UNIVERSITY OF READING

How do reading and listening to stories facilitate vocabulary acquisition?

Thesis submitted for the degree of Doctor of Philosophy

School of Psychology and Clinical Language Sciences

Alessandra Valentini

August 2017

DECLARATION OF ORIGINAL AUTHORSHIP

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Alessandra Valentini

Table of contents

List of Figures

List of Tables

Acknowledgements

My heartfelt thanks to my supervisors Carmel Houston-Price, Jessie Ricketts and Rachel Pye, who have been wonderful mentors, and have provided me with lots of advice and guidance throughout my PhD. Special thanks to Julie Kirkby also, who helped me immensely for the third study. Thanks to Ludovica Serratrice as well, who has been very supportive in my final stages of writing up.

Throughout this PhD I worked with many schools and many children, and I'd like to thank them all for helping me in this project. I'd also like to thank Rhona Stainthorp, who agreed to be the 'voice' of my stories.

I'd like to thank Evelyn Egger, my office-mate, for all the chats, suggestions and encouragement, and Laura Bozicevic, for talking through all the ups and downs of this journey over coffees and lunch. I'd also like to thank Julia Hofweber, George Pontikas, and all those in the department, for all the insightful chats and advice, and especially Abhijeet Patra and Louis Williams for their help, and for sharing this journey with me to the very end. Huge thanks to Nicky Dawson, for helping me with recordings and for all the input, and Jo Johnston, Hazel Mycroft and Emily Hancock for their support and for being excellent sources of wisdom throughout the last year of my PhD. Thanks to Davide Crepaldi and Claudio Luzzatti, who encouraged me to start this journey in the first place.

Finally I'd like to thank all my friends and family for all their support throughout this journey. Many many thanks to Andrea as well, who endured all my stresses, and always reminded me to take it all one step at a time.

To my grandmother, Irene Pedrazzi, who taught me to read.

Abstract

Reading and listening to stories foster vocabulary development (Elley, 1989; Nagy, Anderson, & Herman, 1987; Wilkinson & Houston-Price, 2013). Studies of single word learning in literate children suggest that new words are more likely to be learnt when both their oral and written forms are provided, compared to when only one form is given (Ricketts, Bishop, & Nation, 2009). This thesis explores children's learning of phonological, orthographic and semantic information about words encountered in a story context, comparing performance in different story presentation modalities. Specifically, Year 4 children were exposed to new words embedded within stories in three possible conditions: listening (Studies 1 & 2), reading (Studies 2 & 3), and simultaneous listening and reading ('combined condition' - Studies 1, 2 & 3).

Children learnt the orthographic forms of the words only when exposed to them (reading & combined conditions), but showed reliable semantic and phonological learning in all conditions. Children showed similar phonological learning in all conditions, demonstrating that phonology is automatically generated from orthography. In contrast, some measures revealed better semantic learning in the combined condition, showing both phonological and orthographic facilitation effects.

In the third study we explored the nature of the advantage of the combined condition for semantic learning, examining children's eye-movements to compare their allocation of attention to the text in the combined and the reading conditions. In the combined condition children spent less time reading the new words, as well as learning more new word meanings, compared to the reading condition. This suggests that presenting words in two modalities simultaneously confers a learning advantage by freeing attentional resources.

In conclusion, Year 4 children learn word meanings better when able to listen to stories while reading them. The advantage of the dual modality of presentation may partly be due to this condition freeing attentional resources.

Chapter 1: Literature review

A person's vocabulary is the set of words that they know. Learning new words is an unconstrained skill (Paris, 2005) and it is a process that starts in infancy, and continues throughout a person's life (Duff, Reen, Plunkett, & Nation, 2015; McGregor & Duff, 2013). Considerable research has focused on vocabulary acquisition in the early years of a child's life, while comparatively fewer studies have explored how vocabulary acquisition evolves with the changing linguistic environment that children encounter at school, particularly when they start to read. This thesis explores vocabulary acquisition from stories in school-aged children. Specifically, the main aim of this work was to elucidate how presenting stories in different modalities (i.e. oral, written and oral and written modalities simultaneously) would affect the acquisition of word meanings, and phonological and orthographic forms. This literature review will start by considering the concept of vocabulary knowledge, describing theoretical approaches to word knowledge and vocabulary acquisition. It will then summarise the main findings on vocabulary acquisition from oral and written modalities, first separately, and then in comparison. The effect of combining written and oral presentation will then be explored, in comparison to single modality presentations. Section 1.6 will then briefly explore other factors which, in addition to presentation modality, affect vocabulary acquisition. Finally Section 1.7 will introduce the present research.

1.1 – Vocabulary

1.1.1 - Defining vocabulary knowledge

Defining what we mean when we talk about "word knowledge" is not easy: learning a word is not all or nothing, but an incremental process that takes place in many steps (Nagy & Scott, 2000). For example, Carey (1978) distinguished between fast mapping, which is the initial ability of children to form a partial representation of a word's meaning from context, and slow mapping, which is the process through which children are able to improve and perfect their existing knowledge of a word with repeated exposure in different settings. Most research considers fast-mapping the product of a mutual exclusivity or contrastive process, where children select the referent or the meaning of a new word by considering, and excluding, all the other referents present at encounter for which they already have a name (c.f. Booth, McGregor, & Rohlfing, 2008; Smith & Yu, 2008). Slow-mapping, on the other hand, is often considered as the process which consolidates knowledge over time, irrespective of further exposure (Horst, Parsons, & Bryan, 2011), and helps to integrate the word within the lexicon (Tamura, Castles, & Nation, 2017). Some researchers (Dale, 1965; Paribakht & Wesche, 1997) have proposed a stage-like model of word learning, where the first step is the very first encounter with a new word. In the second stage there is familiarity with the word’s form without explicit knowledge of the meaning, while in the third there is a partial representation of meaning and recognition in context. The fourth stage is characterised by explicit knowledge, and the fifth by appropriate knowledge and use. The first three stages might be included in the term ‘fast-mapping’, as they require relatively few exposures (Durso & Shore, 1991). The following refining of initial knowledge in stages four and five corresponds to slow mapping, and requires much longer (Beck, Perfetti, & McKeown, 1982). Given these differing frameworks for describing the process of acquiring a new word, in the present work we consider the terms ‘fast’ and ‘slow’ mapping to indicate more generally the initial creation of a representation of a word at first encounter, irrespective of the learning mechanism, and the further refining of that word’s representation through subsequent presentations.

Taking into account the incremental property of vocabulary learning is particularly important when considering incidental learning, the ability to infer new word meanings from context quickly and with no direct instruction (Rice, 1990; Schatz & Baldwin, 1986). The concepts of breadth and depth of vocabulary knowledge are clearly linked to the concept of incrementality. Breadth of vocabulary knowledge can be defined as the number of words in the lexicon, while depth of vocabulary knowledge corresponds to the amount of semantic information retained regarding a single entry (Ouellette, 2006). It could be supposed that, while vocabulary breadth could be increased through fast mapping of new words, slow mapping would be required to increase the depth of vocabulary knowledge.

Models from the memory literature, such as the complementary learning system account (McClelland, McNaughton, & O'Reilly, 1995) are also relevant to the description of vocabulary knowledge and the vocabulary acquisition process. This model describes two memory systems: the neocortical memory system, characterised by overlapping representation of word forms, where new input triggers existing nodes, and the hippocampal system, which features sparse and context sensitive representations, not well interconnected with each other. Information first enters memory through the hippocampal system, and is only integrated into the neocortical systems and existing vocabulary knowledge through sleep consolidation (James, Gaskell, Weighall, & Henderson, 2017). Several studies have shown that sleep has a positive effect on breadth of vocabulary knowledge: words are in fact recalled more successfully after a period of sleep (Dumay & Gaskell, 2007; Williams & Horst, 2014). Sleep also has the effect of increasing the integration of new words within the lexicon: after sleep, new words compete with existing vocabulary, affecting the online processing of known words that are similar to the new one (Dumay & Gaskell, 2007; Henderson, Weighall, Brown, & Gaskell, 2012).

Word knowledge is also multidimensional: for each vocabulary entry, we must learn the word’s spoken and written forms, grammatical aspects, collocation (which words tend to co-occur with it), frequency (how likely it is that the word would be used), stylistic register (in what kind of communication would be the word used), meaning and association with other words (Nation, 2013). Each of these aspects is learned incrementally over multiple exposures, and it is not possible to directly predict a learner’s knowledge of one of these features from his knowledge of another (Schmitt, 1998). Perfetti and Hart (2002) particularly underline the importance of the link between different aspects of word knowledge. The Lexical Quality Hypothesis, proposed by these authors, defines a “high quality representation” of a word as a representation that combines the word’s orthographic representation (its written form), its phonological representation (its oral form) and its semantic representation (its meaning), and that is sustained by strong links between these features. In Perfetti and Hart’s view, having a high quality representation of a word makes it possible to rapidly access any of these features of the word, given only one of them. For example it is possible to access the meaning and the spoken form of a word from its written form, an ability that is fundamental for effective reading (this framework will be described in more detail in Section 1.3.2).

1.1.2 - Representations in the mental lexicon

Most theories assume that words are represented in memory in several units of analysis: each item is stored in a phonological lexicon, where the word's oral form is stored (Storkel & Rogers, 2000), a semantic lexicon, where the word's abstract meaning is stored (Dell, 1986; 1988; Murphy, 2003), and an orthographic lexicon, where a word's written form is represented independently of its oral form (Olson, Forsberg, & Wise, 1994), but is connected to both the phonological and semantic lexicon (Ehri, 1992). Once the orthographic lexicon is formed it can be accessed automatically (Steffler, Varnhagen, Friesen, & Treiman, 1998), without the need for phonological recoding. While the phonological and the semantic lexicons both have their origin in the first years of a child's life, learning to read is fundamental for the creation of a stable orthographic lexicon. The effect that learning new words through different modalities can have on a word's representation will be explored in further sections, while here we will focus on the organisation of the mental lexicon, with a focus on the adult lexicon, and the process of accessing the created mental representation.

Several models have been proposed to describe the organisation of the adult lexicon, and how this lexicon is accessed (e.g. Coltheart, 2005; Forster, 1976; Levelt, 1989; Marslen-Wilson & Tyler, 1980; McClelland and Elman, 1986; Morton, 1979; Seidenberg & McClelland, 1989).

One fundamental difference between the models that describe the organisation of the lexicon is in the nature of the representations stored within them. Models differ in the level of integration proposed between the various aspects that constitute the mental representation of a word. Some models highlight the importance of the separate representation of form and meaning (Levelt, 1989), and these different stores are often assumed to support comprehension and production respectively (Aitchison, 2012). Conversely, other models, especially models of lexical access, emphasise the importance of the integration of the different features of the word's representation (Seidenberg & McClelland, 1989). Furthermore, models of access to the mental lexicon differ regarding the expected level of analysis necessary for word recognition. Some accounts postulate that a word form is mostly accessed through bottom-up processes, through the sequential analysis of its parts (e.g. Forster, 1976), while others postulate that top-down processes participate in item selection, predicting that features of the broader context that surrounds a word might affect its selection (e.g. McClelland and

Elman, 1986). The relative importance of bottom-up and top-down processes in models of lexical access can inform theories of lexical organisation. For example, models that stress the importance of bottom-up processes postulate different levels of representation for word forms and meanings (Forster, 1976). In this review we consider models that specifically deal with word forms, either oral or written, and meanings, since the creation of the link between forms and meanings of a word is fundamental to any study of word learning. Theories of semantic organisation that do not consider word forms, such as the hierarchical network model (Collins & Quillian, 1969), the spreading activation model (Collins & Loftus, 1975), or the componential approach (Smith, Shoben, & Rips, 1974), are therefore not considered in this review.

As a model of the organisation of the mental lexicon, Levelt's model (1989) distinguishes between two components of the lexicon: a semantic component, the 'lemma', which includes information about the word's meaning and collocation, and the 'lexeme', which contains information about word form. In this model the production of a word proceeds in fixed stages from meaning (lemma) to word form (lexeme), and the representation of the two are completely separate. Similarly, Forster's (1976) serial search model, a model of lexical access, postulates the existence of separate access files for each item, corresponding to the word's orthographic and phonological forms, and a master file containing all stored information regarding the word, including its phonology, orthography and meaning. Once the item has been accessed within the access file through phonological or orthographic information, the master file can be retrieved. This model therefore emphasises the importance of bottom-up processes and the completion of an initial phonetic or orthographic analysis before contextual information can be taken into account in the selection of the item. Likewise, Morton (1979), in his latest modifications to the logogen model, proposed the existence of separate logogen systems, where different types of information about a word are stored. For example, he proposed a phonological logogen system separate from an orthographic system.

Unlike the previous models, connectionist models tend to describe the mental lexicon as a network in which lexical access arises from the activation of interconnected nodes. The previously described models typically postulate that a specific level of analysis must be completed before another level of analysis can begin: the search within the lexicon proceeds serially. Connectionist models, however, propose that several kinds of information can be considered in parallel during the word recognition process.

Furthermore, parallel distributed processing models see word representations not as stored in a single symbolic unit, but as distributed across several nodes, such that "knowledge of words is embedded in a set of weights on connections between processing units encoding orthographic, phonological, and semantic properties of words, and the correlations between those properties" (Seidenberg & McClelland, 1989, p. 560). In one of these models, Gaskell and Marslen-Wilson (1997) propose that oral information activates a network that contains a distributed representation of both the word's phonology and its meaning. Two separate stores for phonological and semantic information are proposed, but these are triggered simultaneously while listening to the speech input.

The previously described models account for how the lexicon is analysed to retrieve information regarding the item presented, in other words, recognising a word and retrieving its meaning. Another important issue to consider when exploring the ability to access word mental representations is how a word's multiple representations (phonological and orthographic forms) are integrated and linked to the word's meaning, and what is the nature of the link between form and meaning. There are two main accounts that specifically consider this issue, both of which have been devised to explain the process of reading aloud: the dual-route model (Coltheart, 2005; Forster & Chambers, 1973) and the triangle model (Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989; see Figure 1.1). The dual-route model hypothesises (at least) two routes to reading words aloud, and consequently two routes to word recognition. In one route, the process adopts a series of orthography-phonology recoding rules to translate the written form to its oral counterpart, without any need to retrieve semantic information: in this route orthographic information is translated directly into phonological information, and a word can be read before its meaning is activated. The other, lexical route is used to recognise whole words without any recoding. In this route the mental lexicon is accessed by directly activating semantic information, or by accessing the direct link between orthographic and phonological lexicons without accessing meaning. Non-words and new words can only be read via the sub-lexical route (as there are no lexical entries for these items), while irregular words are read via the lexical route, as they do not conform to grapho-phonemic conversion rules. Regular words, on the other hand, can be read by either route.

In contrast, the triangle model (Plaut et al.,1996; Seidenberg & McClelland, 1989) postulates that the representation of forms and meanings of words is distributed, and mappings between written and oral forms and their meanings are based in a distributed activation of connections between orthographic, phonological and semantic systems. The triangle model, therefore, does not presuppose two different routes to reading, but hypothesises that the choice of the correct pronunciation of a word stems from the activation of connection between the different systems. In this second model phonological and orthographic information are deeply connected, and it is their interaction that leads to lexical access. This second account therefore postulates that, by reading a word by sight, the reader retrieves both its phonological form and its meaning, even when specific access to the phonological form is not required (Liberman, Shankweiler, Liberman, Fowler, & Fischer, 1977).

