Intervention for a lexical reading and spelling difficulty in two Greek-speaking primary age children

Aris R. Terzopoulos (MA)[1,2] ,

Georgia Z. Niolaki (PhD)[1,3],

Jackie Masterson (PhD)[1]

[1]Institute of Education, University College London, 20 Bedford Way, London WC1H 0AL j.masterson@ioe.ac.uk

[2] University of Dundee, Nethergate, Dundee DD1 4HN arterzopoulos@dundee.ac.uk

[3] Coventry University, Priory St, Coventry CV1 5FB ab8542@coventry.ac.uk

'Corresponding author'

Georgia Z. Niolaki (PhD),

Coventry University, Priory St, Coventry CV1 5FB ab8542@coventry.ac.uk

Running head: Lexical intervention targeting spelling

Keywords: dysgraphia; dual-route model; spelling intervention

Abstract

An intervention study was carried out with two nine-year-old Greek-speaking dyslexic children. Both children were slow in reading single words and text and had difficulty in spelling irregularly spelled words. One child was also poor in nonword reading. Intervention focused on spelling in a whole-word training using a flashcard technique that had previously been found to be effective with English-speaking children. Post-intervention assessments conducted immediately at the end of the intervention, one month later and then five months later showed a significant improvement in spelling of treated words that was sustained over time. In addition, both children showed generalisation of improvement to untrained words and an increase in scores in a standardised spelling assessment. The findings support the effectiveness of theoretically-based targeted intervention for literacy difficulties.

Word count: 8605 words

Running head: Lexical intervention targeting spelling

Keywords: dysgraphia; dual-route model; spelling intervention

Introduction

Developmental dyslexia and dysgraphia are reading and spelling disorders in children that can persist into adulthood (Gerber, 2012). They are found in languages other than English, although the vast majority of research with dyslexic and dysgraphic participants has been carried out in English. The manifestation of symptoms has been shown to be related to language characteristics, including the consistency of letter-sound relationships (e.g., Gupta & Jamal, 2007; Hanley, Masterson, Spencer, & Evans, 2004; Landerl, Ramus, Moll, & Lyytinen, 2013). In the present study, we report an intervention conducted with two monolingual Greek-speaking children with reading and spelling difficulty. Although intervention studies can assist in providing information about causes of literacy difficulties (see for a review Nickels, Rapp, & Kohnen, 2015) only a handful have been conducted with children learning transparent orthographies, such as Greek (Niolaki, Masterson, & Terzopoulos, 2014).

The theoretical framework adopted in the present study involved dual route (DR) models of reading and spelling (e.g., Barry, 1994; Coltheart, 1981). These have come to be used extensively for single case and case series reports of literacy difficulties in developmental (e.g., Broom & Doctor, 1995a, 1995b; Brunsdon, Coltheart, & Nickels, 2005; Rowse & Wilshire, 2007) and acquired cases (Rapp & Kane, 2002; Schmalzl & Nickels, 2006). According to the DR models, two sets of processes are used by competent readers and spellers. Whole-word lexical processes deal effectively with irregular or exception words (such as *<vehicle> or <εκκλησία>: /eklisia/ (church)*) and familiar words, using a store of lexical orthographic units. Sublexical processes deal effectively with novel printed letter strings and low-frequency regular words, using stored phoneme-grapheme correspondence rules.

The Greek writing system is considered to be transparent for reading, with almost one-to-one phoneme-grapheme correspondences; for spelling though, the situation is rather different, mainly because although pronunciation has changed from the ancient years, spelling has not followed the changes in pronunciation. For the vowel phoneme /i/ there are six different graphemes <η, ι, υ, οι, ει, υι>, for /o/ there are two <ο, ω>, and for /e/ there are two <ε, αι>. Some of these phoneme-grapheme correspondences are rule-governed, such as the fact that verbs end with <ω> and nouns with <ο>, but there are others where spellings are not rule-governed, and only the etymology of the word can provide an explanation (e.g., <κυνηγός> /kiniyos/ (the hunter)). By the end of Grade 3 children have mastered the majority of morpho-syntactic rules which help them spell consistent words and high-frequency irregular words (Harris & Giannouli, 1999).

Different subtypes of developmental dyslexia (phonological, surface, mixed dyslexia) have been identified in the past for the English language (e.g., Manis et al., 1996; Stanovich, Siegel & Gottardo, 1997), and similarly for Greek (Douklias et al., 2009; Niolaki et al., 2014), using the regression-outlier technique, previously employed by Castles and Coltheart (1993). In addition, in a study with adult Greek-speaking developmental dyslexics, Sotiropoulos and Hanley (2017a&b) reported individuals with phonological, surface and mixed dyslexias. In the present study, we report an intervention study with two of the children who took part in the Niolaki et al. (2014) group study. Both children had a lexical spelling deficit. A training programme that focused on improving whole-word lexical skills was devised, with rationale outlined in the *Background Information* section.

Intervention studies

Snowling (1987, p. 147) stated that ‘strangely, the literature on dyslexia has focused more upon its associated factors than upon its remediation’. Nearly thirty years later Griffiths and Stuart (2013) noted that there was a significant lack of single-case intervention studies,

although these are greatly needed to inform support for students with severe and prolonged literacy difficulties. The importance of single-case intervention studies derives from the fact that they are tailored to the child's specific needs. Thus, they can provide a test for the effectiveness of the intervention, controlling for variables not easily controllable in correlational or longitudinal studies. In addition, if they are theory-based the results can be used to evaluate and refine models of cognition (Howard, Best, & Nickels, 2015; Nickels et al., 2015).

We based the lexical intervention in the present study on programmes reported in previous research with English-speaking children. Broom and Doctor (1995a) carried out an intervention study with an eleven-year-old boy, DF, with surface dyslexia. The training lasted for twenty-four sessions and involved a simultaneous oral spelling method, previously used by Bradley (1981). Specifically, DF named the letters in a word at the same time as he wrote them. A discussion of the word's meaning was included and DF wrote a sentence highlighting the word's meaning in a notebook. Items misspelt at the following session were retrained. The researchers found improvement for trained words, but not untrained, for both spelling and reading.

Brunsdon et al. (2005) carried out an intervention study with a twelve-year-old child, MC, with surface dysgraphia. The spelling intervention targeted the lexical route, using a flashcard technique, which had been successfully employed with acquired surface dysgraphic patients. MC's irregular word spelling improved following a four-week training that involved irregular words. The researchers reported that both trained and untrained irregular words improved over the course of the intervention and many of the untrained words showed a gradual increase in degree of similarity to the correct spelling. Brundson et al. tested the efficacy of using mnemonics as part of the intervention, but the findings did not indicate that

this was more effective than using flashcards without mnemonics. In the present study, we decided to use the flashcard technique without mnemonics.

Kohnen, Nickels, Coltheart, and Brunsdon (2008) followed up the study of Brunsdon et al. in an intervention study with a nine-year-old child with surface dysgraphia. The researchers used the same intervention programme as Brunsdon et al. with the aim of investigating the nature of treatment generalisation. Improvement was again found for both treated and untreated irregular words. Results revealed that untreated words were more likely to show improvement if they had many orthographic neighbours and if they were of high frequency.

The current investigation

A lexically-based spelling intervention was employed in the present study for the two children since the assessments conducted by Niolaki et al. (2014), for which results are given in the next section, indicated that they had difficulty establishing lexical representations for reading and spelling. A flashcard-with-repeated-spelling programme was employed for establishing and strengthening orthographic entries, after Broom and Doctor (1995a), Brunsdon et al. (2005) and Kohnen, Nickels, Coltheart and Brunsdon (2008). The aim of the study was to test the effectiveness of the intervention in a transparent orthography and in order to investigate whether generalisation to untrained words would be observed. Next, we present the background information relevant to the two children. Full details of the tests conducted of reading, spelling and literacy-related processes can be found in Niolaki et al. (2014).

