$r_{(80)}$=.74, p<.001	$r_{(80)}$=.10, p=.340
	DDLD group		
	No. of Switches	No. of Clusters	Cluster size
Phonological fluency	$r_{(63)}$=.78, p<.001	$r_{(63)}$=.73, p<.001	$r_{(63)}$=.21, p=.083

Research Question 5. How specific is the verbal fluency deficit in children with DDLD: Does it extend to a nonverbal task (design fluency)?

Children with DDLD did not differ from TD children on the number of correct designs generated in the design fluency task (Figure 4.8.8.). The results demonstrate that children with DDLD do not have difficulties with design fluency, performing age-appropriately. The specificity of the verbal fluency deficit in children with DDLD is supported by evidence showing that there was still a group difference for the mean number of correct responses produced in semantic and phonological fluency tasks. The result demonstrates that after the effects of age and design fluency performance were controlled, children with DDLD still show lexical retrieval difficulties in semantic and phonological fluency tasks, arguing for the specificity of the verbal fluency deficit in children with DDLD (and not for general speed processing difficulties which might have resulted in lower semantic and phonological fluency performance).

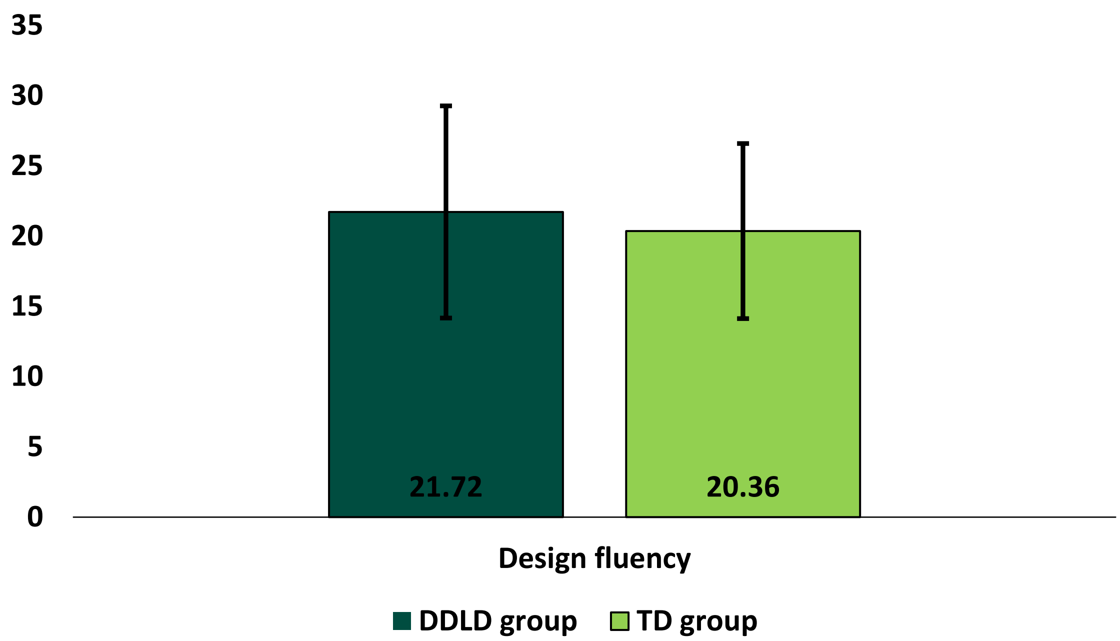


Figure 4.8.8. Raw scores (bars represent *SD*s) of the number of correct designs (design fluency) in the two groups

Overall, the similarities found between the two groups in relation to fluency tasks are the following:

- Word productivity was driven by the number of switches and cluster number but not by cluster size in both verbal fluency categories (Tables 4.8.1. and 4.8.2.)
- Average semantic and phonological cluster size did not differ (Figures 4.8.1. and 4.8.2.)
- Neither group produced many incorrect responses in verbal fluency categories (Tables 4.1.2. and 4.2.2. in the Results chapter)
- Design fluency performance was similar in the two groups (Figure 4.8.8.).

Chapter 5. Discussion

5.1. Objectives of the study

This thesis has set out to answer two questions about the language of Greek children with dyslexia and DLD. Firstly, what is the structure of their lexicon? Lexical organization has been less well-studied in children with dyslexia and DLD than other components of language, such as phonology, morphology and syntax. This investigation sheds light on the underlying cause of lexical difficulties in children with dyslexia and DLD. Secondly, what is the locus of their phonological deficit? Specifically, the second objective of the study was to explain where the phonological deficit in children with DDLD lies, namely, in children's phonological representations or in children's ability to access those representations. Given the comorbidity between dyslexia and DLD, and the interaction between language and literacy during the course of development (Hulme & Snowling, 2014; Marshall & Messaoud-Galusi, 2010), such an empirical investigation is a valuable addition to the literature of developmental language disorders, having theoretical implications with respect to the main causal theories accounting for the contested locus of the phonological deficit in dyslexia and DLD.

The current study had the following objectives: (i) to test the Poor Lexical-Semantic Structure Model against the Slow-Retrieval Model using semantic fluency tasks in order to investigate which of the two models better characterises lexical difficulties in children with DDLD; (ii) to test the two prominent phonological hypotheses of dyslexia and DLD, namely, the Degraded Phonological Representations Hypothesis against the Deficient Phonological Access Hypothesis using phonological fluency tasks, phonological tasks and a spelling-to-dictation task; (iii) to investigate whether semantic and phonological fluency performance can be accounted for by children's language and literacy skills; and (iv) to test the specificity of verbal fluency deficits using a design fluency task not drawing upon verbal processing skills.

Overall, the current study set out to answer the following research questions about semantic and phonological fluency in Greek-speaking children with DDLD:

- What is the structure of the lexicon in children with DDLD compared to TD children? Is poorer semantic fluency performance in children with DDLD better explained by impoverished semantic structure or slower retrieval processes?
- Where does the phonological deficit in children with DDLD lie? Is poorer phonological fluency performance in children with DDLD better explained by degraded phonological representations or by deficient explicit access to (intact) phonological representations?
- Do cluster number and/or cluster size drive productivity in semantic and phonological fluency tasks in TD children and children with DDLD?
- Does semantic and phonological fluency performance relate to children's language and literacy skills?
- How specific is the verbal fluency deficit in children with DDLD: Does it extend to a nonverbal task (design fluency)?

In this study, it was decided to combine the dyslexia group and the DLD group. Firstly, even though dyslexia and DLD are not the same disorders, dyslexia co-occurs with DLD with an overlap of approximately 50%, and accordingly, the probability of showing dyslexia is much higher in children diagnosed with DLD than in those without DLD. Indeed, research has revealed that there is a strong positive correlation between phonological processing skills and nonphonological language skills in TD children, dyslexia, DLD, and in children with dyslexia *plus* DLD (e.g. Ramus et al., 2013). Secondly, there is a lack of Greek standardised tasks and there was therefore an issue how well children with dyslexia could be differentiated from children with dyslexia and/or DLD using currently-available measures. Thirdly, strong evidence originating from a Principal Component Analysis carried out on language and literacy scores within just the children with dyslexia and/or DLD in order to determine whether there were separate loadings onto different components that might justify keeping these children separate. This analysis, presented in section 3.2.1.a., revealed that language and literacy variables did not load on different components in these children, which suggests that it is appropriate to combine them children with dyslexia and/or DLD into a single DDLD group.

5.2. Research question 1. What is the structure of the lexicon in children with DDLD compared to TD children? Is poorer semantic fluency performance in children with DDLD better explained by impoverished semantic structure or slower retrieval processes?

This research question is answered by drawing upon data from analysis of clusters of three semantic categories and computational analysis using the semantic category of 'animals'.

Children with DDLD retrieved significantly fewer correct words in semantic categories than TD children. This finding is not in accordance with some previous studies assessing children with dyslexia which found nonsignificant group differences (e.g. Brosnan et al., 2002; Frith et al., 1995; Griffiths, 1991; Landerl et al., 2009; Marzocchi et al., 2008; Mielnik et al., 2015; Plaza & Guitton, 1997) and adults with dyslexia (Frith et al., 1995; Smith-Spark et al., 2017). Lower productivity in semantic categories is consistent, however, with some other studies of children with dyslexia (Cohen et al., 1999; Korhonen, 1995; Levin, 1990; Menghini et al., 2010; Moura et al., 2015; Plaza et al., 2002; Reiter et al., 2005; Varvara et al., 2014), children with DLD (Henry et al., 2012, 2015; Weckerly et al., 2001), children with WFDs (Messer & Dockrell, 2013), and adults with dyslexia and/or DLD (Hall et al., 2017).

This significant result was accompanied by an extensive overlap in the number of correct words produced between the two groups, in addition to a small amount of variance (3.4%) in semantic fluency accounted for by group after controlling for age and NVIQ. This implies that semantic fluency performance cannot well differentiate children with DDLD from children without DDLD, and it is consistent with the conclusion in the Hall et al.'s (2017) study of adults with dyslexia and/or

DLD, namely, that semantic fluency is a task that could not well differentiate adults with and without dyslexia and/or DLD.

Clustering is related to the integrity of the lexico-semantic network, and switching is related to strategic search and retrieval (Troyer, Moscovitch, Winocur, Alexander, & Stuss, 1998; Troyer et al., 1997; Troyer, Moscovitch, Winocur, Leach, & Freedman, 1998). The number of switches in the two groups was similar, and given that switching is thought to reflect EFs, it can be argued that poorer semantic fluency performance was not limited by EFs involved in the semantic fluency task. The finding of a similar number of switches in the two groups is not consistent with Mielnik et al.'s (2015) study with children with dyslexia, or with Henry et al.'s (2015) and Weckerly et al.'s (2001) study with children with DLD. The finding is also not in accordance with the finding of Smith-Spark et al.'s study (2017) of adults with dyslexia, and Hall et al.'s (2017) study of adults with dyslexia and/or DLD in that in these two studies the disordered groups produced significantly fewer switches than control groups. The number of clusters differed significantly between the two groups, with the TD group producing more clusters than the DDLD group. This finding is in accordance with the studies of Mielnik et al. (2015), Weckerly et al. (2001) and Hall et al. (2017).

Children with DDLD did not differ significantly from TD children on the average size of clusters, suggesting that the organization of the semantic network is similar in the two groups. The results are similar to those from previous studies. The size of semantic clusters did not differ in previous studies of children with DLD (Henry et al., 2015; Weckerly et al., 2001), adults with dyslexia (Smith-Spark et al., 2017) and deaf children with DLD (Marshall et al., 2013). A significantly smaller semantic cluster size was, however, found in Hall et al.'s (2017) study comparing adults with dyslexia and/or DLD with controls.

Overall, the findings imply that children with DDLD have difficulty accessing lexical items and that they do not make it as far through the semantic network in their search for lexical items as TD children in this time-constrained task, but that their semantic representations are typical given that they cluster their responses around subcategories of a similar size to that of TD children. The number of switches was also similar in the two groups. Given that switching is thought to reflect EFs (Troyer, 2000), it can be argued that a significantly lower semantic fluency performance found in the DDLD group cannot be accounted for by poorer EFs involved in the semantic fluency task.

Computational modelling was also used to investigate the semantic network of children with DDLD. This investigation was based on the development of a new metric to capture the likelihood of a speaker producing a particular sequence of lexical items in the fluency task (Davelaar et al., 2015), using the semantic category of animals (Davelaar et al., in preparation). As a reminder to the reader, the metric was computed as follows. Produced sequences were combined into a single network by overlaying the individual paths. The resulting network was directed and weighted. Each individual's sequence had a likelihood of being produced by that reference network. The likelihood of an individual's sequence was the product of the frequency of each word and the conditional probability of all the steps in the sequence. A maximum likely sequence was constructed as a reference path to control for the influence of sequence length. A ratio score was then computed for each individual and was the metric of interest in the study. The closer this number is to 1, the more typical the individual's sequence is. It was predicted that if the Poor Lexical-Semantic Structure Model holds true, children with DDLD would likely have an atypical sequence, and therefore produce a significantly larger ratio score compared to TD children. In contrast, if the Slow-Retrieval Model holds true, the children with DDLD would not differ in their sequence, and the two groups would show a similar ratio score. It was found that the two groups did not differ on ratio score. Just like the findings from analysis of clusters, the finding originating from computational modelling is consistent with the Slow-Retrieval Model, suggesting a similarly well organised semantic network in the two groups. This method is sensitive enough to detect differences in semantic networks and has indeed detected them in other populations: Kenett et al. (2013) found a less developed semantic network structure in children with cochlear implants compared to normal hearing children, and Davelaar et al. (2015) found that adults with MCI differed significantly from healthy adults in the metric used in the current study. The null result in the current study is not due to an inadequate sample size because participant numbers are comparable to the sample size of previous studies which used computational modelling techniques. Kenett et al. (2013) included 54 children in their study (27 with CIs and 27 normal hearing children), and Davelaar et al. (2015), analysing the data originating from Lerner, Ogrocki, & Thomas (2009), included 102 adults (38 controls, 26 with MCI and 38 with AD).

Turning now to the predictions, with respect to semantic fluency, the objective of the study was to investigate whether poorer semantic fluency performance could be attributed to children's poor semantic structure, as proposed by the Poor Lexical-Semantic Structure Model, or to slow retrieval processes while lexical-semantic representations are intact, as proposed by the Slow-Retrieval Model (Lenio et al., 2016). Although both the Poor Lexical-Semantic Structure Model and the Slow-Retrieval Model predict poorer productivity on semantic fluency tasks, the pattern of retrieval, as revealed by an analysis of clusters, is predicted to be different. The former model attributes poorer semantic fluency performance to children's poor semantic structure, and the latter model attributes poorer semantic fluency performance to slow retrieval processes while lexical-semantic representations are intact. It was predicted that if the Poor Lexical-Semantic Structure Model holds true, children with DDLD would show a significantly smaller average cluster size compared to TD children. In contrast, if the Slow-Retrieval Model holds true, children with DDLD would show a similar average cluster size compared to TD children. The findings originating from an analysis of clustering behaviour and from a computational analysis of the semantic category of animals are consistent with the Slow-Retrieval Model in that the two groups did not differ significantly on the average size of semantic clusters produced.

Based on previous research findings, it was predicted that semantic fluency would be driven by the production of more switches and more clusters rather than by the production of bigger clusters.

This was indeed found to be the case. Both in the TD and DDLD groups, the number of correct responses correlated strongly with the number of clusters and the number of switches, but not with average cluster size. The finding that semantic fluency performance was strongly related to switching and clustering behaviour is consistent with previous studies (e.g. Arán-Filippetti & Allegri, 2011; Kosmidis et al., 2004). Automatic processing is linked to the first 15 s, and controlled processing is linked to the subsequent 45 s of the test period in verbal fluency tasks. The rate of production declined over the 60 s of the test period in both verbal fluency categories, as has been reported in previous studies (e.g. Arán-Filippetti & Allegri, 2011; Henry et al., 2015; Marshall et al., 2013). Automatic processing, as indexed by the number of total responses in the first 15 s of the test period, was impaired in children with DDLD, with group accounting for 7.4% of the variance in the number of total responses in the first 15 s. This suggests that children with DDLD have slower access to the lexicon for items considered to be readily accessible. This finding is in accordance with Marshall et al.'s (2013) study of deaf children with DLD who had disproportionate sign-finding difficulties early in the test compared to deaf children without DLD. Marshall et al. interpreted this finding as evidence for access deficits in deaf signers with DLD.