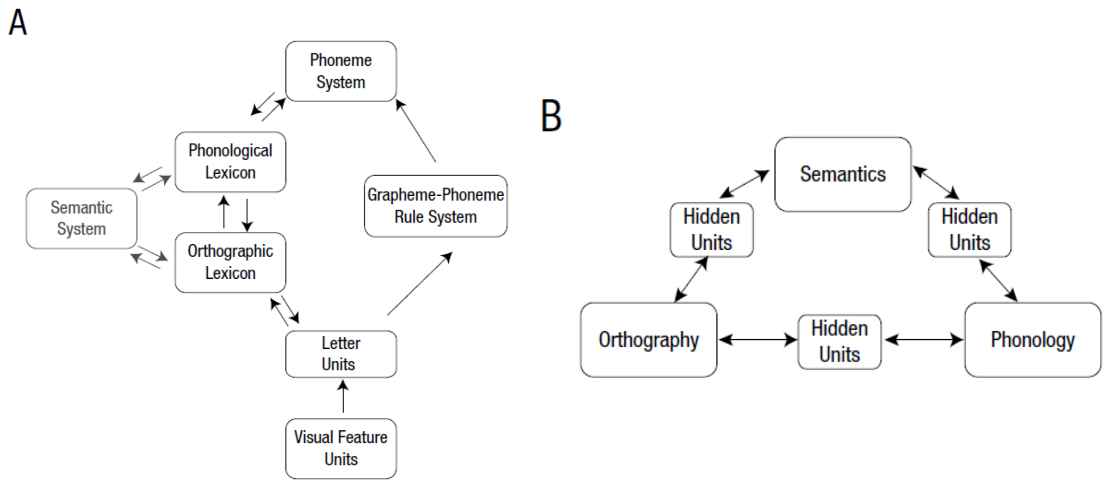


Figure 1.1. Representation of the dual-route (A) and the triangle model (B) adapted from Taylor, Duff, Woollams, Monaghan and Ricketts (2015).

While some of the previously described models (Forster, 1976; Levelt, 1989) support the dual-route model, connectionist and interactive models of the lexicon (Dell, 1986; Dell, Schwartz, Martin, Saffran, & Gagnon, 1996; McClelland & Elman, 1986; Stemberger, 1985) are more akin to the triangle model, in proposing a network of connections between the semantics and forms of a word. The Lexical Quality Hypothesis (Perfetti & Hart, 2002), which emphasises the importance of the link between different representations of a word, also resonates with the triangle model. There is considerable evidence in support of connectionist models. For example, phonological properties influence performance on purely written tasks, and vice-versa

(Bürki, Spinelli, Gaskell, 2012; Seidenberg & Tanenhaus, 1979; Ventura, Morais, Pattamadilok, & Kolinsky, 2004; Waters & Seidenberg, 1985; Ziegler & Ferrand, 1998; Ziegler, Ferrand, & Montant, 2004; Ziegler, Van Orden, & Jacobs, 1997). Bürki et al. (2012) showed that the presentation of the written forms of words after extensive oral training influenced participants' recognition and production of the spoken words. Participants in this study were, for instance, more likely to accept as correct an alternative oral form if this was an acceptable (but incorrect) pronunciation of a given written form. Waters and Seidenberg (1985), on the other hand, showed that phonological regularity (an oral property) can affect written word identification. The triangle model also tends to better account for semantic influences on purely phonological and orthographic processes, such as the impaired recall of phonological forms of *unknown* words (i.e. words for which the meaning was lost) in patients with semantic dementia (Patterson, Graham, & Hodges, 1994) and the facilitating effect of words semantically related to the homophone of the target item in naming (i.e. participants were faster in naming a *ball* – round object, if primed with *dance* – a word semantically related to its homophone: *ball* as a formal dance) (Cutting & Ferreira, 1999).

Whilst the previously presented models of word reading, especially the triangle model, can simulate some aspects of reading development (e.g. Powell, Plaut, & Funnell, 2006), these models do not make direct predictions about how new words are acquired, or what is the easiest way to acquire new words. Nevertheless, the different ways in which the structure of the mental lexicon is conceptualised may have an effect on how the vocabulary acquisition process is envisioned. For example, if the word recognition process is mostly guided by bottom-up processes, we could hypothesize that these processes would be prioritised over the integration of the new word in its context, thus leading to a better representation of the word's form than its meaning at earlier stages of word learning. On the other hand, if we consider the representations of form and meaning as intrinsically linked and not separable during the word learning process, we would expect the acquisition of form and meaning to proceed at the same pace. Similarly, if representations of words in memory are distributed through interconnected phonological, orthographic and semantic nodes, rather than fundamentally separate semantic, phonological and orthographic lexicons, the creation of the different components of the lexical representation might be predicted to proceed in parallel. In this case, learners should acquire phonological, orthographic and semantic representation of increasing and related quality, rather than independent phonological, orthographic and semantic representations. A further prediction of this approach might be that learners need not acquire a good phonological representation of a word before they can acquire a semantic representation, as the two could be learned simultaneously. Connectionist models of lexical development proposed to study the emergence of the link between word meaning and word forms (Li, Zhao, & MacWhinney, 2007; Plunkett, Sinha, Møller, & Strandsby, 1992; Siskind, 1996) could be applied to address questions concerning the relative speed of acquisition of form and meaning for each word. Sections 1.2.2 and 1.3.2 will explore the development of the link between different word features (phonology, orthography and semantics) in vocabulary development.

1.2 - Learning new words from oral language

The previous section explores the organisation of the adult lexicon. We now move on to consider the developmental changes that vocabulary undergoes during childhood. It is important to consider that, while the adult lexicon includes information on orthography, phonology and semantics, for pre-readers, lexical representations contain only phonological and semantic elements and are acquired via oral language. In this section we review how children acquire vocabulary from oral language, and how new words enter the lexicon through the listening modality.

1.2.1 - Early lexical development

One of the first tasks infants face when acquiring language is to segment the speech stream into its components. Before learning the meaning of any word they first need to identify the word's spoken form. Young infants are able to isolate word units from continuous speech, despite the fact that these are rarely presented in isolation (Aslin, 1993). Researchers have identified several features of infant-directed speech that could help infants to identify words and word boundaries, such as emphasis (Messer, 1981) and exaggerated prosody and position, especially final position (Fernald & Mazzie, 1991). In a series of studies, Jusczyk and Aslin (1995) showed that infants as young as 7 and a half months were able to recognise specific words in fluent speech, even distinguishing the same items from similar sounding items. These studies clearly show that children are able to identify and form phonological representations of new items from oral speech from a very young age. Other studies with children (3- to 13-year-olds) have explored the effect of phonological features on word learning, showing that phonotactic probability influences the likelihood of new word learning (Storkel, 2001; Storkel & Rogers, 2000). For example, in a study that used direct instruction, Storkel and Rogers (2000) demonstrated that the meanings of non-words containing rare sound sequences were more difficult to acquire than the meanings of non-words with common sound sequences. These studies suggest that children acquire phonological knowledge concerning new words from oral exposure from a very early age, and that phonological properties influence the ease of vocabulary acquisition.

Infants as young as 6 to 9 months show some initial ability to comprehend common words (Benedict, 1979; Bergelson & Swingley, 2015; Tincoff & Jusczyk, 1999; 2012). Several accounts have been proposed to explain children's ability to associate a new oral form with its referent. The first of these accounts, the constraints approach, postulates that children are directed in their search for the meaning of words by a number of assumptions (Golinkoff, Mervis, & Hirsh-Pasek, 1994; Markman, 1990). These strategies help children generate an intelligent guess about the probable referent of the given word, helping them select the most likely one. One of these is the whole object assumption, according to which a new word will refer to an object, rather than part of it; another is the mutual exclusivity bias, according to which a new label should indicate a new object, not a known one that already has a name (Golinkoff et al. 1994; Woodward & Markman, 1998). Finally, the taxonomic assumption postulates that words applied to objects can also be extended to other objects in the same category. A second account, the socio-pragmatic approach (Baldwin, 1991; Clark, 1993; Tomasello, 2000), on the other hand, proposes that children are able to use joint attention and their understanding of communicative intentions to link a new label to its referent. The child associates the name to the correct object by looking where the adult is looking, and by understanding the underlying meaning of this interaction. A third account, the syntactic bootstrapping hypothesis (Naigles, 1990), proposes that children are aware of the links between specific syntactic and conceptual structures, and they use these links to learn the meaning of words otherwise difficult to learn, such as verbs. For instance, pre-schoolers can use a word's position within a sentence or its morphological features to determine its grammatical category (the construction X-ing would indicate that X is a verb: Dockrell & McShane, 1990). Finally, according to the associative learning account (Smith, Jones, & Landau, 1996; Yu & Smith, 2007), it is not necessary to stipulate the existence of specific principles to explain word learning, since this process can be explained by general learning mechanisms, such as the ability to acquire an association between co-occurring stimuli. Proponents of the associative learning account believe that children are guided by general processes of perception, memory and learning when inferring the meaning of new words. During word learning children's attention is supposed to be drawn to new objects, rather than known ones, making it more likely they will associate new words to new objects since they were paying attention to them. Socio-pragmatic cues can be interpreted as a type of salience cue: parents' attention increases the salience of a specific object in the environment, drawing the attention of the child to it, and increasing the likelihood that the new label is associated with that object. This interpretation discounts any need for the child to understand joint attention (Houston-Price & Law, 2013; Houston-Price, Plunkett, & Duffy, 2006). Several researchers showed that these accounts can be used to explain vocabulary acquisition even in late childhood and adulthood, for example adults seem to apply principles, such as the mutual exclusivity bias when encountering new words amongst known ones (Golinkoff, Hirsh-Pasek, Bailey, & Wenger, 1992), and older children acquire new vocabulary more proficiently when instructed to look for morphemic cues (Baumann et al., 2002).

1.2.2 – Effects of developing an oral lexicon on word representations

Whilst at the beginning of a child's lexical development, vocabulary acquisition can be mostly conceptualised as the acquisition of individual new words, as the lexicon develops it becomes important to consider how new items are integrated within the existing network of words. As seen in Section 1.1.1, sleep facilitates the integration of new words within the lexicon in adults. There is evidence that the effect of sleep is more substantial for children than adults (Weighall, Henderson, Barr, Cairney, & Gaskell, 2017). Some authors (James et al., 2017; Lewis & Durrant, 2011; Wilhelm, Prehn-Kristensen, & Born, 2012) explain this effect by suggesting that the size of existing vocabulary knowledge influences the rate of integration. For these authors, when adults learn new words, there is greater overlap between their new and existing knowledge, which makes lexical integration more efficient and less dependent on consolidation through sleep.

An important difference between children's and adults' lexicons is the number of items stored within them. As children's lexicons are smaller, it is suggested that they have lower lexical neighbourhood density than adults (Charles-Luce & Luce, 1990; but see Coady & Aslin, 2003, for a counter-argument). Since children need to distinguish between fewer words, it is possible that their phonological representations are initially less well specified (Charles-Luce & Luce, 1995). Children with larger vocabularies may develop more specific phonological representation for the words they know, thanks to the need to differentiate similar words stored in their vocabulary (Walley, 1993). This view is sustained by the lexical restructuring hypothesis (Metsala & Walley, 1998), which suggests that children's phonological representations undergo a shift during childhood from holistic to fine-grained segmented representations. This shift supports phoneme awareness, the ability to clearly distinguish and manipulate phonemes in spoken words, which only becomes necessary with the need to distinguish items in an expanding lexicon (Walley, Metsala, & Garlock, 2003). Existing vocabulary and the phonological structure of existing lexical representation also have an effect on new vocabulary acquisition. For example, Storkel and Rogers (2000) showed that older children (10- to 13-year-olds), whose lexical representations are probably more fine-grained, perform better in recognising non-word meanings when the items have high phonotactic probability (large phonological neighbourhoods) than when the non-words deviated from the usual phonotactic pattern of words within their language (sparse neighbourhoods). Seven-year-olds, whose representations might be more holistic, did not show such a difference. Thus, only in middle or late childhood were children more likely to acquire items that were phonologically similar to items already in their lexicon, suggesting that during this period their representations become more fine-grained.

Existing vocabulary has an effect on the acquisition of new words not only in terms of phonological properties, but also in terms of semantics. For example, infants are more likely to acquire words pertaining to categories for which they have deeper knowledge (Borovsky, Ellis, Evans, & Elman, 2016). Words learnt at an earlier age are more interconnected in the semantic network than words learnt later on (Steyvers & Tenenbaum, 2005), leading some authors to suggest that the mechanisms of word learning differ between young children and adults (Hills, Maouene, Maouene, Sheya, & Smith, 2009). Hills et al. hypothesise that adults are more likely to acquire new words that are closely associated with words in their existing lexicon (preferential attachment), while infants and children tend to acquire words with many associates, irrespective of their knowledge of these associates (preferential acquisition). Thus, children learn items that are similar to other words in their linguistic environment earlier than less well integrated words.

In conclusion, as the mental lexicon develops, the representation of the words within it changes: phonological representations become more specified, more words are linked together and included in the semantic network and the specific properties of the items within the lexicon start to influence the word acquisition process.

1.2.3 – Oral vocabulary and oral language comprehension

Oral vocabulary knowledge is essential for listening comprehension: without knowing the meaning of the words that form a sentence, it is impossible to understand the sentence's meaning. Nevertheless, vocabulary knowledge is only one of the skills necessary for good listening comprehension. Knowledge of morphology, the structure of words (e.g. that the morpheme '-s' at the end of countable nouns indicates plurality) and the syntactic rules, that govern sentence structures (such as subject-verb-object word order in English, exemplified by the difference between "the dog chases the cat" and "the cat chases the dog") is also essential for understanding spoken utterances. The distinction between the comprehension of vocabulary and grammar and the importance of each is highlighted by the existence of specific deficits in these skills in neurologically damaged patients (Berndt, Mitchum, & Haendiges, 1996; Druks, 2002).

From a developmental perspective, the acquisition of oral vocabulary and morpho-syntactic knowledge are associated. For example, there is an association between vocabulary and morphological and syntactic knowledge in children between the ages of 8 and 30 months (Fenson et al., 1994), and some researchers suggest that vocabulary is fundamental for the acquisition of grammatical skills (Bates & Goodman, 1999; Marchman & Bates, 1994). Research has also shown that vocabulary knowledge is important for online sentence comprehension, specifically the ability to understand the meaning of a sentence and predict its ending before it has been spoken in its entirety. Borovsky, Elman and Fernald (2012), for example, used eye movements to explore whether the ability of 3- to 10-year-old children and adults to predict the ending of a sentence was related to their vocabulary knowledge. Eye movements towards visual scenes of four items were recorded, while participants listened to a set of sentences. Both children and adults tended to look at the image depicting the last word of the sentence more often than the alternative images, and participants with better vocabulary were faster than participants with lower vocabulary skills at recognising the correct image. This result highlights the importance of vocabulary for online sentence comprehension. Other examples of the link between vocabulary, grammar and more general comprehension abilities come from the literature on specific language impairment (SLI). Children with SLI present problems in morphological and syntactic processing, and a grammatical impairment has been proposed as the core deficit in SLI (van der Lely, 2005). Even though their deficit tends to be mostly associated with grammar, these children also exhibit vocabulary delays and their semantic representations are less well specified than those of their peers (Clarke & Leonard, 1996; McGregor, Newman, Reilly, & Capone, 2002). Children with SLI also seem to need more repetitions than other children to acquire new words (Riches, Tomasello, & Conti-Ramsden, 2005). Morpho-syntactic, oral comprehension and vocabulary difficulties are therefore likely to be associated in their effect on vocabulary acquisition.

In conclusion, although oral vocabulary is only one of the skills required to understand oral language, it is strongly linked with other factors that determine children's comprehension abilities, such as morphological and syntactic knowledge. Vocabulary and syntactic-level sentence comprehension appear to have a bi-directional relationship, with vocabulary facilitating sentence level comprehension (Borovsky et al., 2012), and morpho-syntactic ability facilitating vocabulary acquisition (Clarke & Leonard, 1996; McGregor et al., 2002).

This section has summarised the nature of the links between vocabulary and broader language skills, such as listening comprehension. The next section explores the influence of these factors on vocabulary acquisition within the specific context of listening to a story.