Discussion of the profiles of the case study children and rationale for the type of intervention chosen

The children were TN and OE, both monolingual Greek speakers assessed and diagnosed by regional school education services. TN was a girl aged 9;00 and OE a boy aged 9;08. In Table 1 we present results in background, literacy and cognitive assessments for both case study children and their comparison groups. All tests reported are described to have high reliability and validity (ranging between α =.74-.96). Both children were found to have age appropriate non-verbal reasoning and receptive vocabulary. Text reading fluency was found to be below average for both children. Single word reading latencies were also found to be slow, indicating impaired lexical reading processes. Nonword reading latencies were slower than in typical readers in the case of TN but not OE. This suggests that TN also had a sublexical reading difficulty. OE was also found to be impaired in nonword reading accuracy pointing towards sublexical reading difficulties. This has been associated in the past with a phonological deficit and, indeed, OE was impaired in the phonological awareness task of spoonerisms (Snowling, 1987). For spelling, TN and OE both exhibited poor accuracy with irregular words, and their errors were largely phonologically appropriate. In contrast, non-word spelling was unimpaired indicating, a selective lexical spelling difficulty. In summary then, OE and TN had both lexical and sublexical reading difficulties as well as lexical spelling difficulties.

In cognitive assessments both children were found to have a deficit in letter report tasks, which in the past have been associated with lexical reading and spelling difficulties. The deficit has been interpreted as a difficulty with parallel simultaneous processing of letters during reading (see Valdois et al., 2003), leading to a reliance on small orthographic units in reading, and thereby problems acquiring an orthographic lexicon. Further assessments revealed that OE was also significantly worse than comparison children in visual memory for abstract designs. A deficit of visual memory has been associated with developmental surface

dyslexia and dysgraphia in some cases (see Goulandris & Snowling, 1991), and has been interpreted as a difficulty acquiring an orthographic lexicon.

(Table 1 about here)

This study focussed on a training programme targeting the children's lexical deficit which was the most severe difficulty revealed in the initial assessment. It was decided to target spelling in the intervention programme since previous studies targeting this process have proved to be effective and also they have indicated that intervention for spelling results in improvement in both spelling and reading (e.g., Kohnen et al., 2008) while the reverse is not necessarily found (Caravolas et al., 2001). The focus of the intervention fitted well with the aspirations of the families of TN and OE. The parents of both children had expressed concern, particularly about their spelling performance. A whole-word based flashcard technique was employed for establishing and strengthening orthographic entries, after Broom and Doctor (1995a), Brunsdon et al. (2005) and Kohnen et al. (2008).

Based on previous research and the scarcity of single-case intervention studies on dyslexia in transparent orthographies we aimed to address the following research questions:

1) Would the lexical intervention be effective in a transparent orthography?
2) Would generalisation to untrained words be observed?
3) Would improvement in reading be found?

Intervention study

Method

Pre-intervention assessment

Two pre-intervention baseline assessments involving spelling to dictation were conducted with TN and OE approximately two months following the assessments reported in

the previous section. The baseline assessments were spaced two months apart and were carried out with a set of 90 irregularly spelled words on both occasions. The words consisted of 30 high frequency, 30 medium frequency and 30 low-frequency items. Frequency counts were obtained from HelexKids (Terzopoulos, Duncan, Wilson, Niolaki & Masterson, 2016), a database compiled from textbooks used in Greek primary education. Mean (per million) frequency for the high-frequency words was 452.4 (*SD*=182.2) and mean number of letters was 5.8 (*SD*=1.9). For the medium frequency words, mean frequency was 172.7 (*SD*=64.2) and mean number of letters 6.2 (*SD*=1.6). For the low-frequency words, mean frequency was 41.5 (*SD*=2.8) and mean number of letters was 6.1 (*SD*= 2). Half the items in each frequency band were short, and half were long (words were 3 to 10 letters long). A list of the words and their associated psycholinguistic variables in terms of printed word frequency, word length and orthographic neighbourhood size is given in Appendix A. Spelling accuracy for the 90 irregular words at the two baseline assessments for TN and OE is reported in Table 2.

Two groups of children, who were all typically developing readers and spellers, were recruited to constitute same age comparison groups for TN and OE. They were different children to those recruited in the Niolaki et al. (2014) study. The comparison group for TN comprised 8 boys and 12 girls from Grade 3 (with mean age 9;05, *SD* 0;03, range 8;07-9;05,), and for OE the group comprised 5 boys and 11 girls from Grade 4 (with mean age 9;06, *SD* 0;02, range 9;03 -9;11). The comparison groups were assessed in spelling the irregular word list at the same time that TN and OE completed the first baseline assessment and then when the delayed post-test assessment was conducted, approximately eleven months later (Post-intervention T3). The reason for carrying out the testing twice with the comparison children was to see whether an improvement in spelling scores might be expected over the time that TN and OE were involved in the intervention. In that way we can determine if gains are due to treatment rather than maturation, practice or participation in the intervention (Howard et

al., 2015).

In order to confirm that the comparison groups comprised typically developing learners we assessed them in the same standardised measures of reading and spelling that were employed with the case study children. For spelling the Mouzaki et al. (2007) items, TN's comparison group standard score was 103.3, *SD*=13.6 (range=90-125), and in reading accuracy (using Test A from Panteliadou & Antoniou 2008) it was 108.3, *SD*=16.5 (range=92-125). For OE's comparison group, in spelling the group's standard score was 103, *SD*=14.04 (range=90-118) and in reading accuracy it was 108, *SD*=15.39 (range=95-125).

The results for TN and OE and the comparison groups in spelling the 90 irregular words are given in Table 2. They show that the case study children were considerably impaired in spelling the irregular words at baseline testing. There was no evidence of improvement in scores from the first to the second baseline for OE (he spelled 14/90 words correct on both occasions), and for TN the increase from 20/90 correct at the first baseline to 24/90 correct at the second was not significant (χ^2 (1) = 0.64, *p*=.42). In addition, the results revealed that there was no improvement in spelling scores for the comparison children over the eleven-month period (comparison group for TN: *t*(38)=.21, *p*=.89, comparison group for OE: *t*(30)=.38, *p*=.70).

Qualitative analysis revealed that spelling errors of the case study children were predominantly phonologically appropriate, (TN: 100%, OE 91.1%).

(Table 2 about here)

Selection of words for the intervention

Based on the baseline testing results for TN and OE with the 90-word list, items were selected for the intervention. Words that were misspelt at both baseline assessments (N=61 for TN and N=71 for OE) were included. These were randomly assigned to trained and untrained word sets using the =rand() command in excel, and we ascertained that

psycholinguistic variables did not differ across the two sets for each child (analysis with t-tests across the wordsets for TN revealed: word frequency $t(88)=.26$, $p=.788$, orthographic neighbourhood size $t(88)=.15$, $p=.878$, number of letters $t(88)=.66$, $p=.512$, number of phonemes $t(88)=1.06$, $p=.289$, number of syllables $t(88) =.53$, $p=.596$, and for OE: word frequency $t(88)=.99$, $p=.324$, orthographic neighbourhood size $t(88)=.50$, $p=.615$, number of letters $t(88)=1.01$, $p=.315$, number of phonemes $t(88)=1.4$, $p=.152$, number of syllables $t(88)=.87$, $p=.386$). For TN, 36 words were in the trained set and 25 in the untrained set. For OE 36 words were in the trained set and 35 in the untrained set.

Trained words were divided into six equal sized sets, with one set for each of six weekly intervention sessions. The words sets did not differ across sessions for printed frequency, orthographic neighbourhood size, number of letters, number of phonemes or number of syllables (analysis with Kruskal-Wallis test across the wordsets for TN revealed: frequency $\chi^2(5)=.64$, $p=.986$, neighbourhood size $\chi^2(5)=9.96$, $p=.076$, letters $\chi^2(5)=.69$, $p=.837$, phonemes $\chi^2(5)=1.4$, $p=.914$, syllables $\chi^2(5) =2.3$, $p=.794$, and across the word sets for OE: frequency $\chi^2(5)=2.5$, $p=.772$, neighbourhood size $\chi^2(5)=2.8$, $p=.720$, letters $\chi^2(5)=3.1$, $p=.684$, phonemes $\chi^2(5)=2.1$, $p=.842$, syllables $\chi^2(5)=3.2$, $p=.656$).

