COMPLEX WORD PROCESSING IN TEENAGE POOR READERS

COMPLEX WORD PROCESSING IN TEENAGE POOR READERS - DOES MORPHOLOGICAL KNOWLEDGE HELP OR HINDER?

By REGINA HENRY, B.A.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment
of the Requirements for the Degree Master of Science

McMaster University © Regina Henry, October, 2012

MASTER OF SCIENCE (2012) McMaster University

(Cognitive Science of Language) Hamilton, Ontario

TITLE: Complex Word Processing in Teenage Poor Readers - Does Morphological Knowledge Help or Hinder?

AUTHOR: Regina Henry, B.A. (McMaster University)

SUPERVISORS: Dr. Elisabet Service, Dr. Victor Kuperman

NUMBER OF PAGES: 98

Abstract

This longitudinal study addressed development of morphological awareness in fourteen-to-seventeen-year-olds reading disabled (RD) high school students enrolled in the Wilson Reading Program (Wilson, 1989). Our lexical decision experiment and reading fluency assessment took place in the first (session 1) and last months (session 2) of the school year that included training with morphologically complex English words. The lexical decision stimuli were composed of derived (*critical*), compound (*bathtub*) and pseudo-complex (*postpone*) words from the training program (trained words), matched complex words not in the training program (untrained words), and nonwords. Accuracy and response times were compared between sessions, and with a comparison group of age-matched typical readers. The RD group did not demonstrate large post-training gains in reading fluency, but, there were significant improvements in accuracy and speed in visual lexical decision. These improvements did not extend to auditory lexical decision, suggesting that the observed improvements in visual word recognition were a result of the training, and not a practice effect due to multiple testing sessions. Additionally, there was post-training improvement in both trained and untrained words implying that the RD students were able to generalize their acquired knowledge of grapheme-phoneme mappings and morphological processing to novel words. Both the RD and comparison group demonstrated the same hierarchy of accuracy and response time patterns for complex words suggest a processing advantage for visually presented derived and compound words that is not skill dependent.

Acknowledgements

Primary Supervisory: Dr. Elisabet Service
Co-Supervisor Dr. Victor Kuperman

Committee Members: Dr. Elisabet Service, Dr. Victor Kuperman
Dr. Catherine Anderson

Undergraduate Research Assistants: Danielle Moed, Allana Stevenson, Priscilla Ally, Aaron Araldo and Nora Al-Zanoon

NSERC Grant awarded to Dr. Elisabet Service

Table of Contents

Tables and Figures

Introduction

Learning to read is a complex process that requires symbol-to-sound mapping, symbol/sound-to-meaning mapping and integration of the activated meaning into a mental representation of the current sentence or text. Although most Canadians have mastered this multifaceted skill by the time they enter high school, a significant number have not. In fact according to the latest published figures from Statistics Canada (2005) over 15 million Canadians are unable to read at a level needed to be successful at work and in day-to-day life. Illiteracy is a problem that affects the day-to-day life of adults world-wide (Eme, 2010) and has been the subject of much research (e.g. Castles & Coltheart, 2004; Coltheart, 1981; Denckla & Rudel, 1976; Ehri, 1992; Goswami, 1993; Nation & Snowling, 1998; Norton & Wolf, 2012; Stanovich, 1986; Wagner & Torgesen, 1987; Ziegler & Goswami, 2005). However, high school students with reading deficits have rarely been the subjects of reading research. Yet these adolescents are at a critical juncture in their lives when their future economic stability depends on sufficient reading ability for job training and employment.

This eight-month longitudinal study examined changes in reading fluency and word recognition in fourteen-to-seventeen-year-old reading disabled (RD) high school students after six months of intensive remedial reading intervention with the Wilson Reading Program (Wilson, 1988). The reading disabled students were in the second year of the Wilson Reading Program and the main emphasis during the eight-month period under study was complex word training. In particular, these students were trained on orthographical, phonological and semantic changes resulting from the addition of affixes to form a derived word, or an additional stem to form a compound word. One goal of the thesis was to establish whether the training program would result in measurable improvements in word recognition and reading fluency. To anticipate the discussion, the results showed that the reading disabled students improved in both speed and accuracy on a visual lexical decision task. There was also a numerical trend towards increased reading fluency. There was no significant change in either response time or accuracy in an auditory lexical decision task. This suggests that the observed improvements in visual word recognition were a genuine result of the training, and not a practice effect due to multiple testing sessions. The second major goal of this thesis was to explore differences in the relative ease of recognition across multiple morphological types (compound, derived words and simple words), processing modalities (auditory vs visual) and across skill levels (pre- and post-training reading disabled and typical readers). Skill level was not a factor in speed or accuracy of processing different morphological types. This unexpected result will be discussed later, in light of existing theories of morphological processing. In what follows, I provide the theoretical background for both lines of the present inquiry.

Reading Intervention

Much of the research on intervention effectiveness has centered on American (e.g., Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998; Olson, Wise, Ring, & Johnson, 1997; Stanovich, 1981; Torgesen, Wagner, & Rashotte, 1997; Vellutino et al., 1996) or Canadian (Lovett, Lacerenza, Borden, et al., 2000; Lovett, Steinbach, & Frijters, 2000) students in the primary grades. There is far less research on adolescents (Lovett, Lacerenza, De Palma & Frijters, 2008) and the existing studies report conflicting results. One recently published American report by the National Center for Education Evaluation and Regional Assistance did evaluate the effect of two different literacy interventions (Reading Apprenticeship Academic Literacy and Xtreme Reading) on high-school-aged poor readers. The 5595 participants in this study were at a reading level of grade seven or lower upon entering grade nine (Somers et. al., 2010). Both literacy interventions included instruction in motivation, reading fluency, vocabulary, comprehension, phonics and phonemic awareness and writing. Although the participants demonstrated improvements in comprehension and academic performance during the intervention year, the improved academic performance was not sustained in the subsequent year and there was no increase in vocabulary size (Somers et. al., 2010). In a smaller study of Canadian high school students enrolled in a reading comprehension intervention, Penney (2002) found that students who received additional instruction in word decoding, not only improved in word decoding and word reading but also demonstrated significantly larger improvement in text comprehension than the students only receiving comprehension training. This result contrasts with the United States National Reading Report on Teaching Children to Read (NICHD, 2000) finding that teaching phonemic awareness only had a positive effect up until grade six, The authors suggest that after 6th grade the emphasis should be switched to vocabulary (NICHD, 2000). However even though this report was a meta-analysis based on an extensive review of the literature, an on-going problem with ascertaining the effectiveness of different strategies of improving the reading skills is the variation in participant reading levels, ages, type of intervention and method of evaluating outcomes (e.g. Foorman, Francis, Fletcher, Schatschneider & Mehta, 1998; Lovett, Lacerenza, Borden, Frijters, Steinback, & De Palma, 2000).

The students in our study were enrolled in the Wilson Program which is recommended for adolescent students with deficiencies in phonemic awareness and phonics by the National Literacy Project 2006 and The Secondary Literacy Instruction and Intervention Guide (McPeak, & Trygg, 2007). Significant improvements have been shown for students of a similar age group to the ones in the current study after participation in a Wilson program. For instance in a study of 30 college students with dyslexia, the ones enrolled in a Wilson intervention showed significant growth in reading and spelling ability after one semester of intervention using the Wilson program (Guyer, Banks & Guyer, 1993). Twenty other students, 10 receiving no intervention and 10 receiving only phonological instruction, did not demonstrate significant improvements (Guyer, Banks & Guyer, 1993). The multi-faceted approach of the Wilson program may have been a significant factor in these results, so concentrating on the effects of one training facet is of importance in teasing out the elements that make it a successful intervention. Our focus is on learning and recognition of morphology. Specifically, the students in the current study were receiving explicit morphological instruction as part of their intervention, and were tested on auditory and visual comprehension of morphologically complex words. Previous research has found that explicit morphological instruction positively impacted the spelling of derived words by 13-year-olds with dyslexia (Tsesmeli and Seymour, 2009).

Reading Acquisition Theories

According to the Simple View of reading (e.g., Hoover & Gough, 1990; Catts, Adlof, & Weismer, 2006), both word decoding and text comprehension are required for successful reading. Shankweiler (1999) advocates that word decoding ability should be the first step because text comprehension is not possible without the ability to accurately identify words, but both comprehension and word decoding are required for successful reading. Decoding interventions instruct students on awareness of syllables and rhyming (phonological awareness), phonemic awareness (perception of sounds in a language, often considered a part of phonological awareness), and phonics (grapheme-phoneme mappings). The rationale for this type of intervention is based on the generally acknowledged role of phonological awareness in reading acquisition (e.g., Nation & Cocksey, 2009; Norton & Wolf, 2012; Adlof, Perfetti & Catts, 2011; Ziegler et. al., 2010). The phonological deficit hypothesis (Vellutino et. al., 2004) asserts that word recognition problems are based on difficulties in grapheme-phoneme mapping caused by a deficit in phonological awareness (Liberman & Shankweiler, 1979; Share, 1995). Substantiation for this view is provided by evidence that in contrast to typical readers, poor readers exhibit difficulty in acquiring phonological awareness skills and phonological analysis continues to be a problem into adulthood (Bruck, 1992; Liberman, Shankweiler, Fischer, & Carter, 1974; Liberman & Shankweiler, 1979, 1991; Snowling, 2000a; Wagner & Torgesen, 1987).

Although there is strong evidence for phonological deficits as a cause of reading disabilities, the cognitive processes related to naming ability may also factor into reading problems. Denckla and Rudel (1976) reported that poor readers were impaired in the naming of colours, common objects, letters and digits in comparison to good readers. Based on the findings of Denckla and Rudel (1976), Bowers and Wolf (1993) further investigated rapid naming and determined that it accounted for a unique amount of variance in reading fluency over and above the contribution of phonological awareness. They suggested that some individuals with dyslexia may have a timing deficit that interrupts the automatic processing of symbols (Bowers & Wolf, 1993). In 1999, they put forward the *double deficit hypothesis* to account for the unique portions of variance in reading fluency accounted for by rapid naming and phonological awareness, respectively (Wolf & Bowers, 1999). This hypothesis argues that naming deficits are separate from phonological deficits and that there are three subtypes of poor readers: one with only phonological deficits, a second with only naming deficits and the third group with both (Wolf & Bowers, 1999). This third group is considered to have the most serious reading disability.

The results of studies examining evidence for the double-deficit are mixed, some studies have found that readers with double-deficits are more disabled (e.g., Katzir, Kim, Wolf, Morris, & Lovett, 2008; Kirby, Parrila, & Pfeiffer, 2003; Lovett, Steinbach, & Frijters, 2000; Wolf et al., 2002) but others have found no difference between the double deficit and single deficit or no-deficit groups (e.g., Ackerman, Holloway, Youngdahl, & Dykman, 2001). The double deficit hypothesis was originally developed to explain different subtypes in dyslexia. However, as serial naming speed has been found to predict reading fluency also in unselected samples it may also generalize to other poor readers.

Another important and relevant perspective of reading acquisition is psycholinguistic grain size theory. According to this theory, reading acquisition starts with an understanding of the way in which symbols in a language are mapped to phonological units of different sizes (grains). This process begins with awareness of the large grains like syllables and rimes and progresses to smaller grains, like phoneme and letter units, as the child becomes more familiar with the orthographic/phonological relationship in the language (Ziegler & Goswami, 2005). This transition to understanding the mapping between smaller grain sized orthographic and phonological units varies in difficulty based on the relative transparency of the language. The degree of transparency or opaqueness is dependent on the regularity of the symbol-sound correspondences in a language. In transparent languages like Finnish, Italian, Greek and Spanish with mainly one-to-one letter-sound correspondences, children are quick to acquire knowledge of mappings between letters and phonemes (Ziegler & Goswami, 2005). English, however, is an opaque language with numerous homophones such as *pear* and *pair* that sound the same but are spelled differently. Additionally there are variations in the way letters and sounds map. For instance, several letters may represent a single phoneme ("*t*" in cite sounds the same as "*ght*" in *light*) and so the letter-sound mappings are inconsistent (McDougall, Brunswick & de Mornay Davies, 2010). The students in the current study were receiving training on morphemes which can be either whole word (base morphemes) or a smaller grain size (affixes) roughly equivalent to the large grain syllables and rimes that Ziegler and Goswami (2005) refer to. Based on the grain size theory (Ziegler & Goswami, 2005) the students in our study should be able to acquire knowledge of stems and affixes (the larger grains) faster than letter blends (smaller grains) and this result should not be modality specific.

Complex Word Reading

In successful reading acquisition learners progress from relatively simple monosyllabic words to words containing multiple morphemic constituents. Inflectional morphology allows us to communicate details such as whether events are occurring in the past, present or future, whether there are one or more items, and in many languages gender. Derivational morphology produces new words by adding elements to existing words. This can change word meaning, for instance by adding “un” to “do” (undo), or change a syntactic category as in the change of *pave* from verb to noun by the addition of “ment” (pavement). Morphology is also largely responsible for the development of new words (neologisms). This can occur through the development of new affixes; for instance *Watergate* and *Climategate* (Merkx, Rastle & Davis, 2011), new combinations such as unquenchable, or through new word compound combinations, i.e., *eco-friendly* and *breakdance*.

An open question in reading research is the processing cost of recognizing simple words containing only a single morpheme, and often only one syllable, compared to complex words with more than one syllable or morpheme. The main theoretical division in theories of complex word processing is whether words are accessed as wholes (full-listing) or decomposed into individual morphemes for storage and then recombined during production (decompositional). More specifically, the full-listing accounts propose that words are stored as whole units and individual morphemes do not have much of a role in word recognition (Butterworth, 1983; Lukatela, Carello & Turvey, 1987; Manelis & Tharp, 1977). According to full-listing accounts "*rose*", "*bud*" and "*rosebud*" would all be stored and accessed separately.

Word processing models that are compatible with decomposition include: distributed processing (PDP) models (i.e. Morris et. al. 2007; Perfetti, 1992; Seidenberg & McClelland, 1989) and theories in which word recognition occurs through more than one route, for instance, multiple route (Kuperman et. al., 2008, 2009) and dual route theories (i.e., Baayen, Dijkstra & Schreuder, 1997; De Jong, Schreuder, & Baayen, 2000; Grainger et. al., 2012; Schreuder & Baayen, 1997). In PDP models, meaning plays an integral part in visual word processing, at a relatively early stage, just after primary sensory processing. All words including complex ones are broken down into component features and encoding and retrieval take place through distributed patterns of repeated activation (Seidenberg & McClelland, 1989). Multiple route theory models propose that both whole word and decomposition can take place jointly (i.e., Baayen, Dijkstra & Schreuder, 1997; Kuperman et. al., 2008, 2009; Schreuder & Baayen, 1997).

Taft and Forster (1975, 1976) published the first evidence of the influence of meaning in polymorphemic word recognition over thirty-five years ago. The studies investigated recognition of polymorphemic words during a lexical decision task (Taft & Forster, 1976). The technique they developed has been used extensively since the paper's publication (e.g. Beyersmann, Castles & Coltheart, 2011; El Yagoubi et. al., 2008; Lavric, Rastle & Clapp, 2011; Marslen-Wilson & Tyler, 1997; Morris et. al., 2007; Nation & Cocksey, 2009). In the Taft and Forster (1975, 1976) experiments, most common version of this paradigm, the participants' task is to decide if the stimulus presented on a screen is a real word or not and press a yes button for a real word and a no button for a non-word. Taft and Forster (1976) found that if a real word such as *foot* (the prime) was embedded in a non-word (*footmilge*), reaction times were longer than if there was no embedded word (*mowdflisk*). In subsequent versions of this paradigm, with the target word embedded in primes that are morphologically transparent, opaque or form-based reaction times were longer for all three types. A semantic relationship between the prime and target with transparent primes, for instance: *agreement* (prime) and *agree* (target). In opaque primes, the affix is legal for the particular language but there is no semantic relationship, for instance: *importance* (prime) and *import* (target). A form-based prime contains a pseudo-affix not used in the language: *dialog* (prime) and *dial* (target).