1.2.4 - Vocabulary acquisition from listening to stories

While infants and children learn many words from conversations (Weizman & Snow, 2001), reading and listening to stories enriches children's vocabulary by providing more varied sets of words, including rare or uncommon words that are unlikely to be used in conversation (De Temple, & Snow, 2003). Many studies have shown that listening to stories fosters vocabulary development in children of all ages and abilities (Brett, Rothlein, & Hurley, 1996; Elley, 1989; Houston-Price, Howe, &

Lintern, 2014; Penno, Wilkinson, & Moore, 2002; Robbins & Ehri, 1994; Sénéchal, 1997; Sénéchal & Cornell, 1993; Walsh & Blewitt, 2006; Wilkinson & Houston-Price, 2013; Williams & Horst, 2014), including those 'at risk' of educational failure (Justice, Meier, & Walpole, 2005; Vadasy & Sanders, 2015). Specifically, studies with preschool-aged children have shown that children are able to recognise word meanings even from one exposure to the story (Sénéchal & Cornell, 1993), although they learn word meanings from stories best when they are exposed to the same story more than once (Sénéchal, 1997). They also learn words best given multiple repetitions of the same words; for example, words that are heard four times are more likely to be learnt than words heard only twice (Robbins & Ehri, 1994). Furthermore, several studies have found that 3- and 4-year-olds tend to learn new words best through repeated exposure to the same book, rather than the same number of exposures to words in different contexts (Horst et al., 2011; McLeod & McDade, 2011), an effect that was evident both in immediate recall and delayed retrieval. Overall these effects are more evident in tests of receptive vocabulary (which assess the ability to comprehend words) rather than expressive vocabulary (which assess the ability to produce words, using them meaningfully); children tend to perform very poorly on the latter kind of tests, especially with limited exposure to new words (Sénéchal & Cornell, 1993).

Studies of word learning in school-aged children have similarly found evidence of vocabulary acquisition from listening to stories. For example, Elley (1989) reported substantial incidental vocabulary acquisition by 7- and 8-year-old children when a story was read three times over seven days in the classroom. Similar gains in semantic knowledge of the presented words were found in children up to the age of 11 (Brett et al., 1996; Penno et al., 2002). Wilkinson and Houston-Price (2013) also found that 7- and 9-year-old children made significant gains in their knowledge of words presented orally within a story context in a naturalistic classroom environment. They also compared the effect of hearing the same story three times, versus three different stories; while they found no difference between these for the older group of children, the younger group benefited from presentation of the same story repeatedly, as previously reported in pre-readers (Horst et al., 2011; McLeod & McDade, 2011). Compared to younger children, school-aged children show both receptive and expressive vocabulary learning after story exposure. For example, Houston-Price et al. (2014) found that both 4- and 6-year-old children showed significant learning of new words when listening to stories, both in a recognition task where they were asked to select the right picture for each word, and in a definition task, where they were asked to produce a definition for each word. Listening to stories has also been highlighted as an effective tool to teach new vocabulary to 10-year-old children learning English as a second language (Lin, 2014).

These studies therefore show that, from at least 3 years of age onwards, when listening to stories children have the ability to acquire new vocabulary, acquiring new phonological forms and linking these with the correct meanings. The next section (1.3) explores how children learn to read, and how they acquire an orthographic lexicon. Considering that school-aged children and adults already have a written lexicon, it is worth considering whether readers are able to acquire orthographic representations from a purely oral presentation when listening to stories. The studies previously reviewed suggest that the creation of an orthographic representation is not necessary for successful word learning, since children learn new words efficiently years before they learn to read, and oral language precedes the advent of writing systems. However, some studies have shown that, in good readers, oral tasks, including rhyme judgement and oral priming, are influenced by the orthographic properties of the stimuli (e.g. Chereau, Gaskell & Dumay, 2007; Muneaux & Ziegler, 2004; Seidenberg & Tanenhaus, 1979; Slowiaczek, Soltano, Wieting, & Bishop, 2003; Ziegler, Ferrand & Montant, 2004). It has therefore been proposed that an orthographic recoding mechanism works to create an initial orthographic representation of new oral words in able readers (McKague, Davis, Pratt, & Johnston, 2008). Studies involving single word presentations have shown that adults form initial orthographic representation for items presented orally (Johnston, McKague, & Pratt, 2004; McKague et al., 2008; Nelson, 2004); visual priming effects are found for words that were previously encountered orally, but not visually. However, these effects have not been consistent across tasks. For instance no effect of orthography has been shown in word production tasks: naming latencies are the same for lists of words that share the same initial phoneme and letter, and lists where the same phoneme is represented by different letters (i.e. *c* and *k*) (Alario, Perre, Castel, & Ziegler, 2007). Furthermore, it must be noted that the orthographic representations created from oral presentations are not particularly well specified (Johnston et al., 2004), and it is unclear whether children would be as likely as adults to form orthographic representations of new spoken forms. Nevertheless, it has been proposed that, once children have mastered the reading process, they may create an orthographic skeleton for orally-encountered words (Wegener et al., 2017): specifically they may establish an initial orthographic representation of the oral words they know, which generates expectations regarding the written form of orally known words. These predictions have been tested in adults: McKague et al. (2008) taught adult readers the meaning of new oral words, manipulating the orthographic consistency of the items. Adults performed significantly better in the written lexical decision task for words spelled consistently than inconsistently. Similarly, Wegener et al. (2017) showed that 9- to 10-year-olds were faster in silently reading new, orally trained words embedded within sentences when these had more predictable spelling patterns. The difference between predictable and less predictable spelling patterns was less pronounced for words that were not orally trained, thus showing that the effect was not due to easier processing of predictable compared to unpredictable orthographic forms. This demonstrated that children had formed expectations regarding the written form of the words, and were facilitated in reading them, if these expectations were met.

In conclusion, children have the ability to acquire new vocabulary by listening to stories, acquiring new phonological forms and linking these with the word meanings. There is also some evidence that, once children have learnt to read, they start to build an orthographic skeleton of words they are exposed to only orally. The efficiency with which new vocabulary is acquired from oral context depends on children's age and abilities, with older children showing the ability to capitalise on presentations of the word in diverse contexts, and younger children facilitated by repetition of the word in the same context.

1.3 - Learning new words from written language

While both infants' and older children's vocabulary is characterised by the presence of a phonological and semantic lexicon, in the early school years children also develop an orthographic lexicon. In the present section we will briefly review how children learn to read, the changes to the mental lexicon associated with reading acquisition, and how new words enter the lexicon through the reading modality.

1.3.1 - Learning to read words

At the start of primary school, typically at the age of 5 or 6 in the UK, children first receive formal literacy instruction, and start to learn the orthographic forms for words they previously knew only orally. In alphabetic languages, children need to master the alphabetic principle to learn to read: they need to learn that letters stand as a symbolic representation of the spoken sounds that form a word. In Ehri's (2005a; 2005b) phase theory of sight word reading, children start the pre-alphabetic phase recognising specific words by their visual features and context. They then proceed to the partial alphabetic phase where they can use partial connections between letters and phonemes to recognise words. In the full alphabetic phase, children have developed decoding skills, the ability to translate the orthographic forms of words into their pronunciation, using phonological recoding. An important skill for the transition to this stage is therefore the ability to segment words into their constituent phonemes. Children in the final, consolidated alphabetic phase have acquired not only the correspondence between letters and sounds, but are also able to memorise larger units, including some frequent words, without the need to recode them. Blending of larger grapho-phonemic units is a fundamental skill to reach this stage.

Four main strategies have been proposed to describe the reading process, or how readers recognise written words and how they map orthography onto meaning (Cain, 2010). A first strategy is predicting the word from context. Young readers tend to rely on context more than more skilled readers, such as older children and adults (Stanovich, 1986). This strategy may lead to inaccuracies and errors in word reading (Ehrlich, 1981), since context predictability is generally very low (Gough & Wren, 1999), making the viability of this strategy for word reading controversial (Share, 1995). The use of context is thought to shift during childhood from a strategy to read words, to a way to enhance the speed and accuracy of word identification.

A more reliable strategy for word reading is phonological recoding, the skill that marks the start of the alphabetic phase; to use this strategy, children must be able to employ grapheme-phoneme correspondence rules to predict how a written word sounds, and then recognise the word from its phonological form. The phonological recoding strategy is the basis of decoding skills; its use strengthens the link between orthographic patterns and their associated sounds, helping children to acquire orthographic representations of specific words, thus functioning as a self-teaching mechanism (Share, 1995). In his self-teaching hypothesis, Share (1995, p. 168) suggests that the orthographic lexicon "develops as a result of the self-teaching opportunities provided by successful decoding". As the child becomes a more experienced reader, a printed vocabulary store is created, and this allows the child to avoid the need to access phonology to link orthography and meaning. The developmental shift hypothesis (Doctor & Coltheart, 1980), for example, claims that, while novice readers tend to rely on phonological recoding to access the meaning of words, skilled readers abandon this strategy as they are able to access meanings directly from print, or at least from a combination of phonology and direct access (Barron, 1986).

Another strategy that can be used, once some rudimentary grapheme-phoneme correspondence rule has been acquired (Ehri & Robbins, 1992) is reading by analogy. Thanks to this strategy words are recognised more quickly when the reader knows other words sharing similar grapheme patterns (e.g., a word like *fog* is read more easily if the reader knows how to read *dog*). The use of this strategy naturally increases as the written vocabulary of the child increases (Bowey & Hansen, 1994; Bowey & Underwood, 1996). Even very young readers are able to apply this strategy in specific situations (Goswami, 1986; Pick, Unze, Brownell, Drozdal Jr, & Hopmann, 1978), but older and more proficient readers use it more efficiently (Backman, Bruck, Hebert, & Seidenberg, 1984). As the number of entries in the orthographic lexicon increases, phonological recoding transforms from a process that maps single graphemes into phonemes, to a process in which larger units and their regularities are lexicalised.

A final method for reading words is sight word reading (Cain, 2010), where the reader recognises the word from memory, without the need to recode it phonologically or read it by analogy. Skilled readers use this strategy with many words (Ehri, 1992); they have encountered some words so many times that these words are stored in their memory as whole units. High-frequency words are all recognised by sight, without the need for phonological recoding, while low-frequency words often require recoding, a longer and more effortful process (Weekes, 1997).

1.3.2 - Effects of developing a written lexicon on word representations

The dual-route and the triangle model postulate a link between the orthographic lexicon and phonological and semantic lexicons. There is evidence of a direct link between orthographic and phonological forms: with the development of reading skills, children become increasingly able to automatically activate the phonological forms of words directly from their written form (Rayner, Pollatsek, Ashby, & Clifton, 2012). Words' phonological properties affect the reading process, even in skilled readers (Liberman et al., 1977; Mark, Shankweiler, Liberman, & Fowler, 1977; Van Orden, 1987). For example, when presented with written words that are homophones of words in their other language, bilinguals activate the phonological representation of the homophone in their other language, even though this activation is unnecessary and unhelpful for the task (Haigh, & Jared, 2007). Furthermore, learning to read is linked to the development of well specified phonological representations. For example, the lexical restructuring hypothesis (Metsala & Walley, 1998) suggests that learning to read is supported by the development of more specific phonological representations, which, in turn, are prompted by vocabulary growth (Metsala, 1999).

It has been proposed that the most important consequence of learning to read is the acquisition of word representations (Perfetti, 1992). Through reading children learn many new words, and enhance the lexical quality of words they already know (Perfetti, 1992). The creation of an orthographic lexicon, and heightened phoneme awareness and decoding skills enable the development of more precise, fully-specified phonological representations. When learning to read, children acquire an orthographic lexicon, and also the ability to build the phonological forms of words from their orthographic forms. This creates a redundancy between stored phonological representations, and the phonological representations created through recoding, which overdetermines and therefore facilitates lexical access. Lexical access can happen without redundancy, as in younger readers, but the process is more effortful. These ideas underpin the Lexical Quality Hypothesis (Perfetti & Hart, 2002), which proposes that a representation of high quality is characterised by fully specified orthographic representations and a redundant phonological representation (which can be accessed directly or reproduced from orthography). A representation of high quality, in this view, is a representation that is coherent and reliable in both form and meaning. Good readers and able comprehenders have lexicons characterised by the integration of orthographic and phonological representation, while less skilled readers have representations that are less integrated (Perfetti & Hart, 2002). By this account, efficient lexical access frees attentional resources from the process of word reading, to be allocated to reading comprehension. On the other hand, less skilled comprehenders, whose lexical representations are less reliable, rely more on context for the identification of the words.

Similarly, Ehri (1992; 2014) proposed that, by becoming experienced readers, children acquire a visuo-phonological route into lexical memory. In contrast to the dual-route model, which proposes that the connection between orthography and meaning is direct, but arbitrary, Ehri (2014) argued that, while reading, the reader accesses an amalgam of semantic, phonological and orthographic information, where the link between orthography and meaning is not arbitrary, since it is established due to the link with phonology. Ehri suggested that the correlation between non-word and irregular word reading supports this hypothesis.

Further evidence that acquiring an orthographic lexicon influences existing representations of words comes from studies showing that, in good readers, oral tasks, such as rhyme judgement and oral priming, are influenced by the orthographic properties of the stimuli (e.g. Chereau et al., 2007; Muneaux & Ziegler, 2004; Seidenberg & Tanenhaus, 1979; Slowiaczek et al., 2003; Ziegler et al., 2004). For example, in the classic study by Seidenberg and Tanenhaus (1979) participants were faster in making rhyme judgements (considered a purely phonological task) for words with similar spellings, compared to more different spellings. Chereau, Gaskell and Dumay (2007) also showed that the orthographic similarity between a prime and target facilitated responses in an auditory priming experiment. Given these findings, it has been proposed that an orthographic recoding mechanism works to create an initial orthographic representation of new oral words in able readers (McKague et al., 2008). This mechanism may help to make the new phonological representation more stable, and thus easier to retrieve and link with semantic information (Ventura et al., 2004).

In conclusion, learning to read and acquiring an orthographic lexicon influences a word's semantic and phonological representations.

1.3.3 – Vocabulary depth and breadth and reading comprehension

The ultimate goal of learning to read is comprehending the written text. The Simple View of Reading proposes that the ability to comprehend written text is the product of decoding skills and oral language abilities (Hoover & Gough, 1990). Once the child is able to readily decode or recognise words, the majority of his or her processing capacity is left free to tackle the task of comprehending the written text (Perfetti & Hart, 2002; Samuels & Flor, 1997). Reading comprehension is therefore likely to be highly correlated with the child's comprehension of oral language, once decoding skills have been taken into account (Catts, Fey, Zhang, & Tomblin, 1999). In fact, the relationship between oral and written language comprehension abilities grows stronger over time, while the relationship between decoding and reading comprehension tends to weaken as children grow older. As decoding skills reach their peak, their power to predict reading comprehension diminishes (Gough & Tunmer, 1986; Landi, 2010; Roth, Speece, & Cooper, 2002; Vellutino, Tunmer, Jaccard, & Chen, 2007). Thus, the more automatized the process of reading becomes, the more reading comprehension can be explained by oral comprehension.

Having broad and deep vocabulary knowledge fosters reading ability, especially reading comprehension (Biemiller, 2003; Beck et al., 1982; Brinchmann, Hjetland, & Lyster, 2015; Muter, Hulme, Snowling, & Stevenson, 2004; Nation & Snowling, 2004; Priya & Wagner, 2009). It has been suggested that depth of vocabulary knowledge is particularly associated with reading comprehension (Ouellette, 2006; Tannenbaum, Torgesen, & Wagner, 2006). Vocabulary knowledge predicts reading comprehension, and children with smaller vocabularies show poorer comprehension skills than their peers (Durand, Loe, Yeatman, & Feldman, 2013; Nation & Snowling, 2004; Ricketts, Nation, & Bishop, 2007; Wise, Sevcik, Morris, Lovett, & Wolf, 2007). Children with larger vocabularies are advantaged when reading the words they know, because they are able to retrieve their meanings and the meanings of associated words more quickly and more proficiently, thus understanding the written text more thoroughly than children who know fewer words in less depth (Ouellette, 2006). The lexicons of children who have deeper word knowledge are characterised by a greater number of connections between items, and these children may be more efficient and faster in activating all the necessary features of words, enabling them to comprehend the text with less effort (Nation & Snowling, 1999). Unsurprisingly, given the effect of vocabulary size on reading ability (see Section 1.3.2 for its effect on single word reading), explicit vocabulary instruction practices can help struggling readers (Beck et al. 1982; Elleman, Lindo, Morphy, Compton, 2009). Incidental vocabulary acquisition could be just as important, however, since it is possible to teach strategies to acquire new words while reading (Beck & McKeown, 1991; McGregor & Duff, 2013).