Intervention procedure

At each of the six weekly intervention sessions, a new set of words was introduced. Each new word was written in large letters on an A4 card, and the child was asked to trace it with their finger. Then a discussion of the word's meaning followed in an attempt to connect the orthographic form with pre-existing semantic knowledge (Ouellette, 2010). Children were asked to copy the word and, after a ten-second delay, write the word (this time the word was not in view). In the case of an error, children were asked to look at the correct spelling of the word again, and the whole process was repeated.

Following each session children practised the items at home daily with their parents. Practice lasted approximately 20 minutes per day. Each week a letter with the new words and an outline of the procedure was given to the parents. The procedure involved the following steps: the parent dictated the word to the child, the child wrote the word, feedback was provided (i.e., 'yes, it is correct' or 'no, you have made an error, let's try again'), in case of an error the child looked at the flashcard, the parent covered the word and after a 5 second delay asked his/her child to write the word from memory, if it was correct they moved to the next item. A re-test of previous week's trained items took place at each weekly session. Children were not always 100% correct, and the erroneously spelled words were not retrained.

Results

Post-intervention assessments were conducted for spelling the trained and untrained irregular words. Assessment of the untrained words served as a test for any generalisation of improvement to untreated items. A further measure that we used in looking for any improvement as a result of the intervention comprised re-administration of the Mouzaki et al. (2007) standardised spelling assessment that had been used by Niolaki et al. (2014), as reported in the *Background Information* section. We also re-assessed reading and spelling of the words and nonwords that had been administered in the study of Niolaki et al. (2014) to assess lexical and sublexical processes, as reported in the *Background Information* section. The items were from an earlier study of spelling development in Greek-speaking children by Loizidou-Ieridou et al. (2009) and consisted of 40 words and 40 nonwords. Twenty of the words were regular for spelling, and 20 were irregular for spelling. None of the irregular for spelling words in the Loizidou-Ieridou et al. set appeared in the training words for TN or OE, thus assessment of spelling with these words comprised a further test for generalisation of improvement to untrained items.

Assessment of the specificity of any effects of the intervention was carried out by testing reading comprehension pre- and post-intervention using the Test Alpha (Panteliadou & Antoniou, 2008). We also obtained a measure of arithmetic skill pre- and post-intervention, using the arithmetic subtest from the WISC-IV (Georgas, Paraskevopoulos, Bezevengis, & Giannitsas, 1997) in order to ascertain that any improvements we might observe were not due to an improvement in academic abilities per se.

1. Testing for improvement in trained and untrained words

Words in both the trained and untrained wordsets were those which had been misspelt by the children at both baseline assessments. Post-intervention assessments were conducted at three timepoints: immediately following the end of training (post-test Time 1, T1), one month after T1 (post-test Time 2, T2) and approximately five months after T2 (post-test Time 3, T3). The results are given in Table 2. McNemar's tests were used to look for the significance of any changes across the timepoints. For both children, all contrasts were conducted with Baseline 2 performance as it can carry over any test-retest effects (from Baseline 1 to Baseline 2) and it is the most recent pre-intervention assessment.

i. Trained words

Testing for change in accuracy across time points for TN revealed that between Baseline 2 and T1 there was a significant increase ($\chi^2(1)=30.03$, $p<.001$), between T1 and T2 there was a significant decrease ($\chi^2(1)=6.7$, $p=.009$), and between T2 and T3 there was no further significant change ($p=.22$). Between Baseline 2 and T2, the improvement in accuracy was significant ($\chi^2(1)=20.05$, $p<.001$). Between Baseline 2 and T3, the improvement in accuracy was also significant ($\chi^2(1)=15.6$, $p<.001$). For OE, similarly, there was a significant increase in accuracy between Baseline 2 and T1 ($\chi^2(1)= 14.1$, $p=.0002$), between T1 and T2 there was a significant decrease ($\chi^2(1)= 4.08$, $p=.043$), and between T2 and T3 there was no

further significant change (p=.39). Between Baseline 2 and T2 the improvement in accuracy was significant (χ^2 (1)=6.12, p<.01). Finally, between Baseline 2 and T3 the improvement in accuracy was also significant (χ^2 (1)=7.03, p=.008).

We also examined whether the treatment effect could be differentiated from a possible regression to the mean. We followed the procedure used by Brunsdon et al. (2005). This involved calculating the improvement in accuracy from Baseline 1 to Baseline 2 for all the items in the 90-word list (i.e., the likely effect of regression to the mean) and comparing this with the improvement in the trained items from Baseline 2 to Post-test 1. For TN there was an improvement in accuracy between Baseline 1 and Baseline 2 of 4.4% for the entire set of words in the 90-word list, but the improvement in the trained items from Baseline 2 to T1 was 89%, and this change was significantly larger than that observed for all the items between the two baselines (χ^2(1)=21.7, p<.001). For OE, there was no increase in accuracy between Baseline 1 and Baseline 2 for the entire set of words in the 90-word list. However, OE showed an improvement between Baseline 2 and T1 of 44.4%, and this difference was significant (χ^2(1)=11.1, p<.001).

ii. Untrained words

For TN the increase in number correct between Baseline 2 and T1 was significant (χ^2(1)=4.2, p=.041), while the change from T1 to T2 and T2 to T3 was not significant (p=.1 and p=.1 respectively). Between Baseline 2 and T2 the improvement in accuracy was marginally significant (χ^2 (1)=3.2, p=.06). For OE, between Baseline 2 and T1 there was no improvement, but between Baseline 2 and T2 the improvement in accuracy was significant (χ^2 (1)=4.2, p=.03). Between T1 and T2, a significant improvement was also observed (χ^2(1)=4.2, p=.041). There was no change between T2 and T3 (p=1).

To try to establish whether improvement on the untrained words was a true treatment generalisation effect, rather than the result of regression to the mean, we again followed the procedure used by Brunsdon et al. (2005). This involved calculating the improvement in accuracy from Baseline 1 to Baseline 2 for all the items in the 90-word list (i.e., the likely effect of regression to the mean) and comparing this with the improvement in the untrained items from Baseline 2 to Post-test 1. For TN there was an improvement in accuracy between Baseline 1 and Baseline 2 of 4.4% for the entire set of words in the 90-word list, but the improvement in the untrained items from Baseline 2 to T1 was 24%, and this change was significantly larger than that observed for all the items between the two baselines ($\chi^2(1)=6.7$, $p=.009$). For OE, there was no increase in accuracy between Baseline 1 and Baseline 2 for the entire set of words in the 90-word list, or for the untrained words between Baseline 2 and T1. However, OE showed an improvement between Baseline 2 and T2 of 17.1%, which was significantly larger than the improvement for all the items between the two baselines ($\chi^2(1)=15.1$, $p<.001$).

iii. Bi-gram analysis

In addition to the binary analysis (correct/incorrect) used for the calculation of the effect of the intervention for the two children we used bi-gram analysis which is considered to be a more sensitive measure of improvement, capturing individual differences (Vaughn, Schumm & Gordon 1993; White & Haring, 1980; Apel & Masterson, 2001). Bi-gram analysis allows similarities between target and misspelt words to be explored. Specifically, a point is assigned for correct letter-pairs in sequences of letters and for the correct initial and final letters. For example, the word < χορα> /hora/:country has three bi-grams correct out of five because /o/ is spelt with the wrong grapheme. An analogous example in English is the misspelling <tran> for *train* which has four bi-grams correct out of six because <i> is

omitted. The number of bi-grams correct was calculated for the 90-item word list spelled at the second baseline and each post-test assessment for both TN and OE. This calculation provided the following bi-gram values: for TN the percentage bi-grams correct achieved at Baseline 2 was 81% and the post-tests was T1=91%, T2=88% and T3=91%, for OE the percentage bi-grams correct achieved at Baseline 2 was 72%, and the post-tests was T1=82%, T2=80% and T3=83%.