Nation and Cocksey (2009) observed similar semantic interference at a morphemic level in children learning to read. When seven-year-old children were asked to make category decisions about visually presented objects they were slower and less accurate in classifying pictures if the label contained an embedded word stem suggesting early and automatic influence of morphemic meaning even in beginning readers. For instance when categorizing *ship*, they were slower to reject body-part (*hip*) than animal (Nation & Cocksey, 2009).

Development of Morphological Awareness

Morphological awareness appears to accumulate gradually during the course of language development and investigations of the development of this ability inform us about complex word processing. According to Ziegler & Goswami (2005) children first learn to encode larger phonological units heard in speech, prior to understanding the mapping of smaller grapheme-phoneme units. The free morphemes (stems) and bound morphemes (affixes) that comprise complex words are larger units that children gradually learn to encode as they encounter them in speech (Carlisle, 2010). Language experience contributes to the formation of morphemic representations through acquisition of the morphemes' individual meanings. Carlisle (2010) suggests that children show evidence of separate processing of morphemes through novel complex word creation. She provides two examples from Clark (1982) of *winder* for an ice-cream machine and *flyable* in reference to a cocoon (Carlisle, 2010). During child language development, complex words, such as *forward* are initially stored as whole words. Once the child has had experience with the bound morpheme "*ward*" in other words such as *upward*, *outward*, *inward* and *backward*, the morphemes are processed separately. This allows for novel word creations for instance *bedward* was a novel word my own children used. Neologisms suggest awareness of morphemic composition. The unique combinations created suggest that multi-morphemic words are decomposed; stems and affixes are stored separately and this enables unique recombination during complex word production.

Role of Morphology in Reading Acquisition

In addition to the ability to analyse words into syllables, rimes and smaller phonological units, successful word decoding also requires morphemic analysis. Research into the relationship between morphology and literacy is a relatively neglected field in comparison to research on the role of phonology (Carlisle, 2010; Deacon & Kirby, 2004;

Mann, 2000). Goswami and Ziegler (2006) admitted to “the benign neglect of morphology” when developing their theoretical framework on reading acquisition and psycholinguistic grain size. But morphological changes affect phonological word properties: for instance, “*ph*” is a blend in *phone* but it is not in *uphill* and morphological processing is necessary for skilled reading (Carlisle & Nomanbhoy, 1993; Fowler & Liberman, 1995; Shankweiler, et al., 1995; Singson, Mahoney, & Mann, 2000; Windsor, 2000). Deacon and Kirby (2004) provided evidence of the unique contribution of morphological awareness to reading comprehension over and above that of phonological awareness in longitudinal study assessing reading ability in children from grades 2-5. Siegel (2008) also identified morphological awareness as a potentially significant contributor to reading and spelling deficits over and above phonological and verbal language skills (speech production & comprehension) in a study of 1238 grade 6 students. Arnbrak and Elbro (2000) found that training on the semantic aspects of morphemes increased morphological awareness for grade four and five students with dyslexia and resulted in significant reading comprehension and complex word spelling improvement compared to a similar control group who did not receive morphological instruction. In another study, adolescents with dyslexia used knowledge of root morphemes as a compensatory strategy when reading both single words and text (Elbro & Arnbak, 1996). Nagy, Berninger and Abbott (2006) evaluated the separate contributions of morphological awareness, phonological memory and phonological decoding to various reading measures in students at different grade levels: 4th& 5th; 6th& 7th and 9th& 10th. They found unique contributions of morphological awareness to vocabulary and reading comprehension for all three groups. Additionally morphological awareness contributed significantly to decoding accuracy in grades 8 and 9. Tsesmeli and Seymour (2006) found that the reading deficits displayed by adolescents with dyslexia were related to morphological deficits rather than vocabulary size.

Lexical Quality

The Lexical Quality Hypothesis (Perfetti & Hart, 2002) predicts the contribution of morphology to word recognition and is compatible with decompositional models of complex word processing. The hypothesis posits that the lexical representation of a word is dependent upon the strength of its morphological, phonological and orthographic representations and the strength of the connections between them (Perfetti & Hart, 2002). Repeated exposure to written words results in higher quality lexical representations. As discussed earlier, the addition of affixal morphemes to a base in complex word creation can change the orthographic and phonological representation of the word. Thus, knowledge of morphological units and their combinatorial rules has an important place in reading acquisition (Perfetti, 2011). The addition of an affix often changes the vowel alternation pattern as in *nation* – *national* but the distinction between morpheme units is maintained in the written form. Conversely, written words do not always help to identify either morpheme stems or affixes. *Mansion* and *revision* do not share a morpheme although they do share a spelling -‘ion’ (Perfetti, 2011). Likewise; *hipbone* and *trombone* share the spelling of *bone* but do not share the morpheme. In these cases, the pseudo-base or pseudo-affix needs to be inhibited for correct word identification, resulting in increased processing effort for the written word form. However, when these words are heard instead of read the pseudo morphemes may not be identified and, as a result increased processing costs may not be incurred.

The relative contribution of morphemes to reading may vary based on skill level. It has been shown that relatively skilled readers rely less on morphological structure, and for the most skilled readers, salient (frequent) morphemes embedded in a complex word may impede its recognition (Burani, Marcolni, De Luca & Zoccolotti, 2008; Carlisle & Stone, 2005; Kuperman & Van Dyke, 2011). However, it is not necessarily the case that more skilled readers have less access to morphemic content, but instead that they use a combination of word knowledge obtained from spelling, pronunciation and meaning (Perfetti, 2011). Since each of these components contribute to the quality of the lexical representation, less skilled readers may rely more on the meaning of individual morphemes to identify a word than skilled readers. This was shown for instance in Kuperman and Van Dyke (2011). In comprehension and word segmentation tasks, skilled readers had slower response times for derived words with high frequency stems, while poor readers showed the opposite pattern: their responses to derived words with high frequency stems were faster than their responses to derived words with low frequency stems (Kuperman and Van Dyke, 2011). According to dual-route models of complex word processing (De Jong, Schreuder, & Baayen, 2000; Grainger et. al., 2012) skilled readers can identify a word either by its constituent morphemes or through its whole word form, or use both routes. The exposure to numerous examples of orthographic, phonological and semantic word features experienced by skilled readers is thought to produce higher quality lexical representations resulting in easier and faster processing of complex words (Perfetti & Hart, 2002; Perfetti, 2011).

If the lexical quality of a word is strengthened by exposure to both its whole word form and the individual morphemes of which it is composed, then the frequency of both the base and affixes should influence the lexical quality of a given word. Carlisle and Katz (2006) provided evidence that this is, indeed, true. They looked at the relationship between a measure of word familiarity and derived word reading accuracy in 4th and 6th grade students. The familiarity measure came from a sampling of texts for 3rd through 9th graders (Carroll, Davies, & Richman, 1971) and included both base and whole word frequencies, family size, average family frequency, and word length (number of letters in a word). Frequency was the number of times the derived words or one of their constituent morphemes appeared in the text. Family size was calculated as the number of words with the same base. For instance there are ten family members with the base word *intense intensive, intensely, intensity, intensify, intensified, intensifiers, intensifying, intensities, intensive, and intensively*) in the sampling (Carlisle & Katz, 2006). Some word families may appear more frequently than others with the same number of members. For instance, words from the *activity* family appear more often than words from the *intense* family in the sampled texts (Carlisle & Katz, 2006). Carlisle and Katz (2006) reported that derived words from larger families and with higher frequencies were easier to read. Additionally there was an interaction between grade level and family frequency. There was no difference between good readers in grade 4 and 6 for the high frequency words but 6th grader good readers were better than 4th grader good readers at reading the words with low family frequencies. The results suggest that the grade 6 good readers had better quality lexical representations (Carlisle & Katz, 2006).

Decomposition – Is there a Processing Cost or Benefit

Although the constituent morphemes may help to create high quality lexical representations, morphological dual-route accounts often assume that direct access to full forms is a less costly way to process complex words (e.g., Bertram, Laine, & Karvinen, 1999; Ji, Gagné and Spalding , 2011). However as Carlisle and Katz (2006) demonstrated, derived words seem to provide an advantage at least for poor and beginning readers. In six experiments, Ji, et. al., (2011) investigated the costs/benefits of morphological decomposition by comparing processing of English compounds (*rosebud*) and monomorphemic words (*giraffe*) matched on word length and frequency. All six experiments involved a lexical decision task but the stimulus composition changed. In the first experiment, the stimuli were sixty pairs of monomorphemic and compound words of which 44 were transparent (both parts contributed to meaning – *rosebud*) and 16 were opaque (*hogwash* – neither constituent contributes to the meaning of nonsense). There were an additional 120 non-words: 30 compound-like items with a base word as the final constituent (*rostpepper*); 30 compound-like non-words with a base word as the first constituent (*chivesonse*); and 60 monomorphemic-like non-words (*arithmutia*). The results showed a processing advantage for the compound words over simplex words. In the second experiment, the composition of the compound-like non-words was changed so that both constituents contained an English base word (*restpepper*). Once again compounds were processed faster. In experiment 3, the stimuli were composed of 30 opaque compounds, 30 transparent compounds and 30 monomorphemic words. In addition, there were 30 compound-like nonwords with a base word as the final constituent; 30 compound-like non-words with a base word as the first constituent and 60 monomorphemic-like non-words. Both types of compounds were faster than the monomorphemic words and the transparent ones were also more accurate. In experiment 4, the compound stimuli varied in opaqueness from totally opaque to totally transparent. As in experiment 2, the non-words were formed using two legal base words (*restpepper*). However, in this experiment, all the words had spaces between their constituent parts: i.e. *rose bud*; *sopho more*; *flint pub*. The insertion of a space removed the previously seen processing advantage for opaque compounds over monomorphemic words. Moreover, high frequency initial constituents reduced RT for transparent compounds while increasing RT for opaque ones. Experiment 5, was the same as experiment 4 including the pattern of results, except colour instead of a space was used to divide constituents. To investigate if the spacing or colour was responsible for the results in experiments 4 and 5, experiment 6 used the same word and nonword stimuli but without spaces or colour variations. The results replicated the findings in experiment 4 and 5. The combined results of these experiments suggest that when the meaning of a constituent is relevant to the whole word and decomposition is beneficial rather than costly. However, if the constituent -- especially the modifier in the word initial position -- is not related to the whole word meaning, then a processing cost is incurred.

Preference for Suffixes

Another class of theories that gives a special place to the word initial position is the *Suffixing Preference* account. According to this account, many more languages have predominantly suffixes or a preference for suffixes in comparison to the number of languages with primarily prefixes or a preference for prefixes (Cutler et al., 1985; Hawkins & Cutler, 1988). A look at the data available from the World Atlas of Language Structures Online (WALS) on suffixing and prefixing preferences confirms this (Dryer, 2011). Of the 971 languages investigated for affix preferences, 54.58% have either predominantly suffixing or a preference for suffixing but only 14.70% have a preference for or are principally prefixing.

In syntax, it is the head not the modifiers that determine the phrase type. Similarly, in morphology derivational words can be separated into bases and affixes and since it is the affix that determines the word category it is the affix which is the head (Cutler et al, 1985; Hawkins & Cutler, 1988). For instance adding *ment* to *move* changes it from a verb to noun. According to the head ordering principle (HOP), the affix is on the same side of the base as a preposition (P) would be relative to the noun phrase (NP) within a prepositional phrase or the same order as a verb (V) relative to a direct object NP (Cutler et al., 1985; Hawkins & Cutler, 1988). Therefore based on the HOP there should be prefixes in languages with the pattern of P –NP and VO word order and suffixes in languages with NP – P and /or SOV word order. An examination of languages reveals that it is only VO and P - NP languages that may have exclusive prefixing. Additionally, there are numerous affix types that are predominately suffixes, including plurals on nouns, tense and mood on verbs, but only person marking on direct objects shows a preference for prefixes.

Cutler et al. (1985) also cite psycholinguistic research to bolster the suffixing preference argument. For instance, they discuss tip-of-the-tongue states in which the person may remember the word onset or may be able to produce the word when the onset is given. The authors also argue that the first part of a spoken word to be heard is the onset and that this provides a temporal processing constraint. But endings are also highly salient. Nooteboom (1981) found that participants successfully identified Dutch words 60% of the time when they were given only the word endings. This may be partially due to the stimuli used; there are few words that end in “vark” so the ending may be particularly salient. Cutler et al. (1985) additionally argue that stems and affixes are processed separately. They present data from lexical decision studies where participants had longer RTs for nonwords containing a legal affix as evidence of decomposition. The end result of their arguments is that it is beneficial for auditory comprehension to receive most of the lexical information at the word onset: this is achieved in stem+affix combinations or suffixed words. Prefixed words are disprefered as they place the most informative part at the less salient position in the word, i.e. the word ending. Circumfixes are even rarer, as in the prefix-stem-suffix combination, the stem is in the least salient position in the word. To sum up, languages and language processors (i.e. readers and listeners) should prefer suffixed words based on syntactic patterns, temporal order of speech sounds, salience of the word onset and salience of actual suffixes (Cutler et al., 1985; Hawkins & Cutler, 1988). In addition to testing the effectiveness of a morphological reading intervention, the present study explored morphological effects on lexical decisions. Specifically, it asked to what extent hypothesized processing preferences, such as affixes after stems, and morphological priming in compounds interacts with improvements in word recognition.

Current Study

The present study investigated the contribution of morphological training to increased word recognition and reading fluency in adolescent poor readers with poor vocabulary knowledge. Previous studies have looked at morphological awareness training and visual word recognition but to our knowledge none have looked to see if auditory processing was also affected. The use of both visual and auditory lexical decision tasks allowed us to investigate the effect of the training on both modalities. We expected significant increases in visual word recognition. Additional auditory effects would suggest a general strengthening of lexical quality irrespective of modality. However, it is also possible that morphological knowledge for speech is acquired predominantly during oral language acquisition (see Clark, 1982) and is encapsulated from orthographic familiarity. According to the Lexical Quality Hypothesis (Perfetti & Hart, 2002) knowledge about affixes, base words and their combinatorial patterns should further additively strengthen the quality of a word’s representation. We investigated the accuracy of this prediction for reading disabled adolescents by using compound, derived and monomorphemic words in the lexical decision task. Based on the Lexical Quality Hypothesis we expected that derived words would demonstrate a processing advantage over monomorphemic words after the morphological training due to the strengthening of affix representations after multiple exposures.

We also hypothesized that the strengthening of orthographic, phonological and semantic representations of both bases and affixes would result in a generalization of the new knowledge to untrained words. Further, based on the findings of Ji, et.al (2011), we expected that compound words would demonstrate the greatest improvement due to the lexical boost from two base words. We were also interested in whether the pattern of processing difficulty based on morphological type was skill or modality dependent. Based on the *Suffixing Preference* account of Cutler and colleagues, and on the finding from Ji et. al. (2011) we expected that at least in the auditory domain compounds would have the largest processing advantage, followed by suffixes then prefixes with monomorphemic words showing the greatest processing costs.

Methods

Participants

It was difficult to find a comparison group with the same socio-economic and cognitive profile as the students with reading disabilities. To match on reading level we would need to compare high school students with reading disabilities to 2nd grade typical readers which is not advisable due to differences in life experience. Matching on general cognitive ability was also problematic because we did not have a large access to a large enough sample population or alternatively data from the school board enabling a controlled population sampling based on school assessments. The only local high school that had students with similar socioeconomic status had a population that was primarily English as Second Language students. Although not ideal, we chose a comparison group from a local high school that matched our reading disabled students in age and native language but had typical grade ten reading skills.