In contrast, controlled processing, as indexed by the number of total responses in the subsequent 45 s of the test period, was similar between children with DDLD and TD children. Henry et al. (2015) found that children with DLD showed weaker performance throughout the one-minute test period compared to controls. This finding was not replicated in the current study in that children with DDLD differed in automatic but not in controlled processing compared to TD children. The findings are also not consistent with those of Smith-Spark et al. (2017) who showed nonsignificant group differences between the group of adults with dyslexia and controls for any quartile, and with those of Hall et al. (2017) in adults with dyslexia and/or DLD who showed a similar change over the course of the test period in that the disordered adults were not less fluent than controls after the initial 15 s of the test period.

Verbal fluency tasks are governed by certain rules and require children to inhibit certain responses. Children should inhibit inappropriate responses that come readily to mind, for example, responses that have already been produced in the sequence (in order to avoid repeated responses), and out-of-category responses (in order to abide by the rules of the task). Any impairments in EFs might result in an increased number of incorrect responses, with a high proportion of errors suggesting difficulties in word search and retrieval processes, impaired executive control over semantic search and retrieval strategies. Group accounted for 4.3% of the variance in the number of incorrect responses. The finding of more incorrect responses in the DDLD group compared to the TD group is consistent with Henry et al.'s (2015) finding for proportionally more incorrect responses in children with DLD than in TD children. Semantic fluency categories seemed to be prone mainly to out-of-category and repeated responses. The error ratio in semantic categories was very low, however, in the DDLD group, and its distribution was strongly skewed because many children produced no errors at all. As such, a low proportion of errors in children's responses might be associated with a normal function of frontal brain systems which resulted in normal monitoring of item search and retrieval processes.

5.3.a. Research question 2a. Where does the phonological deficit in children with DDLD lie? Is poorer phonological fluency performance in children with DDLD better explained by degraded phonological representations or by deficient explicit access to (intact) phonological representations?

This research question is answered by drawing upon data from analysis of clusters of three phonological categories.

Children with DDLD retrieved significantly fewer correct words in phonological categories than TD children. This finding is in accordance with previous studies assessing children with dyslexia (e.g.

Brosman et al., 2002; Frith et al., 1995; Griffiths, 1991; Levin, 1990; Marzocchi et al., 2008; Menghini et al., 2010; Moura et al., 2015; Plaza & Guitton, 1997; Plaza et al., 2002; Reiter et al., 2005; Varvara et al., 2014), children with DLD (e.g. Henry et al., 2012, 2015; Weckerly et al., 2001), and adults with dyslexia (Frith et al., 1995; Smith-Spark et al., 2017). In contrast, some previous studies have reported nonsignificant group differences in phonological fluency performance between children with dyslexia and controls (Cohen et al., 1999; Korhonen, 2005; Landerl et al., 2009). These differences might be accounted for by the selection criteria of each study. The issue of inconsistent findings across studies is discussed in a later section of the discussion.

Previous studies have reported that phonological fluency rather than semantic fluency places higher demands on EFs because lexical items are not arranged alphabetically in semantic memory, so retrieving lexical items beginning with a particular letter is more effortful. The number of switches was significantly different in the two groups, with the TD group outperforming the DDLD group. Given therefore that switching draws upon EFs, broader difficulties with EFs cannot be excluded as a factor accounting for poorer phonological fluency performance in the DDLD group. This finding is not consistent with Mielnik et al.'s (2015) study who found that children with dyslexia did not differ from controls on the number of switches. This finding is consistent, however, with Smith-Spark et al. (2017) who tested English adults with dyslexia, and with Henry et al. (2015) who studied English children with DLD.

The number of phonological clusters was similar in the two groups. As presented in the literature review, phonological clustering provides a more implicit measure of the quality of children's phonological representations on the basis that it is the phonological similarity (or overlap) in successively produced responses that might aid word retrieval (e.g. in the phonological cluster '<u>fl</u>ag-<u>fl</u>ower', the retrieval of 'flag' might have facilitated the retrieval of 'flower' because their phonological representations partly overlap since the two words share the first two phonemes). The finding that the two groups did not differ on the number of phonological clusters is not in accordance with the study of Weckerly et al. (2001) with children with DLD who found that children with DLD produced a significantly smaller number of phonological clusters than TD children.

Moreover, the average size of phonological clusters did not differ between children with DDLD and TD children suggesting that implicit access to phonological representations is intact in children with DDLD compared to TD children. This finding is consistent with Smith-Spark et al.'s (2017) study in adults with dyslexia who found a nonsignificant difference between adults with dyslexia and controls with respect to cluster size, and with Weckerly et al. (2001). The lack of group difference for phonological cluster size is also consistent with the argument of Frith et al. (1995) that dyslexia-related verbal fluency problems are due to difficulties with accessing words based on their phonological characteristics rather than due to differences in vocabulary size. This finding is not in accordance, however, with the study by Henry et al. (2015), who assessed children with DLD, and showed that the number of words they produced per cluster was smaller compared to TD children.

Based on previous research findings, it was predicted that semantic fluency would be driven by the production of more switches and more clusters rather than by the production of bigger clusters. This was indeed found to be the case. Both in the TD and DDLD groups, the number of correct responses correlated strongly with the number of clusters and the number of switches, but not with average cluster size. The finding that phonological fluency performance was strongly related to switching and clustering behaviour is consistent with previous studies (e.g. Arán-Filippetti & Allegri, 2011; Kosmidis et al., 2004). These findings suggest that lexical organization is similar in children with DDLD and TD children, and that children move from one subcategory to another, that is, they switch among subcategories, in order to retrieve as many lexical items as possible. Given that switching is seen as a measure of EFs (Kavé et al., 2008; Troyer, 2000), it can be argued that phonological fluency tasks involve EFs to a certain extent.

Troyer (2000) argued that switching draws upon EFs, and Luo et al. (2010) reported that EFs are more important in the later than the earlier parts of the task. Given that EFs are more important in later (controlled processing) than in early (automatic processing) stages of the phonological fluency task, automatic versus controlled processing was also investigated. Automatic processing, as indexed by the number of total responses in the first 15 s of the test period, was similar between children with DDLD and TD children. However, controlled processing, as indexed by the number of total responses in the subsequent 45 s of the test period, was marginally impaired in children with DDLD in that group accounted for 2.3% of the variance in the number of total responses in the subsequent 45 s. The findings suggest that given that it is thought that phonological fluency tasks are more strategic than semantic fluency tasks and therefore place higher demands upon EFs, impaired controlled processing indicates difficulties with EFs in children with DDLD. The findings are not consistent with Henry et al.'s (2015) study in which children with DLD showed weaker performance in all four quartiles of the phonological fluency tasks, nor with Smith-Spark et al.'s (2017) study in which adults with dyslexia produced fewer total responses in the first, second and fourth quartile in the phonological condition than controls.

Group accounted for 3.4% of the variance in the number of incorrect responses, suggesting that children with DDLD were less able to monitor their responses for accuracy and rule violations, and suppress incorrect responses. The finding of more incorrect responses in the DDLD group compared to the TD group is consistent with Levin's (1990) finding for children with dyslexia, and with Henry et al.'s (2015) finding for children with DLD. Phonological fluency categories seemed to be prone mainly to made-up, repeated and out-of-category responses. The mean error ratio in phonological categories was very low, however, in the DDLD group, and its distribution was strongly skewed to the right. This finding indicates that a low proportion of errors in children's responses might be associated with a normal function of frontal brain systems which resulted in normal monitoring of item search and retrieval processes (Marslen-Wilson, 1987).

Turning now to the predictions, with respect to phonological fluency, it was investigated whether poorer phonological fluency performance could be attributed to children's impaired phonological representations as proposed by the Degraded Phonological Representations Hypothesis (Goswami, 2000), or alternatively, to impaired explicit access but intact implicit access to phonological representations as proposed by the Deficient Phonological Access Hypothesis (Ramus & Szenkovits, 2008). Although both the Degraded Phonological Representations Hypothesis and the Deficient Phonological Access Hypothesis predict poorer productivity on phonological fluency tasks, the pattern of retrieval as revealed by an analysis of clusters is predicted to be different. The former hypothesis attributes poorer performance to children's degraded phonological representations, and the latter hypothesis attributes poorer performance to impaired explicit access but intact implicit access to phonological representations. If the Degraded Phonological Representations Hypothesis holds true, children with DDLD would show a significantly smaller average cluster size compared to TD children. In contrast, if the Deficient Phonological Access Hypothesis holds true, children with DDLD would show a similar average cluster size compared to TD children. The findings are consistent with the Deficient Phonological Access Hypothesis in that the two groups did not differ significantly on the average size of phonological clusters produced.

Overall, the number of phonological clusters and average phonological cluster size did not differ significantly in the two groups. These findings imply, respectively, equal subordinate categories available in children with DDLD as those in TD children, and that DDLD children's implicit access to phonological representations is intact. The number of switches was significantly different in the two groups. Given that switching is thought to reflect EFs (Troyer, 2000), it can be argued that the DDLD group's significantly lower phonological fluency performance can be partly accounted for by poorer EFs involved in the phonological fluency task.

5.3.b. Research question 2b. Which hypothesis better characterises the locus of the phonological deficit in children with DDLD in phonological tasks—namely, phoneme deletion and RAN tasks, and a spelling-to-dictation task: The Degraded Phonological Representations Hypothesis or the Deficient Phonological Access Hypothesis?

With respect to phoneme deletion tasks, the TD group significantly outperformed the DDLD group in accuracy and speed performance in phoneme deletion of CCV and CVCVCV items, but the two groups did not differ in accuracy and speed performance in phoneme deletion of the simpler CVC items. However, it should be noted that the TD group were at ceiling on the phoneme deletion task of the simpler CVC items, and therefore the below presented interpretation of the data is limited by the task's ceiling effect.

It was predicted that if the Degraded Phonological Representations Hypothesis holds true, the TD group would outperform the DDLD group in accuracy and speed performance in all three phoneme deletion tasks. In contrast, if the Deficient Phonological Access Hypothesis holds true, accuracy and speed performance in the phoneme deletion task of monosyllable items with simple CVC syllable structure would be equivalent for the two groups; however, the two groups would differ on accuracy and speed performance in the phoneme deletion task of trisyllable items with simple CVCVCV syllable structure and in the phoneme deletion task of monosyllable items with complex CCV syllable structure. For the former hypothesis, the phonological deficit in dyslexia and DLD makes itself manifest in significantly poorer performance in every phoneme deletion task due to impaired phonological representations, irrespective of the load of additional cognitive processes involved in the tasks. For the latter hypothesis, the phonological deficit in dyslexia and DLD makes itself manifest in significantly poorer performance in phoneme deletion tasks loading on additional cognitive processes but in equivalent performance in phoneme deletion tasks not loading on additional cognitive processes. The findings are consistent with the Deficient Phonological Access Hypothesis since group did not account for any variance in accuracy or speed performance in the phoneme deletion task of monosyllabic items with simple CVC syllable structure but did account for significant amounts of variance in accuracy and speed when items had trisyllabic CVCVCV or monosyllabic CCV structure.

With respect to the RAN task, analyses showed that the DDLD group named the pictures in the two cards significantly slower than the TD group, and produced significantly more semantic errors, but that the two groups did not differ in their number of phonological errors and omissions. It was predicted that if the Degraded Phonological Representations Hypothesis holds true, the DDLD group would make phonological errors when naming pictures. If the Deficient Phonological Access Hypothesis holds true, however, the DDLD group would name pictures significantly slower than the TD group but the two groups would show no difference on phonological accuracy. The former hypothesis attributes phonologically inaccurate performance to inaccurate phonological representations. The latter hypothesis attributes slower naming performance to a phonological access deficit in the DDLD group which renders performance on tasks requiring speeded access to phonological representations particularly slow. Phonologically accurate performance is explained by intact access to phonological representations in a task not requiring metalinguistic manipulation. The findings are consistent with the Deficient Phonological Access Hypothesis since the DDLD group showed a significantly slower naming performance than the TD group but the two groups did not differ on the number of phonological errors produced.

The types of spelling errors in Greek children with dyslexia and/or DLD in Grades 2-6, and also in their TD peers in Grades 1-6 in order to investigate firstly whether phonologically correct spelling in dictation is a challenging task and secondly to specify what underlies poor spelling accuracy performance in Greek children with DDLD. Children with DDLD spelled correctly fewer words than TD children, with a large effect size being found. This finding was expected (e.g. Joye et al., 2019; Protopapas et al., 2013). In addition, spelling errors were classified into each of the following three categories of errors, namely, phonological, grammatical and orthographic. Classification of errors into categories allowed the experimenter to identify DDLD but also TD children's impaired processes which reflect on the specific error types. Group comparisons revealed that children with DDLD produced significantly more phonological, grammatical and orthographic errors than TD children, proportionally to the number of total words spelled. However, phonological errors were negligible both in TD children and children with DDLD (respectively, 1.22 versus 3.61%). The findings suggest that difficulties with inflectional morphology and word families account for poor spelling accuracy in Greek children with DDLD.

For the phonological error category, the experimenter analysed DDLD and TD children's substitution, insertion, omission, and inversion errors. It was predicted that if the Degraded Phonological Representations Hypothesis holds true, qualitative analysis of spelling errors would reveal that the DDLD group would display a significantly higher proportion of phonological spelling errors than the TD group. In contrast, if the Deficient Phonological Access Hypothesis holds true, qualitative analysis of spelling errors would reveal that in the DDLD group the majority of spelling errors would be phonologically correct. The former hypothesis attributes phonological spelling errors to children's inaccurate phonological representations, and the latter hypothesis attributes phonological spelling errors to accurate phonological representations but inappropriate orthographic encoding of words using phoneme-to-grapheme mappings that are inappropriate for a particular context in the DDLD group. The findings are consistent with the Deficient Phonological Access Hypothesis in that the majority of spelling errors found were nonphonological errors.

In both groups, most of spelling errors (96%) were phonologically correct. Phonologically correct spelling implies that poor phonological processes cannot be considered as an explanation for poor spelling accuracy in Greek children with DDLD. This is despite the DDLD group having poorer phonological processing skills as measured by tasks of phoneme deletion, NWR and RAN. The finding that phonologically correct spelling is not challenging for Greek children with DDLD is consistent with other studies in the Greek orthography which found that the majority of spelling errors were phonologically correct in children with dyslexia (Niolaki & Masterson, 2013; Porpodas, 1999; Protopapas et al., 2013). This finding is not consistent with other studies in the Greek orthography, however, which revealed that children with dyslexia produced a substantial amount of phonological errors (Diamanti, 2006; Niolaki, Masterson, & Terzopoulos, 2014). Niolaki et al. (2014) in their case study, also reported a rate around 88% of phonologically appropriate errors in Greek children with very poor nonword reading and spelling. This finding is also not consistent with the study of Caravolas and Volín (2001) who assessed Czech children with dyslexia, and reported that children with dyslexia produced a high percentage of phonological errors compared to age-matched controls (respectively, 18.81 versus 3.98). Daigle, Costerg, Plisson, Ruberto, and Varin (2016) who assessed French children with dyslexia, reported that phonological errors were as numerous as visual-orthographic or lexical errors. Likewise, Critten et al. (2014) found that English children with DLD aged 9-10 years, who write in the deep English orthography, produced a greater proportion of phonologically implausible errors compared to both age-matched and spelling-matched controls.