We used Wilcoxon tests to compare bigram accuracy at the different testing points. The difference was significant for both children between the different time points, indicating a steady increase in bigram accuracy (for TN, the difference between Baseline 2 and T1 was z=3.8, p<.001, r=.28, between Baseline 2 and T2 was z=2.8, p=.006, r=.21 and between Baseline 2 and T3 was z= 4.1, p<.001, r=.31; for OE, the difference between Baseline 2 and T1 was z=4.4, p<.001, r=.33, between Baseline 2 and T2 was z=4.2, p<.001, r=.31 and between Baseline 2 and T3 was z=4.5, p<.001, r=.34. These findings support our previous analysis using the binary scores (correct/incorrect). We also calculated the effect-size for bi-gram accuracy in order to determine the effectiveness of the intervention. We applied Kromey and Foster-Johnson's (1996) equation to determine the effect size of the standardised mean difference in bi-gram accuracy at the three post-test assessments:

$$\frac{X\ baseline\ 2 - X\ post-test\ 1}{SD\ post-test\ 1} = d$$

$$\frac{X\ baseline\ 2 - X\ post-test\ 2}{SD\ post-test\ 2} = d$$

$$\frac{X\ baseline\ 2 - X\ post-test\ 3}{SD\ post-test\ 3} = d$$

The formulas gave us medium effect sizes for both children, TN T1: d=.42, T2: d=.32, T3: d=.42 and OE T1: d=.39, T2: d=.28, T3: d=.40. The findings support the previous analyses indicating that the children improved as a result of the training.

iv. Further qualitative observations

We investigated which factors may have led to improvement in the untrained items exploring specifically orthographic and morphological similarities between trained and untrained items. Due to the characteristics of the items (all items selected were irregular words) with great variability in the root (orthographic patterns) and suffix spelling (morphological patterns), it is hard to quantify whether similarities between trained and untrained items led to an improvement in spelling for the untrained words. However, examination of the items could tentatively suggest that occurrence of low-frequency vowels in the root of the items could have given TN and OE new spelling strategies, such as that it is not always optimal to use the most frequent vowel grapheme when spelling a word. For example, the most frequent /o/ grapheme in Greek is <o> (Zipf frequency= 7.22), the least frequent is <ω> (Zipf frequency= 4.55), and the latter usually occurs at the end of a verb.

Teaching TN and OE words that have <ω> /o/ in the root could have helped them develop a new strategy that one should also use the least frequent vowel when spelling the word's root. Specifically, TN could have spelled correctly the word <χρώμα> /hroma/: colour because she was trained on the item <γλώσσα > /glosa/: tongue, which also has in the root the vowel <ω> /o/. Similarly, OE was trained on the items <χρώμα> /hroma/: colour, <μωρό> /moro/: baby and <χώρα> /hora/: country. These three items also have in the root the least frequent <ω> /o/, and they have the same number of syllables as <φωνή> /foni/: voice and <φωτιά> /fotia/ fire, which improved without training, while generalisation did not occur for words with three syllales (e.g., <ζωγράφος>, /zografos/: artist and <δωρεάν> /ðorean/: costless). Similar observations can be made for the other items OE improved in without receiving training.

For TN and OE, we also noticed that they improved in the spelling of suffixes, which could have occurred due to the training of similarly spelled suffixes. For example, TN

improved in spelling the untrained word <βοήθεια> /voithia/: help, and this could have occurred because she was trained in the items <συνέχεια> /sinexia/: continuity and <ενέργεια> /enerjia/: energy. Thus, orthographic similarity between trained and untrained items may enhance generalisation, as for TN, of the nine words she spelled correctly during the three post-test assessments, seven shared some orthographic similarities (in the root or the suffix) to the trained items. For OE, of the seven items spelled correctly during Post Test 2 and 3, three shared similar orthographic patterns to trained words. If we had used more items in the intervention, we might have been able to identify improvement ina greater number of items due to orthographic similarity.

Although our results are based on qualitative observations, they corroborate findings reported by Kohnen et al. (2008). The researchers investigated possible aetiologies for generalisation to untrained items. They were able to identify a robust orthographic similarity effect to untrained items, and they attributed this to the positive feedback from the grapheme level to the orthographic lexicon. Our observations could indicate that findings based on the English orthography can also apply to the more consistent Greek orthography and be used in designing future intervention studies.

We also carried out a correlational analysis involving the word characteristics (word frequency, orthographic neighbourhood size, number of phonemes, letters and syllables) and the untrained items which improved and those that did not. We found a strong association between neighbourhood size and improvement (utilised as a dichotomous variable) at Post-test 1 for TN and Post-test 2 for OE (ρ=.56, p=.002 & ρ=.35, p=.04, respectively*)*. This supports findings in Kohnen et al.'s study in which they suggested that items high in neighbourhood size predict generalisation, as feedback from the graphemic buffer will benefit words with many orthographic neighbours. Phoneme length was negatively associated (p=-.48, p=.013 & ρ=-.33, p=.047, respectively*)* to generalisation at the same post-test points for

both children. Word frequency was not found to relate to generalisation. This could be due to the small number of items used in the current study or to the characteristics of Greek orthography (see, for example, a discussion on the effect of word frequency in Greek in Loizidou-Ieridou et al., 2009).

2. Standardised spelling test performance

As reported in the *Background Information* section, Niolaki et al. (2014) recorded standardised scores in the Mouzaki et al. (2007) single word spelling test of 74 (Confidence Interval (CI) 95%=67-81) for both TN and OE. The test was re-administered for the present study after the intervention at T3. On this occasion, the standardised score for TN was 82 (CI 95%=75-89), and for OE it was 81 (CI 95%=74-88), indicating an increase to just within the average range for both children. Inspection of the confidence intervals indicate that the children's performance at baseline assessment ranged between well below average (>69) to below average (70-84), whereas during T3 assessment it was between below average and low average (85-89).

3. Reading and spelling of words and nonwords from Loizidou-Ieridou et al. (2009)

The items from Loizidou-Ieridou et al. were administered for reading and spelling at T2 and T3. The results are given in Table 3, together with the pre-intervention results and those for comparison children, as reported in Niolaki et al. (2014).

(Table 3 about here)

a. *Irregular word spelling*

Pre-intervention assessment had revealed that for spelling the irregular words (but not the regular words or nonwords) TN and OE were significantly less accurate than comparison children, as outlined in the *Background Information* section. At the post-intervention

assessments, irregular word spelling accuracy was not significantly different from the comparison group mean for TN on either testing occasion (T2, p=.34 and T3, p=.45). For OE at T2 accuracy was marginally significantly worse than that of the comparison group (t(9)=1.8, p=.047, r=0.51), but at T3 the difference was not significant.

b. *Word and nonword reading*

Pre-intervention assessment had indicated that for both TN and OE word reading latencies were significantly slower than those of the comparison children. This was also found to be the case after the intervention (TN: T2 t(78)=5.1, p<.001, r=0.50, T3 t(78)=6.1, p<.001, r=0.57; OE: T2 t(78)=7.6, p<.001, r=0.65, T3 t(78)=4.3, p<.001 r=0.43). For nonwords, latencies had been significantly slower than those of comparison children for TN prior to the intervention, and this was also the case at the post-intervention assessments (T2 t(78)=4.9, p<.0001, r=0.48, T3 t(78)=7.5, p<.001, r=0.64). For OE, nonword latencies had been reported to be slower, but not significantly slower than those of the comparison children prior to the intervention. At both post-intervention assessments OE was found to have significantly slower nonword latencies than the comparison children (T2 t(78)=8.3, p<.001, r=0.68, T3 (t(78)=4.6, p<.001, r=.46).

For OE, pre-intervention nonword reading accuracy was significantly worse than the mean for the comparison children. However, there was no significant difference at either T2 or T3. We are unable to ascertain whether the improvement in nonword reading accuracy for OE was due to the intervention or to other factors since we did not have multiple baseline measures for this task or comparison children's scores over time.

Testing for specificity of the effects of intervention

In order to investigate the specificity of the effects of the intervention, the arithmetic subtest from the WISC-IV battery (Georgas et al., 1997) and reading comprehension subtest from Test Alpha (Panteliadou & Antoniou, 2008) were administered at the outset of the intervention and the end of training. Neither child's results showed indication of improvement (arithmetic subtest TN pre-intervention standard score = 109, post-intervention standard score = 109, OE pre-intervention standard score = 89, post-intervention standard score = 89, for the reading comprehension test TN pre-intervention standard score = 112, post-intervention standard score = 112, OE pre-intervention standard score = 104, post-intervention standard score = 96).