Reading Disabled Participant Group. Twenty-two students were recruited from Parkview Secondary School. All of these students had been diagnosed with reading deficits and were participating in an intensive remedial reading program (Wilson Remedial Reading Program). Since two students left the school before testing was complete, the resulting data pool contained data from twenty participants (10 male and 10 female; at session 1: mean age = 15.09, SD = 0.81, range = 14.08 – 17 years), for analysis. The students received two $10.00 gift certificates for a local coffee shop; one for participating in the fall sessions and the second for participation in the spring sessions. All subjects were informed of their rights as experimental participants, written consent was obtained from parents/guardians and written assent was also obtained from each student. For the purpose of this study, reading disabled is defined as being previously diagnosed as poor decoders and having a reading level greater than four grade levels below typical age-matched readers.

Comparison Participant Group. Twenty-one students with age-typical reading skills (6 male and 15 female; mean age = 15.79, SD = 0.91, range = 15.5 – 16.25 years), were recruited from Waterdown Secondary School. All subjects were informed of their rights as experimental participants, written consent was obtained from parents/guardians and written assent was also obtained from each student.

Cognitive Tests and Lexical Decision Task

Wechsler Abbreviated Scale of Intelligence™ (WASI; Wechsler, 1999). Four subtests measuring verbal (Vocabulary and Similarities tests) and non-verbal (Block Design and Matrix Reasoning) cognitive abilities were administered to each participant in both the experimental and comparison group. The WASI (Wechsler, 1999) subtests provided a tool to assess any differences in cognitive abilities between the two groups and for comparisons with standardized scores from the general population. Given the low reading level of our experimental group, the non-verbal matrix reasoning and block subtests were particularly important for assessing their general cognitive performance(Table 3 summarizes the performance of each group). The WASI was used to characterize the participants as a group rather than as individuals.

Grey Oral Reading Tests (GORT- 4; Wiederholt & Bryant, 2001). Measures of overall reading fluency and comprehension were obtained through administration of the GORT-4 (GORT- 4; Wiederholt & Bryant, 2001). Raw rate (the amount of time taken to read each story) and accuracy (assessed by ability to correctly pronounce each word) were recorded for each story. Comprehension scores based on accuracy of the students' responses to the five questions asked after each story were also recorded during the first session. The Oral Reading Quotient for each participant was calculated for the first GORT-4 administration by summing up individual fluency and comprehension scores and converting them to standard scores based on GORT- 4 (GORT- 4; Wiederholt& Bryant, 2001) normative values.

RAN. Naming speed was assessed by the rapid automatized naming (RAN) letter, digit, colour and object subtests of the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgeson & Rashotte, 1999). Each of these tasks consisted of the speeded naming of a letter, a digit, a picture object or a coloured patch, where the objects are arranged in a grid. The score for each subtest was the time taken to name each of the 36 items in a 4*9 grid.

Lexical Decision Task

Stimuli. The lexical decision task was used to assess changes in the speed and accuracy of auditory and visual word processing before and after six months of reading remediation. For this experimental procedure, 240 words and 240 non-words were used as target stimuli. One hundred and twenty words (trained words, Appendix Table 1) were selected from workbooks 3and 4of the Wilson Reading program. These two workbooks were selected because the students in the experimental group would be completing them between the first and second testing sessions. The words were selected based on four morphological types: 40 compound (baseball), 20 prefixed (misuse), 20 suffixed (directive) and 40 pseudo-complex (stubborn) words.

An additional one hundred and twenty words matched on word length, frequency, age of acquisition and morphological type were selected from the 50-million word SUBTLEX corpus (Brysbaert & New, 2009), based on subtitles to US films and media (untrained words, Appendix Table 2). None of the untrained words were in Wilson workbooks 3and 4. All of the compound words chosen were composed of monosyllabic, singular nouns as morphemes (e.g., baseball), and the words with affixes contained only a prefix or suffix, not both (e.g., *disagree* but not *disagreement*). The pseudo-complex words appeared to either have a stem and affix or to be compound words but the meaning of either morpheme was not related to morphological components (e.g., *trombone*). Two separate lists of 200 non-words each that matched the trained and untrained words on syllable length and word length (in phonemes) were generated with the pseudoword generator software Wuggy (Keuleers & Brysbaert, 2010): The nonwords followed the rules of English phonotactics. In order to mimic compounds, pseudocompounds were also matched on the length of each syllable (in phonemes) and had word-like structures in that each syllable contained at minimum a nucleus plus either onset or coda and most had the structure of onset-nucleus-coda. Examples of words and corresponding pseudowords by type are: *bookcase – boakwase* (compound), *entire – entase* (pseudo-complex), *misfit – ponbit* (prefixed) and *placement – plangdent* (suffixed).To summarize, manipulations in our experiment comprised: 2 modalities (auditory vs. visual) x 4 morphological types (compound, prefix, suffix and pseudo-complex) x 2 levels of lexicality (words vs. nonwords).

Both lists were examined for any non-words that contained actual free morphemes and these were eliminated. The final breakdown of each one hundred and forty non-word list (A2) was similar to that of the word list and contained: 40compound-like non-words, 40pseudo- complex-like non-words, 20 non-words containing a prefix-like structure, and 20 non-words containing a suffix-like structure for a total of 240 non-words. Word length was matched between trained/untrained and word/non-word categories but varied within each type (Table 1). Table 1summarizes word length and frequency by morphological type and Table 2 summarizes non-word length by the morphological type of the corresponding existing words.

Table 1. Descriptive statistics (mean, standard deviation and range) of word length and frequency by morphological type, for words.

Words				
Statistics	**Compound**	**Pseudo-complex**	**Prefixed**	**Suffixed**
Length				
Range	7 – 10	5 – 11	4 -13	5 – 12
Mean	8.25	7.16	7.71	8.85
SD	0.1	1.44	2.02	1.30
Frequency				
Range	1 - 3145	1 - 12288	1 - 10775	8 - 3117
Mean	164.21	952.49	504.68	499.63
SD	405.12	1947.31	1862.22	674.63

Table 2.Non-word descriptive statistics (mean, standard deviation and range) of string length (number of letters) by morphological type.

Non-Words				
Statistics	**Compound**	**Pseudo-complex**	**Prefixed**	**Suffixed**
Length Range	7 – 10	5 – 11	4 – 13	5 - 12
Mean	8.163	7.17	7.18	8.80
Length SD	1.02	1.46	1.99	1.29

The auditory stimuli lists were recorded to digital files by a native speaker of English using Praat software (Boersma & Weenink, 2012) and then segmented into individual waveform files with Audacity 1.3.11software. To enable counterbalancing (see general design below), both the auditory and visual files were randomly split into groups of odd numbered and even numbered files resulting in two visual versions: V1 (odd), V2 (even), and two auditory versions: A1 (odd) and A2 (even), each containing 120 words and 120 nonwords. Each version began with practice trials of six words and one non-word that were not from the target stimulus lists. Additionally, to minimize fatigue, four breaks were programmed into each experiment resulting in five blocks per modality (visual or auditory). On-line randomization was used so that the stimuli were randomized within each block, resulting in different orders of stimulus presentation for each participant.

General Design and Procedure

In this longitudinal study, teenage participants with known reading deficits enrolled in the Wilson Reading Program, an intensive remedial reading intervention, were tested during the first term of the school year and again at the end of the second term (6 months apart on average) in the same school year. They will be referred to as the reading disable or experimental group. The second group of participants, referred to as the comparison group, were all typical readers and were not undergoing reading remediation.

Both the experimental and comparison participants completed standardized assessments of cognitive functioning and reading (WASI and GORT- 4) during the same school year. The experimental group also repeated the GORT- 4 fluency assessment towards the end of the second school term. Both groups further performed experimental tasks (RAN and visual and auditory lexical decision). For the experimental group, the lexical decision tasks were presented twice: in the first term of the school year and again at the end of the second term. For the comparison group, the tasks were presented only once. This design made it possible to do both between-groups comparisons of the experimental and referent group and examine within-group changes in the group with reading deficits. To minimize practice effects, the two versions: Form A and Form B, of the GORT- 4 were used. This made it possible to administer different versions in session one and session two for the experimental group. The referent group completed the GORT- 4 only once, during the same school year. Half of these students completed Form A and the other half completed Form B.

Students completed two different-modality versions of the lexical decision task. If a participant had started with visual stimuli in the earlier testing, they began with auditory stimuli in the second. Therefore, participants who completed V1 then A2 during the first term, completed A1 then V2 during the second term and participants who completed V2 then A1 during term one, completed A2 then V1 in the second term. This design counterbalanced the order of the two modalities, and ensured that the participants made decisions on each of the 480 stimulus items in both auditory and visual presentations. To summarize, manipulations in our lexical decision experiment comprised: 2 modalities (auditory vs. visual) x 4 morphological types (compound, prefix, suffix and pseudo-complex) x 2 levels of lexicality (words vs. nonwords)

Procedure

The reading disabled group participated in three sessions and the comparison group participated in two. Each session lasted approximately 45 minutes and the tasks were completed in the same order.

Session 1. The experimenter introduced herself and gave a general overview of the tasks to be completed in both sessions prior to the participant signing a written assent to participate in the study. The session began with administration of the RAN tasks followed by GORT- 4and lexical decision tasks. For both the RAN and GORT- 4, audio recordings were made of participant responses to enable the researcher to later analyze production errors.

RAN. For this task, participants were asked to name each of the items in a nine by four grid composed of objects, letters, digits or colours. The experimenter timed and recorded the responses. The colour, letter, digit and one of the objects grids were from the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgeson & Rashotte, 1999). Participants also named items in four additional grids with experimental manipulations (these manipulations are the subject of a different experiment reported elsewhere).

GORT- 4. There are fourteen stories in total in the GORT test. For this study participants from the experimental group began at story three, based on their reading grade level, rather than their current secondary school level. The comparison participants began at story nine, the starting point for their current grade level. After the experimenter provided the participant with a sentence summarizing the main story idea, the participant read each story while being recorded and timed. Recording of the participants allowed for a more in-depth analysis of the type of deviations made (for instance substitution of semantically or phonologically similar words). During the reading of the story, the experimenter provided the correct pronunciation of words that the participant was unable to produce. After the participant finished reading a story, the experimenter read the related comprehension questions and five answer choices, one at a time, and recorded the participant’s response. If the participant was able to correctly answer three of the five questions and the experimenter provided correct pronunciation for less than 20% of the words, the same procedure was followed for subsequent stories. The assessment ended when the participant needed prompting for more than 20% of the words or was unable to correctly answer three of the five comprehension questions.

Lexical Decision. The participant sat in a comfortable chair in front of a computer. The experimenter read the instructions out loud while the participant viewed them on the monitor screen. They were told to decide as quickly as possible if the stimulus they saw or heard was a real word. If they thought it was a real word they were to press a key labelled ‘R’ (for “real”) and if it was a non-word, to press the key labelled ‘A’ (for “alien”). The stimuli were presented using SR Research Experiment Builder software (version 1.10.1). The 17 inch monitor had a resolution of 1600 x 1200 and the visual stimuli were presented in lower case Courier New 30 point black font on a white screen. In the visual condition a fixation cross was presented in the centre of the screen for 500 milliseconds followed by the target (word or non-word). The target remained on the screen until the participants indicated their choice by pressing a key for a word or a non-word on the computer keyboard. Reaction times (RTs) from stimulus onset to key press were registered. The auditory stimuli were presented through external speakers with the sound adjusted to a comfortable level. In the auditory condition the fixation remained on the screen for 500 milliseconds prior to the stimulus presentation and then a blank screen was displayed until the participant indicated his/her choice by pressing either ‘R’ or ‘A’. Half of the participants were asked to (a) first listen to a series of auditorily presented words interspersed with auditory presented non-words, and then (b) read a series of visually presented words interspersed with visually presented non-words. The other half completed these tasks in reverse order. Each word was presented to a participant only once – either in the visual or the auditory modality. After completion of the lexical decision task, the student was thanked for participating and reminded that there would be a subsequent session or sessions.

Session 2. After a brief explanation of the tasks to be completed that day, the WASI subtests were administered in the same order for all participants; Vocabulary, Block Design, Similarities and Matrix Reasoning using the same age-relevant starting points for both experimental and comparison groups. In the Vocabulary subtest the participant was asked to give the meaning of a word. The Block Design subtest is a time-limited task in which the participant views a picture of a geometric pattern and has to replicate the design using two-colour blocks. For the Similarities subtest, participants told the experimenter the way in which two word items were alike and the Matrix Reasoning subtest required the participant to view an incomplete pattern grid and chose one of five items to complete the grid.

Session 3. This session occurred approximately six months after the first session and only the reading disabled group participated in it. Different versions of the lexical decision experiment and GORT- 4were repeated. Additionally the GORT- 4 administration procedure differed for the second assessment. We were primarily interested in changes in fluency rather than comprehension so only one randomly selected question was asked to encourage participants to pay attention while reading. The RAN and WASI tests were not repeated.

Results and Discussion

Statistical Considerations

The statistical package *R* was used for data analysis of 28,134 observations in the lexical decision task. Inferential statistics were calculated using linear mixed effects regression models (lmer) implemented in the lme4 library. Regression models with the Gaussian link function were fitted to numeric dependent variables (e.g. lexical decision RT), while logistic regression models with the binomial link function were fitted to binary dependent variables (e.g. lexical decision accuracy). Regression modeling was chosen over the more standard ANOVA analyses because our stimuli were not matched across morphological types on a number of lexical dimensions, including word frequency and length. For this reason, the influence of control variables like word frequency and length had to be controlled for statistically, when estimating the magnitude and significance of critical effects. Moreover, mixed models allow for a simultaneous consideration of multiple covariates, while taking into account the between-subjects and –

items variability. Additionally, logistic regression has been shown to be a more accurate method for analysis of binomial factors (see Jaeger, 2008) like the lexical decision accuracy measure. Regression coefficients for categorical predictors estimate contrasts between the reference level (e.g., “compound”) and each of the remaining levels (e.g., “prefix”, “suffix”, “pseudo-complex”). A negative regression coefficient suggests a lower likelihood of “success” (here, a correct response); a positive coefficient suggests its higher likelihood. For numeric predictors the regression coefficient estimates the slope of the regression line associated with the predictor. The p-values (or t-values in models fitted to numeric dependent variables) enable evaluation of whether the contrast or the regression lines are significantly different from zero, based on the estimated regression coefficient and the associated standard error. Importantly, since the conditions are not matched on word frequency or length, we do not base our analyses on the raw mean accuracy percentages or response times (e.g. mean response times may be faster for prefixed words because these words are shorter than suffixed ones). Instead, we base our interpretations on the mean response times and accuracies as predicted by the regression models. For a detailed description of the method see Baayen (2008) and Jaeger (2008).

Each of the standardized tests (WASI, GORT- 4 and RAN) and experimental (Visual and Auditory Lexical Decision) tasks was analysed separately to obtain descriptive statistics and determine between-groups differences. Between-sessions changes in the lexical decision tasks and fluency measures of the GORT- 4 were also calculated for the reading disabled group as differences in per-subject means (RD2 – RD1). Lexical decision accuracy and response time measures were correlated with lexicality, testing session, and morphological type (*derived*, *compound* and *pseudo-complex*) and their interactions to investigate whether morphological training increases complex word recognition. The contribution of reading training was assessed by comparing the results of visual and auditory correlations of these measures.

Inferential analyses were used to develop models to investigate the relationship between standardized test scores and performance on lexical decision tasks. Additional multiple regression mixed effect models predicted GORT- 4 fluency changes based on lexical decision tasks, WASI and RAN scores and performance on the initial GORT- 4 assessment.

Results

WASI. Table 3 summarizes the IQ scores for each group and demonstrates that the typical readers performed better on both the verbal and performance tests resulting in higher overall IQ scores than the reading disabled participants. Linear regression models fitted to verbal and performance IQs separately with group as predictor and participant as a random effect, confirm that typical readers had significantly higher IQ scores than students in the reading disabled group, (verbal $\beta = 45.58$, SE = 0.26, t = 380.90; performance $\beta = 29.18$, SE = 0.17, t = 169.20; all *p*s < 0.0001).