The heterogeneity of dyslexia and DLD (see for dyslexia: Castles & Friedmann, 2014; see for DLD: Bishop, 2006) means that different samples might show different profiles. In addition, the use of a spelling task with a discontinue rule means the words attempted by the DLDD group were the easier ones since they were at the start of the graded test, and phonological errors may be more apparent when the children are spelling less familiar and more complex words. Importantly, the finding that phonologically correct spelling is not challenging for Greek children with DDLD does not imply that phonological processes are not crucial for spelling ability, and that interventions focused on improving children's phonological processing skills do not have a positive effect on their spelling ability. In fact, it has been reported that children rely heavily on phonological processing skills to spell words in dictation, independently of age and orthographic consistency (e.g. Joye et al., 2019). Having said that, however, as proposed by Daigle et al. (2016), emphasis should be put on features which are not processed phonologically by children, features which can be learned, stored in memory and retrieved during spelling production, such as multi-graphemic phonemes, silent letters, homophones.

The proportion of the other two major categories of errors, namely, grammatical and orthographic errors, in the two groups was also investigated. Children with DDLD produced significantly more grammatical errors proportionally to the number of total words spelled than TD children (respectively, 15.30 versus 7.51%) and more orthographic errors (respectively, 41.32 versus 36.52%). The study proposes the importance of morphological awareness in spelling ability in a consistent orthography. This has been previously reported in Greek (e.g. Desrochers et al., 2017; Diamanti et al., 2017, 2018; Georgiou et al., 2012; Grigorakis & Manolitsis, 2016; Pittas & Nunes, 2014), and in languages using less consistent grapheme-to-phoneme mappings than those of Greek (e.g. Caravolas et al., 2012; Ziegler et al., 2010).

A strength is that a dictated word list was used to shed light on the types of knowledge children with DDLD have difficulty with. In a spelling-to-dictation task, spellers are asked to spell fixed words, not words that they might know better as in narrative productions, implying that a dictation task can reveal better than a narrative production task the strategies used by children to spell words (Daigle et al., 2016). In order to correct for the fact that children stopped after 6 misspelled words in a row, a data analysis method developed by Dr Christian Hennig (October 2017, personal discussion) can be adopted to compare the two groups on the three types of spelling errors considered in error analysis. The major advantage of this data analysis method would be to compute a score for each of the three types of errors, reflecting how a child performed relative to how the other children performed on the same words attempted. Another advantage of this method is that it involves a measurement of the difficulty of a word. It can be argued that this method would better differentiate performance of children who spelled more words and produced fewer errors from those who spelled fewer words but produced more errors, for example. These analyses are beyond the scope of this thesis, however, even though they are considered for future analyses. A methodological limitation of the method suggested to prepare the spelling data is that the score does not take into account the fact that a word which was attempted by fewer children was probably a more difficult word in terms of spelling, and concomitantly, that the children who attempted more difficult words were probably better spellers. However, this is hard to measure given that many children did not try the most difficult words, and considering this, we cannot know how well or badly they had have performed on these words.

5.4. Research question 4. Does semantic and phonological fluency performance relate to children's language and literacy skills?

The Language and Literacy variable is the variable generated by the PCA based on *z* scores of five language and two literacy tasks. *Z* scores were computed relative to the control group's mean and standard deviation for each task. The PCA analysis revealed that all five language measures and the two literacy measures loaded on just one component rather than two separate language and literacy components. This finding supports the decision to combine the dyslexia and DLD groups.

It has been reported that greater productivity in verbal fluency categories is associated with better performance on measures of vocabulary (Ardila et al., 2006; Henry et al., 2015; Marshall et al., 2018). It was investigated whether verbal fluency performance in Greek children with DDLD and TD is predicted by language and literacy skills. It was found that language and literacy skills significantly predict semantic and phonological fluency performance in the overall sample. With respect to language skills, this finding might in turn suggest that a smaller expressive and receptive vocabulary size (as measured with WISC Vocabulary subtask and PPVT-R, respectively), but also poorer performance on measures drawing upon a range of language processing skills, namely, sentence repetition, verbal comprehension, and syntax comprehension, results in poorer semantic fluency performance.

With respect to literacy skills, it has been reported that spelling draws heavily on phonological representations, and the fact that spelling predicted phonological fluency reflects the well-known association between literacy skills and the quality of phonological representations. Interestingly, however, the spelling task predicted semantic fluency which implies an association between literacy skills and the quality of semantic representations. This finding is consistent with Nation's (2017) argument that there is a close relationship between reading and semantics in that semantic representations (in addition to orthographic representations) have an effect on word reading skills as any word has a phonological and an orthographic form, but also a meaning. Nation (2017) proposes that according to the Lexical Legacy Hypothesis, the development of word reading is achieved via the experience of words in diverse and meaningful language environments: it is because reading experience allows a reader to read words in different semantic contexts that leads to a rich and nuanced database about a word and its connections to other words. This hypothesis therefore states that word knowledge is based on lexical co-occurrence in the sense that a word is known as it is related in meaning with other words. Accordingly, it has been reported that aside from children's phonological processing skills, children's knowledge of word meanings is also a predictor of reading skills (Nation & Snowling, 2004). The current study, however, also suggests that children's knowledge of word meanings predicts spelling skills, in addition to reading skills.

Henry et al. (2015) found that better language ability supports performance in earlier parts of the semantic fluency task, when items are more readily available, but seems to cease to be important during the more effortful searching required in later parts of the task, and that better language ability supports performance both in earlier and later parts of the phonological fluency task. Luo et al. (2010), however, revealed that language ability is a more important predictor of semantic than phonological fluency. They argued that this is because more integrated semantic knowledge was needed for the semantic task and because language ability was relevant throughout the semantic fluency task. In the current study, regression analyses revealed that the contribution of the LangLit variable was relevant both in automatic and controlled processing in both verbal fluency categories. After controlling for age and NVIQ, the LangLit variable accounted for 8.2 and 3.3% of the variance in the number of total responses in the first 15 s and the subsequent 45 s in semantic categories, respectively. The LangLit variable accounted for 6% of the variance both in the number of total responses in the first 15 s and the subsequent 45 s in phonological categories. Thus, children's language and literacy skills support performance in earlier and later parts of the test period in semantic and phonological fluency tasks.

5.5. Research question 5. How specific is the verbal fluency deficit in children with DDLD: Does it extend to a nonverbal task (design fluency)?

It was investigated whether poorer semantic and phonological fluency being evident in the DDLD group can be attributed to semantic or phonological difficulties, as proposed by the two phonological hypotheses and the two lexical-semantic models, or to broader difficulties with visuospatial EFs. This was tested by using a design fluency task that measures visuospatial EFs without requiring children's semantic or phonological representations and semantic or phonological processing skills. It was found that group was a nonsignificant predictor of design fluency performance. This suggests that children with DDLD do not have difficulties with design fluency, performing age-appropriately. Furthermore, children with DDLD showed poorer semantic and phonological fluency performance relative to their TD peers even after design fluency performance was controlled, demonstrating the specificity of their verbal fluency deficit. This finding therefore establishes the specificity of the verbal fluency deficit in dyslexia and DLD as the two prominent phonological hypotheses and the two lexical-semantic models claim. Nonsignificant group differences in the design fluency task might suggest that poorer performance on semantic and phonological fluency conditions found in the DDLD group compared to the TD group cannot be accounted for by difficulties with visuospatial EFs in children with DDLD, at least as measured with a design fluency task.

However, given that switching draws upon EFs (Troyer, 2000), and the two groups differed on the number of switches in the phonological condition, with children with DDLD switching significantly less often than their TD peers, broader difficulties with EFs cannot be excluded as a factor accounting for poorer phonological fluency performance in the DDLD group. The finding that children with DDLD were less able to switch between subcategories in phonological categories than TD children but similarly able to TD children to switch between subcategories in semantic categories might indicate that phonological fluency tasks are more strategic than semantic fluency tasks in that they rely more on EFs, and therefore reveal disordered children's weaknesses related to strategic search and retrieval.

Smith-Spark et al. (2017) found a similar pattern of results testing English adults with dyslexia and controls. Their interpretation was that using phonological fluency tasks, the deficits associated with dyslexia are more obvious in the phonological domain. However, the researchers argued that some weaknesses in EFs were also evident and might explain poorer phonological fluency performance in the clinical group of their study. Smith-Spark et al. (2017) also argued that short-form IQ was a predictor of the number of switches in phonological fluency tasks and a stronger predictor of phonological than semantic fluency. According to their interpretation, this might be because phonological fluency tasks are cognitively more complex tasks than semantic fluency tasks. This was the case in the current study too in that a stronger positive correlation was observed between the NVIQ test with phonological than semantic fluency. Smith-Spark et al. (2017) further reported that design fluency showed the stronger correlation than verbal fluency with short-form IQ. This was also the case in the current study in that a stronger positive correlation was observed between the NVIQ test and design fluency than semantic and phonological fluency.

In addition, the fact that the two groups showed a similar design fluency performance does not assume that the DDLD group of children or the population of children with DDLD do not suffer from poor executive skills in general. This testament requires the administration of a wide range of tasks assessing EFs and not a single task. Henry et al.'s (2012) study and Botting et al.'s (2017) study have investigated EFs using a sizeable battery of different tasks. Henry et al. (2012) used ten measures of EFs in children with DLD, in children with low language/cognitive functioning and in TD children, and found evidence for difficulties with EFs in language disordered children in most of the tasks used. Lower performance was evident in tasks measuring verbal and nonverbal executive-loaded working memory, verbal and nonverbal fluency, inhibition and planning. Recently, Botting et al. (2017) assessed a large sample of deaf and hearing children on a comprehensive battery of tasks measuring nonverbal EFs, and they also found evidence for a significantly lower performance on EF tasks in deaf children who were signers and in those who used an oral language. The researchers concluded that EFs build upon language skills, and not vice-versa. The two studies therefore suggest that language disordered children show lower performance on EFs and that poor language skills result in poor EFs than vice versa (see also Jones et al., 2019, for longitudinal data showing the same picture longitudinally as Botting et al. (2017) found concurrently).

5.6. Theoretical implications

To the best of the author's knowledge, this is the first study in Greek children with dyslexia, DLD and TD which investigated lexical organization and lexical retrieval skills using semantic and phonological fluency tasks. The study attempted to tease apart in a child sample two proposed models accounting for lexical difficulties in semantic fluency tasks, namely, the Poor Lexical-Semantic Structure Model versus the Slow-Retrieval Model, and also two prominent phonological hypotheses of dyslexia and DLD using phonological fluency tasks, namely, the Degraded Phonological Representations Hypothesis versus the Deficient Phonological Access Hypothesis. Results from this study revealed that lexical difficulties using semantic and phonological categories were evident in Greek children with dyslexia and DLD.

With respect to semantic categories, results confirmed the hypothesis that lexical difficulties can be attributed to slow retrieval processes while lexical-semantic representations are intact as proposed by the Slow-Retrieval Model (Lenio et al., 2016). This is supported by the experimenter's analysis of semantic categories and by an independent analysis based on the development of a new metric (Davelaar et al., 2015) offering the opportunity to compare semantic networks across different groups in an alternative way.

With respect to phonological categories, results confirmed the hypothesis that the phonological deficit in dyslexia and DLD is attributed to children's impaired explicit access to phonological representations while implicit access to phonological representations is intact consistently with the Deficient Phonological Access Hypothesis (Ramus & Szenkovits, 2008). This finding is interpreted on the basis that children's phonological representations are accurate, however, difficulties in phonological representations become evident whenever high demands are placed upon them. This evidence is also supported by the findings originating from (i) equivalently accurate performance between the two groups in phoneme deletion tasks not loading on phonological STM but significantly poorer performance in the DDLD group compared to the TD group in phoneme deletion tasks loading on phonological STM; (ii) naming lexical items in the RAN task significantly slower for the DDLD group relative to the TD group but equivalently accurate performance in the number of phonological errors produced between the two groups; and (iii) the spelling-to-dictation task in which children with DDLD produced a very low raw number of phonological spelling errors, a low proportion of phonological spelling errors and a majority of spelling errors which was nonphonological.

There were also similarities between the two groups. Neither group produced many errors, average semantic and phonological cluster size did not differ, and word productivity was driven by the number of switches and cluster number, but not by semantic average cluster size. These findings suggest that DDLD children's lexical organization is similar to that of TD children. Findings originating from both verbal fluency categories support the view that retrieving fewer items is related to access to intact representations, be those representations semantic or phonological. Concomitantly, the two hypotheses about impaired representations, be those representations semantic or phonological, are not supported. Of course, only well-designed intervention studies are able to offer firm conclusions about causal links. To this end, Ebbels et al. (2012) revealed a positive effect of a semantic therapy on word-finding difficulties in children with language disorders aged 9-15, as exemplified in a word-finding test. It remains to be investigated though whether a semantic or phonological therapy designed to improve children's lexical retrieval processes has a positive effect on semantic or phonological fluency tasks in particular. Similarly, using either phonological or semantic approaches to intervention for word-finding difficulties, as devised by Best et al. (2017), might help identify where the locus of the word retrieval difficulty is and whether focused intervention improves semantic and phonological fluency performance.

5.7. Strengths, limitations and further directions

An obvious strength of this study is that a large overall sample of children and a large sample of children in each participating subgroup were recruited. West, Vadillo, Shanks, and Hulme (2018) argued that inconsistent results in the literature can be accounted for by the fact that because dyslexia and DLD are two heterogeneous and often co-morbid disorders, language-disordered groups from different studies may not reflect the same behavioural symptoms or underlying cognitive impairments. West et al.'s study provided an excellent critique as to why inconsistent results have been reported in the literature. They argued that inconsistent findings could be accounted for by studies with extreme group designs (a technique consisting of selecting individuals on the basis of extreme scores) with small samples and therefore with low statistical power yielding many false positive results; that is, results showing significant correlations between measures when there are no such correlations. The researchers instead embraced individual variation. They also suggested that a large overall sample of children, but also a large sample of children in each participating subgroup of children, allows one to investigate the associations between measures of interest without over-estimating the size of any association between measures of interest, as an extreme groups design might. Further, Chuard et al. (2019) argued that earlier studies often used univariate tests (for example, *t* tests or Mann-Whitney U tests) that ignored major confounding variables (i.e. implicitly assumed that they did not differ between two groups, for example, between a control group and a clinical group). In contrast, recent studies usually add potential confounding variables as covariates that are 'corrected for' before examining the effect of a variable. The current study is such a study in that confounding variables (i.e. age, NVIQ) were added as covariates that were 'corrected for' before examining the effect of a variable. According to Chuard et al. (2019), this approach renders unlikely the interpretation of the current study to be affected by group differences in confounding variables.

For many children, language disorder occurs as a major difficulty without any obvious explanation, and for these children, the term DLD has been recommended. In the past, there has been a lot of confusion about the use of exclusionary factors which have meant that children are only offered help if they have a very specific problem just with language and nothing else. However, it has become increasingly obvious that many children with language problems do have additional coexisting conditions, such as attentional or motor problems. In line with the CATALISE consortium (Bishop et al., 2017), it seems inappropriate to rule those children out from getting help with their language since research literature gives no support to the idea that we should treat those entirely separate. As such, including in this study children in the DDLD group who had comorbid disorders can be also considered as a strength of the study on the basis that the sample is more representative of the children seen in clinics (Bishop, 2018).