General Discussion

We noted in the Introduction that the majority of studies, and in particular intervention studies, investigating literacy difficulties in children had been carried out with English speakers. Studies of Greek-speaking children are extremely rare, even though Greek is an interesting language in which to conduct such research due to its consistent grapheme-phoneme correspondences but inconsistent phoneme-grapheme correspondences. The present study involved an intervention study conducted with two nine-year-old Greek-speaking dyslexic children, TN and OE.

Testing previously carried out by Niolaki et al. (2014) had revealed that OE's predominant impairment was in lexical processing, as indicated by slow word reading and poor, irregular word spelling, and for TN this seemed to be the major deficit. A training that targeted lexical processes was therefore conducted since previous single case studies that have targeted the impaired process have proved successful (e.g., Broom & Doctor, 1995a, b; Brunsdon, Hannan, Nickels, & Coltheart, 2002; Kohnen et al., 2008; Niolaki & Masterson, 2013). The focus of the training was spelling, and a flashcard technique was employed, which has been used successfully in the past with children of similar ages to TN and OE (see, for

example, Brunsdon et al., 2005). The intervention aimed at strengthening orthographic representations, achieved through repeated exposure to correct spellings and delayed copying, according to Rapp and Kane (2002). We discuss the results of the intervention in terms of the pre-intervention reading and spelling profiles for the children.

Irregular word spelling

Post-intervention improvement was observed for both children for irregular word spelling. Although, we must also highlight that for OE but not for NT, a late generalisation effect only after T2, not immediately after T1, was found. The assessments indicated that the significant gains were sustained over time. Re-assessment of comparison children in spelling the irregular words from which the intervention items were selected revealed no significant change in performance, suggesting that the improvement in accuracy shown by TN and OE was due to the training and not general maturation or test-retest effects. In addition, bi-gram correct calculations indicated medium effect sizes, indicating the effectiveness of the spelling intervention and corroborating findings reported above.

Evidence of generalisation was observed, as the children's scores improved in a standardised spelling assessment, as well as in spelling the set of irregular words from Loizidou-Ieridou et al. (2009) that were not targeted in training. Improvement was also observed in the untrained irregular words that were matched to the training words. Broom and Doctor (1995a) did not find any evidence of generalisation of improvement in their study. This may be because in Broom and Doctor's study the researchers asked for oral report of the letter names during spelling, which might have improved recall of the letters sequence of treated words. This letter specific training might have hindered generalisation to larger letter chunks or even words. There are also a few other explanations proposed by the researchers, for example, treated and untreated words were not matched for psycholinguistic

properties, therefore comparisons between the two word-groups were not possible. In addition, they suggested that the items targeted were of low frequency; therefore, DF had less opportunities to utilise the new learning in other reading and spelling tasks. Our results are in line with those of Brunsdon et al. (2005) and Kohnen et al. (2008) who found improvement in spelling untrained words. This could be the case as words (trained and untrained) in our study were carefully matched on important psycholinguistic variables (frequency, neighbourhood size, phonemes, syllables) and words were selected which varied in frequency (high, medium and low). In addition, findings from our qualitative analysis of the items which improved indicated that orthographic similarities between trained and untrained items could have led to improvement. Some of the generalisation, could be also due to the children exploiting rule-based knowledge, for example in relation to the suffixes in <βοήθεια> /voithia/, <συνέχεια> /sinexia/, <ενέργεια> /enerjia/. Our results also support Kohnen et al.'s (2008) observation that items with many orthographic neighbours have a higher probability to improve after training. However we did not find an association between improvement and printed word frequency, and this could be due to the transparency of the orthography (see Loizidou-Ieridou et al., 2009 for a discussion of the association between printed word frequency and spelling in Greek-speaking children).

In the current study, by conducting detailed pre-intervention assessment, we were able to identify that TN and OE had a lexical spelling deficit (irregular word spelling difficulties) with better performance in regular word and nonword spelling. This knowledge of the children's strengths and weaknesses (Ellis, 2016) helped us tailor an appropriate intervention to their specific needs.

Word and nonword reading latencies

Word reading latencies were reported to be slow for both case study children prior to the intervention, and we found that this was unchanged at post-test. It might be expected that an improvement in lexical skills, as reflected in improvement in irregular word spelling and generalisation to untrained items, should also result in faster word reading latencies. The lack of improvement could be due to the focus in the training on examining and retaining word spellings. This may have led the children to be more attentive to letter sequences within words when reading. This issue seems worthy of investigation in a future study, involving a comparison of the effect on reading of different training techniques.

As well as word reading latencies, nonword reading latencies had been reported to be slow for TN prior to intervention, and we found that this was also the case at post-test. For OE, nonword latencies were slow but not significantly slower than those of comparison children prior to intervention; at T2 and T3 they were significantly slower. This change could be, as discussed above, a result of the focus of the intervention on analysis and retention of word spellings. The children did not practise reading the words trained in the spelling study. If our focus had been on reading then an improvement in reading latencies may have been observed. Alternatively, the change reported may have been due to fluctuation in measurement.

Nonword reading accuracy

OE had shown difficulty in both lexical and sublexical processes prior to the intervention that we carried out with him, as he also had impaired nonword reading accuracy. His accuracy in reading nonwords was observed to improve following the intervention. However, it is not possible to know whether the improvement was related to the intervention since we did not carry out multiple baseline testing for nonword reading and we did not collect relevant data from comparison children. It will be informative to systematically

investigate whether an improvement in lexical processes might generalise to non-word reading accuracy in future studies.

Conclusions and Educational Implications

According to Kohnen and Nickels (2010), an intervention is considered to be successful if improvement in scores on standardised tests is observed, as well as improvement in trained items. About the former, as noted above, we found improvement in standardised spelling test scores for TN and OE to just within the average range following the intervention. Van der Leij (2013) in a recent reflection on intervention studies targeting reading achievement stressed the importance of prolonged interventions for children at risk of dyslexia. It will be important to offer continued support to TN and OE for their literacy skills.

Limitations of the study include the fact that, it would have been informative to assess reading of the trained words, as in the study of Brunsdon et al. (2005), and also to carry out long-term follow-up assessments. With these caveats in mind, it can be argued that the intervention carried out with the two
Greek-
speaking children with overlapping but slightly different reading, and spelling difficulties was effective, since spelling of trained words improved significantly. Post-intervention follow-up testing showed that the improvement for the trained words was not stable since there was a significant decrease in accuracy for spelling the trained words at T2 testing for both children. This indicates that continuous help and support is necessary (see also Torgesen, 1998, 2002 for similar arguments for English-speaking children).

We found evidence of improvement for untrained words in our study, and it seems important to carry out further research to specify the conditions under which it is observed. The intervention described can easily be carried out in schools. It is not too time-consuming,

and the children were motivated to complete the programme and discussed its benefits at the end. According to a recent UK report of identification of literacy difficulties and interventions (Rose, 2009), targeted training, such as that in the present study, is vital for children with reading and spelling difficulties, and at the same time it does not contradict the philosophy of support and inclusion in the mainstream classroom (Norwich & Lewis, 2007; Reid, 2013). Overall, results from the pre- and post-test assessments revealed improvement in irregular word spelling as a result of the intervention. There was also an improvement in standardised spelling test results for both children. The children mentioned that the technique provided them with a strategy which could help them learn the spellings of difficult words more easily. Finally, our data speak against the suggestion that whole-word training can compromise phonics. If this were the case, nonword spelling and nonword reading accuracy should have decreased, but this was not the case (also see McArthur et al., 2015 for similar results). There were fluctuations with nonword reading latencies which warrant further investigation, but cannot be clearly linked to the training at this point in time. Overall, for a transparent language like Greek, focus on whole word training techniques seem an effective strategy for irregular word spelling and for older spellers who have grasped the alphabetic principle as was the case for TN and OE.

Acknowledgements:

We are grateful to the children for donating their time and to the Editor and Reviewers for their valuable comments. We would like to thank Saskia Kohnen for her valuable contribution to this manuscript and a second Anonymous Reviewer.