Table 3. Descriptive IQ statistics: means, SDs (in brackets) and ranges for WASI sub-tests in the reading disabled and typical reading groups.

WASI Results

	Verbal IQ		**Performance IQ**		**Full IQ**	
Group	**M (SD)**	**Range**	**M (SD)**	**Range**	**M (SD)**	**Range**
RD1	70.88(12.61)	50 - 90	76.32(11.67)	57 – 70	71.61(9.89)	53 - 90
TYP	113.22(14.12)	82 - 136	105.553(11.87)	70 -	110.56(13.01)	82 - 132

RAN. Table 4 summarizes the performance (time in seconds to name grid) on RAN tasks for both the RD1and TYP groups. Linear regression models fitted individually to colour, digit, letter an object RAN tasks with group as predictor and participant as a random effect, confirm that typical readers were significantly faster at naming than the reading disabled group: colour β = -7.06, SE = 0.08, t = -92.60; digit, β = -2.60, SE = 0.05, t = 48.50; letter β = -3.25, SE = 0.05, t = 65.1; object, β = -5.97, SE = 0.08, t = 77.30; all *p*s < 0.0001).

Table 4. Descriptive statistics of naming times in seconds: means, SDs (in brackets) and ranges of RAN results for reading disabled and typical reading groups.

RAN Task

Group	**Colour**		**Digit**		**Letter**		**Object**	
	M (SD)	Range	M (SD)	Range	M (SD)	Range	M (SD)	Range
RD1	27.55(6.36)	15 - 36	17.53(3.87)	13 - 26		15 -		22 -
TYP	20.05(3.76)	14 - 30	15.03(3.43)	10 - 24	16.04(2.62)	12 -	23.37(3.12)	18 -

GORT- 4. Table 5 summarizes the GORT- 4 descriptive statistics for accuracy, rate and fluency for all three groups and comprehension statistics for RD1 and TYP groups. Comprehension was not assessed for RD2.

Table 5. Descriptive statistics: means (M), SDs and ranges of grade equivalencies for GORT- 4 sub-tests in reading disabled and typical reading groups.

GORT- 4

Group	Accuracy		Rate		Fluency		Comprehension	
	M (SD)	Range	M (SD)	Range	M (SD)	Range	M (SD)	
RD1	3.18(1.47)	1.4-8	3.64(1.58)	1.4-8	3.131(1.36)	1-7.7	4.22(2.41)	1 – 9
RD2	2.86(1.42)	1.2-7	3.79(1.91)	1-8	3.15(1.59)	1-7.4	-------------	--------
TYP	9.27(2.94)	3.2-12.7	9.92(2.36)	3-12.7	9.75(2.75)	3-12.7	9.95(2.47)	6 –

Linear regression models fitted individually to grade equivalency score for accuracy, rate, fluency and comprehension with group as predictor and participant as a random effect confirmed that typical readers were faster (rate), more accurate and had better comprehension (all *p*s < .0001). The combination of increased speed (significant) but decreased accuracy (non-significant) resulted in a non-significant increase in overall fluency in RD2 as compared to RD1, mean = 3.15, SD = 1.60; RD1, mean = 3.10, SD = 1.39 (see Table 6 for results of the regression model). This result will be discussed later in relation to similar findings for the visual lexical decision task, i.e., a significant decrease in visual response times with a non-significant decrease in accuracy.

Table 6. Summary of linear regression models of grade equivalency score for accuracy, rate, fluency and comprehension with group (RD1 & RD2) as predictor. Participant was a random effect for each of these models (SDs not reported). B stands for the estimated regression coefficient, SE is the standard error, and t and p report the t- and p-values of respective tests.

GORT -4

Test	B	SE	t	p
Rate	0.19	0.03	7.16	<.0001
Accuracy	-0.18	0.02	-8.29	<.0001
Fluency	0.02	0.02	1.1	=0.37

We further explored individual differences in how much fluency was gained between sessions RD1 and RD2 as a function of the individual's fluency in the first session RD1. Figure 1 plots the change in fluency, defined as the difference between the individual score in RD2 and RD1 session, against the individual score in RD1. Both the regression line (solid) and the scatterplot smoother lowess trend line (dotted) reveal, that in general, students with better pre-training fluency scores showed a larger post-training fluency improvement. A positive reliable Spearman's correlation between pre- and post-training fluency scores $\rho = 0.35$, $p<.05$ supported this finding. This suggests that the intensive training in the framework of the Wilson program shows differential benefits in reading fluency, depending on the pre-training levels: this finding may be useful when forming the cohorts to participate in the program. There are two reading interventions currently used in the school district, it may be that the students in the higher fluency range would derive more benefit from the academic program (PHAST: Lovett et.al., 2000) than from the Wilson.

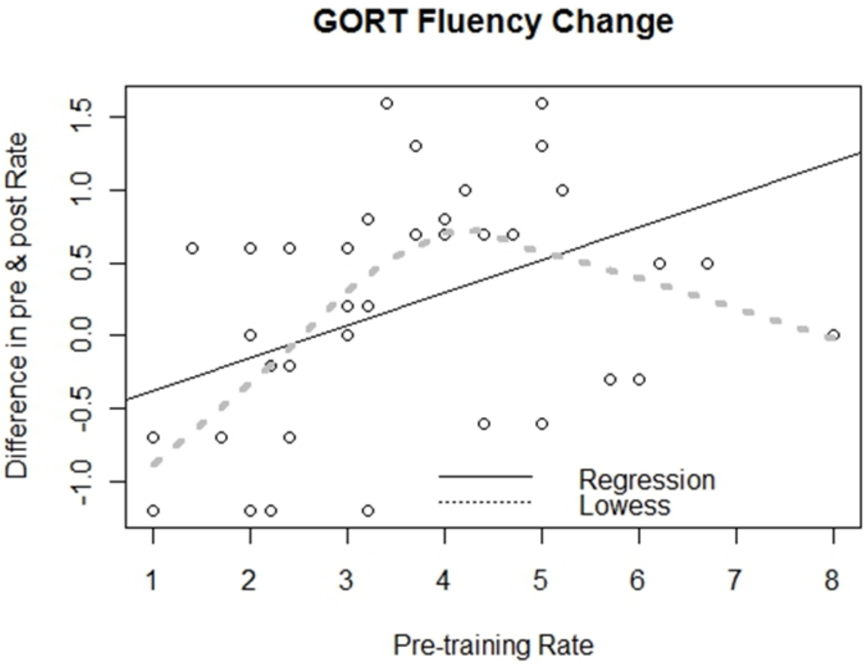

Figure 1. Change in fluency after training as a function of pre-training fluency, with the best fit regression line (solid and the scatterplot smoother lowess trend line (dotted).

Correlation analyses predicting GORT- 4 fluency changes from GORT- 4 comprehension, WASI and RAN tasks revealed that verbal IQ scores were the only predictors of an increase in fluency that approached significance. WASI verbal IQ was positively correlated with the increase in fluency, ρ = .44, p = .08. As with the correlation between better pre-training fluency scores and magnitude of post-training gain in fluency, students who had a higher verbal IQ benefited more from the training than the students with low verbal IQs. The non-significant probability value may be due to the small number of data points (20), and thus reduced statistical power. It is possible that with a larger sample the correlation between higher verbal scores on WASI and increased fluency would reach significance. To sum up, the training was most effective for the students who scored in the second and third quartile than for the ones who scored in the upper or lower quartile in pre-training fluency.

Lexical Decision.

Analyses were performed on data after the removal of any unexpected responses (hitting a key without the label R or A). This ‘clean-up’ resulted in a data loss of less than 1% in each modality. Linear regression models were fitted individually to visual and auditory accuracy scores (number of correct responses) and response times to examine the effects of group (*RD1, RD2 & TYP*), morphology, word length, word frequency (log value) and word training as fixed effects, with word and participant as random effects. Residuals of the mixed-effects models for lexical decision accuracy and response times were almost always skewed. To reduce skewness, we removed outliers from the respective datasets, i.e., points that fell outside the range of -2.5 to 2.5 units of SD of the residual error of the model. Once outliers were removed, the models were refitted with the same predictors for fixed effects, interactions and random effects. Less than 4% of the data were lost in each modality through this trimming (auditory = 3.6%, visual = 3.2%).

Accuracy. Table 7 summarizes the means and SDs for visual and auditory lexical decision accuracy (hits and correct rejections) and demonstrates that participants in all groups had a higher percentage of accurate scores in auditory compared to visual lexical decision for both trained and untrained words. The means for trained versus untrained words in RD1 and TYP were not significantly higher (see below).

Table 7. Descriptive statistics: percent correct means and SDs (in brackets) of auditory and visual accuracy by group.

Lexical Decision Accuracy by Modality

Modality	**RD1**		**RD2**		**TYP**	
	Trained	Untrained	Trained	Untrained	Trained	Untrained
Auditory Accuracy	90(30)	83(38)	88(33)	78(41)	94(24)	92(27)
Visual Accuracy	78(41)	70(46)	84(37)	75(43)	92(27)	89(31)

Logistic regression models were used to assess the likelihood that participants would be more or less accurate based on group (RD1, RD2 & TYP) and modality (visual or auditory) of the lexical decision task (Table 8). Analysis revealed that modality was a significant predictor of accuracy, with all readers being less accurate in visual than in auditory lexical decision $\beta = - 1.70$, SE = 0.12, z = - 14.70, $p < .0001$.Typical readers were significantly more likely to be accurate than the RD1 group, $\beta = 1.44$, SE = 0.53, z = 2.69, $p < .007$ in both modalities. All interactions between student group and lexical decision modality were significant at $p < .001$. The negative coefficient for RD2 indicates the non-significant decrease in post-training auditory accuracy as compared to the pre-training level.

Table 8. Summary of the logistic regression model fitted to the accuracy of auditory lexical decision. SD of participant as a random effect was 0.65 and SD of word as a random effect was 0.85.

Lexical Decision Accuracy by Modality Model

Predictor	Coefficients	Standard Error	Wald's Z	Pr(>|z|)
Intercept (RD1)	3.88	0.40	9.80	= .000
RD2	-0.97	0.54	-1.80	= .07
TYP	1.44	0.53	2.69	= .007
Visual modality	-1.70	0.12	- 14.70	= .000
RD2: Visual modality	1.17	0.16	7.46	= .000
TYP: Visual modality	0.62	0.20	3.19	= .001

Auditory Accuracy. The means and SDs of number of correct responses for the four morphological types of stimuli in the auditory lexical decision task are presented in Table 9.

Table 9. Descriptive statistics: percent correct means and SDs (in brackets) by morphological type, for auditory lexical decision accuracy in reading disabled and typical reading groups.

Auditory Accuracy

Groups	**Prefixed**	**Compound**	**Suffixed**	**Pseudo-**
RD1	0.90(0.37)	0.88(0.34)	0.86(0.37)	0.84(0.38)
RD2	0.87(0.39)	0.84(0.36)	0.82(0.39)	0.80(0.39)
TYP1	0.96(0.28)	0.91(0.26)	0.95(0.29)	0.91(0.31)

A logistic regression model fitted to trimmed auditory accuracy data (see above for the explanation of trimming procedures), with the multiple predictors word length, frequency, word training, group, session and morphological type confirmed that there were no significant differences in auditory accuracy between the groups (see Table10). Furthermore, a stable pattern of accuracy based on morphological type was revealed in the RD group at both tests: compound = prefixed > pseudo-complex = suffixed. Recognition of pseudo-complex and suffixed words was significantly less accurate than recognition of compound words and prefixed words (ps < .0001). Word length was not a significant factor, while log frequency did predict accuracy. The words in the program (trained words) may have been easier as there was a significant effect of word training type (trained or untrained words) on auditory accuracy (trained were more accurate). However, there was no interaction between RD2 and untrained words indicating that the training generalized to words not encountered in the reading program rather than only benefitting the recognition of words seen during the training program. It is difficult to interpret the interaction between word and morphological type for the typical readers because their accuracy was at or close to ceiling.

Table 10. Summary of the logistic regression model fitted to the accuracy of the auditory lexical decision. SD of participant as a random effect was 1.57 and SD of word as a random effect was 2.08. The intercept is responses for RD1 on trained, compound words.

Lexical Decision Auditory Accuracy Model

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.52	1.22	2.07	= .039
Word Length	0.00	0.12	0.03	= .973
log(FREQcount)	0.96	0.10	9.57	= .000
Untrained	-1.40	0.38	-3.67	= .000
SchoolSessionRD2	-0.58	0.60	-0.97	= .335
SchoolSessionTYP1	0.16	0.61	0.26	= .795
Morph-pseudocomplex	-2.75	0.51	-5.36	= .000
Morph-suffix	-2.34	0.61	-3.83	= .000
Morph-prefix	-0.87	0.56	-1.56	= .118
Trainedu:SchoolSessionRD2	-0.29	0.29	-1.00	= .318
Trainedu:SchoolSessionTYP1	0.87	0.37	2.34	= .019
SchoolSessionRD2:Morph-pseudocomplex	0.27	0.34	0.77	= .439
SchoolSessionTYP1:Morph-pseudocomplex	1.14	0.40	2.85	= .004
SchoolSessionRD2:Morph-suffix	0.05	0.41	0.12	= .904
SchoolSessionTYP1:Morph-suffix	2.11	0.61	3.45	= .001
SchoolSessionRD2:Morph-prefix	0.43	0.42	1.02	= .309
SchoolSessionTYP1:Morph-prefix	2.20	0.87	2.54	= .011

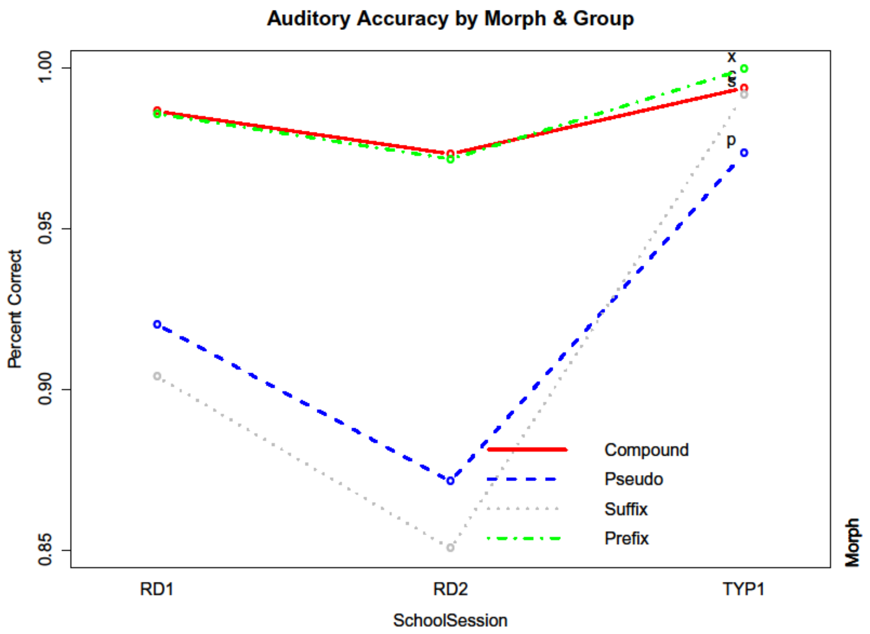

Figure 2. Auditory accuracy by group and morphological type based on the Logistic regression model reported in Table 10.

Visual Accuracy. Table 11summarizes means and standard deviations for each morphological type by modality and group. In the RD groups, a visual accuracy pattern for compound and derived words emerges of compound > prefixed > suffixed > pseudo-complex. Both TYP and RD1 groups are least accurate at processing pseudo-complex words. The typical readers are at or near ceiling on compound and the RD group are more accurate for compounds versus derived words.