In conclusion, possessing an extensive oral vocabulary makes the process of reading easier: extensive vocabulary knowledge has a positive effect on reading comprehension. Reading comprehension, in turn, can support the acquisition of new words from the written text. The next section explores how written texts support the learning of new vocabulary, and the important role played by comprehension of the story context for word learning.

1.3.4 - Vocabulary acquisition from reading stories

When the process of reading has become more automatized, it becomes a new learning tool that can be applied, automatically, to word learning (Cain, Oakhill, & Lemmon, 2004; Jenkins, Stein, & Wysocki, 1984; Nagy, Anderson, & Herman, 1987; Ricketts, Bishop, Pimperton, & Nation, 2011; Schwanenflugel, Stahl, & McFalls 1997; Swanborn & de Glopper, 1999). Investigations into children's reading habits have shown that children who read extensively may be exposed to 80 times as many words as children who rarely read (Anderson, Wilson, & Fielding, 1988). Reading habits at age eight years, including the amount of time spent reading, predict vocabulary size several years later (Cain & Oakhill, 2011). There is a bidirectional relationship between reading and vocabulary knowledge. While reading provides children with a way to learn new words, vocabulary knowledge affects the development of reading skills. Several studies, for example, have found a positive relationship between vocabulary size and single word reading ability, especially exception word reading (Nation & Snowling, 2004; Ricketts et al., 2007), while others have shown effects of vocabulary on reading comprehension (section 1.3.3). The reciprocal effects of vocabulary on reading and reading on vocabulary can be described as a bootstrapping process that proceeds without help and results in a Mathew Effect (Duff, Tomblin, & Catts, 2015; Stanovich, 1986), whereby children who start school with bigger vocabularies show greater increases in vocabulary knowledge and reading comprehension abilities than children who start school with lower vocabularies. This effect is also known as the Beginner Paradox in the second language literature: while the most effective way to build a vocabulary is through reading extensively, beginner students of a second language do not have sufficient vocabulary to read texts to access new words (Coady, 1997). However, other research shows a compensatory development pattern, where children with lower abilities experience greater increase in both vocabulary and reading (Pfost, Hattie, Dörfler, & Artelt, 2014), thus questioning whether the Matthew Effect can be generalised to all reading and vocabulary skills, and all children.

Several studies have examined the learning of specific word meanings from reading. Jenkins, Stein and Wysocki (1984) found that American fifth-grade children (between 10- and 11-year-olds) were able to learn new words while reading passages. Learning was measured by a definition task, where children were asked to define the word, and a sentence completion task, where children were asked to complete a sentence with the correct word. Nagy et al. (1987) also found that students as young as eight years old can acquire new words from a single exposure in text. Although the gains in vocabulary knowledge after a single exposure were relatively low, the finding has been replicated repeatedly (Nagy, Herman, & Anderson, 1985; Shefelbine, 1990). In a seminal study of partial word knowledge, Schwanenflugel, Stahl, and McFalls (1997) studied the effect of presenting both unknown and partially known words to American fourth graders in a story context, finding small amounts of learning for both types of word. The fact that children in this study were able to acquire at least some of the words presented seems particularly important, given that the stories were designed for older children, and story comprehension was quite low. This study also showed that word features, such as imageability and part of speech, influenced word learning; non-nouns were learnt better than nouns in this study, which contradicts the majority of findings in the opposite direction (Gentner, 1982). In their meta-analysis, Swanborn and de Glopper (1999) make a useful distinction between studies whose aim is to determine the extent of incidental learning (e.g. Nagy et al., 1987) and studies that aim to explore the effect of text variables on word learning, such as number of repetitions (e.g. Durkin, 1990, cited in Swanborn & de Glopper, 1999) or level of text comprehension (Diakidoy, 1993, cited in Swanborn & de Glopper, 1999). They found that texts containing a smaller number of unknown words are easier to read and understand, which facilitates vocabulary acquisition: participants learned more new words from texts containing fewer unknown words.

Studies have also shown that children learn the orthographic forms of words from written presentation (e.g. Bowey & Miller, 2007), even if these are not associated with a meaning (Landi, Perfetti, Bolger, Dunlap, & Foorman, 2006). Learning orthographic forms and meanings are distinguishable processes, and context has a differential effect on the two (see Section 1.6.1 for discussion). Children and older readers are also able to learn phonological forms from written presentations, as the phonological properties of written words are activated during reading (see Section 1.3.2). Furthermore, studies have found that silent reading of short stories boosts children’s ability to read new words aloud (Bowey & Muller, 2005), and that phonological abilities correlate with orthographic learning from silent reading (Bowey & Miller, 2007). Share’s (1995) self-teaching hypothesis postulates that new orthographic forms are encoded thanks to the concurrent creation of phonological forms. Therefore children activate and learn phonological forms when reading silently through phonological recoding. This hypothesis fits with claims that good decoders are better at learning phonological forms of new words while reading, thanks to their knowledge of the link between orthography and phonology, which enables them to generate phonology automatically while reading (Ehri, 2014).

In conclusion, these studies show that reading boosts vocabulary knowledge, and that children can acquire new meanings, phonological forms and orthographic forms by reading stories.

1.4 - Comparison between reading and listening to stories

The previous sections established that vocabulary acquisition can be fostered both by reading and listening to stories, but have not considered which modality better supports word learning. Suggate, Lenhard, Neudecker and Schneider (2013) compared learning following exposure to new words in monolingual German children (8- to 10-year-olds) in three story presentation conditions: independent reading, listening to an adult reading the story, and listening to an adult telling the story in his own words. In the listening conditions, children were exposed to spoken forms (phonology) and meanings (semantics) of new words while in the reading condition they encountered written forms (orthography) and semantics. All words were meaningfully embedded in the story, and half of the words were accompanied by images depicting their meaning. Vocabulary acquisition was assessed through an oral picture-matching task. Children who listened to the stories were more likely to demonstrate knowledge of the meanings of the new words than children who read the stories themselves, suggesting that oral presentation is more beneficial for vocabulary learning in primary-school-aged children. Although older children acquired more new words than younger ones, there was no interaction between age and presentation modality, thus showing that even older children learnt more words from listening to stories.

In contrast with Suggate et al.'s (2013) findings, studies of adults learning English as a second language typically show that these participants acquire new words more easily when presented with material in written, rather than oral form (Brown, Waring,

& Donkaewbua, 2008; Sydorenko, 2010; Vidal, 2011). Studies exploring memory for word lists and verses also show better performance for written than oral material, even in children (Dean, Yekovich, & Gray, 1988; Hartman, 1961; Menne & Menne, 1972). For example, in Menne and Menne's study (1972), 8- to 9-year-olds exposed to a written presentation learnt more verses than those exposed to the text orally. While learning verses is certainly different from learning new words, this research showed how written presentation can improve memory performance for linguistic material, even in children. Research on adult vocabulary acquisition also claims an advantage for visually presented material over orally presented material (Nelson, Balass, Perfetti, 2005).

Several accounts have been offered for the advantage of written presentation among skilled readers. One hypothesis is that orthographic memory traces are encoded in memory more distinctly than phonological memory traces (Gallo, McDermott, Percer, & Roediger, 2001). This account hypothesizes that readers may require multiple encounters with an oral form to create a clear representation of that form in memory, but can form a clearer orthographic form in fewer encounters, and with less effort (Nelson et al., 2005). Another hypothesis is that the superiority of written presentation stems from the reader's greater ability to derive phonology from a word's written form, than to derive orthography from an oral presentation. It is known that reading aloud (i.e. forming a phonological representation from a written presentation) is easier than spelling (i.e. forming an orthographic representation from a spoken presentation), even in highly regular languages (Bosman & Van Orden, 1997; Cossu, Gugliotta, & Marshall, 1995). This hypothesis is supported by evidence previously reported, that an oral presentation creates only an orthographic skeleton (Wegener et al., 2017), not a full orthographic representation (Johnston et al., 2004). Conversely, the presentation of a written form activates phonology automatically, either by prompting phonological recoding strategies (Share, 1995), or by activating the word form (Ehri, 2014). Following this line of reasoning, it would be hypothesised that the written presentation modality should foster the creation of a better and more stable representation of new words in memory, including both phonological and orthographic properties, and therefore more easily accessible than representations formed by listening only (Perfetti & Hart, 2002).

In conclusion, since the first medium for vocabulary acquisition is oral language, it is likely that oral presentation would be the preferred and easiest method to acquire vocabulary early on, but, with the increase of reading abilities over time, children might come to rely more on written texts to learn new vocabulary. While it seems likely that the orthographic representation of a new word formed during reading would be of higher quality than that formed during listening (Johnston et al., 2004; Wegener et al., 2017), it is less clear whether the phonological representation created through reading and recoding would equal that created through listening. On the other hand, semantic learning should be stronger following reading than listening, at least in good readers. Assuming that semantic learning depends on the quality of the lexical entry in terms of high quality form and meaning representations, a written presentation should be superior to an oral presentation, because the written presentation provides both orthographic and phonological information, while the oral presentation provides only limited information regarding orthography. Nevertheless, as discussed above, the evidence relating to the superiority of the reading modality for vocabulary acquisition is mixed, and likely depends on the reading abilities of the participants (Brown et al., 2008; Sydorenko, 2010; Suggate et al. 2013).

1.5 - Learning new words while reading and listening

1.5.1 - Dual presentation modality

Previous sections have established that listening to and reading stories foster vocabulary development. A different stream of research has explored the positive effect of presenting oral and written texts simultaneously. Children do not particularly attend to written texts before learning to read (Evans, Williamson, & Pursoo, 2008; Justice, Pullen, & Pence, 2008); nevertheless, studies by Apel and colleagues have shown that even in the very early stages of reading development, children can acquire a mental graphemic representation of new words, when they are exposed to written storybooks read aloud by adults (Apel, 2010; Apel, Brimo, Wilson-Fowler, Vorstius & Radach, 2013; Apel, Thomas-Tate, Wilson-Fowler, & Brimo, 2012; Wolter & Apel, 2010). In these studies, children were exposed to new words through stories, both orally and written, multiple times, and their ability to spell and recognise the new orthographic forms was assessed. Collectively these studies showed that even young children are able to acquire an initial orthographic representation of new words by looking at written text, and do so more easily for more predictable forms (i.e. words with higher phonotactic and orthotactic probabilities). Reading and spelling skills predicted children's orthographic learning.

Similar effects have been found when teachers or parents engage in a technique called "read-aloud", which involves reading aloud a story while allowing the child to look at the printed story, with or without further questioning or active participation by the child or adult (Fien et al., 2011; Ivey, 2003; Leung, 2008). Most of this research is carried out in the educational field to improve teaching practices, and most of the studies employ other methods to engage students with the text, as well as reading aloud. These studies show that reading aloud activities can promote vocabulary and reading comprehension, especially for struggling readers and children with low vocabulary. This modality of presenting texts is often used in schools, and it has been found to have a positive effect on reading ability (Chomsky, 1976; Rasinski, 1989; 1990) and on vocabulary development (Schneeberg, 1977). In addition, text-to-speech translation software, which provides the oral forms of difficult words within a text while children read, has proven effective in improving reading abilities (word decoding & text reading) in poor readers (Elbro, Rasmussen, & Spelling, 1996). These studies show that a combined presentation that includes both oral and written modalities can foster vocabulary acquisition, with some positive effects even at early stages of reading development.

A second situation that might be considered dual presentation is when children read aloud by themselves. In such a condition, children are forced to create an oral representation of the text, while the written representation is directly provided (e.g., Cunningham, 2006; Cunningham, Perry, Stanovich, & Share, 2002; Nation, Angell, & Castles, 2007; Ricketts et al. 2011). Cunningham et al. (2002) asked 6- to 7-year-olds to read aloud a series of short stories containing pseudowords repeated six times. Three days later their orthographic knowledge of the items was assessed through an orthographic recognition task, in which children had to choose the correct form out of four alternatives, and a spelling task. Children were more likely to choose the correct spelling than a homophonic spelling, and were more likely to produce the correct written form rather than an alternative, showing that they had acquired an orthographic representation of the target pseudowords. In a similar study where 7- to 8-year-olds read aloud short stories containing new words repeated four times, Ricketts et al. (2011) found that children learnt orthographic forms (assessed through an orthographic forced choice task and a spelling task) and meanings (assessed through a picture choice task). In these studies, decoding and orthographic skills predicted learning of orthographic forms, and existing vocabulary predicted semantic learning. More proficient readers were more likely to both read the words aloud correctly, and show better orthographic learning.

Recent studies have explored children's acquisition of new words through e-books (Allen, Hartley, & Cain, 2015; Bus, Takacs, Kegel, 2015; Korat, 2010; Mol, Bus, & de Jong, 2009; Shamir, Korat, Fellah, 2012; Smeets & Bus, 2015; Takacs, Swart & Bus, 2014). These studies show that exposure to e-books has a positive effect on vocabulary acquisition and reading in children from kindergarten onwards. One of the features of e-books is that children can listen to the narration of the written text while reading, providing a dual modality of presentation. Some studies have shown that this presentation type supports vocabulary acquisition (Smeets & Bus, 2015). However, e-books also have other features that may drive positive effects in vocabulary acquisition, including general multimedia effects (i.e. sounds and animation) and specific features designed to improve comprehension (i.e. animated cues to complex vocabulary, vocabulary games or dictionary options). Smeets and Bus (2015) found that after four sessions of e-book reading, vocabulary acquisition, as tested through a sentence completion task, was higher for animated e-books with vocabulary hot-spots compared to e-books without hot-spots or other multimedia elements. Similarly to studies with second language learners (e.g., Silverman & Hines, 2009; Verhallen & Bus, 2010), this work shows that vocabulary acquisition is promoted by multimedia presentation that includes features other than supporting oral text. It is therefore important to carefully consider the presentation modality employed when comparing the learning resulting from e-books versus more traditional storybooks, since the two activities have many differentiating features.

In conclusion, presenting a story in the oral and written forms simultaneously can promote word learning, both for semantic learning and form (orthographic and phonological) learning. Nevertheless, when considering instances of dual presentation it is important to consider whether the effects are driven by the combination of written and oral modality, or whether other effects may be influencing the results. For instance, multimedia features may influence learning from e-books, while children's reading ability may have an impact on learning when reading aloud is required.

1.5.2 - Comparison between single and dual modality presentations

While the previous section established that presenting stories in two modalities simultaneously promotes vocabulary acquisition, the present section directly compares the effects of dual and single modality presentation on vocabulary acquisition.

Experimental evidence for a benefit of dual modality has primarily come from studies of single word learning. One body of work has demonstrated that having access to orthographic forms promotes oral vocabulary learning, an effect referred to as 'orthographic facilitation' (e.g., Miles, Ehri, & Lauterbach, 2016; Ricketts, Bishop, & Nation, 2009; Rosenthal & Ehri, 2008). Rosenthal and Ehri (2008) compared new word learning in two conditions: a condition where spoken words were accompanied by spelling (orthography present), and a condition where only oral forms were presented (orthography absent). They found that 7- to 8-year-old children learnt the phonological form of the words (assessed by a word production task), their written form (spelling task) and their meaning (word-definition matching task) better in the orthography present condition. They replicated the result with 10- to 11-year-olds and different tasks (definition production), and showed that the presence of spelling was particularly helpful for learning the phonological forms of words among better readers, who reached the learning criterion faster in this condition. In a similar study, Ricketts et al. (2009) found that 8- to 9-year-old children learnt the meanings (assessed via a word-picture matching task) and orthographic forms (spelling task) of new words better when words were presented both orally and written than when only an oral presentation was provided. Interestingly, in both studies, this effect was significant even when children were not alerted to the presence of the orthography. The orthographic facilitation effect has also been reported in children with learning disabilities (Autism, Down Syndrome and SLI; Lucas & Norbury, 2014; Mengoni, Nash, & Hulme, 2013; Ricketts, Dockrell, Patel, Charman, & Lindsay, 2015) and in children learning English as a second language (Hu, 2008). For second language learners, phonological awareness interacts with the presence of orthography, such that children with higher abilities benefit more from the dual presentation.