References

Apel, K., & Masterson, J. J. (2001). Theory-Guided Spelling Assessment and InterventionA Case Study. *Language, Speech, and Hearing Services in Schools*, 32(3), 182-195.

Barry, C. (1994). Spelling routes (or roots or rutes). In G.D.A. Brown & N.C. Ellis (Eds.), *Handbook of spelling: theory, process and intervention* (pp 27-49). Chichester: Wiley.

Bradley, L. (1981). The organization of motor patterns for spelling: An effective remedial strategy for backward. *Developmental Medicine and Child Neurology, 23*, 83-91.

Broom, Y.M. & Doctor, E. (1995a). Developmental surface dyslexia: A case study of the efficacy of a remediation programme. *Cognitive Neuropsychology, 12, (1),* 69-110.

Broom, Y.M. & Doctor, E. (1995b). Developmental phonological dyslexia: A case study of the efficacy of a remediation programme. *Cognitive Neuropsychology, 12, (7),* 725-766.

Brunsdon, R., Hannan, T.J., Nickels, L. & Coltheart, M. (2002). Successful treatment of sublexical reading deficits in a child with dyslexia of a mixed type. *Neuropsychological Rehabilitation, 12, (3),* 199-229.

Brunsdon, R., Coltheart, M., & Nickels, L. (2005). Treatment of irregular word spelling in developmental surface dysgraphia. *Cognitive Neuropsychology, 22, (2),* 213-251.

Caravolas, M., Hulme, C., & Snowling, M. J. (2001). The foundations of spelling ability: Evidence from a 3-year longitudinal study. *Journal of memory and language*, 45(4), 751-774.

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

Coltheart, M. (1981). Disorders of reading and their implications for models of normal reading. *Visible Language, 15,* 245-286.

Crawford, J. R., & Howell, D. C. (1998). Comparing an individual's test score against norms derived from small samples. *The Clinical Neuropsychologist, 12*, 482-486.

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

Dubois, M., Lafaye de Micheaux, P., Noël, M.P., & Valdois, S. (2007). Pre-orthographical constraints on visual word recognition: Evidence from a case study of developmental surface dyslexia. *Cognitive Neuropsychology*, 24, 1-38.

Ellis, A. W. (2016). *Reading, Writing and Dyslexia (Classic Edition): A Cognitive Analysis*. Psychology Press.

Georgas, D. I. Paraskevopoulos, I.., Bezevengis, & Giannitsas, N. (1997). *Guidelines for the Greek WISC III*. Athens: Ellinika Grammata.

Gerber, P. J. (2012). The Impact of Learning Disabilities on Adulthood A Review of the Evidenced-Based Literature for Research and Practice in Adult Education. *Journal of Learning Disabilities*, *45*(1), 31-46.

Goulandris, N. K., & Snowling, M. (1991). Visual memory deficits: a plausible cause of developmental dyslexia? Evidence from a single case study. *Cognitive Neuropsychology, 8*(2), 127-154.

Gupta, A., & Jamal, G. (2007). Reading strategies of bilingual normally progressing and dsyslexic readers in Hindi and English. *Applied Psycholinguistics, 28*(1), 47-68.

Hanley, J.R., Masterson, J., Spencer, L., & Evans, D. (2004). How long do the advantages of learning to read a transparent orthography last? An investigation of the reading skills and reading impairment of Welsh children at 10-years of age. *The Quarterly Journal of Experimental Psychology, 57*, 1393-1410.

Harris, M., & Giannouli, V. (1999). Learning to read and spell in Greek: the importance of letter knowledge and morphological awareness. In M. Harris & G. Hatano (Eds.),

Learning to read and write: A cross-linguistic perspective (pp. 51–69). Cambridge: University Press.

Howard, D., Best, W., & Nickels, L. (2015). Optimising the design of intervention studies: Critiques and ways forward. *Aphasiology*, 29(5), 526-562.

Kohnen, S., Nickels, N., Coltheart, M., & Brunsdon, R. (2008). Predicting generalisation in the training of irregular-word spelling: Treating lexical spelling deficits in a child. *Cognitive Neuropsychology*, *25*(3), 343–375.

Kohnen, S. & Nickels, L. (2010). Teaching children with developmental spelling difficulties in a one-to-one context. *Australasian Journal of Special Education*, *34*, 36-60.

Kromrey, J. D., & Foster-Johnson, L. (1996). Determining the efficacy of intervention: The use of effect sizes for data analysis in single-subject research. *The Journal of Experimental Education*, *65*(1), 73-93.

Landerl, K., Ramus, F., Moll, K., Lyytinen, H., et al. (2013). Predictors of developmental dyslexia in European orthographies with varying complexity. *Journal of Child Psychology and Psychiatry, 54*, 686-694.

Leij, A. (2013). Dyslexia and early intervention: what did we learn from the Dutch Dyslexia Programme?. *Dyslexia, 19*(4), 241-255.

Loizidou-Ieridou, N., Masterson, J., & Hanley, J.R. (2009). Spelling development in 6-11year-old Greek-speaking Cypriot children. *Journal of Research in Reading, 33*, 247-262.

McArthur, G., Castles, A., Kohnen, S., Larsen, L., Jones, K., Anandakumar, T., & Banales, E. (2015). Sight word and phonics training in children with dyslexia. *Journal of Learning Disabilities*, *48*(4), 391-407.

McDougall, P., Borowsky, R., MacKinnon, G. E., & Hymel, S. (2005). Process dissociation of sight vocabulary and phonetic decoding in reading: A new perspective on surface and phonological dyslexias. *Brain and Language, 92*, 185–203.

Mouzaki, A. Protopapas, A. Sideridis, P. & Simos, G. (2007). Psychometric properties of a new test of spelling achievement in Greek. *Epistimes tis Agogis, 1* 129-146.

Naglieri, J. A. (1985). *Matrix Analogies Test (short form) (MAT–SF)*. San Antonio: The Psychological Corporation.

Nickels, L., Rapp, B., & Kohnen, S. (2015). Challenges in the use of treatment to investigate cognition. *Cognitive Neuropsychology*

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.

Niolaki, G.Z., Terzopoulos, A. & Masterson, J. (2014). Varieties of developmental dyslexia in Greek children. *Writing Systems Research.* http://dx.doi.org/10.1080/17586801.2014.893862.

Norwich, B., & Lewis, A. (2007). How specialised is teaching children with disabilities and difficulties? *Journal of Curriculum Studies, 39*,127-150.

Panteliadou, S., & Antoniou, F. (2007). *Test Alpha (Reading Test)*. ΕΠΕΑΕΚ, ΥΠΕΠΘ. (Standardised test for reading disabilities in Greek).

Paraskevopoulos, I.N., Kalantzi-Azizi, A., & Giannitsas, N.D. (1999). *Αθηνά Τεστ διάγνωσης δυσκολιών μάθησης. (Athena Test: Diagnosis of Learning Difficulties).* Athens: Ellinika Grammata.

Peyrin, C., Lallier, M., Demonet, J., Pernet, C., Baciu, M., Le Bas, J. F., & Valdois, S. (2012). Neural dissociation of phonological and visual attention span disorders in

developmental dyslexia: FMRI evidence from two case reports. *Brain & Language, 120*, 381–394. doi:10.1016/j.bandl.2011.12.015

Rapp, B., & Kane, A. (2002). Remediation of deficits affecting different components of the spelling process. *Aphasiology, 16,* 439–454.

Reid, G. (2013). *Dyslexia and Inclusion. Classroom approaches for assessment, teaching and learning.* New York: Routledge.

Rose, J. (2009). Identifying and teaching children and young people with dyslexia and literacy difficulties, HMG UK, /publications/eOrderingDownload/00659-2009DOM-EN.pdf, accessed September 2013

Rowse, H.J., & Wilshire, C.E. (2007). Comparison of phonological and whole-word treatments for two contrasting cases of developmental dyslexia. *Cognitive Neuropsychology, 24*(8), 817-842.

Schmalzl, L. & Nickels, L. (2006). Treatment of irregular word spelling in acquired dysgraphia: Selective benefit from visual mnemonics, *Neuropsychological Rehabilitation, 16,* 1-37.