However, some researchers are more ambivalent about the term “DLD”. Recently, Marshall (2019) discussed some of the advantages and disadvantages of using “DLD” versus “SLI” as a label but also a diagnostic category in child language research. One of the advantages discussed was that by embracing the term “DLD” it is easier to generalize the findings to many children seen in clinical settings who have co-occurring difficulties. However, is this really the case if there is greater heterogeneity in the sample? One of the disadvantages discussed was that by embracing the DLD term results in bigger samples but also in greater heterogeneity of the sample. This in turn implies that (if the sample of a study is so heterogeneous) there is an issue of how the results are generalised. Marshall (2019) concluded by arguing that despite the proposed change on terminology and conceptualization of DLD by the CATALISE consortium (Bishop et al., 2017), it is still an empirical question whether including in the same group children with DLD who also have co-occurring conditions (e.g. ADHD) and children with DLD in the absence of co-occurring conditions is the right choice for researchers in the field. Bigger samples, however, as discussed above (West et al., 2018), can also result in greater replicability of the results than smaller samples. With respect to the NVIQ criterion proposed as a cut-off point by the CATALISE consortium (Bishop et al., 2018), it is the case that IQ scores are unstable over time, and children with intellectual disability are still excluded from receiving the diagnosis of DLD.

In this study, even though the main hypotheses are supported by empirical evidence, only well-controlled clinical trials involving random assignment to treatment and control groups can reach conclusions about causal links. Well-designed intervention studies would offer firm conclusions about whether poorer semantic fluency performance in children with DDLD can be attributed to children’s retrieval difficulties while lexical-semantic representations are intact. Likewise, intervention studies would offer firm conclusions about whether poorer phonological fluency performance in children with DDLD can be attributed to children’s impaired explicit access to phonological representations while implicit access to them is intact. This could be achieved by identifying children's difficulties and conducting interventions based on theoretical accounts of lexical difficulties. For example, Best et al. (2017) conducted an intervention study designed to enhance metacognitive awareness and word retrieval skills in children with WFDs aged 6-8 years. They found that after the intervention, children could retrieve more lexical items than controls. The researchers concluded that this intervention study shows an effect of a word-finding therapy in children with language disorders.

Further, in this study, lexical-semantic representations were found to be intact in children with DDLD using semantic fluency tasks. However, Sheng and McGregor (2010) used the association task in children with DLD and TD children in which children were asked to say the first word that comes to mind when hearing a prompt word, such as, the word *birthday*. The idea behind this task was that children will provide words which will be semantically associated to the prompt words. For example, a semantic association to the word *birthday* will be the words *cake*, *candles*, or *years*. The researchers found that there was a group of children with DLD who also had a deficit in expressive vocabulary, and a gap between receptive and expressive vocabulary (referred to as children with WFDs in the literature). This group of children showed deficits in lexical-semantic organization, as indexed by producing fewer semantically-related responses. However, a lot of variability was reported in the DLD group. Sheng and McGregor's study therefore showed that deficits in lexical-semantic organization do not appear in all children with DLD but in a subgroup of children who also show a smaller/delayed vocabulary size. As such, even though in the current study, the two groups did not differ on the organization of the semantic network using semantic fluency tasks, group differences might have been evident if another task assessing semantic organization of the lexicon had been used.

Another limitation is that the concept of access to representations is underspecified in the literature. As Mirman and Britt (2014) argue in the context of semantic access deficits in adults, it is not clear precisely what researchers mean when they refer to 'access', nor what the nature of the ‘access deficit’ is. Further investigation, using different research methods, is needed to shed light on the origin of access deficits in dyslexia and DLD. To this end, Boets et al. (2013) reported that in adults with dyslexia less coordination was found between brain regions in the bilateral auditory cortex that process basic phonemes and Broca’s region, a region in the brain’s frontal lobe known to be involved in higher-level language processing. The researchers interpreted this evidence as suggesting that deficient access to (intact) phonological representations originates from the above-mentioned disconnection between cortical regions and Broca’s region in adults with dyslexia. It remains to be investigated though whether this finding can be replicated in a sample of children with dyslexia and DLD.

In the current study, the two groups showed a similar design fluency performance, and since the design fluency task was used as a measure of children’s EFs, this finding might be interpreted as evidence that poorer phonological fluency performance cannot be attributed to difficulties with EFs in the DDLD group. However, another measure of EFs, namely, switching, suggested difficulties with EFs in the phonological condition in the DDLD group compared to the TD group as the former group switched significantly fewer times in this task than the latter group. Smith-Spark et al. (2017) argued that EF demands should be equal in semantic, phonological and design fluency tasks, in order for one to reach firm conclusions about any possible effects of EFs on fluency tasks. Following Smith-Spark et al.’s (2017) argument, this can be acknowledged as another limitation of the current study too, since the effect of EFs cannot be separated from the effect of difficulties with phonological processing skills.

Last but not least, the DDLD group was a heterogeneous group of children in that the selection of participants was mainly based on the diagnosis that children had received in order to assign them to the DDLD group. Further, it was decided to combine children with diagnoses of dyslexia and DLD together in a single group, as justified in the *Participants* section. However, careful selection of children who fall into more distinct groups (e.g. Ramus et al., 2013) – assuming that adequate assessments of language and literacy skills exist in the language to differentiate these groups – might reveal that children with different profiles of literacy and language impairments have different loci for their retrieval difficulties. Last but not least, given that this is the first investigation of the lexical organization in Greek children with dyslexia and DLD, the findings of the current study need to be replicated.

5.8. Conclusions

The current study is the first developmental study designed to investigate the structure of the lexicon in Greek children with dyslexia and/or DLD using semantic fluency tasks by teasing apart the Poor Lexical-Semantic Structure Model and the Slow-Retrieval Model. Clustering and switching behaviour associated with productivity in the TD and DDLD groups, revealing that the semantic lexicon is organized in a similar way in both groups. Importantly, children with DDLD retrieved significantly fewer correct items in semantic fluency tasks than their TD peers, but the two groups did not differ on the size of semantic clusters. Given that a similarly-sized average cluster was found in the two groups, the findings indicated that slower retrieval processes of lexical items from the mental lexicon resulted in poorer semantic fluency performance in children with DDLD relative to TD children. It is concluded that the lexical retrieval difficulties experienced by children with DDLD in semantic fluency tasks are better explained by the Slow-Retrieval Model (Lenio et al., 2016; Mengisidou, Marshall, & Stavrakaki, under review) than by the Poor Lexical-Semantic Structure Model. The Slow-Lexical Retrieval Model is further supported by data originating from applying computational modelling techniques (Davelaar et al., 2015) used to investigate any group differences in children's semantic network using solely the semantic category of animals (Davelaar et al., in preparation). The results showed that DDLD children's semantic network, as measured with the computed ratio score, was similar to that of their TD peers.

The second objective of the study was to investigate the contested locus of the phonological deficit in Greek children with DDLD using phonological fluency tasks by teasing apart two prominent phonological hypotheses of dyslexia and DLD, one which considers that the phonological deficit in dyslexia and DLD lies in children's impaired phonological representations (the Degraded Phonological Representations Hypothesis), and another which considers that the phonological deficit in dyslexia and DLD lies in children's difficulty to access (intact) phonological representations (the Deficient Phonological Access Hypothesis). The children with DDLD retrieved fewer correct items in phonological fluency tasks than did TD children, and they also switched less often between clustered and/or non-clustered responses. However, a similarly-sized average cluster was found in the two groups. Given that the size of phonological clusters was considered to be an implicit phonological measure of the quality of phonological representations, the above-mentioned finding suggested that in children with DDLD phonological representations were as robust and distinct as those of TD children. This is consistent with the Deficient Phonological Access Hypothesis (Mengisidou & Marshall, 2019; Ramus & Szenkovits, 2008).

Children's language and literacy skills predicted semantic and phonological fluency performance suggesting that poorer semantic and phonological fluency performance in children with DDLD is partly attributed to their inferior language and literacy skills (Mengisidou & Marshall, 2019; Mengisidou et al., under review). Further investigation is, however, needed to shed light on the underlying causes of slow retrieval processes in children with dyslexia and DLD. Furthermore, children with DDLD showed poorer semantic and phonological fluency performance relative to TD children even after controlling for the effect of design fluency performance. This finding supports the specificity of the verbal fluency deficit in children with DDLD on the basis that only semantic and phonological processing difficulties, and not general processing speed difficulties, underlie poorer semantic and phonological fluency performance in children with DDLD. Considering that the verbal fluency data support the view that an access deficit (and not poor semantic and phonological representations themselves) is responsible for poorer semantic and phonological fluency performance in children with DDLD, the above-mentioned finding in turn suggests that the access deficit is restricted to verbal material.

Overall, data originating from both verbal fluency categories supports the view that retrieval processes in children with DDLD are slower relative to TD children, with the quality of DDLD children's semantic and phonological representations themselves being intact. As Lenio et al. (2016) argued, poorer verbal fluency performance is not deterministic to its cause but rather *multifactorial*. Therefore, for both verbal fluency categories, slower retrieval processes originating from deficient access to (intact) semantic and phonological representations, in addition to difficulties with EFs and inferior language and literacy skills, influence verbal fluency performance in Greek children with dyslexia and DLD (Mengisidou & Marshall, 2019; Mengisidou et al., under review).

References

Al-Yagon, M., Cavendish, W., Cornoldi, C., Fawcett, A. J., Grünke, M., Hung, L. Y., Jiménez, J. E., Karande, S., van Kraayenoord, C. E., Lucangeli, D., Margalit, M., Montague, M., Sholapurwala, R., Sideridis, G., Tressoldi, P. E., & Vio, C. (2013). The proposed changes for DSM-5 for SLD and ADHD: International perspectives—Australia, Germany, Greece, India, Israel, Italy, Spain, Taiwan, United Kingdom, and United States. *Journal of Learning Disabilities, 46*, 58-72.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Anastasiou, D., & Protopapas, A. (2015). Difficulties in lexical stress versus difficulties in segmental phonology among adolescents with dyslexia. *Scientific Studies of Reading, 19*(1), 31-50.

Anastasiou, D., & Polychronopoulou, S. (2009). Identification and overidentification of special learning disabilities (dyslexia) in Greece. *Learning Disability Quarterly, 32*(2), 55-69.

Arán-Filippetti, V. A., & Allegri, R. F. (2011). Verbal fluency in Spanish-speaking children: analysis model according to task type, clustering, and switching strategies and performance over time. *Clinical Neuropsychology, 25*(3), 413-436.

Ardila, A., Ostrosky-Solís, F., & Bernal, B. (2006). Cognitive testing toward the future: the example of semantic verbal fluency (animals). *International Journal of Psychology, 41*, 324-332.

Barry, J. G., Yasin, I., & Bishop, D. V. (2007). Heritable risk factors associated with language impairments. *Genes, Brain, and Behavior, 6*(1), 66-76.

Barry, D., Bates, M. E., & Labouvie, E. (2008). FAS and CFL forms of verbal fluency differ in difficulty: a meta-analytic study. *Applied Neuropsychology, 15*(2), 97-106.

Beckage, N., Smith, L., & Hills, T. (2011). Small worlds and semantic network growth in typical and late talkers. *PLoS One, 6*, e19348.

Benjamin, D. J., Berger, J., Johannesson, M., Nosek, B. A., Wagenmakers, E., ... Johnson, V. (2018). Redefine statistical significance. *Nature Human Behaviour, 2*, 6-10.

Bertola L., Cunha Lima M. L., Romano-Silva M. A., de Moraes E. N., Diniz B. S., & Malloy-Diniz L. F. (2014). Impaired generation of new subcategories and switching in a semantic verbal fluency test in older adults with mild cognitive impairment. *Frontiers in Aging Neuroscience, 6*(141).

Best, W. (2005). Investigation of a new intervention for children with word-finding problems. *International Journal of Language and Communication Disorders, 40*, 279-318.

Best, W., Hughes, L. M., Masterson, J., Thomas, M., Fedor, A., Roncoli, S., Fern-Pollak, L., Shepherd, D. L., Howard, D., Shobbrook, K., & Kapikian, A. (2017). Intervention for children with word-finding difficulties: a parallel group randomised control trial. *International Journal of Speech-Language Pathology, 31*, 1-12.

Bird, J., Bishop, D. V., & Freeman, N. H. (1995). Phonological awareness and literacy development in children with expressive phonological impairments. *Journal of Speech and Hearing Research, 38*(2), 446-462.

Bishop, D. V. M., & Snowling, M. J. (2004). Developmental dyslexia and specific language impairment: Same or different? *Psychological Bulletin, 130*, 858-888.

Bishop, D. V. M. (2006). What causes specific language impairment in children? *Current Directions in Psychological Science, 15*(5), 217-221.

Bishop, D. V. M. (2007). Curing dyslexia and attention-deficit hyperactivity disorder by training motor co-ordination: Miracle or myth? *Journal of Paediatrics and Child Health, 43*(10), 653-655.

Bishop, D. V. M., & Norbury, C. F. (2008). Speech and language disorders. In M. Rutter, D. Bishop, D. Pine, S. Scott, J. Stevenson, E. Taylor, & A. Thapar (Eds.), *Rutter's Child and Adolescent Psychiatry* (5th ed., pp. 781-801).

Bishop, D. V. M., MacDonald, D., Bird, S., & Hayiou-Thomas, M. E. (2009). Children who read words accurately despite language impairment: Who are they and how do they do it? *Child Development, 80*(2), 593-605.

Bishop, D. V. M. (2010). Which neurodevelopmental disorders get researched and why? *PLoS One, 5*(11), e15112-e15112.

Bishop, D. V., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & CATALISE consortium. (2016). CATALISE: A multinational and multidisciplinary Delphi consensus study. Identifying language impairments in children. *PLoS One, 11*(12), e0168066.

Bishop, D. V. M. (2017). Why is it so hard to reach agreement on terminology? The case of developmental language disorder (DLD). *International Journal of Language and Communication Disorders*, 52, 671-680.

Bishop, D. V., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & CATALISE-2 consortium. (2017). Phase 2 of CATALISE: a multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology and Psychiatry, 58*(10), 1068-1080.

Bishop, D. V. (2018, June 9). Developmental language disorder: the need for a clinically relevant definition [Web log post]. Retrieved from deevybee.blogspot.com/2018/06/developmental-language-disorder-need.html

Boada, R., & Pennington, B. F. (2006). Deficient implicit phonological representations in children with dyslexia. *Journal of Experimental Child Psychology, 95*, 153-193.

Boets, B., De Smedt, B., Cleuren, L., Vandewalle, E., Wouters, J., & Ghesquière, P. (2010). Towards a further characterization of phonological and literacy problems in Dutch-speaking children with dyslexia. *British Journal of Developmental Psychology, 28*, 5-31.

Boets, B., Op de Beeck, H. P., Vandermosten, M., Scott, S. K., Gillebert, C. R., Mantini, D., Bulthé, J., Sunaert, S., Wouters, J., & Ghesquière, P. (2013). Intact but less accessible phonetic representations in adults with dyslexia. *Science, 342*, 1251-1254.