Table 11. Descriptive statistics in percent correct: means and SDs (in brackets) by morphological type, for visual lexical decision accuracy in reading disabled and typical reading groups.

Lexical Decision Visual Accuracy

Visual Accuracy	Compound	Suffixed	Prefixed	Pseudo-complex
RD1	0.78(0.41)	0.76(0.43)	0.74(0.45)	0.70(0.46)
RD2	0.83(0.37)	0.80(0.41)	0.77(0.44)	0.77(0.42)
TYP1	0.90(0.28)	0.94(0.29)	0.93(0.33)	0.88(0.33)

A logistic regression model fitted to trimmed visual accuracy data (see statistical considerations), with multiple predictors confirmed the morphological accuracy pattern (Table 12). Responses to compound words were more likely to be accurate than the responses to pseudo-complex, suffixed and prefixed words (all *ps.* < .05) (the low mean visual accuracy in compounds for the TYP group resulted from outliers that were trimmed in the regression model). Word length was not a significant factor, but log frequency did predict accuracy. The model also demonstrated that typical readers were more likely to be accurate. There was a significant fixed effect of word training type (trained or untrained words) on visual accuracy, but no interaction between RD2 and untrained words indicating again that the morphological training generalized to the untrained words even in the critical visual modality.

Table 12. Summary of the logistic regression model fitted to visual accuracy of participant as a random effect was 0.87 and SD of word as a random effect was 1.11.

Lexical Decision Visual Accuracy Model

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.88	0.78	2.41	= .016
Word Length	-0.09	0.06	-1.46	= .145
log(FREQcount)	0.55	0.05	10.82	= .000
Untrained	-0.69	0.21	-3.37	= .001
SchoolSessionRD2	0.44	0.71	0.62	= .537
SchoolSessionTYP1	2.38	0.73	3.26	= .001
Morph-pseudocomplex	-2.10	0.27	-7.73	= .000
Morph-suffix	-1.03	0.33	-3.16	= .002
Morph-prefix	-0.63	0.31	-2.03	= .042
Trainedu:SchoolSessionRD2	-0.26	0.21	-1.25	= .211
Trainedu:SchoolSessionTYP1	0.22	0.26	0.82	= .412
SchoolSessionRD2:Morph-pseudocomplex	0.22	0.25	0.88	= .379
SchoolSessionTYP1:Morph-pseudocomplex	0.65	0.30	2.14	= .032
SchoolSessionRD2:Morph-suffix	-0.42	0.31	-1.34	= .181
SchoolSessionTYP1:Morph-suffix	0.84	0.43	1.96	= .050
SchoolSessionRD2:Morph-prefix	-0.30	0.31	-0.98	= .327
SchoolSessionTYP1:Morph-prefix	0.66	0.44	1.52	= .128

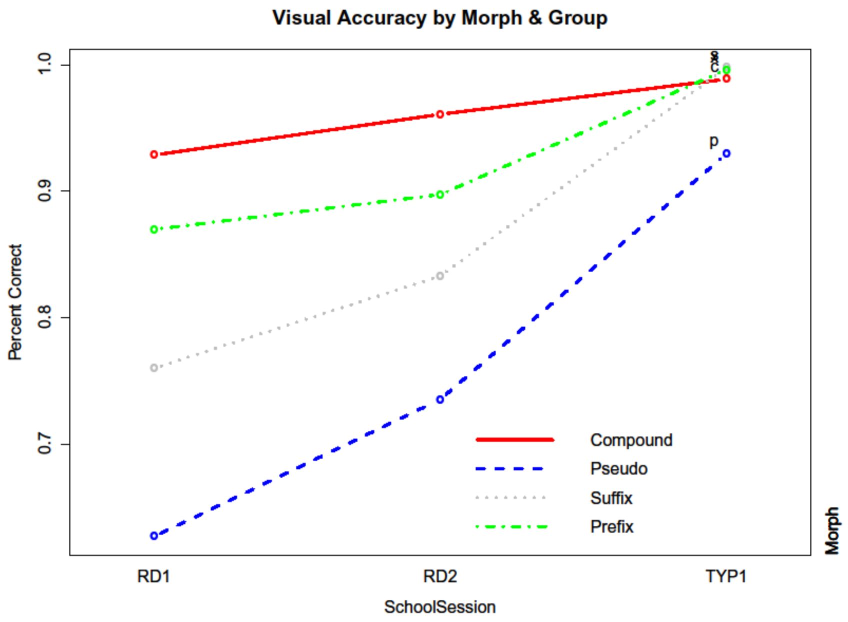

Figure 3. Visual accuracy by group and morphological type based on the Logistic regression model reported in Table 12.

A comparison of Figure 2 (auditory accuracy) and Figure 3 (visual accuracy) reveals that the increased accuracy for RD2 only occurred in the visual modality not in the auditory. This suggests that the change in visual accuracy was not due to a test practice effect, i.e. the repeated exposure to similar testing conditions. Additionally, Figures 2 and 3 demonstrate that the processing difficulty of each morphological type relative to other types remained largely stable, even though the hierarchies of processing difficulties varied between visual and auditory modalities.

Response times

Auditory RT. Table 13 summarizes means and SDs in milliseconds for auditory response times. The students in the RD groups were an average of 275 ms faster after training. Words with suffixes showed the biggest decrease in response times.

Table 13. Descriptive statistics in auditory RT: means and SDs (in brackets) by morphological type in reading disabled and typical reading groups.

Auditory Response Times

Group	Compound	Pseudo-complex	Suffixed	Prefixed
RD1	1496.43(912.75)	1528.55(991.34)	1577.01(995.98)	1511.36(883.17)
RD2	1243.46(707.18)	1237.66(783.24)	1224.79(778.51)	1309.40(911.74)
TYP1	866.61(444.10)	880.50(428.78)	849.65(455.67)	916.84(489.17)

A model fitted to trimmed auditory RT data with multiple predictors (see Table 14) showed that RD2 readers were faster than RD1 (β = -0.06, SE = 0.03, p = 0.004). There was no significant difference between TYP and RD1 (β = -0.018, SE = 0.03, p = 0.73) indicating that the RD group was not disabled in auditory processing.

As expected, less frequent and longer words had longer response times ($p < .0001$) and untrained words were processed more slowly than trained words (β= 0.01, SE = 0.01, p = 0.03). Additionally there was a pattern of faster response times for pseudocomplex and suffixed words: compound < suffixed = pseudo-complex < prefixed (all *ps* < .01). This finding may illustrate the preference for salient word onsets, i.e. suffixed over prefixed words, in auditory processing described in Hawkins and Cutler (1988): this point is discussed below.

Table 14. Summary of the logistic regression model fitted to auditory response times.SD of participant as a random effect was 0.29 and SD of word as a random effect was 0.83.

Lexical Decision Auditory RT Model

Predictors	Estimate	Std. Error	t value	p-value
(Intercept)	7.02	0.08	82.9	= .000
Word Length	0.03	0.01	6.33	= .000
log(FREQcount)	-0.05	0.00	-11.39	=.000
Untrained	0.07	0.02	3.33	= .001
SchoolSessionRD2	-0.16	0.10	-1.68	= .037
SchoolSessionTYP1	-0.45	0.09	-4.74	= .000
Morph-pseudo-complex	0.14	0.03	5.21	= .000
Morph-suffix	0.11	0.03	3.45	= .000
Morph-prefix	0.08	0.03	2.61	= .007
Trainedu:SchoolSessionRD2	-0.01	0.03	-0.45	= .668
Trainedu:SchoolSessionTYP1	-0.08	0.02	-3.48	= .001
SchoolSessionRD2:Morph-pseudocomplex	-0.06	0.03	-1.9	= .055
SchoolSessionTYP1:Morph-pseudocomplex	-0.01	0.03	-0.35	= .719
SchoolSessionRD2:Morph-suffix	-0.06	0.04	-1.71	= .087
SchoolSessionTYP1:Morph-suffix	-0.08	0.03	-2.15	= .031
SchoolSessionRD2:Morph-prefix	-0.03	0.04	-0.81	= .417
SchoolSessionTYP1:Morph-prefix	0.02	0.04	0.58	= .572

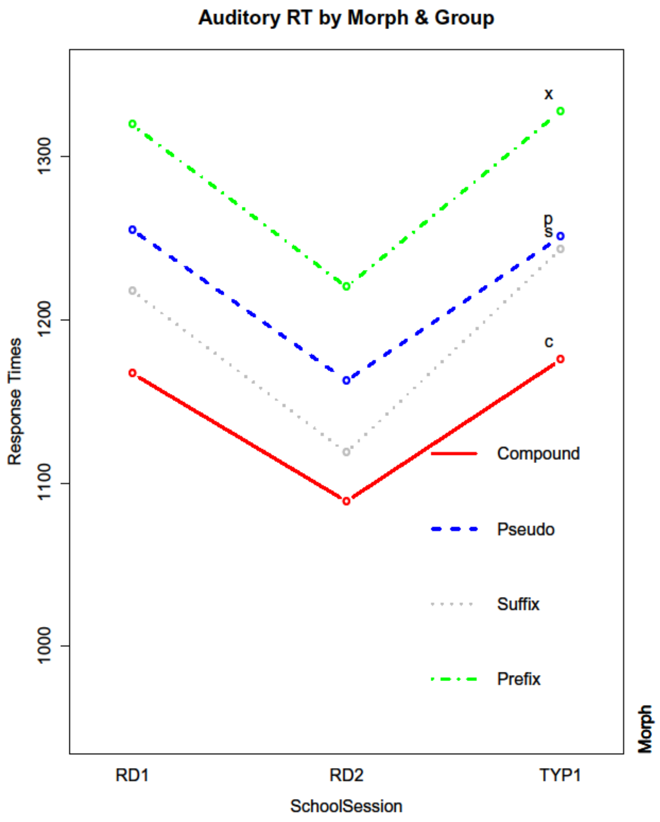

Figure 4. Auditory RT by group and morphological type based on the Logistic regression model reported in Table 14.

Visual RT. Table 15 summarizes means and SDs in milliseconds for visual response times. The students in the RD groups were an average of 311 ms faster after training. Words with suffixes showed the biggest decrease in response times.

Table 15. Descriptive statistics in visual RT: means and SDs (in brackets) by morphological type in reading disabled and typical reading groups.

Lexical Decision Visual RT

	Compound	**Suffixed**	**Prefixed**	**Pseudo-complex**
RD1	1470.55(871.84)	1598.41(1028.44)	1599.06(927.36)	1483.44(979.35)
RD2	1213.62(681.65)	1239.83(805.01)	1279.54(834.60)	1174.26(777.51)
TYP1	842.61(415.09)	835.59(448.07)	893.79(498.27)	843.10(372.95)

Table 16 shows the results of a linear mixed effects regression model fitted to visual response times with outliers greater than 2.5 SDs removed. Typical readers were faster than RD1 ($\beta = -0.47$; SE = 0.10; p = 0.000), and RD2 readers were faster than RD1 ($\beta = -0.02$, SE = 0.10, p= .037). Untrained words were processed more slowly but there was no interaction between the RD1 and RD2 groups, indicating this difference did not depend on testing session. The students had generalized the morphemic knowledge acquired in the training program to novel words. If there had been no generalization, the untrained words would have been slower than the trained words for RD2 but not RD1. Longer and less frequent words were processed more slowly, as expected (both ps <.0001).

The pattern of processing times: compound < suffixed = prefixed < pseudo-complex (see Figure 5) reveals that in the visual modality, compounds and derived words had a processing advantage over pseudocomplex words.

Importantly, RTs for compounds and affixed words decreased by the same amount across the two training sessions, and showed no sign of interaction with group.

Table16. Summary of the logistic regression model fitted to visual response times. SD of participant as a random effect was 0.30 and SD of word as a random effect was 0.09.

Lexical Decision Visual RT Model

	Estimate	Std. Error	t value	p-vals
(Intercept)	6.98	0.09	81.01	= .000
Word Length	0.04	0.01	7.96	= .000
log(FREQcount)	-0.05	0.00	-13.13	= .000
Untrained	0.05	0.02	2.55	= .010
SchoolSessionRD2	-0.20	0.10	-1.99	= .006
SchoolSessionTYP1	-0.47	0.10	-4.90	= .000
Morph-pseudo-complex	0.17	0.03	6.68	= .000
Morph-suffix	0.10	0.03	3.21	= .001
Morph-prefix	0.13	0.03	4.45	=.000
Trainedu:SchoolSessionRD2	0.04	0.02	1.88	=.070
Trainedu:SchoolSessionTYP1	-0.07	0.02	-3.33	= .001
SchoolSessionRD2:Morph-pseudocomplex	-0.10	0.03	-3.74	= .000
SchoolSessionTYP1:Morph-pseudocomplex	-0.01	0.02	-0.42	= .680
SchoolSessionRD2:Morph-suffix	-0.06	0.03	-1.77	= .078
SchoolSessionTYP1:Morph-suffix	-0.07	0.03	-2.42	= .017
SchoolSessionRD2:Morph-prefix	-0.06	0.03	-2.02	= .041
SchoolSessionTYP1:Morph-prefix	-0.04	0.03	-1.32	=.188

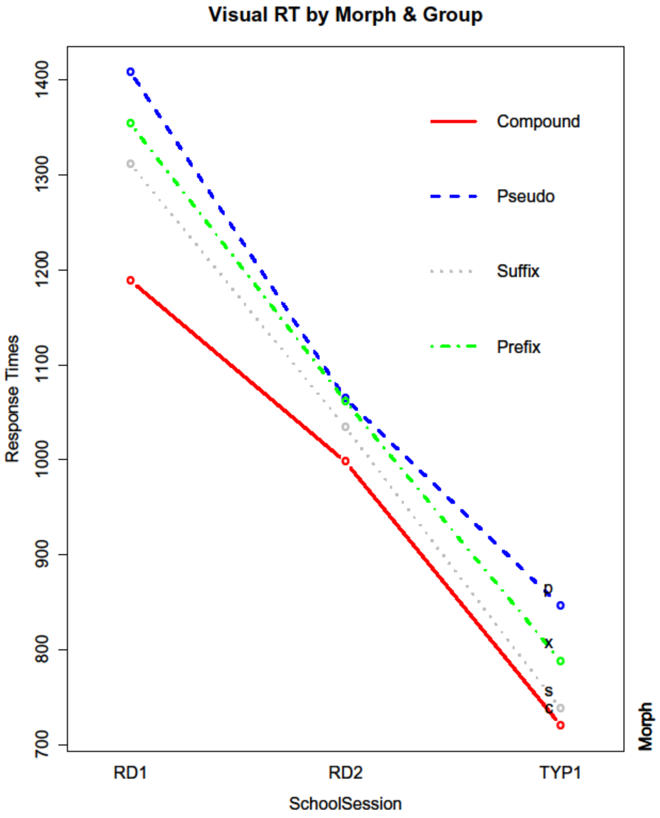

Figure 5. Visual RT by group and morphological type based on the linear multiple regression model reported in Table 16.

The lexical decision results will be discussed in more detail below but the following is a brief summary. The reading disabled group showed significant improvements in post-training response times and accuracy for visual but not auditory lexical decision. There were no significant between-group differences in either RT or accuracy based on morphological type.

Correlations between standardized tests and pre & post training lexical decision measures

A trimmed data set looking only at the RD1 group was used for the first set of correlational analyses. It was further split into visual and auditory datasets to enable analyses of factors correlated with increased accuracy and faster response times in each modality. Matrix plotting was performed on components with $r \geq .3$ to investigate relationships between initial accuracy and response time measures and scores on the standardized tests.

Visual. Figure 6 shows that although GORT fluency, colour, digit and letter RAN measures were correlated with pre-training visual accuracy in lexical decision. Only GORT fluency ($r = .91$, $p = .03$) and letter RAN ($r = -.88$, $p = .05$) reached significance. Correlations with colour ($r = -.86$, $p = .06$) and digit RAN ($r = -.87$, $p = .06$) were only marginally significant. None of the standardized measures were significantly correlated with visual response time.