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), 34–49.

Sotiropoulos, A., & Hanley, J. R. (2017a). Developmental surface and phonological dyslexia in both Greek and English. *Cognition*, *168*, 205-216.

Sotiropoulos, A., & Hanley, J. R. (2017b). Lexical decision performance in developmental surface dysgraphia: Evidence for a unitary orthographic system that is used in both reading and spelling. *Cognitive Neuropsychology*, 1-19.

Snowling, M.J. (1987). *Dyslexia: A cognitive developmental perspective,* Oxford: Blackwell

Torgesen, J. K. (1998). Catch them before they fall. *American Educator, 22,* 32-41.

Torgesen, J. K. (2002). The prevention of reading difficulties. *Journal of school psychology*, *40*(1), 7-26.

Valdois, S., Bosse, M.L., Ans B., Zorman, M., Carbonnel, S., David, D. & Pellat, J. (2003). Phonological and visual processing deficits are dissociated in developmental dyslexia: Evidence from two case studies. Reading and Writing, 16, 543-572.

Vaughn, S., Schumm, J. S., & Gordon, J. (1993). Which motoric condition is most effective for teaching spelling to students with and without learning disabilities?. *Journal of Learning Disabilities*, *26*(3), 191-198.

White, 0. R., & Haring, N. G. (1980). *Exceptional teaching*, 2nd ed. Columbus, OH: Merrill.

Appendix A: List of 90 irregular words and their characteristics

Items			Frequency	Nsize	Phonemes	Letters	Syllables
φως	/fos/	light	457.0	.0	3.0	3.0	1.0
ώρα	/ora/	time	695.0	.0	3.0	3.0	2.0
ζώα	/zoa/	animals	875.0	.0	3.0	3.0	2.0
ζωή	/zoy/	life	1402.0	.0	3.0	3.0	2.0
λύση	/lisi/	solution	541.0	4.0	4.0	4.0	2.0
σώμα	/soma/	body	605.0	.0	4.0	4.0	2.0
χώρα	/xora/	country	549.0	.0	4.0	4.0	2.0
σειρά	/sira/	line	674.0	2.0	4.0	5.0	2.0
παιδί	/pethi/	child	710.0	.0	4.0	5.0	2.0
σχήμα	/shima/	form	259.0	1.0	5.0	5.0	2.0
χρόνια	/xronja/	years	1047.0	1.0	5.0	6.0	2.0
γλώσσα	/glossa/	tongue	515.0	1.0	5.0	6.0	2.0
εικόνα	/ikona/	image	998.0	.0	5.0	6.0	3.0

σήμερα	/simera/	today	995.0	.0	6.0	6.0	3.0
ελλάδα	/elatha/	Greece	911.0	.0	6.0	6.0	3.0
δύναμη	/thinami/	power	380.0	.0	6.0	6.0	3.0
θάλασσα	/thalasa/	sea	659.0	.0	6.0	7.0	3.0
σχολείο	/sxolio/	school	864.0	.0	6.0	7.0	3.0
κείμενο	/cimeno/	text	1051.0	.0	6.0	7.0	3.0
αριθμός	/arithmos/	number	859.0	.0	7.0	7.0	3.0
ιστορία	/istoria/	history	885.0	.0	7.0	7.0	4.0
βοήθεια	/voithia/	help	503.0	.0	6.0	7.0	4.0
συνέχεια	/sinexia/	all the time	576.0	1.0	7.0	8.0	3.0
ενέργεια	/energia/	energy	550.0	.0	7.0	8.0	4.0
άνθρωπος	/anthropos/	human	529.0	.0	8.0	8.0	3.0
εκκλησία	/eklisia/	church	430.0	.0	7.0	8.0	4.0
κεφάλαιο	/kefaleo/	chapter	673.0	.0	7.0	8.0	4.0
περιβάλλον	/perivalon/	environment	414.0	.0	8.0	9.0	4.0
παράδειγμα	/parathigma/	example	650.0	.0	9.0	10.0	4.0
δραστηριότητα	/thrastiriotita/	activity	500.0	.0	10.0	10.0	5.0
ήχος	/ihos/	sound	150.0	.0	4.0	4.0	2.0

ψυχή	/psihy/	soul	147.0	4.0	4.0	4.0	2.0
φωνή	/foni/	voice	351.0	.0	4.0	4.0	2.0
νησί	/nisi	island	251.0	.0	4.0	4.0	2.0
πρωί	/proi/	morning	258.0	.0	4.0	4.0	2.0
πλοίο	/plio/	ship	120.0	.0	4.0	5.0	2.0
χρώμα	/xroma/	colour	391.0	.0	5.0	5.0	2.0
φωτιά	fotja/	fire	237.0	.0	5.0	5.0	2.0
κλάσμα	/clazma/	fraction	308.0	.0	6.0	6.0	2.0
καιρός	/keros/	weather	123.0	.0	5.0	6.0	2.0
άνοιξη	/anicsi/	spring	143.0	.0	5.0	6.0	3.0
μήνυμα	/minima/	message	133.0	.0	6.0	6.0	3.0
ειρήνη	/irini/	peace	262.0	.0	6.0	6.0	3.0
υλικό	/iliko/	material	221.0	.0	6.0	6.0	3.0
δοχείο	/thoxio/	can	95.0	.0	5.0	6.0	3.0
μαθητής	/mathitis/	pupil	128.0	1.0	7.0	7.0	3.0
αλήθεια	/alithia/	truth	221.0	.0	6.0	7.0	3.0
πρόσωπο	/prosopo/	face	343.0	.0	7.0	7.0	3.0
πλατεία	platia/	square	104.0	.0	6.0	7.0	3.0

γυναίκα	/yineka/	woman	230.0	.0	6.0	7.0	4.0
μυστικό	/mistiko/	secret	89.0	.0	7.0	7.0	3.0
διάρκεια	/thiarkia/	duration	390.0	.0	7.0	8.0	4.0
ανάπτυξη	/anapticsi/	development	213.0	.0	8.0	8.0	4.0
συζήτηση	/sizitisi/	conversation	245.0	.0	8.0	8.0	4.0
τηλεόραση	/tileorasi/	television	232.0	.0	8.0	8.0	5.0
κατασκευή	/kataskevy/	construction	231.0	.0	9.0	9.0	4.0
καλοκαίρι	/calokeri/	summer	185.0	.0	8.0	9.0	4.0
οικογένεια	/ikoyenia/	family	357.0	.0	8.0	10.0	5.0
συγγραφέας	/siyrafeas/	writer	120.0	.0	9.0	10.0	4.0
φωτογραφία	/fotografia/	photo	190.0	.0	10.0	10.0	5.0
λύπη	/lipi/	sadness	59.0	2.0	4.0	4.0	2.0
όγκος	/ogos/	volume	61.0	2.0	3.0	4.0	2.0
πύλη	/pili/	gate	47.0	.0	4.0	4.0	2.0
φυτό	/fito/	plant	84.0	.0	4.0	4.0	2.0
ξύλα	/csila/	wood	68.0	.0	4.0	4.0	2.0
μωρό	/moro/	baby	74.0	.0	4.0	4.0	2.0
πίεση	/piesi/	pressure	58.0	.0	5.0	5.0	3.0

κοιλιά	/kilja/	belly	57.0	.0	4.0	6.0	2.0
αιώνας	/eonas/	century	45.0	.0	5.0	6.0	3.0
σφαίρα	/sfera/	bullet	54.0	.0	5.0	6.0	2.0
ποίηση	/piisi/	poetry	57.0	.0	5.0	6.0	3.0
δωρεάν	/thorean/	free of charge	41.0	.0	6.0	6.0	3.0
είδηση	/ithisi/	news	83.0	.0	5.0	6.0	3.0
γήπεδο	/yipetho/	field	58.0	.0	6.0	6.0	3.0
έλληνας	/elinas/	Greek	67.0	.0	6.0	7.0	3.0
σύννεφα	/sinefa/	clouds	84.0	.0	6.0	7.0	3.0
γεωργία	/yeorgia/	agriculture	55.0	.0	7.0	7.0	4.0
χειμώνας	/himonas/	winter	38.0	1.0	7.0	8.0	3.0
στοιχείο	/stixio/	element	80.0	.0	6.0	8.0	3.0
κάτοικος	/catikos/	resident	21.0	1.0	7.0	8.0	3.0
ζωγράφος	/zografos/	painter	63.0	.0	8.0	8.0	3.0
εργαλείο	/eryalio/	tool	40.0	.0	7.0	8.0	4.0
μυστήριο	/mistirio/	mystery	66.0	.0	8.0	8.0	4.0
μοναστήρι	/monastiri/	monastery	51.0	.0	9.0	8.0	4.0
συμμαθητής	/simathitis/	classmate	40.0	.0	8.0	9.0	4.0