Boets, B. (2014). Dyslexia: reconciling controversies within an integrative developmental perspective. *Trends in Cognitive Sciences, 18*(10), 501-503.

Botting, N. (2014). Specific Language Impairment (Overview). In P. J. Brooks & V. Kempe (Eds.), *Encyclopedia of Language Development* (pp. 567-573). Thousand Oaks, CA: Sage.

Botting, N., Jones, A., Marshall, C., Denmark, T., Atkinson, J., & Morgan, G. (2017). Nonverbal Executive function is mediated by language: a study of deaf and hearing children. *Child Development, 88*(5), 1689-1700.

Briscoe, J., Bishop, D. V., & Norbury, C. F. (2001). Phonological processing, language, and literacy: a comparison of children with mild-to-moderate sensorineural hearing loss and those with specific language impairment. *Journal of Child Psychology and Psychiatry, 42*(3), 329-340.

Brizzolara, D., Chilosi, A., Cipriani, P., Di Filippo, G., Gasperini, F., Mazzotti, S., Pecini, C., & Zoccolotti, P. (2006). Do phonologic and rapid automatized naming deficits differentially affect dyslexic children with and without a history of language delay? A study of Italian dyslexic children. *Cognitive and Behavioral Neurology, 19*(3), 141-149.

Brooks, P. J., & Kempe, V. (2012). *Language development*. United Kingdom: Blackwell.

Brosnan, M., Demetre, J., Hamill, S., Robson, K., Shepherd, H., & Cody, G. (2002). Executive functioning in adults and children with developmental dyslexia. *Neuropsychologia, 40*, 2144-2155.

Caravolas, M., & Volín, J. (2001). Phonological spelling errors among dyslexic children learning a transparent orthography: The case of Czech. *Dyslexia, 7*, 229-245.

Caravolas, M., Volín, J., & Hulme, C. (2005). Phoneme awareness is a key component of alphabetic literacy skills in consistent and inconsistent orthographies: Evidence from Czech and English children. *Journal of Experimental Child Psychology, 92*(2), 107-139.

Caravolas, M., Lervåg, A., Mousikou, P., Efrim, C., Litavský, M., Onochie-Quintanilla, E., Salas, N., Schöffelová, M., Defior, S., Mikulajová, M., Seidlová-Málková, G., & Hulme, C. (2012). Common patterns of prediction of literacy development in different alphabetic orthographies. *Psychological Science, 23*(6), 678-686.

Cassar, M., Treiman, R., Moats, L., Pollo, T. C., & Kessler, B. (2005). How do the spellings of children with dyslexia compare with those of nondyslexic children? *Reading and Writing, 18*(1), 27-49.

Castles, A., & Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition, 47*(2), 149-180.

Castles, A., & Friedmann, N. (2014). Developmental dyslexia and the phonological deficit hypothesis. *Mind and Language, 29*(3), 270-285.

Catts, H. W. (1993). The relationship between speech-language impairments and reading disabilities. *Journal of Speech, Language, and Hearing Research, 36*(5), 948-958.

Catts, H. W., Adlof, S. M., Hogan, T. P., & Weismer, S. E. (2005). Are specific language impairment and dyslexia distinct disorders? *Journal of Speech, Language, and Hearing Research, 48*(6), 1378-1396.

Chami, S., Munro, N., Docking, K., McGregor, K., Arciuli, J., Baker, E., & Heard, R. (2018). Changes in semantic fluency across childhood: normative data from Australian-English speakers. *International Journal of Speech-Language Pathology, 20*(2), 262-273.

Chertkow, H., & Bub, D. (1990). Semantic memory loss in dementia of Alzheimer's type. What do various measures measure? *Brain, 113*, 397-417.

Chuard, P. J. C., Vrtílek, M., Head, M. L., & Jennions, M. D. (2019). Evidence that nonsignificant results are sometimes preferred: Reverse *P*-hacking or selective reporting? *PLoS Biology, 17*(1): e3000127. doi: 10.1371/journal.pbio.3000127

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Cohen, M. J., Morgan, A. M., Vaughn, M., Riccio, C. A., & Hall, J. (1999). Verbal fluency in children: developmental issues and differential validity in distinguishing children with attention-deficit hyperactivity disorder and two subtypes of dyslexia. *Archives in Clinical Neuropsychology, 14*(5), 433-443.

Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review, 82*, 407-428.

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *Journal of Child Psychology and Psychiatry, 42*(6), 741-748.

Critten, S., Connelly, V., Dockrell, J. E., & Walter, K. (2014). Inflectional and derivational morphological spelling abilities of children with Specific Language Impairment. *Frontiers in Psychology, 5*(948).

Crowe, S. F. (1998). Decrease in performance on the verbal fluency test as a function of time: evaluation in a young healthy sample. *Journal of Clinical and Experimental Neuropsychology, 20*, 391-401.

Daigle, D., Costerg, A., Plisson, A., Ruberto, N., & Varin, J. (2016). Spelling errors in French-speaking children with dyslexia: Phonology may not provide the best evidence. *Dyslexia, 22*(2), 137-157.

Davelaar, E. J., Vrontissis, A., & Fullgrabe, C. (2015, September). *Graphical network analysis on verbal fluency data*. Poster session presented at the 19th Conference of the European Society for Cognitive Psychology (ESCOP), Paphos, Cyprus.

de Bree, E., Rispens, J., & Gerrits, E. (2007). Non-word repetition in Dutch children with (a risk of) dyslexia and SLI. *Clinical Linguistics and Phonetics, 21*(1112), 935-944.

Desrochers, A., Manolitsis, G., Gaudreau, P., & Georgiou, G. (2017). Early contribution of morphological awareness to literacy skills across languages varying in orthographic consistency. *Reading and Writing, 31*(8), 1695-1719.

Di Filippo, G., Zoccolotti, P., & Ziegler, J. C. (2008). Rapid naming deficits in dyslexia: a stumbling block for the perceptual anchor theory of dyslexia. *Developmental Science, 11*(6), 40-47.

Diamanti, V. (2006). *Dyslexia and dysgraphia in Greek in relation to normal development: Cross-linguistic and longitudinal studies*. Unpublished PhD Thesis. University College London.

Diamanti, V., Goulandris, N., Stuart, M., & Campbell, R. (2014). Spelling of derivational and inflectional suffixes by Greek-speaking children with and without dyslexia. *Reading and Writing, 27*(2), 337-358.

Diamanti, V., Mouzaki, A., Ralli, A., Antoniou, F., Papaioannou, S., & Protopapas, A. (2017). Preschool phonological and morphological awareness as longitudinal predictors of early reading and spelling development in Greek. *Frontiers in Psychology, 8*(2039).

Diamanti, V., Goulandris, N., Stuart, M., Campbell, R., & Protopapas, A. (2018). Tracking the effects of dyslexia in reading and spelling development: A longitudinal study of Greek readers. *Dyslexia, 24*(2), 170-189.

Diamanti, V., Goulandris, N., Campbell, R., & Protopapas, A. (2018). Dyslexia profiles across orthographies differing in transparency: An evaluation of theoretical predictions contrasting English and Greek. *Scientific Studies of Reading, 22*(1), 55-69.

Dockrell, J. E., & Marshall, C. R. (2015). Measurement issues: assessing language skills in young children. *Child and Adolescent Mental Health, 20*(2), 116-125.

Douklias, S. D., Masterson, J., & Hanley, J. R. (2010). Surface and phonological developmental dyslexia in Greek. *Cognitive Neuropsychology, 26*(8), 705-23.

Dunn, L. M., & Dunn, D. M. (1981). *Peabody Picture Vocabulary Test-Revised.* Circle Pines, MN; American Guidance Service.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test* (4th Ed). Oxford, UK; Pearson.

Dunn, L. M., Douglas, M., & Dunn, D. M. (2009). *British Picture Vocabulary Scale* (3rd Ed.). London, UK; National Foundation for Educational Research in England and Wales, GL Assessment.

Dyson, H., Best, W., Solity, J., & Hulme. C. (2017). Training mispronunciation correction and word meanings improves children's ability to learn to read words. *Scientific Studies of Reading, 21*(5), 392-407.

Elbro, C., Borstrøm, I., & Petersen, D. K. (1998). Predicting dyslexia from kindergarten: The importance of distinctness of phonological representations of lexical items. *Reading Research Quarterly, 33*(1), 36-60.

Eisenmajer, N., Ross, N., & Pratt, C. (2005). Specificity and characteristics of learning disabilities. *Journal of Child Psychology and Psychiatry, 46*(10), 110-1115.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149-1160.

Field, A. (2013). Discovering statistics using IBM SPSS statistics (4th Edition). Sage, London: Sage.

Frederickson, N., Frith, U., & Reason, R. (1997). *Phonological assessment battery*. Windsor: NFER Nelson.

Frith, U. (1992). Cognitive development and cognitive deficit. *The Psychologist: Bulletin of the British Psychological Society, 5*, 13-19.

Frith, U., Landerl, K., & Frith, C. (1995). Dyslexia and verbal fluency: More evidence for a phonological deficit. *Dyslexia, 1*, 2-11.

Frith, U., & Happé, F. (1998). Why specific developmental disorders are not specific: On-line and developmental effects in autism and dyslexia. *Developmental Science, 1*(2), 267-272.

Frith, U. (1999). Paradoxes in the definition of dyslexia. *Dyslexia, 5*, 192-214.

Gabrieli, J. D. E. (2009). Dyslexia: A new synergy between education and cognitive neuroscience. *Science, 325*, 280-283.
Gallagher, A., Frith, U., & Snowling, M. J. (2000). Precursors of literacy delay among children at genetic risk of dyslexia. *Journal of Child Psychology and Psychiatry, 41*(2), 203-213.
Gathercole, S. E., & Baddeley, A. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language, 29*, 336-360.
Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The Children's Test of Nonword Repetition: a test of phonological working memory. *Memory, 2*(2), 103-27.
Gathercole, S. E., & Baddeley, A. D. (1996). The Children's test of nonword repetition. San Diego, CA: The Psychological Corporation.
Georgas, D. D., Paraskevopoulos, I. N., Bezevegis, E. G., & Giannitsas, N. D. (1997). *Greek WISC-III: Wechsler Intelligence Scales for Children*. Athens: Ellinika Grammata.
Georgiou, G. K., & Parrila, R. (2013). Rapid naming and reading: A review. In H. L. Swanson, K. L. Harris, & S. Graham (Eds.), *Handbook of Learning Disabilities* (pp. 169-185). New York, NY: Guilford.
Georgiou, G. K., Ghazyani, R., & Parrila, R. (2018). Are RAN deficits in university students with dyslexia due to defective lexical access, impaired anchoring, or slow articulation? *Annals of Dyslexia, 1*-19.
Georgiou, G. K., Parrila, R., & Papadopoulos, T. C. (2008). Predictors of word decoding and reading fluency in English and Greek: A cross-linguistic comparison. *Journal of Educational Psychology, 100*(3), 566-580.
Gooch, D., Hulme, C., Nash, H. M., & Snowling, M. J. (2014). Comorbidities in preschool children at family risk of dyslexia. *Journal of Child Psychology and Psychiatry, 55*(3), 237-246.
Goswami, U. (2000). Phonological representations, reading development and dyslexia: Towards a cross-linguistic theoretical framework. *Dyslexia, 6*, 133-151.
Goswami, U. (2003). Why theories about developmental dyslexia require developmental designs. *Trends in Cognitive Sciences, 7*(12), 534-540.
Graf Estes, K., Evans, J. L., & Else-Quest, N. M. (2007). Differences in the nonword repetition performance of children with and without specific language impairment: a meta-analysis. *Journal of Speech Language and Hearing Research, 50*(1), 177-95.
Griffiths, P. (1991). Word-finding ability and design fluency in developmental dyslexia. *British Journal of Clinical Psychology, 30*, 47-60.
Grigorakis, I., & Manolitsis, G. (2016). The contribution of morphological awareness on the early stages of spelling development. *Preschool and Primary Education, 4*(1), 128-148.
Grigorenko, E. L. (2001). Developmental dyslexia: An update on genes, brains, and environments. *Journal of Child Psychology and Psychiatry, 42*(1), 91-125.
Guez, A., Panaïotis, T., Peyre, H., & Ramus, F. (2018). Predictors of the IQ-achievement gap in France: A longitudinal analysis. *Intelligence*, 69, 104-116.
Hall, J., McGregor, K. K., & Oleson, J. (2017). Weaknesses in lexical-semantic knowledge among college students with specific learning disabilities: evidence from a semantic fluency task. *Journal of Speech, Language, and Hearing Research, 60*(3), 640-653.
Harris, J. (1994). *English Sound Structure*. Oxford: Blackwell.
Hazin, I., Leite, G., Oliveira, R. M., Alencar, J. C., Fichman, H. C., Marques, P. d. N., & Berlim de Mello, C. (2016). Brazilian normative data on letter and category fluency tasks: effects of gender, age, and geopolitical region. *Frontiers in Psychology, 7*, 684.
Henry, L. A., Messer, D. J., & Nash, G. (2012). Executive functioning in children with specific language impairment. *Journal of Child Psychology and Psychiatry, 53*(1), 37-45.
Henry, L., Messer, D. J. & Nash, G. (2015). Executive functioning and verbal fluency in children with language difficulties. *Learning and Instruction, 39*, 137-147.

Henry, L. A., & Botting, N. (2017). Working memory and developmental language impairments. *Child Language Teaching and Therapy, 33*(1) 19-32.
Hesketh, A., & Conti-Ramsden, G. (2013). Memory and language in middle childhood in individuals with a history of specific language impairment. *PLoS One, 8*(2), e56314.
Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. B. (2009). Categorical structure among shared features in networks of early-learned nouns. *Cognition, 112*, 381-396.
Hoenig, J. M., & Heisey, D. M. (2001). The abuse of power. The pervasive fallacy of power calculations for data analysis. *The American Statistician, 55*(1), 19-24.
Hulme, C., Snowling, M., Caravolas, M., & Carroll, J. (2005). Phonological skills are (probably) one cause of success in learning to read: A comment on Castles and Coltheart. *Scientific Studies of Reading, 9*(4), 351-365.
Hulme, C., & Snowling, M. J. (2009). *Developmental disorders of language learning and cognition*. Oxford, UK: Wiley-Blackwell.
Hulme, C., & Snowling, M. J. (2014). The interface between spoken and written language: Developmental disorders. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*, 2-8.
Hurks, P. P. M., Vles, J. S. H., Hendriksen, J. G. M., Kalff, A. C., Feron, F. J. M., Kroes, M., van Zeben, T. M., Steyaert, J., & Jolles, J. (2006). Semantic category fluency versus initial letter fluency over 60 seconds as a measure of automatic and controlled processing in healthy school-aged children. *Journal of Clinical and Experimental Neuropsychology, 28,* 684-695.
Hurks, P. P. M., Schrans, D., Meijs, C., Wassenberg, R., Feron, F. J. M., & Jolles, J. (2010). Developmental changes in semantic verbal fluency: analyses of word productivity as a function of time, clustering and switching. *Child Neuropsychology, 16*, 366-387.
John, S., Rajashekhar, B., & Guddattu, V. (2016). Word retrieval ability on phonemic fluency in typically developing children. *Applied Neuropsychology: Child, 5*(4), 252-263.
Jones, A., Atkinson, J., Marshall, C., Botting, N., St Clair, M. C., & Morgan, G. (2019). Expressive vocabulary predicts nonverbal executive function: A 2-year longitudinal study of deaf and hearing children. *Child Development*. doi: 10.1111/cdev.13226
Joseph, H., & Nation, K. (2018). Examining incidental word learning during reading in children: The role of context. *Journal of Experimental Child Psychology, 166*, 190-211.
Joye, N. (2018). *Spelling with a developmental language disorder: predictors, strategies and error patterns in French- and English-speaking students at the end of primary school.* Unpublished PhD thesis, UCL Institute of Education, University College London.
Joye, N., Broc, L., Olive, T., & Dockrell, J. E. (2019). Spelling performance in children with developmental language disorder: A meta-analysis across European languages. *Scientific Studies of Reading*. doi: 10.1080/10888438.2018.1491584
Kaiser, H. (1974). An index of factorial simplicity. *Psychometrika, 39*(1), 31-36.
Kalnak, N., Peyrard-Janvid, M., Forssberg, H., & Sahlén, B. (2014). Nonword repetition - a clinical marker for specific language impairment in Swedish associated with parents' language-related problems. *PLoS One*, *9*(2), e89544.
Kamhi, A., & Catts, H. (1986). Toward an understanding of developmental language and reading disorders. *Journal of Speech, Language, and Hearing Research, 51*, 337-47.
Kavé, G., Kigel, S., & Kochva, R. (2008). Switching and clustering in verbal fluency tasks throughout childhood. *Journal of Clinical and Experimental Neuropsychology, 30*(3), 349-359.
Kenett, Y. N., Wechsler-Kashi, D., Kenett, D. Y., Schwartz, R. G., Ben-Jacob, E., & Faust, M. (2013). Semantic organization in children with cochlear implants: Computational analysis of verbal fluency. *Frontiers in Psychology, 4*, 1-11.