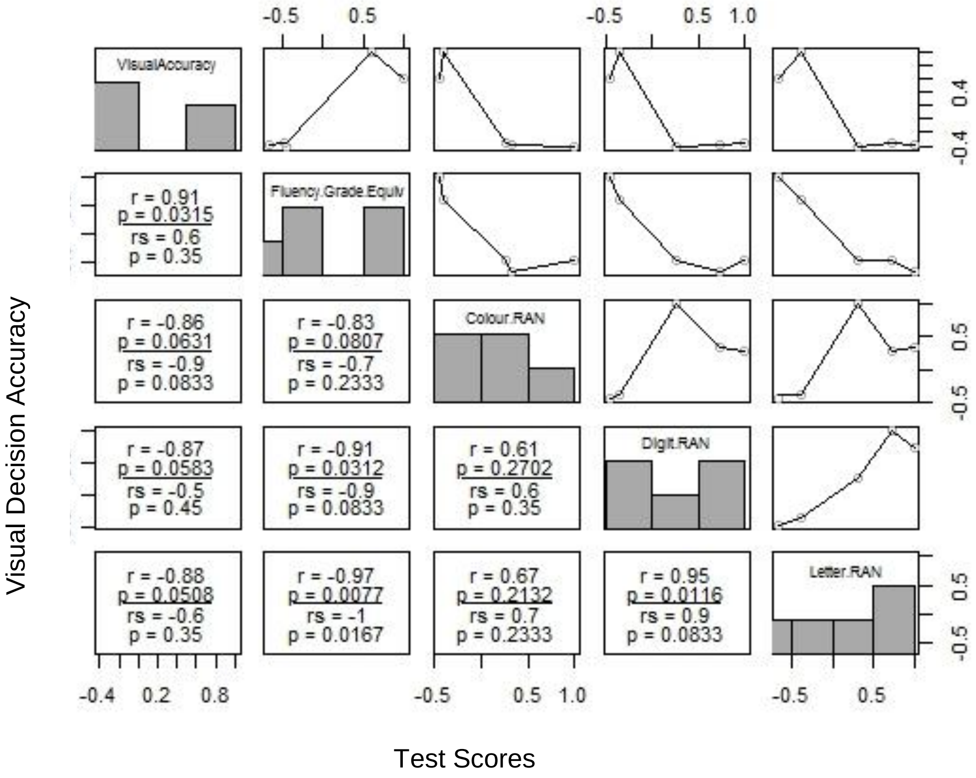

Figure 6. Along the diagonal are histograms of the following distributions: percent correct in visual lexical decision, GORT accuracy grade equivalency, RAN colour and RAN letter. Above the diagonal of the plot matrix are scatterplots of all pairs of variables along with lowess smoothing lines. Panels below the diagonal report the Pearson’s correlation coefficient r, the Spearman’s coefficient rho and the respective probability values.

Auditory. Figure 7 shows that although GORT fluency(r = .56), GORT comprehension (r = -.24) and WASI performance (r = .98) measures were correlated with pre-training auditory accuracy, only the WASI performance IQ value significance (p =.02). None of the standardized measures was significantly correlated with auditory response time.

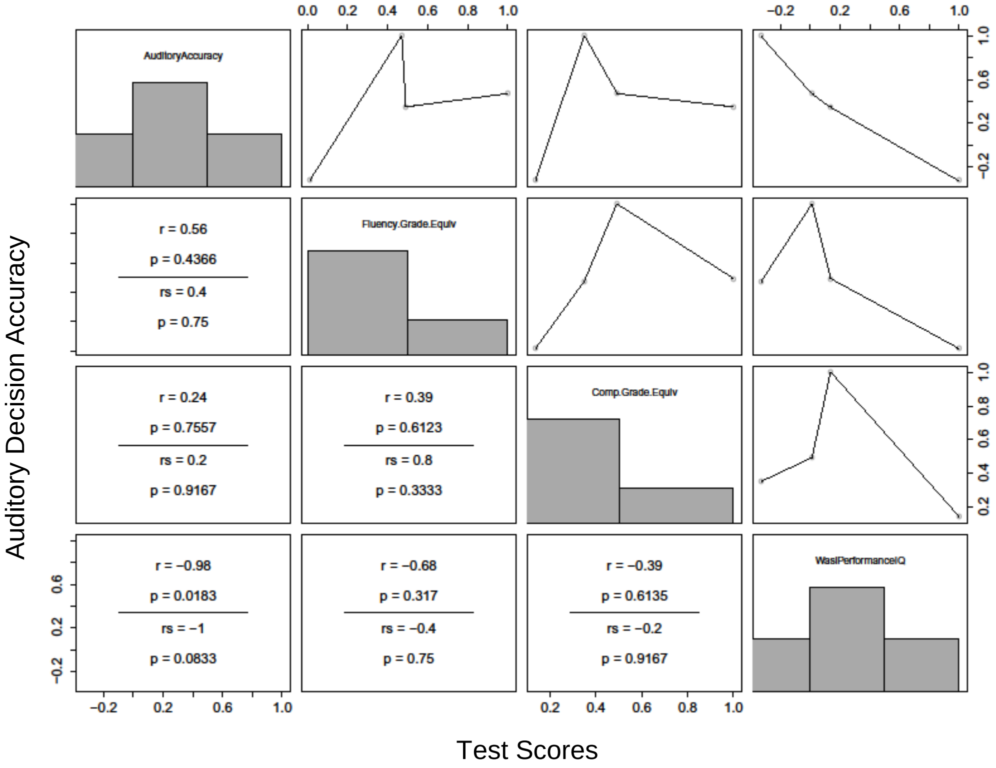

Figure 7. Along the diagonal are histograms of the following distributions: percent correct in visual lexical decision, GORT fluency grade equivalency, GORT comprehension grade equivalency and WASI performance IQ. Above the diagonal of the plot matrix are scatterplots of all pairs of variables along with lowess smoothing lines. Panels below the diagonal report the Pearson's correlation coefficient r, the Spearman's coefficient rho and the respective probability values.

We further explored individual differences in the impact of training (measured as the difference between the pre and post-training auditory and visual response accuracy and latency) as a function of the individual’s pre-training RAN, WASI and GORT-4 scores. There were no significant correlations between GORT- 4, WASI and RAN scores and differences between pre and post-training lexical decision accuracy and response times.

General Discussion

There were two main goals for this study. One was to evaluate the effects of eight-months of intensive morphological training on the reading fluency as well as visual and auditory processing of morphologically complex words in high-school students whose reading skills had remained at elementary school level. A second goal was to investigate the differences between types of morphological complexity and how they affect processing at different stages of the developmental trajectory, from the pre-training reading disabled students to the same students post-training performance and the performance of the typical students. To be able to characterize general learning capacity in the reading disabled intervention group, the cognitive level of the students was estimated with the short-form test WASI. A specific cognitive correlate of reading fluency was tested by variants of the rapid naming task RAN. Text reading fluency and comprehension were measured with GORT-4. The impact of training on single word processing was tested in auditory and visual lexical decision tasks.

Intervention Success

First, and most importantly for the students, there was improvement in both word recognition and fluency as a function of training. As discussed above, these students are at a point in their lives when their ability to function independently and enjoy success in life is to a large extent dependent on their reading skills. The findings that even as late as high-school, measureable improvements in reading fluency are possible means that, with additional intervention, the life path of these adolescents could be positively altered. There were no systematic relationships between prior IQ or naming and the effect of the 6-month training. In other words, the RD group demonstrated positive effects from the training regardless of the individual starting point. The results of the WASI, GORT and RAN tasks confirm the reading deficits of the students enrolled in the reading intervention program, but even students with the lowest IQ estimates showed post-intervention improvements in lexical decision speed and accuracy.

Reading fluency and comprehension. The RD group GORT-4 passage reading performance both pre- and post-training was, on average, six grade levels below their age-matched peers. Although most of the students enrolled in the reading intervention demonstrated increased post-training fluency, as Figure 1 demonstrates, the participants with either very low or relatively high scores did not show an improvement in fluency. A review of individual scores did not point to any obvious reasons – the students who did not improve in fluency did not have either the lowest or highest scores on any of the other measures. With such a small sample size it is difficult to pinpoint the reasons for this non-significant trend. There was greater improvement in response time than in accuracy and this may relate to automatization and generalization of letter-phoneme mapping and morpheme knowledge discussed below in the discussion of lexical decision results.

Lexical Decision. The results of this study suggest that morphological training leads to improved word processing but only in the visual modality. In Figure 2 (auditory accuracy) the slope of the lines between RD1 and RD2 is very shallow, illustrating the lack of significant change between pre and post training auditory accuracy. Accuracy for compounds and prefixes was close to ceiling, but this was not the case for suffixes and pseudo-complex words. Although RD2 groups was faster in auditory lexical decision, the finding that there was no significant improvement in accuracy indicates that the participants got better at planning and executing the motor response (pressing the response key), not at auditory word recognition. However, Figure 3 which represents visual accuracy shows a different picture. In Figure 3, the lines between RD1 and RD2 have a steep positive slope demonstrating the significant increase in post-training visual accuracy. Also, the largest gains were in visual response times, indicating increased automaticity of word recognition. Figures 4 (auditory response times) and 5 (visual response times) both show a decrease in response times, but closer examination reveals a 300 ms average decrease in visual but only a 200 ms average decrease in auditory decision times. LaBerge and Samuels (1974) described reading automaticity as the ability to build internal representations or codes needed to perform some of the tasks needed for fluent reading. These codes in turn are thought to free up attentional resources, resulting in more efficient processing. In the current study, it appears that the participants in the reading program developed internal representation of letter groups, reducing the amount of attention needed to perform part of the word recognition task resulting in an overall reduction in resources needed and consequently faster word processing. In other words, the word recognition process became more automatic in the visual modality. This increased automaticity may have been a factor in the improved post-training reading rates for the GORT-4 texts, too. Alternatively, the increased rates may have been a result of reading practice.

The improvement in visual word processing was not limited to the words encountered in the training program. Table 12 illustrates that there was no significant interaction between pre and post-training visual accuracy for the RD group and whether the words were trained or not. A post-training improvement in both trained and untrained words implies that the RD students were able to generalize their acquired knowledge of grapheme-phoneme mappings and morphological processing to novel words. Table 16 illustrates that although the untrained words were processed more slowly than the trained words, response times for both decreased after training, providing evidence that the RD group were able to apply their new knowledge to novel words. The training that the RD group was enrolled in emphasized the phonological and orthographical changes resulting from the addition of an affix or additional base word (in the case of compounds). Wolf and Segal (1999) found that participants trained on word meanings in context were able to use their acquired knowledge to understand the meaning of semantically related associates encountered in similar contexts. Given that a particular affix often has the same meaning in one derived word as in another it may be this semantic relatedness that is responsible for the generalization across words.

The findings of the current study and the semantic relatedness results of the Wolf and Segal (1999) study are in line with the lexical quality hypothesis (Perfetti & Hart, 2002). According to this hypothesis, the strength of a word's representation or lexical quality is dependent on the strength of, and connections between, a word's morphological, phonological, orthographic and semantic representations. Additionally, improving lexical quality results in decreased response time and increased accuracy of word recognition. The intervention program provided information about the meanings of individual morphemes and the orthographic and phonological changes that occur in derivational and compounding processes. This kind of training would increase the lexical quality of both the free and bound morphemes encountered during the intervention, resulting in improved accuracy and reduced time to respond when they are encountered again. The untrained words contained either the same bound morphemes or used similar connections and so the students were able to use their lexical knowledge to recognize the novel words, resulting in a generalization of the acquired knowledge. The ability to generalize morphemic knowledge to novel words may have been a factor in the increased fluency found in the post-training GORT-4 results.

Predictors of Improvement

Cognitive capacity and reading improvement. The RD groups low verbal IQ scores could be due to the symbiotic nature of verbal IQ as measured by WASI verbal subtests and the reading deficits diagnosed through GORT-4 and Woodcock–Johnson III (WJ III; Woodcock, 2001), previously administered through the school. All these measures are closely related to vocabulary size. This has resulted in an on-going argument in the learning disabilities literature about the reciprocal relationship between verbal IQ and reading ability (Stage et al., 2003). In particular, children with smaller vocabularies tend to do poorly on both verbal IQ and reading tasks. However, the present intervention group did not have markedly better performance compared to verbal scores. Yet, only the verbal WASI scores showed a trend towards predicting improvement in reading fluency. Thus, IQ as such did not appear to modulate intervention success in this study but vocabulary size may have. I can only speculate on the reasons for the low performance IQ estimates. It could be that the RD group are cognitively impaired or it may be partly related to test anxiety. Stanovich (1986) found that as children with reading problems became more anxious about reading they reduced the amount that they read, leading to increased reading difficulty and increased anxiety in a vicious circle. Stanovich (1986) argued that this negative cycle can generalize beyond the original problem of reading into other areas of cognitive function. Fluency improvement was moderately related to pre-training fluency.

That the RD group had smaller vocabularies than the typical group is evident in their performance on the auditory lexical decision task. If the students in the RD group had vocabularies similar to the referent group we would expect similar scores in the auditory lexical decision task because orthographical knowledge is not needed for this task. However, as discussed earlier in the section on auditory lexical decision results, the RD readers were less likely to accurately recognize an auditorily presented word in comparison to the typical readers. It would be interesting to see if there is an increase in

IQ scores after successful reading interventions. This would suggest that low verbal IQ scores may be related more to reading deficiencies than to actual cognitive impairment.

It has been shown that low verbal scores in can be caused by retrieval deficits rather than small vocabulary (Wolf & Segal, 1999). Retrieval deficits occur, for instance, during tip-of-the-tongue states, when a person knows a word, but cannot access or retrieve it. They also occur when a person *knows* a word well enough to recognize the meaning, but cannot produce the meaning out of context. In a study of reading intervention outcomes for a group of severely reading disabled 13-year-olds, Wolf and Segal (1999) presented evidence of retrieval deficits rather than lack of word knowledge. They discuss results in which children are unable to give the correct meaning of a word in expressive vocabulary tests like the WASI verbal subtests but demonstrate knowledge of the same words during a multiple choice task.

RAN and reading improvement. As discussed earlier in section on reading acquisition theories, slow naming times, particularly for letters and digits, are highly correlated with reading deficits (Denckla & Rudel, 1976; Savage et al., 2007; Wolf & Bowers, 1999). Additionally, some research points to object naming being more strongly correlated with reading comprehension (Wolf & Obregón, 1992) while letter and digit naming may be more predictive of reading rate. In the current study, the RD participants had low scores on all measures of RAN and reading. However, RAN performance did not predict improvement in reading fluency or lexical decision performance, suggesting that the effectiveness of the intervention program was not affected by cognitive factors underlying serial naming speed. This is compatible with views that assume rapid naming impairments to be unrelated to the phonological processing impairments typically found in dyslexia. For instance, a recent eye-tracking study found a relationship between the ability to efficiently program eye-movements and reading skill level (Kuperman, Van Dyke & Henry, 2012). When both the articulatory and semantic components of RAN were removed (by varying the task), leaving only the programming of eye-movements to rapidly engage and disengage from the grid items, the results still explained 13% of the unique variance in natural reading. It may be that in addition to a lack of phonological awareness, the RD students also have difficulty in engaging and disengaging from sequentially presented items.

Morphology

When accuracy and response times were viewed in relation to morphological types, previously unreported results were revealed about the pattern of processing difficulty. First, visual lexical decision accuracy and response time results showed a stable ordering of processing difficulty across morphological types, irrespective of reading level. Second, there was a distinct advantage for compound words over derived and pseudo-complex words in the visual lexical decision task, in that these words were processed more quickly and accurately than words from other morphological types.