The studies reviewed so far presented new word forms and meanings in isolation, outside of a story context. Vadasy and Sanders (2015) explored the benefit of orthographic presentation when words were provided in a story context; English-speaking children in kindergarten (5-year-olds) were presented with stories orally, with one group also exposed to the written forms of the words after they had been encountered orally. This group outperformed the listening-only group in spelling the words, and showed a trend towards better performance in word comprehension and definition production tasks. The facilitative effect of orthography seems most prominent for alphabetic languages (Li et al., 2016) and children. Adults, in fact, may experience facilitation in form learning, but not meaning learning, when provided with input in more than one modality (Miles et al., 2016). In summary, when children have been taught phonological forms and semantics either with or without orthography, they show more learning of phonology, semantics and orthography for items that are taught with orthography present. To date, all studies investigating orthographic facilitation have employed a direct instruction approach to teaching new words. Generalisation to an incidental learning context, where children's attention is not directly drawn to the new words, remains to be explored.

In line with these results, studies investigating the impact of multimodal presentation on memory for lists of words or verses have found superior performance when items are presented orally and written at the same time, or only written, compared to when items are presented only orally. These results have been replicated in adults and 8-year-old children (Hartman, 1961; Menne & Menne, 1972).

Another source of evidence that simultaneously listening to and reading stories leads to better learning comes from studies that test the existence of a 'phonological facilitation' effect, the superiority of a combined presentation modality to the written presentation modality. In a study conducted by Rosenthal and Ehri (2011), children in fifth grade (10- to 11-year-olds) read stories that contained novel words. They were asked to pronounce half of the new words aloud when they encountered them. Phonological learning (assessed by asking children to pronounce the word corresponding to a given definition), orthographic learning (assessed through a spelling task) and semantic learning (a definition recognition task) were all greater for words that had been pronounced, confirming the presence of phonological facilitation. However, phonological learning was similar in the two conditions when pronunciation approximations (i.e. words similar to the target word, containing the same initial phoneme, and sharing at least one other phoneme or syllable) were allowed. It is therefore unclear whether there was phonological facilitation for phonological learning. Even though, in this study, words were embedded in passages, simulating incidental learning from reading, attention was drawn to the target words by underlining them, and only these words were pronounced aloud.

Other evidence of the positive effect of the presence of phonology comes from the literature on self-teaching. For example de Jong and Share (2007) found that 8- to 9-year-old Dutch children were faster in reading aloud words they had encountered in a passage compared to homophones (i.e. they were faster in reading aloud *slouk* than *slauk*, if they had previously encountered *slouk*, even though reading both items required to produce the same oral forms), and that this effect was particularly prominent when they had read the passage aloud, rather than silently. Similar effects were not found in an orthographic choice task or spelling task. While this study compares the effect of presence or absence of phonology of words encountered in a story context, it must be noted that phonology was not directly provided, but was built by the children through phonological recoding. Similarly, Kyte and Johnson (2006) tested the effect of phonological suppression by presenting single words to 10-year-olds, and asking them either to read them aloud or silently while repeating a nonsense syllable (reading silently with concurrent articulation). In this study children learnt orthography significantly better in the reading aloud condition, when tested in an orthographic choice and a spelling task. However, differences between the reading aloud and reading silently with concurrent articulation conditions were not consistent across analyses. When orthographic knowledge was assessed using a reading aloud (word naming) task, children were slightly faster in the reading aloud condition, but no more accurate. In summary, these studies show that orthographic, phonological and semantic learning for new words may be enhanced by the presentation of phonology in context, albeit in all these studies the phonology was derived by the participants, rather than directly provided.

No study has so far compared children's learning of words in listening, reading, and combined listening and reading conditions when words are presented incidentally in stories in children's first language. Some studies of multimedia story presentation (e.g. Takacs et al., 2014) have compared conditions where the written and oral texts were present to conditions where only one of these was present, but in general studies involving e-books compare conditions that differ in multiple ways, as discussed above, making their interpretation in terms of presentation modality alone difficult. For example, Terrell and Daniloff (1996) presented new words in a story context in three modalities: in a narrated story on television, in a story on a computer, or in a picture book that included the written text and was read aloud by an adult. Preschool children's word recognition in this study was better in the picture book condition, but it is unclear whether it was the written presentation or the adult interaction that was the key factor in enhancing performance in this study.

Specific comparisons have been made between dual and single presentation modalities in adults learning English as a second language (L2 learners): these studies typically find superior learning when material is presented in written and dual modality versus spoken formats (Brown et al., 2008; Chang, 2011; Sydorenko, 2010). For example, Chang (2011) showed that L2 English learners who were exposed to books in both the oral and written modalities experienced significant gains in their vocabulary knowledge compared to a control group. In another study, Chang (2009) reported that second language learners showed better comprehension of stories when these were presented in a combined oral and written modality compared to when they were presented in the listening modality only. In a study that compared all three presentation modalities (listening, reading and combined listening and reading), Brown et al. (2008) found greater learning of word meanings (assessed by multiple choice and translation tasks) in reading and combined conditions, compared to a listening condition, in Japanese adults. Similar results were obtained in studies where videos with captions (which could be considered a dual modality of input) were used to assess word learning, especially the acquisition of word meanings (Neuman & Koskinen, 1992; Sydorenko, 2010). In these studies, L2 learners and bilingual 11- to 13-year-olds learnt vocabulary (assessed via a written recognition task and a meaning recognition task) from videos with captions better than from videos without captions, but they learnt the same number of word meanings from videos with captions with or without the presence of the audio (Sydorenko, 2010). These results are not surprising for adult second language learners, since they tend to rely on the written text more than the oral input for comprehension (Lund, 1991). For them, the combined listening while reading modality allows the written text to support their poorer listening abilities. Reading may be the preferred medium of language learning for second language learners because many aspects of the reading process have already been acquired and can be transferred across languages. For example, decoding abilities and cognitive strategies can be partly transferred from one language to the other, if the two languages share the same writing system (e.g., the alphabetical system; Bialystok, Luk, & Kwan, 2005; Koda, 1990).

In contrast to adult L2 learners, children learning to read in their first language tend to have stronger listening than reading comprehension skills, and may rely more on spoken language than the written medium. If so, the generalisation of results from studies of adult second language learning is not straightforward. We might, in fact, expect children to acquire vocabulary from listening while reading to stories as well as from listening to them, and from both these conditions to a greater extent than from a reading only condition, especially if their reading abilities are not well developed. On the other hand, given the hypotheses illustrated in Section 1.4, written (and combined written and oral) presentations may have an advantage over oral presentations due to the superiority of orthographic memory traces, or the better representation elicited by the written modality. Given the existence of both an orthographic and a phonological facilitation effect in children, a dual presentation modality might be expected to facilitate vocabulary learning compared to either reading or listening presentations.

In summary, evidence of orthographic and phonological facilitation for vocabulary acquisition has so far come from studies presenting words in isolation, rather than in context (Ricketts et al., 2009; Rosenthal & Ehri, 2008), and from studies that confine the dual modality of presentation to the words of interest, rather than to the narrative as a whole (Rosenthal & Ehri, 2011). Evidence for a beneficial effect of dual presentation has also come from studies of adult second language learners (Brown et al., 2008). Questions therefore remain about whether such facilitation effects are maintained when the dual presentation modality is extended to the entire story in which words are embedded, and when children are learning words in their first language.

1.5.3 - Explaining the advantage of a dual modality presentation

Why might there be an advantage for a dual presentation modality over a single presentation modality? One possibility is that providing both oral and written forms for items frees up attentional resources during encoding, meaning that resources can be allocated to story comprehension and to the encoding of word meanings. This idea might be compared to how good decoders can use their robust knowledge of the link between orthography and phonology to generate phonology automatically while reading, leaving more resources for text comprehension and the learning of word meanings (Ehri, 2014). It also resonates with cognitive load theory in multimedia learning (Mayer, 2014; Mayer, Moreno, Boire, & Vagge, 1999; Paas, Renkl, & Sweller, 2003), which states that situations that reduce cognitive load are more conducive to learning. This framework has been used to identify types of instruction that facilitate learning by reducing the effect of extraneous cognitive load. Extraneous cognitive load is caused by elements that require the use of working memory but do not increase knowledge per se (Sweller, 2010). One of the tasks that children have to perform when encountering a new written word is to recode it to create a phonological representation; if this task is performed for them, via the direct provision of an oral form, we might assume that the extraneous cognitive load has been reduced. This idea is supported by Paivio's dual coding theory (1986), which considers the existence of two separate channels to process visual and auditory information and suggests that engaging both channels, creating cross-channel (dual) representations, is at the heart of successful learning (see also Baddeley's model of working memory: Baddeley, Eysenck, & Anderson, 2009).

Another possibility is that, after a dual presentation, multiple cues to the same item (phonological, orthographic and semantic) are stored in memory, and all these cues can be exploited to access the word in memory. This idea is in line with episodic theories of word identification. For example, Reichle and Perfetti (2003), who base their hypotheses on an episodic memory framework (MINERVA 2: Hintzman, 1984), proposed that the information provided by each new encounter with a word is represented via three different features, phonology, orthography and semantics. The ease of retrieval of any of these features given any other (e.g. pronunciation or meaning from spelling) would depend on how well the information was encoded and on the overall content of the long-term memory trace of the word. As Perfetti and Hart (2002) proposed in the Lexical Quality Hypothesis, lexical representations that include both phonological and orthographic information are of 'higher quality', and therefore more easily retrieved than words represented with only phonological or orthographic information. The simultaneous presentation of phonology and orthography might result in a memory trace that is more informative about the word than a single modality presentation, creating a representation of better quality in memory. The Lexical Quality Hypothesis proposes that when words are of higher quality in memory, their processing is easier, and processing capacity can be freed for other tasks, such as reading comprehension.

There are therefore two, non-mutually exclusive accounts of how a combined (dual) presentation modality might facilitate vocabulary acquisition. One states that a dual presentation facilitates the process of word learning online by freeing resources, either from the need to decode the text (compared to a written only presentation) or from the need to attend to the presentation over time (compared to a listening presentation). Alternatively, dual presentation may facilitate the creation of a representation of higher quality in memory, which in turn would free attentional resources after the representation has been created. The nature of the outcome of the word learning process might differentiate these two ideas.

It could be hypothesised that, if the first account is correct, and a dual modality of presentation facilitates the process of word learning by freeing attentional resources online at the point of encoding the new word, children might acquire representations of similar quality in this condition, compared to single modality conditions. On the other hand, they should acquire more information regarding word meanings from dual presentations. For example, if we compare the reading and combined conditions, we might expect no difference in children's orthographic learning, since the two conditions provide the same information. The same is possible with phonological learning: children might acquire phonology via recoding in the written modality, and via direct presentation in the dual modality. This would be especially true for good readers, who are able to create phonological representations from reading with ease. However, learning phonology using decoding could be considered more effortful than acquiring phonology via direct presentation of the oral form, thus all children in the dual modality presentation condition will have more resources free to understand the text, and to encode the new word meanings, thus obtaining stronger semantic representations than the single modality presentation group.

On the other hand, if the second account is correct, and dual presentation modality enhances vocabulary acquisition by creating a better representation of the items in memory, and strengthening the links between phonology, orthography and semantics for a particular word, we might expect a better performance in all tasks (phonological, orthographic as well as semantic) among children in the combined condition.

Although the idea that a dual presentation modality frees resources at encoding is appealing, whether bimodal presentation does in fact reduce cognitive load might be questioned. The need to process two forms simultaneously might actually be more cognitively demanding than processing one form. In addition, is the process of recoding written forms into oral forms really more demanding than paying attention to two channels at the same time? This assumption violates the redundancy principle in multimedia learning, which states that students learn more when the same information is not presented in more than one modality (i.e. the information given by two different modalities should be related, but not identical; Kalyuga & Sweller, 2014). Although we might argue that oral and written presentations of the same text provide different representations, it is clear that the underlying message is identical. A second issue is whether phonological recoding should be considered extraneous cognitive load. While it is true that recoding is an effortful process, especially when reading has not been fully mastered, previous research has shown that recoding new words is automatic: phonology is activated even during fundamentally written tasks (e.g. Haigh, & Jared, 2007; Liberman et al., 1977). A third difficulty in applying the theory of cognitive load to vocabulary acquisition is that this theory was built to explain secondary learning, the acquisition of culturally-specific knowledge such as scientific learning, rather than biologically primary learning, the acquisition of information that we have evolved to process, such as language (Mayer, 2014). For these reasons, it is unclear whether we should consider the dual modality presentation as less cognitively demanding than a single modality presentation, at least in the framework of multimedia cognitive load.

1.6 - Other factors influencing vocabulary acquisition

1.6.1 - Learning new words from context: the effects of context variability and definitions

Although both pre-school and school-aged children can derive the meaning of new words from discourse and oral texts even when very scarce information is provided (Rice, Buhr, & Nemeth, 1990; Oetting, Rice, & Swank, 1995), learners need several high-quality presentations of new vocabulary items, in informative contexts, to be able to learn words in depth, and form semantic associations between the words (Beck, McKeown, & Kucan, 2002; Frishkoff, Perfetti, & Collins-Thompson, 2011; Schwanenflugel et al., 1997). Repeated presentations of a word increase the probability of encoding it in memory, and diverse contextual presentations provide varied information about its meaning, facilitating deeper learning (Blachowicz & Fisher, 2000; Nation, 2017).

As discussed above, the probability of learning a word increases with the number of encounters (e.g. Durkin, 1990, cited in Swanborn & de Glopper, 1999). For example, children exposed to a written word ten times are more likely to learn it, compared to those who read the words only four times (Stanley & Ginther, 1991). The frequency effect demonstrates the consequences of repeated encounters: words that occur more frequently in the language are identified faster in lexical decision tasks (Whaley, 1978), have faster naming speed (Forster & Chambers, 1973), and faster tachistoscopic identification (Solomon & Howes, 1951). Thanks to repeated presentation to high frequency words, participants are likely to have formed a better representation of these words in memory, including more detailed information regarding their meaning and their forms. High frequency words would therefore be better integrated in the semantic network, and easier to read by sight.

While children can learn words incidentally even when the context is not particularly informative (Nagy et al., 1987; Shatz & Baldwin, 1986), several studies support the idea that encountering a word in several different contexts or providing a definition (which increases context informativeness) foster vocabulary learning in both oral and written presentations (Coyne, Simmons, Kame'enui, & Stoolmiller, 2004; Dickinson, 1984; Justice et al., 2005; Wilkinson & Houston-Price, 2013). Ricketts et al. (2011) presented new words in specific or general written contexts to 7- to 8-year-old children. Children learnt more precise semantic information, as assessed by a word-picture matching task, when provided with a more specific context (that a giraffe is a specific type of animal, rather than just an animal). However, they were able to infer the superordinate category of the word even from the more general contextual presentation. In this research, therefore, the specificity of the context had an effect only when specific knowledge of a word meaning was required. Other studies exploring word learning from oral story presentations found that definitions embedded in story context were more beneficial for word learning than context alone (Wilkinson & Houston-Price, 2013), while research involving written story presentations has shown that providing both context and a definition for new words fosters learning to a greater extent than definitions alone (Stahl, 1983). From these studies, it seems that children who were provided with more information regarding words were able to form richer semantic representations of those words, compared to children who received less information, and they were therefore more likely to learn and remember them.