Klenberg, L., Korkman, M., & Lahti-Nuuttila, P. (2001). Differential development of attention and executive functions in 3- to 12-year-old Finnish children. *Developmental Neuropsychology, 20*(1), 407-428.

Korkman, M., Kirk, U., & Kemp, S. (1998). *NEPSY: A developmental neuropsychological assessment*. San Antonio, TX: The Psychological Corporation.

Korkman, M., Kemp, S. L., & Kirk, U. (2001). Effects of age on neurocognitive measures of children ages 5 to 12: a crosssectional study on 800 Children from the United States. *Developmental Neuropsychology, 20,* 331-354.

Kosmidis, M. H., Vlahou, C. H., Panagiotaki, P., & Kiosseoglou, P. (2004). The verbal fluency task in the Greek population: Normative data, and clustering and switching strategies. *Journal of the International Neuropsychological Society, 10*, 164-72.

Kronbichler, M., Hutzler, F., & Wimmer, H. (2002). Dyslexia: Verbal impairments in the absence of magnocellular impairments. *NeuroReport, 13*(5), 617-620.

Ladefoged, P. (2001). *Vowels and consonants: An introduction to the sounds of languages*. Oxford: Blackwell Publishers.

Landerl, K., & Wimmer, H. (2000). Deficits in phoneme segmentation are not the core problem of dyslexia: Evidence from German and English children. *Applied Psycholinguistics, 21*, 243-262.

Landerl, K., Fussenegger, B., Moll, K., & Willburger, E. (2009). Dyslexia and dyscalculia: two learning disorders with different cognitive profiles. *Journal of Experimental Child Psychology, 103*, 309-324.

Landerl, K., Ramus, F., Moll, K., Lyytinen, H., Leppänen, P. H., Lohvansuu, K., … & Schulte-Körne, G. (2013). Predictors of developmental dyslexia in European orthographies with varying complexity. *Journal of Child Psychology and Psychiatry, 54*(6), 686-694.

Lefavrais, P. (1967). *Test de l'Alouette*. Paris: Les Editions du Centre de Psychologie Appliquée.

Lenio, S., Lissemore, F. M., Sajatovic, M., Smyth, K. A., Tatsuoka, C., Woyczynski, W. A., & Lerner, A. J. (2016). Detrending changes the temporal dynamics of a semantic fluency task. *Frontiers in Aging Neuroscience, 8*(252).

Leonard, L. (1998). *Children with specific language impairment*. Cambridge, MA: MIT Press.

Leong, V., Hämäläinen, J., Soltész, F., & Goswami, U. (2011). Rise time perception and detection of syllable stress in adults with developmental dyslexia. *Journal of Memory and Language, 64*, 59-73.

Lerner, A. J., Ogrocki, P. K., & Thomas, P. L. (2009). Network graph analysis of category fluency testing. *Cognitive and Behavioural Neurology, 22*(1), 54-52.

Levin, B. E. (1990). Organisational deficits in dyslexia: possible frontal lobe dysfunction. *Developmental Neuropsychology, 6*, 95-110.

Loucas, T., Baird, G., Simonoff, E., & Slonims, V. (2016). Phonological processing in children with specific language impairment with and without reading difficulties. *International Journal of Language and Communication Disorders, 51*(5), 581-588.

Lundberg, I. (2002). The child's route into reading and what can go wrong. *Dyslexia, 8*, 1-13.

Luo, L., Luk, G., & Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition, 114*(1), 29-41.

Lyytinen, H., Aro, M., Eklund, K., Erskine, J., Guttorm, T., Laakso, M. L., Leppänen, P. H., Lyytinen, P., Poikkeus, A. M., & Torppa, M. (2004). The development of children at familial risk for dyslexia: Birth to early school age. *Annals of Dyslexia, 54*(2), 184-220.

Marinelli, C. V., Romani, C., Burani, C., & Zoccolotti, P. (2015). Spelling acquisition in English and Italian: A cross-linguistic study. *Frontiers in Psychology, 6*(1843).

Marslen-Wilson, W. (1984). Function and processing in spoken word recognition: a tutorial review. In H. Bouma, & D. G. Bouwhuis (Eds.), *Attention and Performance X: Control of Language Processing* (pp. 169-185). Hillsdale, NJ: Lawrence Erlbaum Associates.

Marshall, C. R., & van der Lely, H. K. J. (2009). Effects of word position and stress on onset cluster production: Evidence from typical development, SLI and dyslexia. *Language, 85*, 39-57.
Marshall, C. R., & Messaoud-Galusi, S. (2010). Editorial introduction. Developmental disorders of language and literacy: Special Issue. *British Journal of Developmental Psychology, 28*, 1-4.
Marshall, C. R., Ramus, F., & van der Lely, H. (2010). Do children with dyslexia and/or specific language impairment compensate for place assimilation? Insight into phonological grammar and representations. *Cognitive Neuropsychology, 27*(7), 563-586.
Marshall, C. R., Rowley, K., Mason, K., Herman, R., & Morgan, G. (2013). Lexical organization in deaf children who use British sign language: Evidence from a semantic fluency task. *Journal of Child Language, 40*, 193-220.
Marshall, C. R. (2014). Word production errors in children with developmental language impairments. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1634): 20120389.
Marshall, C., Rowley, K., & Atkinson, J. (2014). Modality-dependent and -independent factors in the organisation of the signed language lexicon: Insights from semantic and phonological fluency tasks in BSL. *Journal of Psycholinguistic Research, 43*(5), 587-610.
Marshall, C. R., Jones, A., Fastelli, A., Atkinson, J., Botting, N., & Morgan, G. (2018). Semantic fluency in deaf children who use spoken and signed language in comparison with hearing peers. *International Journal of Language and Communication Disorders, 53*(1), 157-170.
Marshall, C. (2019, April). What are the advantages and disadvantages of reconceptualising "specific language impairment" as "developmental language disorder"? Presentation, Aristotle University of Thessaloniki, Thessaloniki, Greece. 4th April 2019.
Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition, 25*, 71-102.
Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology, 58,* 25-45.
Marzocchi, G. M., Oosterlaan, J., Zuddas, A., Cavolina, P., Geurts, H., Redigolo, D., Vio, C., & Sergeant, J. A. (2008). Contrasting deficits on executive functions between ADHD and reading disabled children. *Journal of Child Psychology and Psychiatry, 49*(5), 543-552.
Matute, E., Rosselli, M., Ardila., A., & Morales. G. (2004) Verbal and nonverbal fluency in Spanish-Speaking children. *Developmental Neuropsychology, 26*(2), 647-660.
McArthur, G. M., Hogben, J. H., Edwards, V. T., Heath, S. M., & Mengler, E. D. (2000). On the "specifics" of specific reading disability and specific language impairment. *Journal of Child Psychology and Psychiatry, 41*(7), 869-874.
Melby-Lervåg, M., Lyster, S. A., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin, 138*(2), 322-352.
Menghini, D., Finzi, A., Benassi, M., Bolzani, R., Facoetti, A., Giovagnoli, S., Ruffino, M., & Vicari, S. (2010). Different underlying neurocognitive deficits in developmental dyslexia: a comparative study. *Neuropsychologia, 48,* 863-872.
Mengisidou, M., & Marshall, C. R. (2019). Deficient explicit access to phonological representations explains phonological fluency difficulties in Greek children with dyslexia and/or developmental language disorder. *Frontiers in Psychology,* 10:638. doi: 10.3389/fpsyg.2019.00638
Mengisidou, M., Marshall, C. R., & Stavrakaki, S. (submitted). Semantic fluency difficulties in developmental dyslexia and developmental language disorder (DLD): Poor semantic structure of the lexicon or slower retrieval processes?
Messaoud-Galusi, S., & Marshall, C. R. (2010). Introduction to this special issue exploring the overlap between dyslexia and SLI: The role of phonology. *Scientific Studies of Reading, 14*(1), 1-7.
Messer, D., Dockrell, J. E., & Murphy, N. (2004). The relation between naming and literacy in children with word finding difficulties. *Journal of Educational Psychology, 96*, 462-470.

Messer, D., & Dockrell, J. E. (2006). Children's naming and word-finding difficulties: descriptions and explanations. *Journal of Speech, Language, and Hearing Research, 49*(2), 309-324.
Messer, D., & Dockrell, J. E. (2013). Children with word finding difficulties: Continuities and profiles of abilities. *First Language, 33*, 433-448.
Mielnik, A., Lockiewicz, M., & Bogdanowicz, M. (2015). Semantic and phonological verbal fluency in students with dyslexia. *Acta Neuropsychologica, 13*, 253-266.
Mirman, D., & Britt, A. E. (2014). What we talk about when we talk about access deficits. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1634): 20120388.
Moll, K., Ramus, F., Bartling, J., Bruder, J., Kunze, S., Neuhoff, N., … & Landerl, K. (2014). Cognitive mechanisms underlying reading and spelling development in five European orthographies. *Learning and Instruction, 29*, 65-77.
Morton, J., & Frith, U. (1995). Causal modeling: A structural approach to developmental psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), *Wiley series on personality processes. Developmental psychopathology, Vol. 1. Theory and methods* (pp. 357-390). Oxford, England: John Wiley & Sons.
Moura, O., Simões, M. R., & Pereira, M. (2015). Executive functioning in children with developmental dyslexia. *Clinical Neuropsychology, 28*, 20-41.
Mouzaki, A., Sideridis, G. D., Protopapas, A., & Simos, G. (2007). Investigation of psychometric properties of a spelling achievement task in students of Grade 2, 3, 4, and 5 of Primary school. *Education Sciences, 1*, 129-146.
Mouzaki, A. & Protopapas, Th. (2010). *Orthography. Learning and disorders*. Athens: Gutenberg.
Nash, H. M., & Snowling, M. J. (2008). Semantic and phonological fluency in children with Down syndrome: Atypical organization of language or less efficient retrieval strategies? *Cognitive Neuropsychology, 25*, 690-703.
Nash, H. M., Hulme, C., Gooch, D., & Snowling, M. J. (2013). Preschool language profiles of children at family risk of dyslexia: Continuities with specific language impairment. *Journal of Child Psychology and Psychiatry, 54*, 958-968.
Nathan, L., Stackhouse, J., Goulandris, N., & Snowling, M. J. (2004). The development of early literacy skills among children with speech difficulties: a test of the "critical age hypothesis". *Journal of Speech, Language, and Hearing Research, 47*(2), 377-391.
Nation, K. (2017). Nurturing a lexical legacy: reading experience is critical for the development of word reading skill. *Npj Science of Learning, 2*, Article: 3.
Nation, K. (2009). Form-meaning links in the development of visual word recognition. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*, 3665-3674.
Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: broader language skills contribute to the development of reading. *Journal of Research in Reading, 27*(4), 342-356.
Nebes, R. D., Boller, F., & Holland, A. (1986). Use of semantic context by patients with Alzheimer's disease. *Psychology and Aging, 1*(3), 261-269.
Nieto, A., Galtier, I., Barroso, J., & Espinosa, G. (2008). Verbal fluency in school-aged Spanish children: normative data and analysis of clustering and switching strategies. *Revista De Neurologia, 46*(1), 2-6.
Nikolopoulos, D., Goulandris, N., & Snowling, D. (2003). Developmental dyslexia in Greek. In N. Goulandris (Ed.), *Dyslexia in different languages: Cross-linguistic comparisons* (pp. 53-67). London, United Kingdom: Whurr Publishers.
Nikolopoulos, D., Goulandris, N., Hulme, C., & Snowling, M. (2006). The cognitive bases of learning to read and spell in Greek: Evidence from a longitudinal study. *Journal of Experimental Child Psychology, 94*, 1-17.

Niolaki, G. Z., Masterson, J., & Terzopoulos, A. R. (2014). Spelling improvement through letter-sound and whole-word training in two multilingual Greek- and English-speaking children. *Multilingual Education, 4*(20), 1-23.

Niolaki, G. Z., & Masterson, J. (2013). Intervention for a multi-character processing deficit in a Greek-speaking child with surface dyslexia. *Cognitive Neuropsychology, 30*(4), 208-232.

Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., Vamvakas, G., & Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: evidence from a population study. *Journal of Child Psychology and Psychiatry, 57*(11), 1247-1257.

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., … Yarkoni, T. (2015). Promoting an open research culture. *Science, 348*(6242), 1422-1425.

Pallant, J. (2010). *SPSS survival manual: A step by step guide to data analysis using SPSS* (4th ed.). McGraw-Hill, Berkshire, England.

Panteliadou, S., & Antoniou, F. (2007). *Reading Test Alpha*. Athens: YPEPQ-EPEAEK.

Papadopoulos, T. C., Kendeou, P., & Spanoudis, G. (2012). Investigating the factor structure and measurement invariance of phonological abilities in a sufficiently transparent language. *Journal in Educational Psychology,* 104, 321-336.

Papadopoulos, T. C., Spanoudis, G., & Kendeou, P. (2009). The dimensionality of phonological abilities in Greek. *Reading Research Quarterly,* 44, 127-143.

Parrila, R. K., & Protopapas, A. (2017). Dyslexia and word reading problems. In K. Cain, D. L. Compton, & R. K. Parrila (Eds.), *Theories of Reading Development* (pp. 333-358). John Benjamins.

Pennington, B. F., & Bishop, D. V. (2009). Relations among speech, language, and reading disorders. *Annual Review of Psychology, 60*, 283-306.