κυβέρνηση	/kivernisi/	government	70.0	.0	9.0	9.0	4.0
συνάντηση	/sunadisi/	meeting	54.0	1.0	8.0	9.0	4.0
πληροφορία	/pliroforia/	information	42.0	.0	10.0	10.0	5.0
μυθιστόρημα	/mithistorima/	novel	23.0	.0	11.0	11.0	5.0
αριθμητική	/arithmitiki/	mathematic	75.0	.0	10.0	10.0	6.0

Table 1: Scores in background assessments for TN and OE and their comparison groups (standard deviations are in parentheses). Results for modified *t*-tests are reported which examine differences in scores for TN and OE and their respective comparison groups.

Measure	Task description	Skill assessed	TN	OE	Modified *t*-test statistics for TN	Modified *t*-test statistics for OE
Non-verbal reasoning[α] (max correct= 34)			24	12	*p>.05* for TN & OE	
Comparison group mean (TN: Nine comparison children, mean age 9;01, SD=0;02; OE: Nine comparison children, mean age 9;05, SD=0;02)			**18.9 (5.2)**	**19.6 (6.1)**		
Spoken Language						
Vocabulary[b] (max=174)	Oral word picture matching (point to 1 out of 4 alternative pictures)	Receptive vocabulary	119	125	*p>.05* for TN & OE	
Comparison group mean			**115 (5.9)**	**116 (12.2)**		
Reading						
Reading Fluency[c] (Standard Scores)	Number of words read correctly from a passage in 1 minute	Text reading fluency	76 (95%CI[1]: 69-83)	77 (95%CI[1]: 70-84)	Below average performance for both TN & OE	
Word reading accuracy[d] (max=40)	Single word reading test	*Lexical* reading	40	39	*p>.05* for both TN & OE	
Comparison group mean (as reported in Niolaki et al. 2014)			**36.8 (3.2)**			
Word reading latencies[d] (msecs)	Single word reading test (vocal reaction times extracted from sound files	*Lexical* reading	1536	1605	*p<.003.* for both TN & OE	

	using *Checkvocal*, Protopapas, 2007)					
Comparison group mean (as reported in Niolaki et al. 2014)			**911(160)**			
Nonword reading accuracy[d] (max=40)	Nonword word reading test	*Sublexical* reading	35	29	*p>.05*	*p<.05*
Comparison group mean (as reported in Niolaki et al. 2014)			**34.3 (2.4)**			
Nonword reading latencies [d] (msecs)	Nonword reading test (items created by substitution/omission of letter in real word (e.g., <φιλάτροποσ> /filatropos/ created from the word <φιλάνθρωποσ> /filanthropos/)	*Sublexical* reading	1592	1215	*p<.01*	*p>.05*
Comparison group mean (as reported in Niolaki et al. 2014)			**1103 (176)**			
Spelling						
Spelling accuracy[e] (Standard Scores)	Single word spelling Normed referenced standardised spelling test	*Lexical* spelling	74 (95%CI[1]: 67-81)	74 (95%CI[1]: 67-81)	Below average performance for both TN & OE	
Irregular words [d] (max correct=20)	Single word spelling	*Lexical* spelling	3	3	*p<.05* for both TN & OE	
Comparison group mean (as reported in Niolaki et al. 2014)			**10.4 (3.2)**			

Qualitative analysis of spelling errors in irregular items	Phonologically appropriate errors (e.g., <πεδί> /pethi/ instead of <παιδί> /pethi/) has been interpreted as reliance on sublexical processes due to difficulties with lexical processes.	*Sublexical* spelling	100%	100%	*p>.05* for both TN & OE	
Comparison group mean (as reported in Niolaki et al. 2014)			**96.3% (6.6)**			
Nonwords [d] (max correct=40)	Non word spelling test (see nonword reading for test's characteristics)	*Sublexical* spelling	39	36	*p>.05* for both TN & OE	
Comparison group mean (as reported in Niolaki et al. 2014)			**36 (2.7)**			
Cognitive assessments						
Phonological ability[f] (max correct=20)	Spoonerisms task (e.g., swap the first sound from each word *king-John* -> Jing-Kohn)	Phonological abilities	10	2	*p>.05*	*p<.05*
Comparison group mean (as reported in Niolaki et al. 2014)			**12.7 (4.8)**			
Visual Memory[g] (max correct =32)	Visual Memory for abstract designs (testee asked to reproduce order of presentation set	Visual Memory	12	4	*p>.05*	*p<.01*

	after a five sec. interval)					
Comparison group mean (as reported in Niolaki et al. 2014)			**15.1 (4.0)**			
Letter report arrays [h] (max correct=20)	A description of the task can be found in Niolaki et al., 2014.	Multi-character processing	0	2	*p<.05*	*p=.07*
Comparison group mean (as reported in Niolaki et al. 2014)			**7.67 (3.4)**			
Letter report total letters [h] (max correct=100)	A description of the task can be found in Niolaki et al., 2014.	Multi-character processing	54	60	*p<.001*	*p<.01*
Comparison group mean (as reported in Niolaki et al. 2014)			**80.2 (6.7)**			

Note: [a]Matrix Analogies Test (Naglieri, 1985), [b]PPVT (adapted for Greek, Simos et al., 2011), [c] Test Alpha subtest (Panteliadou & Antoniou, 2007), [d]Lozidou-Ieridou, Masterson & Hanley (2009), [e]Mouzaki et al. (2007), [f] Adapted from the Phonological Assessment Battery for the Greek language (PhAB, Frederickson, Frith, & Reason,1997), [g](Paraskevopoulos et al., 1999), [h] An adaptation of the task in Greek developed by Bosse and Valdois (2003)
[1]CI- Confidence Interval

Table 2: Accuracy in spelling the irregular words at baseline and three post-intervention assessments for TN and OE and at Baseline 1 and Post-Time 3 for the comparison children (standard deviations are in parentheses)

		Baseline 1	Baseline 2	Post T1	Post T2	Post T3
TN	Total set/90	20	24	63	52	51
	Trained subset/36	0	0	32	22	17
	Untrained subset/25	0	0	6	5	6
TN comp. group	Total set/90	72.2 (15.4)	-	-	-	72.1 (15.2)
OE	Total set/90	14	14	32	31	33
	Trained subset/36	0	0	16	8	12
	Untrained subset/35	0	0	0	6	6
OE comp. group	Total set/90	73.2 (9.1)	-	-	-	74.4 (8.5)

Table 3: Pre-intervention accuracy scores in spelling and accuracy/latency in reading for items from Loizidou-Ieridou et al. (2009) for TN, OE and comparison children as reported in Niolaki et al. (2014), and scores for TN and OE following the intervention (standard deviations are in parentheses)

		TN			OE	
	Comparison group mean	Pre-interven-tion	Post T2	Post T3	Pre-interven-tion	Post T2	Post T3
Irregular word spelling (max = 20)	10.4(3.2)	3	9	10	3	4	7
Regular word spelling (max = 20)	17.1 (2.2)	19	19	19	14	14	18
Nonword spelling (max = 40)	36 (3.7)	39	40	40	36	36	39
Word reading (max = 40)	36.8 (3.2)	40	37	38	39	37	35
Nonword reading (max = 40)	34.3 (2.4)	35	39	40	29	33	38
Word latency (msecs)	911(160)	1536	1609	2995	1605	2194	1692
Nonword latency (msecs)	1103(176)	1592	1434	1689	1215	1510	1358