The observed advantage for compound words in both modalities may be best explained by the increased amount of semantic information carried by the morphemes and the whole word. Ji, et. al. (2011) suggested that multiple sources of semantic information provided by the presence of two stems boost whole word recognition. They found that both transparent (rosebud) and opaque (hogwash) English compounds were processed faster than simplex (giraffe) words (Ji, et.al., 2011). The results of the current experiment provide additional evidence that compound word processing receives a boost from multiple semantic sources. Theories involving more than one processing route for word recognition (i.e. Baayen, Dijkstra & Schreuder, 1997; De Jong, Schreuder, & Baayen, 2000; Morris et. al. 2007; Perfetti, 1992; Schreuder, & Baayen, 1997), argue that whole word processing is faster and more accurate than decomposition, however, many studies have shown that words are decomposed (e.g. Beyersmann, Castles & Coltheart, 2011; El Yagoubi et. al., 2008; Lavric, Rastle & Clapp, 2011; Marslen-Wilson & Tyler, 1997; Morris et. al., 2007; Nation & Cocksey, 2009; Taft and Forster, 1976). The faster and more accurate processing of compound words across skill levels provides evidence that knowing the meaning of one morpheme in a complex word can assist a reader in accessing the whole word meaning.

Visual. The pattern of visual lexical decision accuracy was compound > prefixed > suffixed > pseudo-complex. Participants were equally fast in responding to words with suffixes than to words with prefixes, resulting in the following response pattern of RT latencies: compound < suffixed = prefixed < pseudo-complex. The pattern of increased accuracy and decreased response times that we found for visually presented complex words dovetails nicely with the growing body of evidence discussed previously (see the sections on lexical quality and decomposition – processing cost or benefit), that both derived and compound words have a processing advantage over simplex (mono-morphemic) words during visual presentation. This processing advantage for complex words over simplex ones has been observed in a number of languages and tasks. For example, in a lexical decision task comparing simplex and derived words, Bertram et al. (1999) reported faster response times for Finnish derived words; Burani and Thornton (2003) for Italian and Fiorentino and Poeppel (2007) for English. In an eye-tracking and naming study, Inhoff, Briihl and Schwartz (1996) found that participants had longer first fixation durations but faster response times for both English compounds and suffixed words in comparison to simplex words.

In the current experiment pseudo-complex words were used instead of simplex words, but these should be even more difficult because they have no structure and contain a foil lexeme. The word *exact*, one of the pseudo-complex word stimuli, provides a good example of the increased processing difficulties. The *ex* in *exact* might lead a reader to the erroneous conclusion that this is a derived word, additionally there is a lexeme foil *act*, and the individual representations of both *ex* and *act* will need to be supressed in order to access the correct meaning of *exact*. The increased possibility of error would cause a decrease in accuracy, while the need to suppress the foil would result in longer response times. This is what we found for both the typical and reading disabled groups. In contrast to pseudo-complex words, the individual morphemes in complex words may provide a processing boost rather than hindrance (Ji, et. al., 2011).

A well-known theory about the relative ease of processing morphological types is called the *Suffixing Preference*. According to this theory developed by Cutler, Hawkins & Gilligan (1985), suffixes have a processing advantage over prefixes in the hierarchy of morphological processing of complex words. This theory is discussed earlier, so the following is just a brief summary of the arguments. Based on typology and syntax (see the preference for suffixing section of the introduction), the *Suffixing Preference* account argues that: 1. complex words are decomposed into morphemes, 2. word stems are more informative than affixes and thus the preference is to have stems in the more salient word onset position, and 3. this preference leads to a processing advantage for suffixed over prefixed words.

To our knowledge, an experimental comparison of responses to suffixes and prefixes has not previously been reported. We provide empirical evidence for the suffix preference in auditory complex word processing but not in visual. Contrary to the results expected from the *Suffixing Preference* account, we found that in the visual lexical decision task, responses to prefixed words were more accurate and equally fast. Since the tendency across languages is to prefer suffixes (the preference for suffixing section of the introduction) the stem+affix combination is more expected than the affix+stem one. It is possible that the results might have differed if suffix frequency and productivity had been taken into account when developing the stimuli. However, this was not practical for this study because the trained word stimuli came from the Wilson program.

While the advantage of prefixed over suffixed words in visual word recognition is a topic for further research, one explanation of the discrepancy may follow from the nature of the task. Auditory processing imposes a unidirectional temporal ordering of access to morphemes, but during visual presentation, the whole word may be accessible at once (subject to the restrictions of visual acuity, Rayner, 2009). Evidence from event related potential studies (e.g., Hauk et. al., 2006; Holcomb & Grainger, 2006) show that word onsets are salient even in visual processing, but in visual word presentation a person may begin reading and processing a word from the beginning, end or middle.

Auditory. We used the auditory condition as a control for practice effects associated with repeated testing of RD, and for the between-groups comparison on a linguistic activity (speech comprehension) that is only partly related to literacy level. For this reason, our stimuli were not matched across groups on specific auditory word properties, such as the position of the uniqueness points (the point in time that the auditory signal is compatible with only one word) or the frequencies of biphones, especially those straddling constituent boundaries (e.g., uphill). Similarly, it is likely that our stimuli contained naturally occurring prosodic cues that differentiate the pronunciation of a stand-alone word ("fuse") and that same word as a segment of a larger word ("fuselage”). Thus, the present results in the auditory modality need to be treated with caution.

The pattern of auditory accuracy was compound = prefixed > pseudo-complex = suffixed. Compound words retained their advantage when measured by response time, but the prefixed and suffixed words changed places in the order of response by morphological type resulting in a pattern of compound < suffixed = pseudo-complex < prefixed.

Given the results of the visual lexical decision task and the evidence of a processing boost provided by multiple stems described in Ji, et.al., (2011) we expected the advantage found for compound word processing in the auditory modality. The responses for compound words were both the fastest and most accurate (on par with prefixed words).

In the auditory lexical decision task, the advantage of suffixed over prefixed words shows up where it is predicted, i.e. in the latencies of correct responses to existing words. This supports Cutler's Suffixing Preference account: since the stem contains more information about the word and appears earlier in suffixed words than in prefixed ones, the former are more common and easier to process. While recognized faster than prefixed words, suffixed words elicit fewer accurate responses, both in the visual and the auditory modality. This appears to run counter to Cutler's claims of the relative processing ease of suffixed words. Note, however, that Cutler's claims are specific to speech comprehension and to the task of comprehending existing words. They need not hold true for the ease of discriminating existing (prefixed or suffixed) words from their non-word counterparts.

Unlike the visual modality, lexical decision to pseudowords was neither the slowest nor the least accurate, among morphological types. A likely explanation for the discrepancy between modalities is that articulatory and prosodic cues associated with pronunciation of foil stems within pseudo-complex words do not allow for their identification as free morphemes. When a compound word like *trombone* is presented as an auditory stimulus, the co-articulation of labial [m] and [b] in *trombone*, does not allow for the identification of *bone* as a free morpheme. Similarly the differing stress pattern in *postpone* versus *postman* prevents the identification of *post* as a separate word stem in *postpone* but not in *postman*. This contrasts with visual stimuli in which no formal cue exists to discriminate between a pseudoword (mister) and a genuine complex word (twister). In the visual domain then, when activated, the meanings of the foil morphemes have to be suppressed in order to arrive at the correct word interpretation.

Conclusions and Future Directions

One of the questions that we sought to answer was whether the processing order of morphological types was skill dependent and/or modality dependent. Our results indicate that the processing order of morphological types is not skill dependent because the order of processing difficulty was stable across morphological types, irrespective of reading levels. There were however, modality based processing differences for derived and pseudo-complex words. Compound words showed a distinct processing advantage in both auditory and visual modalities. Ji, Gagné and Spalding (2011) found a distinct advantage for compound words in visual lexical decision with adult participants and the current study expands these findings to include auditory lexical decision and a younger participant group. Processing of complex words was both more accurate and faster when the words were presented visually. Future research with younger participants is needed to determine when this advantage for derived words begins. Contrary to our findings, Carlisle and Katz (2006) found that accurate reading of derived words was skill dependent. The poor readers in both 4th and 6th grade were significantly less accurate in reading derived words than the better readers in those grades. The poor readers in grade six were also less accurate than the good readers in 4th grade. There was however a significant task difference in the two studies: in our study the participants only had to determine if the stimulus was or was not a word, whereas in the Carlisle & Katz (2006) study the children had to name the word. As noted in our discussion of the morphology results in auditory lexical decision, additional morphemes can change the articulation of the stem and the stress pattern may be on the bound morpheme instead of the stem. The increased errors by the poor readers in the study might have been a result of phonological deficits overcoming the semantic boost from multiple morphemes. Future research with younger readers and multiple tasks can help to determine the point at which the multiple morphemes embedded in derived words becomes a boost rather than a hindrance. Additional research with high school level readers will tell us if these results generalize to more natural tasks. For example a replication of the Carlisle & Katz (2006) with high school instead of primary school readers and eye-movement studies to provide information about the order of processing difficulty during text reading.

The difference in processing order between modalities suggests that lexical decisions are performed in fundamentally different ways based on mode of presentation. It appears from the response time data, that a unidirectional temporal processing of morphemes occurs in auditory but not in visual word recognition. Further research is needed to understand what underlies the difficulty order in the two domains.

Finally, despite the very low verbal IQ and naming scores of some of the individuals in the reading group there did not appear to be any systematic differences that prevented individuals in the poor reading group from benefiting from the training program. The adolescents in this study became more proficient at word recognition and this generalized to words beyond their specific training program resulting in increased reading fluency. This suggests that it is not too late for reading intervention in high school.

References

Ackerman, P. T., Holloway, C.A., Youngdahl, P. L., & Dykman, R. A. (2001). The double-deficit theory of reading disability does not fit all. *Learning Disabilities Research and Practice*, 16, 152–160.

Adlof, S. M. Perfetti, C. A. & Catts, H. C. (2011). Developmental changes in reading comprehension: Implications for assessment and instruction. In S. J. Samuels & A. E. Farstrup (Eds.), *What research has to say about reading instruction* (pp. 186–214). Newark, DE: International Reading Association.

Arnbak, E., & Elbro, C. (2000). The effects of morphological awareness training on the reading and spelling skills of young dyslexics. *Scandinavian Journal of Educational Research*, 44, 229–251.

Baayen, R. H. (2008). Analyzing Linguistic Data. *A Practical Introduction to Statistics Using R*. Cambridge University Press.

Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language*, 37, 94–117.

Bertram, R., Laine, M., & Karvinen, K. (1999). The interplay of word formation type, affixal homonymy, and productivity in lexical processing: Evidence from a morphologically rich language. *Journal of Psycholinguistic Research*, 28, 213–226.

Beyersmann, E., Castles, A. & Coltheart, M. (2011). Early morphological decomposition during visual word recognition: Evidence from masked transposed-letter priming. *Psychonomic Bulletin & Review*, 18, 937-942.

Boersma, Paul & Weenink, David (2012). Praat: doing phonetics by computer [Computer program]. Version 5.3.32, retrieved 17 October 2012 from http://www.praat.org/

Bowers, P. G., & Wolf, M. (1993). Theoretical links among naming speed, precise timing mechanisms and orthographic skill in dyslexia. *Reading and Writing: An Interdisciplinary Journal*, 5, 69 – 85.

Bruck, M. (1992). Persistence of dyslexics' phonological awareness deficits. *Developmental Psychology*, 28, 874–886.

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, 41, 977-990.

Burani, C., Marcolini, S., De Luca, M., & Zoccolotti, P. (2008). Morpheme-based reading aloud: Evidence from dyslexic and skilled Italian readers. *Cognition*, 108, 243-262.

Butterworth, B. (1983). Lexical representation. In B. Butterworth (Ed.), *Language production* (pp. 257–294). San Diego, CA: Academic Press.

Carlisle, J. F. (2010). Effects of instruction in morphological awareness on literacy achievement: an integrative review. *Reading Research Quarterly*, 45, 464-487.

Carlisle, J. F. & Katz, L. A. (2006). Effects of word and morpheme familiarity on reading of derived words. *Reading and Writing*, 19, 669–693.

Carlisle, J. F., & Nomanbhoy, D. M. (1993). Phonological and morphological awareness in first graders. *Applied Psycholinguistics*, 14, 177–195.

Carlisle, J. F., & Stone, C. A. (2003). The effects of morphological structure on children’s reading of derived words in English. In E. Assink & D. Sandra (Eds.), *Reading complex words: Cross language studies*. New York: Kluwer.

Carroll, J. B., Davies, P., & Richman, B. (1971). *Word frequency book*. New York: American Heritage.

Castles, A., & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition*, 91, 77–111.

Catts, H. W., Adlof, S. M., & Weismer, S. E. (2006). Language deficits in poor comprehenders: a case for the simple view of reading. *Journal of Speech, Language, and Hearing Research*, 49, 278-293.

Clark, E. V. (1982).The young word-maker: A case study of innovation in the child's lexicon. In E. Wanner and L. R. Gleitman (Eds.), *Language acquisition: The state of the art*. (pp. 390-425). Cambridge: Cambridge University Press.

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

Cutler, A., Hawkins, J. A. & Gilligan, G. (1985). The suffixing preference: a processing explanation. *Linguistics*, 23, 723-758.

Deacon, S. H., & Kirby, J. R. (2004). Morphological: Is it more than phonological? Evaluating the roles of morphological and phonological awareness in reading development. *Applied Psycholinguistics,* 25, 223–238.

Denckla, M. B., & Rudel, R. G. (1976). Naming of objects by dyslexic and other learning- disabled children. *Brain and Language* 3, 1-15.

de Jong, N. H., Schreuder, R. & Baayen, R. H. (2000). The morphological family size effect and morphology. *Language and Cognitive Processes,15,* 329–365.

Dryer, M. S. (2011). Prefixing vs. Suffixing in Inflectional Morphology. In: Dryer, Matthew S. & Haspelmath, Martin (Eds.) *The World Atlas of Language Structures Online*. (chapter 26). Munich: Max Planck Digital Library. Available online at http://wals.info/chapter/26 Accessed on 2012-10-15.

Ehri, L. C., (1992). Reconceptualizing the development of sight word reading and its relationship to recoding. In P. B. Gough, L. C. Ehri & R. Treiman, Rebecca (Eds.), *Reading acquisition* (pp. 107-143). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

El Yagoubi, R., Chiarelli, V., Mondini, S., Perrone, G., Danieli, M. & Semenza, C. (2008). Neural correlates of Italian nominal compounds and potential impact of headedness effect: An ERP study, *Cognitive Neuropsychology,* 25, 559-581.

Fiorentino, R., & Poeppel, D. (2007). Compound words and structure in the lexicon. Language and Cognitive Processes, 22, 953–1000.Foorman, B. R., Francis, D. ,Fletcher, J. M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology*, 90, 37-55.

Fowler, A. E., & Liberman, Y. I. (1995). The role of phonology and orthography in morphological awareness. In L. Feldman (Ed.), *Morphological aspects of language processing* (pp. 157–188). Hillsdale, NJ: Erlbaum.

Goswami, U. (1993). Toward an interactive analogy model of reading development: Decoding vowel graphemes in beginning reading. *Journal of Experimental Child Psychology*, 56, 443–475.

Grainger, J. & Holcomb, P. J. (2009). Watching the word go by: On the time-course of component processes in visual word recognition. *Language and Linguistics Compass*, 3, 128-156.

Grainger, J., Lété, B., Bertand D., Dufau S., Ziegler, J. C. (2012). Evidence for multiple routes in learning to read. *Cognition*, 123, 280-292.

Guyer, B. P., Banks, S. R., & Guyer, K. E. (1993). Spelling improvements for college students who are dyslexic. *Annals of Dyslexia*, 43, 186 -193.

Hawkins, J. A. & Cutler, A. (1988). Psycholinguistic Factors in Morphological Asymmetry. In J. A. Hawkins, Ed., *Explaining Language Universals*, Basil Blackwell, Oxford.