Other researchers have explored the effect of contextual repetition and of the type of contexts on the representation of words in memory (Bolger, Balass, Landen, & Perfetti, 2008; Frishkoff et al., 2011; Reichle & Perfetti, 2003) within the framework of episodic memory theories. One of these models (Reichle & Perfetti, 2003) considers learning of word forms as the product of a number of episodes where phonological and orthographic forms, as well as contextual information about a word, have been encountered. In this framework the meaning of a word is considered as the sum of the information abstracted from numerous contexts, rather than a static representation in memory (Tenpenny, 1995). Definitions act as specific contexts, and may facilitate the identification of a word's core meaning (set of defining features). Within a memory framework, more similar contexts are characterized by overlapping traces in memory, enabling the stronger acquisition of a small number of features. Conversely, more diverse contexts, characterized by less overlap, allow for a larger number of features to be acquired, but each encounter will leave a weaker memory trace. The Lexical Legacy Hypothesis (Nation, 2017) emphasises the importance of semantic diversity in the contextual input, suggesting that reading does not facilitate the acquisition of lexical quality through repetition, but that written language provides greater contextual variation, which fosters stronger representations.

While context, especially diverse context, seems to facilitate the acquisition of new word meanings, the same is not always observed for learning of word forms. Studies that have explored orthographic learning from reading or reading and listening either find no facilitation from context (Cunningham, 2006; Nation et al., 2007; Ricketts et al., 2011), or even lower levels of orthographic learning from context than from single item presentation (Landi et al., 2006; McKague, Pratt, & Johnston, 2001; Stuart, Masterson, & Dixon, 2000). In the study by Ricketts et al. (2011), differences in the specificity of contextual information did not impact on orthographic learning, in either an orthographic choice task or a spelling task. Context facilitates word recognition, making reading aloud less effortful, by increasing a word's predictability and, possibly, pre-activating its representation (Nation & Snowling, 1998; Stanovich, Nathan, West, & Vala-Rossi, 1985). Nevertheless, the presence of context could impair learning of specific orthographic or phonological representations, by compelling children to allocate more attention to the meaning than to the form of the word. For example, Landi et al. (2006) found that 5- to 8-year-old children were more likely to read new words correctly if these were presented in context, but retention of the ability to read the words aloud over time was greater for words presented in isolation. These findings suggest that words acquired in isolation have a better quality form representation in memory than words acquired in context. These results might be considered problematic for the Lexical Quality Hypothesis, which would predict that learning encounters that facilitate the acquisition of representations of higher quality (i.e. contextual presentation) would affect both form and semantic representations. Some authors have explained these results by suggesting that the negative effect of context is due to its higher demands on attention processes (Landi et al., 2006).

However, the effect might depend on testing modality and item features. Ouellette and Fraser (2009) presented new words in the written and oral (reading aloud) modality to 9- to 10-year-old children. Words were either accompanied by a definition or not. Results showed positive effects of the presence of semantic information when orthographic knowledge was tested in an orthographic recognition task, but not when tested in a spelling task. In a similar study that explored orthographic learning using an orthographic recognition task, 6- to 8-year-olds were presented with words both orally (reading aloud) and in the written form, within stories or word lists (Wang, Castles, Nickels, & Nation, 2011). In this study the context condition enhanced orthographic learning for irregular words but not regular words. Similarly to the study by Ouellette and Fraser, Bond (2014) found that 8-year-olds learnt to read words aloud more easily when training presented words in context rather than on their own. On the other hand, they learnt the spellings of words similarly in the two conditions.

In conclusion, while it is clear that presenting new words in context enhances semantic learning compared to presenting words in isolation, regardless of modality, orthographic learning is not always facilitated by contextual presentation, and the effect of contextual presentation on orthographic learning depends on the task and word type. Furthermore, while diverse contextual presentation may have a positive effect on semantic learning over time, repeated presentation might cement words in memory faster, albeit with less associated semantic information. An open question is whether the provision of specific contextual information supports the acquisition of both specific and more general semantic knowledge of a word.

1.6.2 – Word features

Grammatical category, concreteness and conceptual complexity (i.e. the extent to which the meaning of a new word is similar to the meanings of known words) all affect the ease with which children learn new words in both the oral and written modalities (Durkin, 1990, cited in Swanborn & de Glopper, 1999; Gentner, 1982; Nagy et al., 1987; Schwanenflugel, 1991; Schwanenflugel et al., 1997). In addition, some studies show that increasing knowledge for partially known words is easier than learning words never encountered before (Durso & Shore, 1991), although this result has not always been replicated (Schwanenflugel et al., 1997; Adlof, Frishkoff, Dandy, & Perfetti, 2016). While the effect of familiarity or partial word knowledge on semantic learning is debatable, it seems that familiarity has a clear positive effect on learning word forms (Adlof at al., 2016).

Word-likeness affects the probability of acquisition: words that are less similar to others due to less regular letter or sound patterns are initially easier to acquire, which Storkel, Armbrüster and Hogan (2006) interpreted as due to saliency. Nevertheless, the probability of these more salient words being retained after training is lower than that of more word-like words (Aicher, 2016; Frisch, Large, & Pisoni, 2000; Luce & Large, 2001; Thorn & Frankish, 2005). This effect is usually explained at the level of the lexicon: incorporating new items within the existing framework is easier if their forms are more similar to existing phonological and orthographic patterns. For similar reasons words that can be decomposed into morphological constituents may be easier to acquire than root words, especially if the root of the new morphologically-complex word is known (McBride-Chang, Wagner, Muse, Chow, & Shu, 2005).

In conclusion, word level features affect the probability of learning both word meanings and forms. It is therefore important to consider such features when selecting items for word learning studies. Specifically it is important to take into account their concreteness, grammatical category, conceptual complexity, word-likeness and familiarity.

1.6.3 - Individual differences

Individual differences also constrain whether children successfully learn new words from spoken and written language exposure. When considering learning word meanings from listening to stories, several researchers (Penno et al., 2002; Robbins &

Ehri, 1994; Sénéchal, Thomas, & Monker, 1995) have found that children with higher oral language and vocabulary skills experience greater gains in word learning than lower-skilled children (cf. Justice et al., 2005). Amongst second language learners, children with high proficiency in the second language show the greatest gains in semantic learning of new words (Lin, 2014). When learning words from reading and reading aloud, reading skills, especially reading comprehension (Jenkins et al., 1984; Swanborn & de Glopper, 1999), vocabulary knowledge and working memory have all been highlighted as factors influencing learning of word meanings (Bartolotti & Marian, 2016; Cain, Oakhill, & Elbro, 2003; Cain et al., 2004; Jenkins et al., 1984; Ricketts et al. 2011; Swanborn & de Glopper, 1999). For instance, Swanborn and de Glopper (1999) found that both age and reading comprehension influenced participants' ability to learn words from context, and several studies have shown that intelligence, vocabulary knowledge, working memory and reading comprehension are related to the ability to extract word meanings from written text (Cain et al., 2003; Sternberg & Powell, 1983; Swanborn & de Glopper, 2002). Cain et al. (2003) showed, for example, that seven- and eight-year-olds' reading comprehension predicted their ability to extract meanings of new words in context. In this study, children with low comprehension skills were able to capitalise on information provided near the new word in the text, but they were unable to make a connection between the word and its meaning when this information was further away in the text. In experiments where both the oral and written forms of the words were available, vocabulary knowledge has been found to correlate with semantic learning of the new words (Rosenthal & Ehri, 2008).

When considering learning of the phonological forms of new words, as assessed by word production tasks, decoding skills, measured by word and non-word reading abilities, are a significant predictor of learning, when words are presented orally or both orally and written (Rosenthal & Ehri, 2008). In a similar study, better decoders remembered the phonological forms of the words they had read, and pronounced these better, than less skilled decoders (Rosenthal & Ehri, 2011). Decoding skills have also emerged as a significant predictor of learning of the orthographic form of words both after reading (Bowey & Miller, 2007) and reading aloud (Ricketts et al., 2011; Rosenthal & Ehri, 2011). More controversial is the effect of vocabulary on orthographic knowledge. For example, Ouellette and Fraser (2009) found that vocabulary, as well as decoding skills, predicted both orthographic recognition and spelling after a read-aloud presentation, while other studies have not found this effect (Ricketts et al., 2011).

As discussed in Section 1.5, presenting words simultaneously in their oral and written forms is more beneficial to word learning than presenting spoken words alone (orthographic facilitation). The extent to which children show orthographic facilitation in their semantic learning is associated with their decoding skills, and the same correlation has been found for orthographic learning (Ricketts et al., 2009; Rosenthal & Ehri, 2008). The amount of orthographic facilitation seen in semantic learning is correlated with phonological awareness in children learning English as a second language, with children with higher phonological awareness benefiting more from the accompanying presence of orthography (Hu, 2008). On the other hand differences between groups in non-verbal abilities, vocabulary knowledge, phonological awareness and verbal short term memory do not seem to determine whether orthographic facilitation is seen in semantic and orthographic learning (Lucas & Norbury, 2014; Mengoni et al., 2013; Ricketts et al., 2015).

Similar effects have been explored in studies assessing the facilitation provided by elicitation of the phonological form when reading (phonological facilitation). Rosenthal and Ehri (2011) divided 9- to 11-year-old children into good and poor word readers, and found that both groups produced more correct pronunciations of the target words (phonological learning), when they had phonologically recoded the words, but that this difference was not significant for good readers when approximations to the pronunciations were accepted as correct. Furthermore, no difference between ability groups was found in the results of a spelling task (orthographic learning). Thus, this study shows that the extent of phonological facilitation for learning phonology may depend on children's reading abilities, while the extent of phonological facilitation for orthographic learning does not.

These findings lead to the conclusion that better decoders are more sensitive to the link between orthographic and phonological forms of the words, and can use this ability to secure the new words in memory, making them able, at a later stage, to recall the words' meanings and forms more easily than less skilled readers (Ricketts et al., 2009; Rosenthal & Ehri, 2008; 2011). Given the role of individual abilities in determining word learning, it is important to explore whether children with different abilities benefit to a similar extent from the provision of oral and written information at the same time, in order to discover the optimal presentation modality for children of different abilities. The relationship between decoding skills and word learning seen in single word presentation tasks may generalise to learning from story presentations.

Alternatively, abilities such as oral language or reading comprehension might impact more strongly on word learning in contextual presentations, or in a dual modality presentation, where children need to pay attention to two modalities simultaneously, and integrate the information from both.

1.7 - Introduction to the present research

This work explores the effect of presentation modality on word learning by comparing vocabulary acquisition from listening to stories, reading stories and listening and reading stories at the same time (combined condition), in a population of Year 4 (9 year-old) children. This population was chosen for their relatively well-developed basic reading skills, alongside their large variation in reading abilities. Measures of children's acquisition of phonological, orthographic and semantic knowledge of a set of new words were collected; semantic knowledge tasks were designed to explore the precise nature of children's representation of word meanings.

Study 1 explores the effect of class-based oral and combined oral and written presentations of a story on children's phonological, orthographic and semantic learning of new words contained in the story. This study was designed to study story presentation effects in a naturalistic context, and investigated both immediate and longer term retention of word knowledge. Definitions were embedded in the stories to facilitate vocabulary acquisition. Study 2 compares three story presentation modalities (reading, listening, and reading and listening simultaneously) delivered to small groups. This study also explores the effect of providing a definition, by comparing children's learning of words presented with and without a definition within the story context. Study 3 elucidates the strategies employed by children within the reading and combined reading and listening conditions, by recording eye-movements during story presentation, measuring recognition and production of the new vocabulary within the story, and assessing the link between their phonological and orthographic representations. In each study, measures of individual abilities were also collected, to assess their role in children's learning. To this end, measures of reading accuracy, reading comprehension, vocabulary, oral language comprehension and non-verbal skills were collected.

Studies of incidental vocabulary acquisition from stories have shown that school-aged children acquire new words both from listening to stories and from reading them (sections 1.2.4 and 1.3.4). It is also clear that providing detailed information regarding the meaning, phonological form and orthographic form of a word facilitates vocabulary acquisition (sections 1.5.2 and 1.6.1). Nevertheless, it is unclear whether children acquire words more easily from the oral or written modality, especially at an age when reading abilities are still developing (section 1.4). Although much research has explored the beneficial effect of a dual presentation in single word presentation paradigms, the effect of a simultaneous presentation of oral and written texts has not been systematically compared to single modality presentations in children's word learning (section 1.5.2). The present research aims at filling the gap between the literature on single word presentation, which shows a clear benefit of a dual presentation modality for word learning in school age children, and the literature on vocabulary acquisition from stories, where a direct comparison between the three possible story presentation modalities has not been carried out, at least with children learning their first language, and when exploring both semantic, phonological and orthographic learning.

Reading stories aloud is a common activity within primary school classes, especially in the early years, while older children are often encouraged to read stories silently on their own. Given that both reading, listening, and listening and reading stories at the same time are activities that facilitate vocabulary acquisition, the present research aims to establish which of these modalities of story presentations best promotes word learning, which is of practical importance to inform classroom practice.

From a theoretical perspective, the first two studies explore the hypothesis that different presentation modalities support the acquisition of words of higher or lower lexical quality. The framework provided by the Lexical Quality Hypothesis (Perfetti & Hart, 2002) has previously been applied to evidence that better readers create higher quality representations of new words (Perfetti, 2007), but it has not been used to make clear predictions about the conditions under which words are optimally encoded or retained. The current studies, therefore, explore whether word representations created through different presentation modalities vary in their lexical quality. The third study explores attention allocation in dual presentation and reading only modalities, to assess whether dual presentation facilitates learning by freeing resources online, as suggested by the cognitive load theory in multimedia learning (Mayer et al., 1999; Paas et al., 2003).

Several methodological issues were considered when choosing how to assess vocabulary acquisition. Previous studies (James et al., 2017; Tamura et al., 2017) have distinguished between measures of lexical configuration, which assess knowledge regarding the new words, and measures of lexical engagement, which assess whether the newly acquired item competes with other items in the lexicon. Although lexical engagement is considered a measure of the presence of the items within the lexicon, and tends to appear over time (Tamura et al, 2017), we chose to measure the acquisition of the new item features more directly, by focussing on measures of lexical configuration, more commonly collected in research on orthographic and phonological facilitation (Ricketts et al. 2009; Rosenthal & Ehri, 2011).

When considering semantic learning, learning may be assessed in terms of recognition or production. This distinction is often related in the literature to the distinction between measures of breadth of vocabulary knowledge, such as word-picture matching tasks or category or definition selection tasks (Nagy et al. 1987), and measures of depth of vocabulary knowledge, such as tasks that require the participant to produce a definition for the given word (Ouellette, 2006). However, we consider that depth of knowledge can be assessed using recognition tasks, by asking participants to recognise specific features of words. This allows us to differentiate the contrast between depth and breadth of vocabulary knowledge from the contrast between production and recognition tasks, which might be important, given that one type of task (production), could be intrinsically more difficult that the other (Swanborn & de Glopper, 1999). Following this approach, a semantic task was created for Studies 1 and 2 that was composed of three parts: a category recognition sub-task, which assessed participants' ability to correctly recognise the category of a given item between four alternatives; a sub-category recognition sub-task, which required participants to select the sub-category for a given item; and a definition recognition sub-task, that required participants to recognise the correct definition of the given word.

The use of different measures of learning of word forms can lead to inconsistent results. Studies testing orthographic knowledge, for instance, have used assessments that require the child to either recognise or spell the correct orthographic form, or both; effects of modality of presentation and context have been reported to differ, depending how orthographic knowledge is measured (e.g. Ouellette & Fraser, 2009). Moreover, spelling tasks assess the link between the phonology and orthography of the word, rather than the orthographic representation per se. Further difficulties arise when considering knowledge of the oral form of the word, sometimes tested using a word production task where the correct definition is given and the subject has to produce the correct word (e.g. Rosenthal & Ehri, 2008). The difficulty with this assessment is that it does not test phonological representations per se, but the link between phonology and semantics. Word form knowledge can also be assessed through a reading aloud task, but this assesses the link between the phonological and orthographic form of an item. Thus, both spelling and reading aloud assess the strength of the link between the two forms of a word, rather than the strength of each form independently. Moreover, when completing these assessments children may rely on their knowledge of grapheme-phoneme correspondence rules; their performance might reflect general reading and spelling abilities rather than their knowledge of the target words. Another assessment used in the literature is the lexical decision task, where the participant is asked to recognise whether a specific item is a word or not. While this task can be performed in either oral or written format, enabling independent assessment of access to the written or oral lexicons, the results of semantic priming tasks indicate that, when asked if an item is a word or not, at least partial activation of the underlying meaning of the word occurs (e.g. Neely, 1991).