Perfetti, C. A., & Hart, L. (2001). The lexical basis of comprehension skill. In D. S. Gorfein (Ed.), *Decade of behavior. On the consequences of meaning selection: Perspectives on resolving lexical ambiguity* (pp. 67-86). Washington, DC, US: American Psychological Association. doi.org/10.1037/10459-004

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

Pittas, E., & Nunes, T. (2014). The relation between morphological awareness and reading and spelling in Greek: A longitudinal study. *Reading and Writing, 27*(8), 1507-1527.

Plaza, M., Cohen, H., & Chevrie-Muller, C. (2002). Oral language deficits in dyslexic children: weaknesses in working memory and verbal planning. *Brain and Cognition, 48*(2-3), 505-512.

Porpodas, C. D. (1999). Patterns of phonological and memory processing in beginning readers and spellers of Greek. *Journal of Learning Disabilities, 32,* 406-416.

Porpodas, C. D. (2006). Literacy acquisition in Greek: Research review of the role of phonological and cognitive factors. In R. M. Joshi & P. J. Aaron (Eds.) *Handbook of Orthography and Literacy* (pp. 189-200). London: Lawrence Erlbaum Associates John Benjamins.

Porpodas, C. D. (2007). *Test of Detection and Investigation of Reading Difficulties in Kindergarten and Grades 1 and 2*. Athens: YPEPQ-EPEAEK.

Protopapas, A., & Skaloumbakas, C. (2007). Traditional and computer-based screening and diagnosis of reading disabilities in Greek. *Journal of Learning Disabilities, 40*, 15-36.

Protopapas, A., Skaloumbakas, C., & Bali, P. (2008). Validation of unsupervised computer-based screening for reading disability in the Greek elementary Grades 3 and 4. *Learning Disabilities: A Contemporary Journal,* 6, 45-69.

Protopapas, A., & Skaloumbakas, C. (2008). Η αξιολόγηση της αναγνωστικής ευχέρειας για τον εντοπισμό αναγνωστικών δυσκολιών. *Psychologia, 15*(3), 267-289.

Protopapas, A., & Vlahou, E. L. (2009). A comparative quantitative analysis of Greek orthographic transparency. *Behavior Research Methods, 41*(4), 991-1008.

Protopapas, A., Simos, P. G., Sideridis, G. D., & Mouzaki, A. (2012). The components of the simple view of reading: a confirmatory factor analysis. *Reading Psychology,* 33, 217-240.
Protopapas, A., Altani, A., & Georgiou, G. K. (2013). Development of serial processing in reading and rapid naming. *Journal in Experimental Child Psychology,* 116, 914-929.
Protopapas, A., Fakou, A., Drakopoulou, S., Skaloumbakas, C., & Mouzaki, A. (2013). What do spelling errors tell us? Classification and analysis of errors made by Greek schoolchildren with and without dyslexia. *Reading and Writing, 26*(5), 615-646.
Protopapas, A. (2014). From temporal processing to developmental language disorders: mind the gap. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1634):20130090.
Protopapas, A. (2017). Learning to Read Greek. In L. Verhoeven & C. Perfetti (Eds.), *Learning to read across languages and writing systems* (pp. 155-180). Cambridge, United Kingdom: Cambridge University Press.
Ramus, F. (2001). Outstanding questions about phonological processing in dyslexia. *Dyslexia, 7*, 197-216.
Ramus, F. (2002). Evidence for a domain-specific deficit in developmental dyslexia. *Behavioral and Brain Sciences, 25*(6), 767-768.
Ramus, F. (2003). Developmental dyslexia: Specific phonological deficit or general sensorimotor dysfunction? *Current Opinion in Neurobiology, 13*(2), 212-218.
Ramus, F., Rosen, S., Dakin, S.C., Day, B.L., Castellote, J.C., White, S., & Frith, U. (2003). Theories of developmental dyslexia: Insights from a multiple case study of dyslexic adults. *Brain, 126*, 841-865.
Ramus, F. (2004a). The neural basis of reading acquisition. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences* (3rd ed., pp. 815-824). Cambridge, MA: MIT Press.
Ramus, F. (2004b). Should neuroconstructivism guide developmental research? *Trends in Cognitive Sciences, 8*(3), 100-101.
Ramus, F., & Szenkovits, G. (2008). What phonological deficit? *The Quarterly Journal of Experimental Psychology, 61*(1), 129-141.
Ramus, F. (2010). Le déficit phonologique dans la dyslexie développementale. World Dyslexia Forum, UNESCO, Paris, 3-5/02/2010. Retrieved from http://www.worlddyslexiaforum.org/
Ramus, F., Peperkamp, S., Christophe, A., Jacquemot, C., Kouider, S., & Dupoux, E. (2010). A psycholinguistic perspective on the acquisition of phonology. In C. Fougeron, B. Kühnert, M. d'Imperio & N. Vallée (Eds.), *Laboratory Phonology 10: Variation, Phonetic Detail and Phonological Representation* (pp. 311-340). Berlin: Mouton de Gruyter. Retrieved from http://www.lscp.net/persons/ramus/en/puben.html
Ramus, F., & Ahissar, M. (2012). Developmental dyslexia: The difficulties of interpreting poor performance, and the importance of normal performance. *Cognitive Neuropsychology, 29*(1-2), 104-122.
Ramus, F. (2013). Dyslexia, Phonological processing in. In H. Pashler (Ed.), *Encyclopedia of the Mind* (pp. 269-271). Thousand Oaks, CA: Sage.
Ramus, F., Marshall, C. R., Rosen, S., & van der Lely, H. K. J. (2013). Phonological deficits in specific language impairment and developmental dyslexia: Towards a multidimensional model. *Brain, 136*(Pt 2), 630-645.
Ramus, F., & Altarelli, I. (2014). Developmental Dyslexia. In P. J. Brooks & V. Kempe (Eds.), *Encyclopedia of Language Development* (pp. 139-146). Thousand Oaks, CA: Sage.
Raven J. C. (2008). *Raven's Coloured Progressive Matrices and Vocabulary Scales*. Pearson, UK.
Regard, M., Strauss, E., & Knapp, P. (1982). Children's production on verbal and nonverbal fluency tasks. *Perceptual and Motor Skills, 55*, 839-844.
Reiter, A., Tucha, O., & Lange, K. W. (2005). Executive functions in children with dyslexia. *Dyslexia, 11*(2), 116-131.

Resch, C., Martens, R., & Hurks, P. (2014). Analysis of young children's abilities to cluster and switch during a verbal fluency task. *Clinical Neuropsychology, 28*(8), 1295-1310.

Rice, M., & Brooks, G. (2004). *Developmental dyslexia in adults: A research review*. National Research and Developmental Centre for adult literacy and numeracy.

Riva, D., Nichelli, F., & Devoti, M. (2000). Developmental aspects of verbal fluency and confrontation naming in children. *Brain and Language, 71*, 267-284.

Rogers, T. T., & McClelland, J. L. (2008). Précis of semantic cognition: A parallel distributed processing approach. *Behavioral and Brain Sciences, 31*, 689-714.

Rohrer, D., Wixted. J. T., Salmon, D. P., & Butters, N. (1995). Retrieval from semantic memory and its implications for Alzheimer's disease. *Journal of Experimental Psychology Learning, Memory and Cognition, 21*(5), 1127-1139.

Saksida, A., Iannuzzi, S., Bogliotti, C., Chaix, Y., Démonet, J. F., Bricout, L., … Ramus, F. (2016). Phonological skills, visual attention span, and visual stress in developmental dyslexia: insights from a population of French children. *Developmental Psychology, 52*(10), 1503-1516.

Sauzéon, H., Lestage, P., Raboutet, C., N'Kaoua, B., & Claverie, B. (2004). Verbal fluency output in children aged 7-16 as a function of the production criterion: qualitative analysis of clustering, switching processes, and semantic network exploitation. *Brain and Language, 89*(1), 192-202.

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94*(Pt 2), 143-174.

Shao, Z., Janse, E., Visser, K., & Meyer, A. S. (2014). What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. *Frontiers in Psychology, 5*, 1-10.

Shelton, J. R., & Caramazza, A. (1999). Deficits in lexical and semantic processing: implications for models of normal language. *Psychonomic Bulletin & Review, 6*(1), 5-27.

Sheng, L., & McGregor, K. K. (2010). Lexical-semantic organisation in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 53*, 146-159.

Sideridis, G., Mouzaki, A., Protopapas, A., & Simos, P. (2008). Psychometric investigation of a spelling task for students in primary school. *Psychology, 15*, 290-315.

Sideridis, G., Antoniou, F., Mouzaki, A., & Simos, P. (2015). *Raven's Educational CPM/CVS*. Athens: Motivo.

Simos, P. G., Sideridis, G. D., Protopapas, A., & Mouzaki, A. (2011). Psychometric evaluation of a receptive vocabulary test for Greek elementary students. *Assessment for Effective Intervention, 37*(1), 1-17.

Smith-Spark, J. H., Henry, L. A., Messer, D. J., Edvardsdottir, E., & Zięcik, A. P. (2016). Executive functions in adults with developmental dyslexia. *Research in Developmental Disabilities, 53-54*, 323-41.

Smith-Spark, J. H., Henry, L. A., Messer, D. J., & Zięcik, A. P. (2017). Verbal and nonverbal fluency in adults with developmental dyslexia: phonological processing or executive control problems? *Dyslexia, 23*(3), 234-250.

Snowling, M. J., & Hulme, C. (1994). The development of phonological skills in reading. *Philosophical Transactions of the Royal Society B: Biological Sciences, 346*, 21-26.

Snowling, M.J. (2000). *Dyslexia* (2nd ed.). Oxford: Blackwell.

Snowling, M. J., Gallagher, A., & Frith, U. (2003). Family risk of dyslexia is continuous: individual differences in the precursors of reading skill. *Child Development, 74*(2), 358 373.

Snowling, M. J., & Hulme, C. (2008). Reading and other specific learning difficulties. In M. Rutter, D. Bishop, D. Pine, S. Scott, J. Stevenson, E. Taylor, & A. Thapar (Eds.), *Rutter's Child and Adolescent Psychiatry* (5th ed., pp.802-819). Oxford, UK: Blackwell.

Snowling, M. J. (2014). Dyslexia: A language learning impairment. *Journal of the British Academy, 2,* 43-58.

Snowling, M. (2017). Dyslexia and developmental language disorder: same or different? Retrieved from https://www.acamh.org/blog/dyslexia-developmental-language-disorder-different/?utm_content=buffer07db7&utm_medium=social&utm_source=facebook.com&utm_campaign=buffer

Soroli, E., Szenkovits, G., & Ramus, F. (2010). Exploring dyslexics' phonological deficit III: Foreign speech perception and production. *Dyslexia*, 16(4), 318-340.

Spanoudis, G. C., & Natsopoulos, D. (2011). Memory functioning and mental verbs acquisition in children with specific language impairment. *Research in Developmental Disabilities*, 32, 2916-2926.

Spanoudis, G. C., Papadopoulos, T. C., & Spyrou, S. (2018). Specific language impairment and reading disability: Categorical distinction or continuum? *Journal of Learning* Disabilities, 00, 1-12.

Sprenger-Charolles, L., Cole, P., Bechennec, 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). *Revue Europeenne de Psychologie Appliquee, 55*(3), 157-186.

Stavrakaki, S. (2001). Comprehension of reversible relative clauses in specifically language impaired and normally developing Greek children. *Brain and Language, 77*, 419-431.

Stavrakaki, S. (2005). Greek neurolinguistics: The state of the art. *Journal of Greek Linguistics, 6*, 187-234.

Stavrakaki, S., & Tsimpli, I. M. (2000). Diagnostic Verbal IQ Test for Greek preschool and school age children: Standardization, statistical analysis, psychometric properties, in *Proceedings of the 8th Symposium of the Panhellenic Association of Logopedists* (Athens: Ellinika Grammata), 95-106.

Stavrakaki, S., M.A. Chrysomallis, M. A., & Petraki, E. (2011). Subject-verb agreement, object clitics and wh-questions in bilingual French-Greek SLI: The case study of a French-Greek child with SLI. *Clinical Linguistics & Phonetics, 25*, 339-367.

Sullivan, G. M., & Feinn, R. (2012). Using effect size—or why the p value is not enough. *Journal of Graduate Medical Education, 4*(3), 279-282.

Szenkovits, G., & Ramus, F. (2005). Exploring dyslexics' phonological deficit I: Lexical vs. sub-lexical and input vs. output processes. *Dyslexia, 11*(4), 253-268.

Szenkovits, G., Darma, Q., Darcy, I., & Ramus, F. (2016). Exploring dyslexics' phonological deficit II: phonological grammar. *First Language, 36*(3), 316-337.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn & Bacon/Pearson Education.

Talli, I. (2010). *Linguistic abilities in developmental dyslexia and specific language impairment (SLI): A comparative and cross-linguistic approach*. Unpublished joint PhD Thesis, Université Paris Descartes and Aristotle University of Thessaloniki.

Talli, I., Sprenger-Charolles, L., & Stavrakaki, S. (2015). Is there an overlap between Specific Language Impairment and Developmental Dyslexia? New insights from French. In S. Stavrakaki (Ed.), *Specific language impairment. Current trends in research (LALD) 58* (pp. 57-87). Amsterdam: John Benjamins.

Talli, I., Sprenger-Charolles, L., & Stavrakaki, S. (2016). Specific language impairment and developmental dyslexia: What are the boundaries? Data from Greek children. *Research in Developmental Disabilities*, 49-50, 339-353.

Thompson, P. A., Hulme, C., Nash, H. M., Gooch, D., Hayiou-Thomas, E., & Snowling, M. J. (2015). Developmental dyslexia: Predicting individual risk. *Journal of Child Psychology and Psychiatry, 56*(9), 976-87.

Tomblin, J. B., & Zhang, X. (2006). The dimensionality of language ability in school-age children. *Journal of Speech, Language, and Hearing Research, 49*, 1193-1208.

Troster, A. I., Warmflash, V., Osorio, I., Paolo, A. M., Alexander, L. J., & Barr, W. B. (1995) The roles of semantic networks and search efficiency in verbal fluency performance in intractable temporal lobe epilepsy. *Epilepsy Research, 21*(1), 19-26.

Troyer, A. K. (2000). Normative data for clustering and switching on verbal fluency tasks. *Journal of Clinical and Experimental Psychology, 22*(3), 370-378.

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: evidence from younger and older healthy adults. *Neuropsychology, 11*(1), 138-146.

Troyer, A. K., Moscovitch, M., Winocur, G., Leach, L., & Freedman, M. (1998). Clustering and switching on verbal fluency tests in Alzheimer's and Parkinson's disease. *Journal of International Neuropsychological Society, 4*(2), 137-143.

Troyer, A. K., Moscovitch, M., Winocur, G., Alexander, M. P., & Stuss, D. (1998). Clustering and switching on verbal fluency: the effects of focal frontal- and temporal-lobe lesions. *Neuropsychologia, 36*(6), 499-504.