Hoover, W. & Gough, P.B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal*, 2, 127–160.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, 59, 434–446.

Ji, H., Gagné, C. L. & Spalding, T. L. (2011). Benefits and costs of lexical decomposition and semantic integration during the processing of transparent and opaque English compounds. *Journal of Memory and Language*, 65, 406–430.

Katzir, T, Kim, Y.S., Wolf, M., Morris, R., & Lovett, M.W. (2008). The varieties of pathways to dysfluent reading: comparing subtypes of children with dyslexia at letter, word, and connected text levels of reading. *Journal of Learning Disabilities,* 41, 47-66.

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods 42*(3), 627-633.

Kirby, J. R., Parrila, R., & Pfeiffer, S. (2003). Naming speed and phonological processing as predictors of reading development. *Journal of Educational Psychology*, 95, 453–464.

Kuperman, V., Bertram, R., & Baayen, R. H. (2008). Morphological dynamics in compound processing. *Language and Cognitive Processes*, 23, 1089-1132.

Kuperman, V., Schrueder, R., Bertram, R., & Baayen, R. H. (2009). Reading polymorphemic Dutch compounds: Toward a multiple route model of lexical processing. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 876–895.

Kuperman, V. and Van Dyke, J.A. (2011). Individual differences in visual comprehension of morphological complexity. In L. Carlson, C. Hoelscher, and T. Shipley (Eds.), *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society* (pp. 1643-1648). Austin, TX: Cognitive Science Society.

LaBerge, D., & Samuels, J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 6, 293–323.

Lavric, A., Rastle, K. & Clapp, A. (2011). What do fully visible primes and brain potentials reveal about morphological decomposition? *Psychophysiology*, 48, 676–686.

Liberman, L.Y., & Shankweiler, D. (1985). Phonology and the problems of learning to read and write. *Remedial and Special Education*, 6, 8-17.

Liberman, I. Y., Shankweiler, D., Fischer, F. W., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology*, 18, 201–212.

Lovett, M. W., Lacerenza, L., Borden, S. B. (2000). Putting struggling readers on the PHAST track: a program to integrate phonological and strategy-based remedial reading instruction and maximize outcomes. *Journal of Learning Disabilities*, 33, 458-476.

Lovett, M. W., Lacerenza, L., Borden, S. L., Frijters, J. C., Steinback, K. A., & De Palma, M. (2000). Components of effective remediation for developmental reading disability: combining phonological and strategy-based instruction to improve outcomes. *Journal of Educational Psychology*, 92, 263-283.

Lovett, M. W., Lacerenza, L., De Palma, M., Benson, N. J., Steinbach, K. A., & Frijters J. C. (2008). Preparing teachers to remediate reading disabilities in high school: What is needed for effective professional development? *Teaching and Teacher Education,* 24, 1083–1097.

Lovett M.W., Steinbach, K.A., & Frijters, J.C. (2000). Remediating the core deficits of developmental reading disability: a double deficit perspective. *Journal of Learning Disabilities*, 33, 334-358.

Lukatela, G., Carello, C., & Turvey, M. T. (1987). Lexical representation of regular and irregular inflected nouns. *Language and Cognitive Processes*, 2, 1–17.

Manelis, L., & Tharp, D. A. (1977). The processing of affixed words. *Memory & Cognition*, 5, 690–695.

Marslen-Wilson, W. D. & Tyler, L. K. (1997). Dissociating types of mental computation. *Nature* 387, 592–594.

McDougall, S., Brunswick, N., & de Mornay Davies, P. (2010). Reading and dyslexia in different orthographies: An introduction and overview. In McDougall, S., Brunswick, N., & de Mornay Davies, P. (Eds), *Reading and dyslexia in different orthographies* (pp. 3-21). N.Y., New York, England: Psychology Press.

Merkx, M., Rastle, K. & Davis, M. H. (2011). The acquisition of morphological knowledge investigated through artificial language learning. *The Quarterly Journal of Experimental Psychology, First*, 1- 21.

McPeak, L., & Trygg, L. (2007). The Secondary Literacy Instruction and Intervention Guide. Mill Valley, CA: *Stupski Foundation*. Retrieved from www.centeroninstruction.org.

Morris, J., Frank, T., Grainger, J. & Holcomb, P. J. (2007). Semantic transparency and masked morphological priming: An ERP investigation. *Psychophysiology*, 44, 506-521.

Nagy,W., Berninger, V. W., & Abbott, R. D. (2006). Contributions of morphology beyond phonology to literacy outcomes of upper elementary and middle school students. *Journal of Educational Psychology*, 98, 134–147.

Nation, K., & Cocksey, J. (2009). Beginning readers activate semantics from sub-word orthography. *Cognition*, 110(2), 273-278.

Nation, K. & Snowling, M. J. (1998). Semantic processing and the development of word-recognition skills: Evidence from children with reading comprehension difficulties. *Journal of Memory and Language*, 39, 85–101.

Nooteboom. S. G. (1981). Lexical retrieval from fragments of spoken words: beginnings versus endings. *Journal of Phonetics* 9, 407-424.

Norton, E. S. & Wolf, M. (2012). Rapid Automatized Naming (RAN) and reading fluency: implications for understanding and treatment of reading disabilities. *Annual Review of Psychology,* Vol. 63: 427-452.

Olson, R. K., Wise, B. W., Ring, J., & Johnson, M. (1997). Computer-based remedial training in phoneme awareness and phonological decoding: Effects on the post training development of word recognition. *Scientific Studies of Reading*, 1, 235–253.

Penney, C. G. (2002). Teaching Decoding Skills to Poor Readers in High School. *Journal of Literacy Research*, 34, 99-119.

Perfetti, C. A. (2011). Phonology is critical in reading: But a phonological deficit is not the only source of low reading skill. In S. A. Brady, D. Braze, & C. A. Fowler (Eds.), *Explaining individual differences in reading* (pp. 153-171). New York: Routledge.

Perfetti, C. A. (1992). The representation problem in reading acquisition. In P.B. Gough, L.C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 145-174). Hillsdale, NJ: Erlbaum.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Vehoeven. C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189-213). Amsterdam/Philadelphia: John Benjamins.

Scammacca, N., Roberts, G., Vaughn. S., Edmonds, M., Wexler, J., Reutebuch, C. K., & Torgesen, J. K. (2007), Interventions for adolescent struggling readers: A meta-analysis with implications for practice. Portsmouth, NH: RMC Research Corporation, Center on Instruction.

Shankweiler, D, Liberman, I. Y., Mark, L. S., Fowler, C. A., & Fischer, F. W. (1979). The speech code and learning to read. *Journal of Experimental Psychology: Human Learning and Memory*, 5, 531–545.

Shankweiler, D., Lundquist, E., Katz, L., Stuebing, K.K., & Fletcher, J.M., Brady, S., Fowler, F., et al. (1999) Comprehension and decoding: patterns of association in children with reading difficulties. *Scientific Studies of Reading* 3, 69–94.

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

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of visual word recognition and naming. *Psychological Review*, 96, 523–568.

Schreuder, R., & Baayen, R. H. (1997). How complex simple words can be. *Journal of Memory and Language*, 37, 118–139.

Siegel, L. (2008). Morphological awareness skills of English language learners and children with dyslexia. *Topics in Language Disorders*, 28, 15-27.

Singson, M., Mahony, D., & Mann, V. (2000). The relation between reading ability and morphological skills: Evidence from derivation suffixes. *Reading and Writing*, 12, 219–252.

Somers, M., Corrin, W., Sepanik, S., Salinger, T., Levin, J., Zmach, C., & Silverberg, M. (2010). The enhanced reading opportunities study final report: The impact of supplemental literacy courses for struggling ninth-grade readers (NCEE 2010-4021). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U. S. Department of Education.

Stanovich, K. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21, 360–407.

Statistics Canada. (2003). *Adult literacy skills and life skills survey - 2003*. Ottawa, ON: Statistics Canada.

Swanborn, M., & de Glopper, K. (2002). Impact of reading purpose on incidental word learning from context. *Language Learning, 52*(1), 95–117.

Stanovich, K. E. (1981). Relationships between word decoding speed, general name-retrieval ability, and reading progress in first-grade children. *Journal of Educational Psychology* 73, 809-15.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21, 360–407.

Taft, M., & Forster, K. *I. (1976).* Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior, 15, 607–620.*

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1994). Longitudinal studies of phonological processing and reading. *Journal of Learning Disabilities*, 27, 276–286.

Tsesmeli, S. N., & Seymour, P. H. K. (2006). The effects of training of morphological structure spelling derived words dyslexic adolescents. *Reading and Writing*, 19, 587–625.

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, 2–40.

Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Pratt, A., Chen, R. S., & Denckla, M. B. (1996). Cognitive profiles of difficult to remediate and readily remediated poor readers: early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology*, 88, 601–638.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101*, 192–212.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). Comprehensive test of phonological processing. Austin, TX: PRO-ED.

Wiederholt, J. L., & Bryant, B. R. (2001). Gray Oral Reading Test-IV.Austin; TX: Pro-Ed: 2001a.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI).* San Antonio, TX: Harcourt Assessment.

Wilson, B. (1998). The Wilson Reading Method, Part 1, *Learning Disabilities Journal* 8, 12-13.

Wolf, M., & Bowers, P. G. (1999). The double deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology*, 91, 1–24.

Wolf, M., O’Rourke, A. G., Gidney, C., Lovett, M., Cirino, P., & Morris, R. (2002). The second deficit: An investigation of the independence of phonological and naming-speed deficits in developmental dyslexia. *Reading and Writing: An Interdisciplinary Journal*, 15, 43–72.

Wolf, M. & Segal, D. (1999). Retrieval-rate, accuracy, and vocabulary elaboration (RAVE) in reading-impaired children: A pilot intervention program, *Dyslexia: An International Journal of Theory and Practice,* 5, 1–27.

Ziegler, J. C., Grainger, J. & Brysbaert, M. (2010). Modelling word recognition and reading aloud. *European Journal of Cognitive Psychology*, 22: 5, 641 — 649

Ziegler & Goswami (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. Psychological Bulletin, 1313–29.

Appendices

A1. Word stimuli by training type (trained /untrained) and morphological type.

Trained Words		**Untrained Words**	
Compounds	Pseudo-complex	Compound	Pseudo-complex
bagpipe	exact	aircraft	against
baseball	index	anthill	arsenal
baseline	panic	bankbook	stunt
basketball	admire	bathroom	button
bedtime	advise	beehive	buzzard
campfire	entire	bookcase	candidacy
caveman	escape	bookshop	command
clambake	exotic	bootlace	cryptic
classmate	invite	cheesecake	department
clothespin	ticket	clubhouse	dialog
cupcake	umpire	coffeepot	discern
drugstore	volume	courthouse	electron
fireball	athlete	cowhide	extract
fishline	collect	farmhouse	freeze
fishplate	collide	fieldmouse	irony
flagpole	compare	flowerpot	fuselage
frostbite	compete	folktale	galaxy
gatepost	concoct	footbridge	gluten
grapevine	confuse	glasshouse	heaven
handshake	dictate	gravestone	helmet
landslide	pollute	hairline	early
hemline	reptile	hairpiece	jerkin
catfish	vampire	hairpin	number
hotcake	complete	hornpipe	parenthesis
kingsize	concrete	jawbone	mister
lifetime	cosmetic	lampshade	plaintiff
manhole	costume	lifeboat	planet
muskmelon	district	lighthouse	pulpit
pancake	postpone	motorcycle	rabbit
pinhole	stampede	mousetrap	signet
ringside	trombone	nighttime	smuggle
rosebud	establish	notebook	stubborn
salesman	intellect	penknife	studio

selectman	recognize	porthole	surgeon
software	valentine	racetrack	tactile
springtime	accomplish	roommate	textile
stovepipe	contribute	sandstone	trolley
sunrise	distribute	crabcake	trumpet
whalebone	callisthenic	shoelace	twinkle
wildlife	contemplate	blueberry	villain
Prefix		**Prefix**	**Suffix**
disinfect	active	disagree	basic
dislike	massive	disallow	angelic
imprison	athletic	disbelief	creative
incomplete	basement	indiscreet	metallic
inconsistent	fantastic	intone	movement
inexpensive	magnetic	misuse	assertive
infiltrate	pavement	invaluable	detective
inside	assistant	involuntary	directive
intake	attendant	misbehave	effective
invalid	attentive	mislead	placement
misconduct	combatant	misquote	judgement
misfit	expensive	replace	coolant
recommend	frustrate	transform	defendant
transatlantic	impulsive	undo	attractive
undid	statement	undress	collective
unless	disruptive	unearth	accountant
unravel	expressive	unease	consultant
unsafe	impressive	unequal	contestant
uphill	inhabitant	uneven	destructive
unzip	postponement	unfold	nourishment

A2. Non-words stimuli matched with word stimuli by training type (trained /untrained) and morphological type.

Trained Non-words		**Untrained Non-words**	
Compound Trained	Pseudo-Complex	Compound Untrained	Pseudo-Complex
vaylape	emaft	arlcraid	aginshett
bemtice	inneb	athhime	ancetal
caughan	dinic	bartwook	stuse
cugcawl	apmite	bawsvoom	bottin
huptine	adwint	beazike	bezzand
fottuke	entase	boakwase	callibagy
mauwode	edpape	baokshom	commeeth
lunnarn	etonim	boodloll	creplic
linnole	invose	cheavecale	denintment
russgud	lecket	plubhoose	discove
limmose	umpord	loffiezat	emastron
banchulu	vosuke	coultwouse	explart
bampteck	arelesh	culhews	explave
cassnire	collinx	fassheese	queemp
drombace	collyll	feandcouse	idety
litthace	combire	frowervat	fupelirm
frangpite	compise	folfcale	nalegy
gathbost	concerk	foddbrexed	plulen
gralkhine	conwase	glirenouse	hounen
vandgike	dirbate	blaulstene	hildet
tandscide	mollunt	haitrine	eadly
hedtibe	replive	mairmiews	jowkin
gatlich	vimpind	hustlon	gudber
chotcik	copplene	hothpiff	pamenclufis
prozdize	conprote	jaurane	cisler
limptill	cosronic	larmchade	plairtirl
minroys	costoft	fimloat	sparet
pungbelon	distruke	pightcouse	pudpog
pardake	pothbene	motansygle	pobbit
pibwull	stampids	mouchstap	pidnet
rirenide	chombore	vightleeme	phoggle
rumpbod	estailect	nimedool	stubbect
sursemal	entellorm	yunswike	chusio

plocktume	regoprize	poldhule	surbept
sodgwase	gomentine	raulstrad	tactave
spriretirt	accomflect	poammand	lexlile
stordpiff	confridunt	rairsteene	spolley
fonrere	dengribute	kradzeek	chummet
whowsbine	munichtalic	sheilafe	sminsle
wirtlibe	conteoplent	blofserry	vellark
Prefix	**Suffix**	**Prefix**	**Suffix**
diefincet	actigs	disatroo	cabic
senlich	mammive	desallop	ancosic
etshodod	ameletic	dasmeloof	ancomic
inseystint	bansment	indinclout	sevellic
egcostoonent	fundestic	undord	mongment
invargrent	sutcetic	anfalid	ammertive
unbape	palftine	unquinable	deveotive
unbisk	assessite	infiluntary	dicuctism
unbaft	atteshans	ponbekave	ebbactive
unkvilk	attentics	bisneed	plangdent
sencongolt	compinant	misfrete	jasthment
ponbit	espenrive	replert	compinist
ruksupend	spopplate	trarefown	degastant
stinslupranic	imbukize	unsa	amprective
urmon	stathtine	umchess	cottoctive
ortest	diswoptive	unooked	actooshant
olcoyet	exbinnive	uniant	connollant
olcocet	imbluctive	unedral	montipsate
udjart	inbepitant	unewan	desplectize
onwey	possmitement	unfark	weenment