Assessing phonological and orthographic representation separately from each other and separately from semantic knowledge is therefore very complex. In the present research we attempted to assess the quality of children's phonological and orthographic representation separately, devising two separate tasks to measure phonological and orthographic learning. In the phonological task children were asked to select the correct oral form of the target words among two alternatives, without reference to the written form. Similarly, in the orthographic task, we tested children's ability to recognise the written forms of words between two alternatives, without reference to the oral forms of the words, in a test similar to that devised by Share (1999). Study 3 additionally explored the links formed between orthographic and phonological forms, by asking children to decide whether an oral form corresponded to the given written form.

In conclusion, the main aim of this work was to compare story presentation modalities, to discern whether school-aged children acquire words more easily from reading, listening to, or reading and listening to stories at the same time, when presented with rich contexts and definitions. From previous research we expected that the combination of oral and written presentations would be the most conducive for the acquisition of both forms and meanings (Ricketts et al., 2009; Rosenthal & Ehri, 2008;

2011). A second aim was to explore the source of the benefit of combined presentation in terms of children's attention and resource availability in that modality compared to others (Mayer et al., 1999). Finally, we aimed to elucidate the role of individual differences in vocabulary acquisition, and ascertain whether all children benefit similarly from dual presentation, or whether this effect is moderated by children's abilities. It was expected that reading ability would affect learning of word forms, especially in the reading and reading and listening conditions (Ricketts et al., 2011), while vocabulary knowledge would predict children's ability to learn the meanings of new words (Cain et al., 2004; Lin, 2014; Penno et al., 2002; Ricketts et al., 2011; Sénéchal et al., 1995).

Chapter 2: Study 1

2.1 - Introduction

This study looked at how children learn new words from stories, comparing a listening condition where stories were presented only orally, to a combined condition where stories were simultaneously presented in both the oral and written modalities.[1] Learning of the phonological and orthographic forms of the words was explored, as well as learning of the semantic features of these words. Regarding the meaning of the words, learning of different features of word meanings was explored. The comparison between two conditions allowed us to directly test hypotheses about the additive benefit of orthography for phonological, orthographic and semantic learning.

2.1.1 - Learning new words from stories

As discussed in Chapter 1, many studies have shown that listening to stories fosters vocabulary development in children of all ages and abilities (Brett et al., 1996; Elley, 1989; Houston-Price et al., 2014; Penno et al., 2002; Robbins & Ehri, 1994; Sénéchal & Cornell, 1993; Walsh & Blewitt, 2006; Wilkinson & Houston-Price, 2013; Williams & Horst, 2014). For example Elley (1989) found a significant amount of incidental vocabulary acquisition in 7- and 8-year-old children when a story was read three times over a 7-day period in the classroom. Similar gains in semantic knowledge of the presented words were found for children from kindergarten up to the age of 11 (Brett et al., 1996; Penno et al., 2002; Robbins & Ehri, 1994; Sénéchal & Cornell, 1993).

In a different line of research many studies also showed that children in Year 2 and above can learn new words when reading stories (Jenkins et al., 1984; Nagy et al., 1987; Ricketts et al., 2011; Schwanenflugel et al., 1997). For example Nagy et al. (1987) found that American 8- to 13-year-olds learnt the meanings of new words, tested

[1] In the original design of the study it was expected that each class would be exposed to the stories in three different conditions. The three conditions were: 1) story reading, 2) story listening, and 3) a combined listening plus reading condition. Unfortunately, due to difficulties during data collection, namely the drop out of a class from the study and the misinterpretation of the instructions for one story presentation by one of the teachers, it was, in the end, only possible to collect data for conditions 2 and condition 3.

through a definition-choice task, when reading stories. Similar results were found in more recent studies (Ricketts et al., 2011; Schwanenflugel et al., 1997).

A number of studies have compared the effect of presenting words in different modalities on children's ability to learn the meanings of these words. Suggate et al. (2013), for example, found that children who read the stories by themselves were less likely to learn the meanings of the new words embedded inside these stories than children who listened to the stories read by the adult. Furthermore, a number of studies using single word learning paradigms have shown that presenting both the written and the oral form of a word can be beneficial for learning a new word's meaning, compared to a single modality presentation (Ricketts et al., 2009; Rosenthal & Ehri, 2011; 2008). Second language learners have shown similar advantages of a dual modality presentation over an oral presentation (Brown et al., 2008; Neuman & Koskinen, 1992; Sydorenko, 2010). Therefore, it seems that being exposed to the oral and the written form of new words leads to a better performance in semantic and phonological tasks, compared to being exposed to words only orally. Several possible explanations for an advantage of a dual presentation modality have been explored in Chapter 1. The main hypothesis considered here is that providing both oral and written forms frees up attentional resources, that can be allocated to story comprehension and to the encoding of new word meanings (Ehri, 2014; Mayer et al., 1999; Paas et al., 2003), either by facilitating the process of word learning or by facilitating the creation of a representation of higher quality in memory (Perfetti & Hart, 2002).

Despite the extent of the literature on the advantage of a combined presentation of oral and written word forms for semantic learning, no study has explored how the presentation of entire stories or passages in a combined modality affects word learning, compared to other presentation modalities, especially when considering both forms and meanings of new words. This modality of presenting texts is often used in schools, and it has been found to have a positive effect on reading abilities and vocabulary (Chomsky, 1976; Rasinski, 1990; Schneeberg, 1977). It is therefore of interest to explore if the positive effects highlighted for word learning in single word presentation tasks can be generalized to a story presentation modality.

Furthermore no previous research has explored the extent to which different presentation modalities affect which semantic features are acquired or the depth of semantic learning. Since vocabulary learning is an incremental process, the knowledge regarding each new word could potentially be enhanced at every new encounter, and the nature of the encounter could influence the specific features acquired (Nation, 2017). For example oral and written mediums provide information in different ways: children can revisit specific parts of a written text if they want to, while they cannot revisit information presented only orally, unless the information is repeated. For these reasons, in this study semantic learning was explored in depth, analysing what information children retained of the words presented within the stories.

As seen in Chapter 1 the factors influencing the probability of children learning new words while reading or listening to stories can be divided into three main groups: individual-level factors, word-level factors and context factors. Individual differences appear to determine the extent of vocabulary acquisition. Particularly, reading skills predict both phonological (Rosenthal & Ehri, 2008) and orthographic learning (Ricketts et al., 2011; Rosenthal & Ehri, 2011), as well as the extent of the orthographic facilitation from dual presentation (Ricketts et al., 2009; Rosenthal & Ehri, 2008). Reading skills also affect learning of word meanings from written presentations (Jenkins et al., 1984), and vocabulary affects semantic learning irrespective of presentation modality (Penno et al., 2002; Ricketts et al., 2011; Robbins & Ehri, 1994; Rosenthal & Ehri, 2008; Sénéchal et al., 1995). In the present study we explored the extent to which individual abilities were associated with word learning within the context of our paradigm, and their effect on the extent of any dual presentation advantage.

While exploring the relationship between individual differences and word learning from stories was one of the aims of the present research, exploring the effects of word type and context was not. For this reason we controlled for word type, choosing rare concrete nouns as the target items. To control for contextual factors, all words were repeated the same number of times and relevant definitions were provided for all of them. Furthermore particular attention was paid to make sure all of the words were an integral part of the plot of the stories, to facilitate learning (Kintsch & van Dijk, 1978).

Vocabulary knowledge is incremental (Nagy & Scott, 2000), and offline consolidation plays an important role in the creation of a stable representation of the word in the lexicon (Henderson et al., 2012; Henderson, Weighall, & Gaskell, 2013). Specifically, words learnt from storybooks tend to be retained for several weeks (McGregor, Sheng & Ball, 2007; Penno et al., 2002; Sénéchal, 1997). Furthermore some research has shown a consolidation effect, with words being remembered better in a second testing session than in a first one (Dumay & Gaskell, 2007; Wilkinson &

Houston-Price, 2013). Therefore, for this research, children were tested twice, to assess both immediate learning (at one day delay), and retention (at one week delay).

2.1.2 – Aims and Hypotheses

The principal aim of the present research was to investigate whether different type of story presentation would foster different levels of vocabulary acquisition, in terms of semantic, phonological and orthographic learning. We particularly aimed to explore the effects of different presentation modalities on naturalistic vocabulary acquisition in a classroom environment, to determine the approach that fosters vocabulary growth most effectively in school-aged children. Studies have compared the effect of presenting words in different modalities in native and second language learning (Brown et al., 2008; Chang, 2009; Rosenthal & Ehri, 2008; Rosenthal & Ehri, 2011; Suggate et al., 2013), but the benefit of presenting stories in a combined listening and reading condition has yet to be explored in children learning their first language. Thus, the present study was designed to compare how well children learn the phonological and orthographic forms and meanings of new words encountered in stories, in conditions that do or do not provide the orthographic and phonological forms of these. Vocabulary acquisition was assessed comprehensively, by analysing different features of word meanings acquired by the children. Based on previous research, we predicted that children would acquire phonological, orthographic and semantic information about words more readily in the combined condition compared to the listening condition (cf. Ricketts et al., 2009; Rosenthal & Ehri, 2008). The rationale for this hypothesis is based on the idea that the combined condition provides more information regarding the new words, especially regarding their written form, compared to the listening condition.

Several studies have analysed how individual abilities relate to word learning (Jenkins et al., 1984; Justice et al., 2005; Penno et al., 2002; Ricketts et al., 2009; Ricketts et al., 2011; Robbins & Ehri, 1994; Rosenthal & Ehri, 2008; 2011; Sénéchal et al., 1995), but none of these studies have explored how existing abilities impact on learning shown in different presentation modalities when words are presented in a story context. Therefore, a secondary aim of this research was to determine whether children with varying abilities would learn words to the same extent or whether oral or written language abilities or non-verbal cognitive abilities would impact on vocabulary learning within each condition. Furthermore, with this research, it was possible to explore how children’s abilities would impact on their learning of different features of the words (i.e. whether different abilities impacted more strongly on phonology learning, orthographic learning or semantic learning). It was specifically expected that children with higher level of vocabulary knowledge would learn more word meanings than children with smaller vocabularies, in both presentation modalities (Robbins & Ehri, 1994; Rosenthal & Ehri, 2008). Furthermore, reading abilities were expected to predict phonological and orthographic learning in both conditions (Ricketts et al., 2011; Rosenthal & Ehri, 2008; 2011). It was also hypothesised that the orthographic facilitation effect expected for semantic, orthographic and phonological learning would interact with decoding abilities, with better readers showing a larger facilitation effect than less skilled readers (Ricketts et al., 2009; Rosenthal & Ehri, 2008).

2.2 - Method

2.2.1 - Participants

Participants for this study were recruited from two Year 4 classes in the same primary school; of the 63 children recruited, 3 with a first language other than English were excluded from the analyses, and 6 did not complete all testing sessions due to absence. The final sample consisted of 54 children (M_{age}= 8.99 years; *SD* = .28; range= 8.50 – 9.42), 28 children in class A (16 boys), and 26 in class B (10 boys). All children had English as their first language.

2.2.2 - Design

To explore our hypotheses we designed a study in which two classes of Year 4 children were presented with two purpose-built stories. Each class was exposed to two stories, in two different conditions, using a repeated-measures design. The presentation of the two stories was sequential: after the presentation of the first story, all the testing sessions linked to it were completed, before the presentation of the second story. The two conditions were: 1) story listening, where the teacher read the story to the children in the class, and 2) a combined condition where the teacher read the story in class, and children were invited to read along with booklets. Thus, in the listening condition, children were exposed only to the phonology of the new vocabulary, and in the combined condition they were exposed to both the orthography and phonology of the words. Each story was presented once a week for three consecutive weeks, following a paradigm used in previous research (Wilkinson & Houston-Price, 2013).

The design was a 2 condition (listening *vs.* combined) x 2 order (listening first *vs.* combined first) design, with one class per cell as described in Table 2.1. The order of presentation of the conditions (combined or listening condition first) and the specific story presented in each condition were counterbalanced. Both groups encountered the India story first, followed by the Normans story.

Table 2.1

First Study Design

	First Story	Second Story
Class A	listening (India story)	combined (Normans story)
Class B	combined (India story)	listening (Normans story)

2.2.3 - Experimental Stimuli

The 16 target words used in the study were of low frequency and chosen from secondary school history topics, specifically the Norman Era, and the British Empire in India. All the words used are real English words and can be found in the Oxford English Dictionary. The 16 target words comprised two sets of 8 words each, one set for each of the two stories written for the study. Each set comprised 8 words, one from each of 8 chosen categories: animal, object, part of a house, building, vehicle, job, clothing and food. An additional control set was composed of 8 words (4 from the two historic periods, and 4 from the Roman Era). All the words were rare concrete nouns, and the three lists (India, Normans and control) were matched for word length and frequency using the MRC Psycholinguistic Database (Coltheart, 1981), the CELEX lemma database (Baayen, Piepenbrock, & Gulikers, 1995) and the SUBTLEX-UK (Van Heuven, Mandera, Keuleers, & Brysbaert, 2014). See Table 2.2 for psycholinguistic properties of the target words in the three lists. For this study, we chose rare concrete nouns to avoid variability due to word type and concreteness. It was assumed that the nouns would reveal the greatest learning from the stories (Gentner, 1982), therefore allowing us to compare the conditions of interest. A detailed description of the words can be found in Appendix A. Sixteen additional items, two for each category, were selected to be included in the pre-test. These words were of higher frequency, and included to ensure children's compliance to the task.

Table 2.2

Psycholinguistics properties of the words used in the stories

Parameter	**Set**	**N**	**Mean**	**SD**	**Min**	**Max**	**Set comparison**	
							F	***p***
Length	Normans	8	6.75	1.49	4	8	.02	.984
	India	8	6.88	1.64	5	9		
	control	8	6.88	1.73	5	10		
Freq. MRC	Normans	6	2.33	2.07	0	5	.31	.740
	India	6	2.00	3.63	0	9		
	control	4	1.00	1.41	0	3		
CELEX Total Freq.	Normans	6	4.17	2.86	0	8	.89[a]	.458
	India	6	38.67	61.71	3	163		
	control	4	5.25	5.56	0	11		
Freq. SUBTLEX-UK	Normans	6	44.50	43.84	5	124	.21	.810
	India	8	54.13	57.02	3	138		
	control	6	37.33	37.46	2	105		
Zipf SUBTLEX-UK	Normans	6	2.15	.501	1.47	2.79	.04	.965
	India	8	2.16	.571	1.30	2.84		
	control	6	2.08	.543	1.17	2.72		

Note. Length = number of letters in the words; Freq. MRC = frequency derived from the MRC Psycholinguistic Database (Coltheart, 1981); CELEX Total Freq. = frequency derived from the COBUILD scale in the CELEX lemma database (Baayen et al., 1995 - the total frequency count of the words in the database was used due to the extremely low frequency of the chosen words); Freq. SUBTLEX-UK = frequency derived from the SUBTLEX-UK (Van Heuven et al., 2014); Zipf SUBTLEX-UK = frequency derived using the Zipf scale from the SUBTLEX-UK; N = number of words used to compute the mean frequency (the three sets were each composed of 8 words, but not all the words were present in all the databases used).
[a] Welch's F is reported since the assumption of homogeneity of variance was violated.