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2011). Variation in verbal fluency: a latent variable analysis of clustering, switching, and overall performance. *Quarterly Journal of Experimental Psychology, 64*(3), 447-466.

van Alphen, P., de Bree, E., Gerrits, E., de Jong, J., Wilsenach, C., & Wijnen, F. (2004). Early language development in children at risk of dyslexia. *Dyslexia, 10*, 265-288.

van der Lely, H. K. (1998). SLI in children: Movement, economy and deficits in the computational syntactic system. *Language Acquisition*, *72*, 161-192.

van der Lely, H. K. J., & Marshall, C. R. (2010). Assessing component language deficits in the early detection of reading difficulty risk. *Journal of Learning Disabilities, 43*(4), 357-368.

van der Lely, H. K., & Marshall, C. R. (2011). Grammatical-specific language impairment: A window onto domain specificity. In J. Gouendouzi, F. Loncke, & M. J. Williams (Eds.), *The handbook of psycholinguistics and cognitive processes: Perspectives in communication disorders* (pp. 401-418). New York: Psychology Press.

Vandewalle, E., Boets, B., Ghesquière, P., & Zink, I. (2010). Who is at risk for dyslexia? Phonological processing in five-to seven-year-old Dutch-speaking children with SLI. *Scientific Studies of Reading, 14*(1), 58-84.

Varvara, P., Varuzza, C., Sorrentino, A. C. P., Vicari, S., & Menghini, D. (2014). Executive functions in developmental dyslexia. *Frontiers in Human Neuroscience, 8*, 1-8.

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry, 45*(1), 2-40.

Vonberg, I., Ehlen, F., Fromm, O., & Klostermann, F. (2014). The absoluteness of semantic processing: lessons from the analysis of temporal clusters in phonemic verbal fluency. *PLoS One, 9*(12), e115846.

Vogindroukas, I., Protopapas, A., & Stavrakaki, S. (2009). *Action pictures. Test of grammar and information accuracy*. Chania: Glafki.

Weckerly J., Wulfeck B., & Reilly J. (2001). Verbal fluency deficits in children with specific language impairment: slow rapid naming or slow to name? *Child Neuropsychology, 7*, 142-152.

Wechsler, D. (1991). *WISC-III: Wechsler Intelligence Scale for Children*. San Antonio, TX: Psychological Corporation.

West, G., Vadillo, M. A., Shanks, D. R., & Hulme, C. (2018). The procedural learning deficit hypothesis of language learning disorders: we see some problems. *Developmental Science, 21*(2), e12552.

White, S., Milne, E., Rosen, S., Hansen, P., Swettenham, J., Frith, U., & Ramus, F. (2006a). The role of sensorimotor impairments in dyslexia: A multiple case study of dyslexic children. *Developmental Science, 9*(3), 237-269.

White, S., Frith, U., Milne, E., Rosen, S., Swettenham, J., & Ramus, F. (2006b). A double dissociation between sensorimotor impairments and reading disability: A comparison of autistic and dyslexic children. *Cognitive Neuropsychology, 23*(5), 748-761.

Whiteside, D. M., Kealey, T., Semla, M., Luu, H., Rice, L., Basso, M. R., & Roper, B. (2016). Verbal fluency: Language or executive function measure? *Applied Neuropsychology: Adult, 23*(1), 29-34.

Williams, D. M., & Lind, S. E. (2013). Comorbidity and diagnosis of developmental disorders. In C. R. Marshall (Ed.), *Current Issues in Developmental Disorders* (1st ed., pp. 19-45). New York, USA: Psychology Press.

Wimmer, H. (1996). The early manifestation of developmental dyslexia: Evidence from German children. *Reading and Writing, 8*(2), 171-188.

Wimmer, H., Mayringer, H., & Landerl, K. (1998). Poor reading: A deficit in skill automatization or a phonological deficit? *Scientific Studies of Reading, 2*(4), 321-340.

Woods, D. L., Wyma, J. M., Herron, T. J., & Yund, E. W. (2016a). Computerized analysis of verbal fluency: normative data and the effects of repeated testing, simulated malingering, and traumatic brain injury. *PLoS One, 11*(12), e0166439.

Woods, D. L., Wyma, J. M., Herron, T. J., & Yund, E. W. (2016b). A computerized test of design fluency. *PLoS One, 11*(5), e0153952.

World Health Organization. (2018). *International statistical classification of diseases and related health problems* (11th revision). Retrieved from http://www.who.int/classifications/icd/en/

Zarino, B., Crespi, M., Launi, M., & Casarotti, A. (2014). A new standardization of semantic verbal fluency test. *Neurological Sciences, 35*(9), 1405-1411.

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

Ziegler, J. C., Bertrand, D., Tóth, D., Csépe, V., Reis, A., Faísca, L., Saine, N., Lyytinen, H., Vaessen, A., & Blomert, L. (2010). Orthographic depth and its impact on universal predictors of reading: A cross-language investigation. *Psychological Science, 21*(4), 551-559.

Footnotes

- [1]‘αναπτυξιακή δυσλεξία’, ‘ειδική διαταραχή της ανάγνωσης’, ‘ειδική μαθησιακή διαταραχή (δυσλεξία)’, ‘ενδείξεις ειδικής μαθησιακής δυσκολίας (δυσλεξία)’, ‘ενδείξεις δυσλεξίας’, ‘ειδικές δυσκολίες στη μάθηση (δυσλεξία)’, ‘μαθησιακά κενά με ενδείξεις δυσλεξίας, ‘διαταραχή της ανάγνωσης και του συλλαβισμού’, ‘μαθησιακές δυσκολίες εξαιτίας φτωχής αποκωδικοποίησης, φτωχής αναγνωστικής ευχέρειας, ή φτωχής φωνολογικής επίδοσης, αργής αναγνωστικής ταχύτητας, ή/και ορθογραφικών λαθών’, ‘μαθησιακές δυσκολίες (χαρακτηριστικά ειδικού τύπου εστιασμένα στην ανάγνωση και γραφή)’, ‘μαθησιακές δυσκολίες στο γραπτό λόγο και στην ανάγνωση’.

- [2]‘ειδική γλωσσική διαταραχή’, ‘διαταραχή στη γλωσσική έκφραση και ως εκ τούτου γενικευμένες μαθησιακές δυσκολίες’, ‘μαθησιακές δυσκολίες στο πλαίσιο διαταραχής γλωσσικής έκφρασης’, ‘γλωσσική διαταραχή’, ‘διαταραχή στην έκφραση της γλώσσας’, ‘διαταραχή γλωσσικής έκφρασης’, ‘διαταραχή στην έκφραση και στην αντίληψη της γλώσσας’, ‘δυσκολία στο γραπτό και προφορικό λόγο’, ‘δυσκολίες στην ανάπτυξη του προφορικού λόγου (μέτρια εκφορά και απόδοση του λόγου, σχετικά φτωχό λεξιλόγιο)’, ‘δυσκολίες στην ανάπτυξη του προφορικού και γραπτού λόγου’, ‘ελλείψεις στο λεξιλόγιο, δυσκολίες στη σύνταξη και απόδοση του προφορικού και γραπτού λόγου’, ‘δυσκολίες στην ανάπτυξη των μαθησιακών δεξιοτήτων, του προφορικού λόγου και του φωνολογικού συστήματος’, ‘προβλήματα στη γλώσσα και στο λόγο’, ‘μικτή γλωσσική διαταραχή προσληπτικού και εκφραστικού τύπου’.

Appendix

A. Information sheet

Αγαπητοί γονείς,

Θα θέλαμε να σας ενημερώσουμε σχετικά με μία έρευνα που πρόκειται να διεξαχθεί στο πλαίσιο εκπόνησης διδακτορικής διατριβής. Η έρευνα θα πραγματοποιηθεί μεταξύ άλλων και στο σχολείο φοίτησης του παιδιού σας, σε συνεργασία με το UCL Institute of Education, University College London του Ηνωμένου Βασιλείου, και για αυτό το σκοπό θα θέλαμε να ζητήσουμε τη γραπτή συγκατάθεσή σας προκειμένου να συμμετάσχει το παιδί σας στην έρευνα. Για το σκοπό της έρευνας χρησιμοποιούμε μια σειρά δοκιμασιών που αξιολογούν την ανάπτυξη των γλωσσικών ικανοτήτων των παιδιών, συμπεριλαμβανομένων γνωστικών, αναγνωστικών, ορθογραφικών, φωνολογικών και λεξιλογικών δοκιμασιών. Η διάρκεια της δοκιμασίας είναι περίπου μια ώρα και μισή (χωρίς τα διαλείμματα) για κάθε μαθητή. Η δοκιμασία πραγματοποιείται σε ήσυχο χώρο εντός του σχολικού κτιρίου (ή ακόμα και στο σπίτι του μαθητή εάν αυτή είναι η προτίμηση των γονέων), με διαλείμματα να παρέχονται εάν χρειαστεί.

Βάσει των αρχών δεοντολογίας του Ελσίνκι (Declaration of Helsinki's; World Health Organization, 2011), τα ονόματα των παιδιών που θα συμμετάσχουν στην έρευνα δε θα γνωστοποιηθούν κατά τη δημοσίευση της έρευνας και φυσικά θα παραμείνουν απόρρητα. Όλα τα δεδομένα θα παραμείνουν απόρρητα επίσης. Η συμμετοχή των παιδιών είναι εθελοντική και μπορούν να αποχωρήσουν από την έρευνα ανά πάσα στιγμή και για οποιοδήποτε λόγο. Θα θέλαμε να ζητήσουμε τη γραπτή συγκατάθεσή σας για τη συμμετοχή του παιδιού σας στην έρευνα, καθώς και τη συμπλήρωση ενός ερωτηματολογίου που θα μας βοηθήσει να επιλέξουμε τα παιδιά που θα πάρουν μέρος στην έρευνα. Τέλος, θα θέλαμε να σας ενημερώσουμε ότι το ερωτηματολόγιο καθώς και το σύνολο του ερευνητικού σχεδιασμού έχει λάβει έγκριση από την Επιτροπή του University College London που εγκρίνει τις έρευνες φοιτητών σε ακαδημαϊκό πλαίσιο (Faculty Research Ethics Committee of the UCL Institute of Education), καθώς και από την αρμόδια υπηρεσία του Υπουργείου Παιδείας, Έρευνας και Θρησκευμάτων που εγκρίνει επιστημονικές έρευνες στα Ελληνικά δημόσια σχολεία.

Στόχος μας είναι η άψογη συνεργασία με γονείς και μαθητές, και επιθυμούμε αυτή η συνεργασία να είναι τόσο ευχάριστη για τους μαθητές όσο και κερδοφόρα. Θα είμαστε στη διάθεσή σας για οποιοδήποτε περαιτέρω πληροφορία χρειαστείτε. Σας ευχαριστούμε πολύ για τη συνεργασία!

Με εκτίμηση,

[redacted]

[redacted]

B. Written consent

Συγκατάθεση γονέα

Ημερομηνία …………………….

Εγώ, ο γονέας

Όνομα ………………………..

Επίθετο ………………………..

Τηλέφωνο επικοινωνίας …………………………..

δέχομαι το παιδί μου

Όνομα …………………..

Επίθετο ……………………..

Ημερομηνία γέννησης ………………………..

Τάξη ………………………..

να συμμετάσχει στην έρευνα

Υπογραφή

………………………

C. Questionnaire

Αγαπητέ γονέα,

Χρειάζεται να συλλέξω κάποιες επιπλέον πληροφορίες σχετικά με το παιδί σας. Στη βάση των απαντήσεών σας, θα προσπαθήσω να σχηματίσω μια ομάδα παιδιών που θα συμμετάσχουν στην έρευνα. Αν υπάρχει κάποια ερώτηση που δε θέλετε να απαντήσετε παρακαλώ πηγαίνετε στην επόμενη ερώτηση. **Παρακαλώ διαβάσετε προσεκτικά τις ερωτήσεις και κυκλώσετε τις απαντήσεις σας με ΝΑΙ ή ΟΧΙ**. Όλες οι πληροφορίες που δίνονται θα είναι εμπιστευτικές. Σας ευχαριστούμε για τη συνεργασία!

Μιλάει το παιδί σας τα ελληνικά από τη γέννησή του;	ΝΑΙ	ΟΧΙ
Μιλάει το παιδί σας κάποια άλλη γλώσσα πέραν των ελληνικών από τη γέννησή του;	ΝΑΙ	ΟΧΙ
Φοίτησε το παιδί σας κανονικά στο ελληνικό δημοτικό σχολείο μέχρι τώρα (δηλαδή, εγγράφηκε στην Ά τάξη του ελληνικού δημοτικού σχολείου και έκτοτε φοιτά στο ελληνικό δημοτικό σχολείο);	ΝΑΙ	ΟΧΙ
Υπήρξε κάποια περίοδος στη σχολική ζωή του παιδιού σας που να απουσίασε από το σχολείο για 3 μήνες ή παραπάνω;	ΝΑΙ	ΟΧΙ
Έπασχε το παιδί σας από κάποια νευρολογική διαταραχή (π.χ. επιληψία);	ΝΑΙ	ΟΧΙ
Λαμβάνει το παιδί σας φαρμακευτική αγωγή για κάποια νευρολογική, ψυχιατρική, συμπεριφορική, ή συναισθηματική διαταραχή;	ΝΑΙ	ΟΧΙ
Έχει το παιδί σας ιστορικό επαναλαμβανομένων επεισοδίων μέσης ωτίτιδας ή γνωστά επεισόδια μέσης ωτίτιδας μέσα στην περίοδο των προηγούμενων 12 μηνών;	ΝΑΙ	ΟΧΙ
Έχει το παιδί σας κάποια ανεπίλυτα προβλήματα όρασης (δηλαδή, όχι απλά προβλήματα που επιλύονται με γυαλιά οράσεως);	ΝΑΙ	ΟΧΙ
Έχει λάβει το παιδί σας κάποια επίσημη διάγνωση για μια από τις διαγνώσεις που φέρουν τους ακόλουθους όρους: αναπτυξιακή δυσλεξία, ειδική γλωσσική διαταραχή, διαταραχή ελλειμματικής προσοχής/υπερκινητικότητας, ή αυτισμός;	ΝΑΙ	ΟΧΙ

D. Examples of semantic and phonological coding of responses

An example of the semantic coding for the category of 'foods'

Responses	Correct/Incorrect	Switches	Clusters	Cluster size
Milk	1		Dairy	2
Cheese	1			
Legumes	1	1	Legumes	4
Lentils	1			
Beans	1			
Chickpeas	1			
Fish	1	1		
Meat	1	1		
Bread	1	1		
Yogurt	1	1	Dairy	2
Cheese	0			
Vegetables	1	1	Vegetables	6
Lettuce	1			
Tomato	1			
Cucumber	1			
Onion	1			
Cabbage	1			
Total responses	17/1	6	4	
Cluster size				3.5

An example of the phonological coding of the letter 'chi'

Responses	Correct/Incorrect	Switches	Clusters	Cluster size
Χριστίνα	1		Initial four phonemes	4
Χρυσόψαρο	1			
Χρυσός	1			
Χρυσόστομος	1			
Χελιδόνι	1	1		
Χαρακτήρας	1	1	Initial three phonemes	2
Χάρτης	1			
Χολαίνω	1	1		
Χείλη	1	1		
Χρήση	1	1	Initial three phonemes	2
Χρήμα	1			
Χορός	1	1		
Χειροκρότημα	1	1		
Χαλαρός	1	1	Initial two phonemes	4
Χαλασμένος	1			
Χάρτινος	1			
Χαραγμένος	1			
Total responses	17/0	8	4	
Cluster size				3

E. Examples of design fluency coding for the structured and the random array