To make sure that the chosen items were unlikely to be known by Year 4 children, we piloted them with 26 Year 7 children. Specifically, we asked Year 7 children if they knew the words, and to provide a meaning for them if they did. A response was determined correct if the child identified the word as a known word, and was able to provide a synonym or a related word. The three sets of words did not differ in terms of how many Year 7 children knew the meanings of the words (we report the Kruskal-Wallis test, since our data did not meet the assumption for parametric analysis: H (2) = 2.56, p = .277). The words known by the highest number of children (*verandah* in the India set and *pottage* in the Normans set) were still known by fewer than 25% of the children. The majority of words were not known by any Year 7 pupil, and only 1 word in the control set, 2 words in the India set, and 4 words in the Norman set were known by a few children. From all 24 words, children tended to know between 0 and 2 words, with 9 children not recognising any of the words, 9 recognising one word, and 8 recognising two words. The words recognised varied from child to child.

A second pilot study was carried out with adults, to investigate adults' knowledge of the words, and their ability to produce the correct pronunciation and spelling. The analysis regarding word knowledge with 18 adults yielded a similar result to the previous reported analysis, i.e. the 3 sets of words were known by similar numbers of adults (H (2) = .65, p = .721). Adults tended to know the words more often than Year 7 students, with the highest known words being *verandah* in the India set, known by 18 adults, and *atrium* in the control set known by 17 adults, with 4 words, 2 from the India set and 2 from the control set, known by none of the adults in the sample. The data from adults were not used to exclude words from the study, since our aim was to teach meaningful vocabulary to the children, which could eventually be useful to them. The fact that adults knew some of the target words was considered good evidence that they would eventually meet these words again.

Adults' ability to produce the correct spelling and pronunciation was used to compare the three sets of words for spelling and pronunciation predictability. Part of the adult sample (8 adults) was asked to write down our stimuli and their spellings were scored for correctness. For each word we considered the number of adults who spelled the item correctly; the spelling of the words in the three sets were equally predictable ($M_{\text{India set}}$ = 3.63; $M_{\text{Normans set}}$ = 3.13; $M_{\text{control set}}$ = 2.75; $F(2, 22)$ = .25, p = .779). The remaining 10 adults were asked to pronounce the stimuli, and their pronunciations were scored for accuracy. The pronunciations of the words in the three sets were equally highly predictable (H (2) = 0.04, p = .982) as all the sets had a mean of 8 adults out of 10 pronouncing them correctly , with 19 words in total being pronounced correctly by 8 or more adults. When scores were converted in percentages, word pronunciations seemed more predictable than word spellings for adults, both for all the sets taken together (we report the Wilcoxon-Signed rank test, since our data did not meet the assumption for parametric analysis: T = 8.00, p < .001), and for all the sets separately (Norman set: T = 0.00, p = .012; India set: T = 1.00, p = .017; control set: T = 2.50, p = .030).

The two stories used in the study were written specifically for this research by the researcher, each story containing its specific set of words. Each target word was repeated three times in its story, and a definition was embedded in the story each time a new target word was introduced for the first time. The definitions accompanying each target word are described in Appendix B. The three repetitions of the same word were distributed within the specific story, so that each word was presented once towards the beginning, once towards the middle and once towards the end of the story. The stories were matched for length, for reading ease, and for grade level. Specific details of each story are presented in Table 2.3. The stories written for this study are reported in Appendix C. The following excerpt provides an example of how the target words and their definitions were embedded in the stories (the word and the definition are highlighted and underlined for ease).

"*A few months ago, when Uncle Jack was in India, he was swimming in a pond, and a* ***gavial***, <u>*a type of crocodile with an elongated muzzle, that eats fish*</u>*, tried to bite his feet off. He must have mistaken Uncle Jack's feet for fish! Uncle Jack was lucky that day. A kind shopkeeper, who happened to be passing by, helped him to escape from the animal. To help Uncle Jack recover from his frightening experience with the* ***gavial****, the shopkeeper invited him to his house.*"

Table 2.3

Properties of the Stories

Parameter	Normans Story	India Story
words	1340	1325
characters	5499	5758
paragraphs	39	41
sentences	86	96
sentence per paragraph	2.3	2.4
words per sentence	15.1	13.4
characters per word	3.9	4.1
passive sentences	4%	5%
flesh reading ease	85.1	80.1
flesh-kinkaid grade level	5.1	5.4

2.2.4 - Vocabulary acquisition tasks

Children's knowledge of the words was measured by accuracy scores, and represented children's ability to recognise the right answer in forced-choice tasks. This methodology enabled us to explore children's knowledge in substantial depth in a short time. Forced-choice tasks incur the possibility that the children might choose the right answer by chance, and this is particularly true for the phonological and orthographic tasks, where there is a 50% chance of the children choosing the right answer by guessing. Nevertheless, the two alternatives procedure was chosen in order to minimise the memory load of the phonological and orthographic tasks. All tasks were programmed using E-Prime 2.0 (Schneider, Eschman, & Zuccolotto, 2002), and presented using a laptop.

2.2.4.1 - Phonological task. The phonological task was designed to assess children's phonological knowledge of the words. In this task children saw two dinosaurs on a computer screen, presented sequentially, each one linked to a specific oral stimulus, one corresponding to the correct spoken form of the word, the other corresponding to a distractor. For example, the distractor for hɔːbək (hauberk) was hɑːbək (see Appendix D). The distractors for this task were created, when possible, by selecting the wrong pronunciation with the highest frequency elicited from the pilot testing with adults. Children were instructed to choose the dinosaur that "said the word best" by pressing the corresponding button on the keyboard (the blue button for the blue dinosaur or the yellow button for the yellow one).

Experimental trials were preceded by a set of instructions to introduce the child to the task, lasting approximately 1 minute, followed by four practice trials, where the child had to choose the right spoken form of four common words (HOUSE, SHIRT, DOG and CAR, presented in this fixed order). Each dinosaur was presented in turn for 2 seconds, during which time the associated oral stimulus was heard once. The two slides forming a trial were repeated in an alternating loop until an answer was provided, and the number of repetitions needed to give a response was recorded. A blank slide of 500 ms divided each trial from the next. Finally a second set of instructions were provided, to tell the child that the following words would be "more difficult words" that were in the stories previously presented in class. The child was also instructed to guess the right answer if he/she did not remember the words. A set of example trials are provided in Figure 2.1.

There were 8 experimental trials, each corresponding to one of the words being tested (Story 1, Story 2, or control words). Both the order of presentation of the words and the link between the two dinosaurs and the correct answer were randomised. Responses were recorded both in terms of accuracy, number of word repetitions and reaction times for each word. The task took between 3 and 5 minutes to be completed.

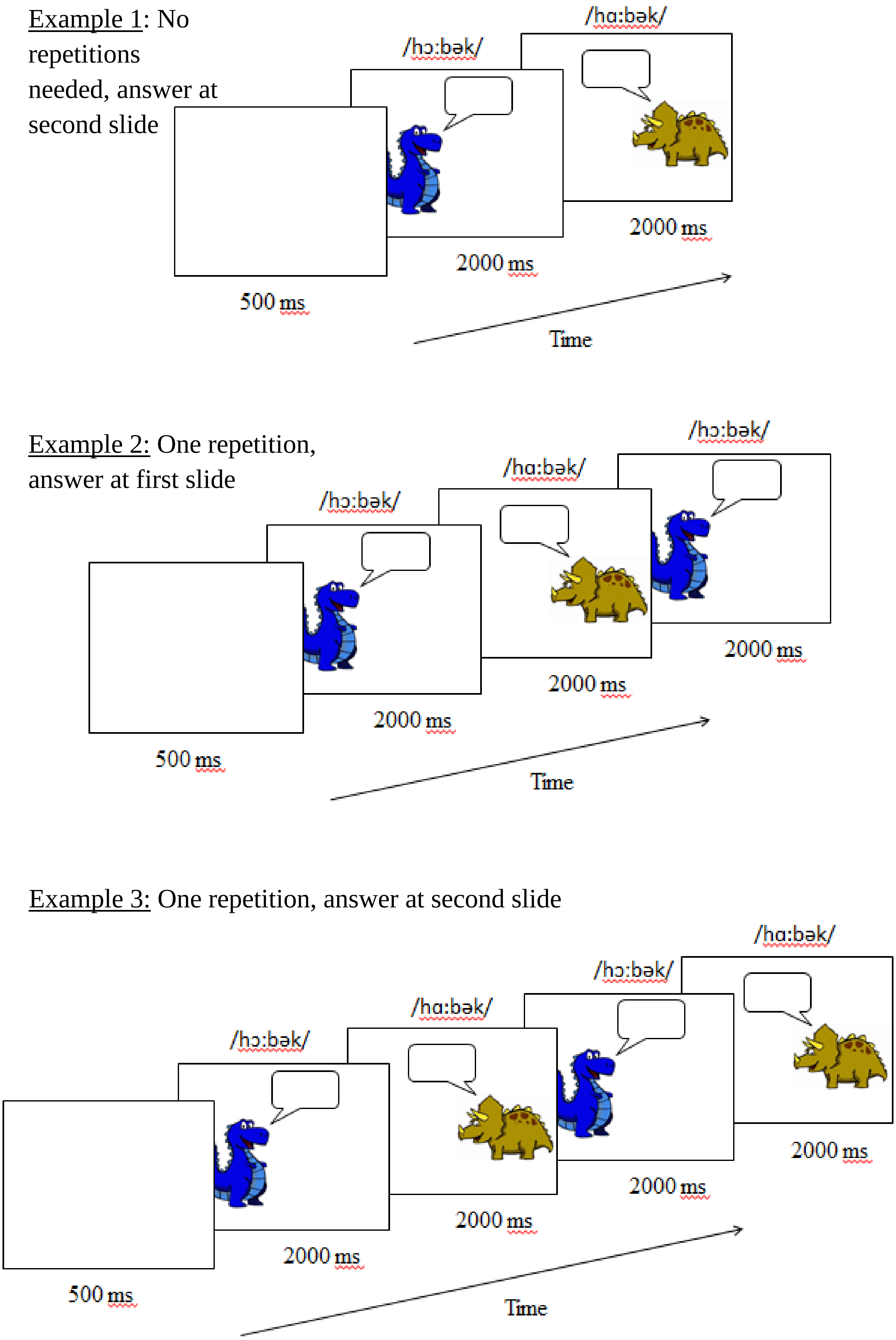


*Figure 2.1.*Three examples of event sequence on a given trial in the phonological task.

2.2.4.2 - Orthographic task. The orthographic task was designed to probe children's orthographic knowledge of the words. The design of this task was very similar to that of the phonological task, but in this case the dinosaurs were paired with two written stimuli, one corresponding to the correct orthographic form of the word, the other corresponding to a distractor. Examples of distractors are horberk for hauberk, palisaid for palisade and mosoleum for mausoleum (see Appendix E). Whenever possible, the distractors for this task were chosen from the incorrect spellings produced by adults during pilot testing. Children were instructed to choose the dinosaur who wrote the word best by pressing the corresponding button on the keyboard.

It was decided not to expose children to the oral form of the words in this task to probe children's pure orthographic knowledge of the words, irrespective of their ability to derive orthography from phonology. Furthermore this testing modality allowed us to counterbalance the order of presentation of this task and the phonological task, which would have not been possible if the correct phonological form of the words were presented during this task (they might have primed correct responses in the phonological task).

The experimental trials were preceded by a set of instructions to introduce the child to the task, lasting approximately 1 minute, followed by four practice trials, where the child had to choose the correct written form of four very common words (HOUSE, SHIRT, DOG and CAR, presented in this fixed order). Each dinosaur corresponded to a slide composed by a picture element and a text element; each slide was presented for 3.5 seconds, to allow time for the children to read the written word. The two slides that formed a trial were repeated in a fixed alternating sequence until an answer was provided, and the number of repetitions needed to give a response was recorded. A blank slide of 500 ms divided each trial from the next. Finally a second set of instructions were provided, to tell the child that the following words would be "more difficult words" that were in the stories previously presented in class. The child was also instructed to guess the right answer if he/she did not remember the words. An example of the event sequence of a given trial is provided in Figure 2.2.

There were 8 experimental trials, one for each word tested. Both order of presentation of the words and the link between dinosaurs and the correct answer were randomised, and we collected accuracy and reaction time for each trial. The task took between 3 and 5 minutes to be completed.

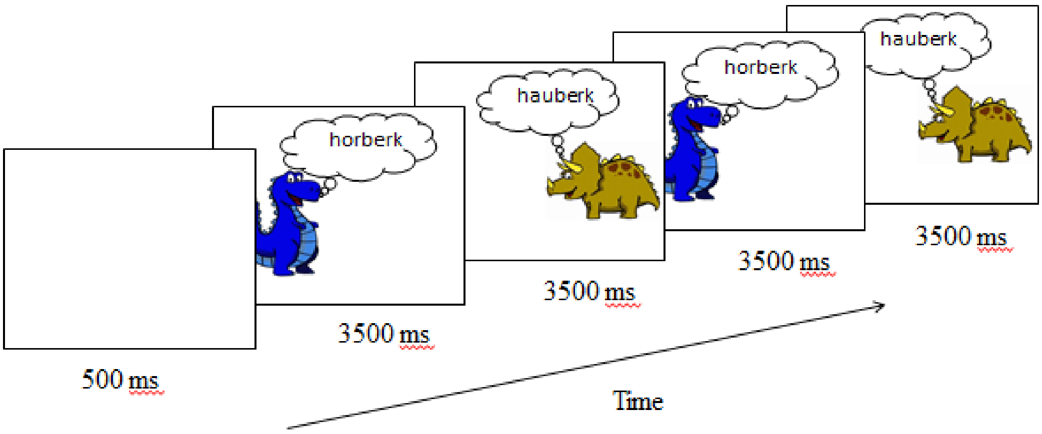


Figure 2.2. Example of event sequence on a given trial in the orthographic task.

2.2.4.3 - Semantic tasks. This section describes the two versions of the semantic task used for this research. The three-step task was used at post-test, while a shorter one-step task was used during pre-test and during the control words post-test session.

Each step of the three-step task required a more detailed knowledge of the words compared to the previous one. In each step, children were instructed to choose the right description of the target word among four alternatives. The words and the definitions were presented in both the auditory and written modality. The first step required only a general knowledge of the word, where the alternatives corresponded to general categories, such as Job, Clothing, Building or Animal. The second step represented a more detailed knowledge of the words, where the four possible sub-categories were all included in the higher category identified in the first step. Finally, the third step involved choosing between four detailed definitions of the word, similar to the ones presented in the stories. The foils for the first step were chosen from the correct categories for different words, and all the categories were presented the same number of times (i.e. they were the correct choice once, and the incorrect choice three times), while the alternatives for the second and the third step were created ad hoc, with the sub-categories of the second step all being part of the main category of step 1 (for example all items of clothing), and the alternative definitions being similar to the correct definition, differing from it by one feature.

At the beginning of the task the child was presented a first set of instructions to introduce him/her to the task and the answering procedure. To choose the right response for each word the child had to press one of four buttons corresponding to the right answer on the screen: four buttons were highlighted in different colours on the keyboard, each corresponding to the position and colour of one of four rectangles on screen (red at top left, green at top right, blue at bottom left and yellow at bottom right). After the first set of instructions that introduced the child to the task, the first practice word was presented (HOUSE), and the categories associated with it appeared one at the time (CLOTHING, JOB, FOOD OR DRINK and BUILDING). Independently of the answer to this first step, further instructions were then presented. These instructions introduced the second step, explaining that, if the child had chosen the correct category, he/she would hear the word again, this time accompanied by another set of categories. In the second step for HOUSE the choice was between CASTLE, CHURCH, SCHOOL and HOME. Independently of the answer to the second step, the child was then introduced to the third step, with further instructions explaining that the same word would be repeated a third time in case of correct answer. In the final step the choice was between A PLACE TO WORK, A PLACE TO SLEEP BUT NOT EAT, A PLACE TO PRAY and A PLACE TO LIVE. This interactive method of giving the instructions, where children were completing each step in turn, and receiving further instructions between steps, was designed to maximise children’s understanding of the task, and minimise errors due to lack of comprehension of the instructions. Following this first practice trial the child was administered the remaining 3 practice trials (SHIRT, DOG and CAR) without further instruction. After the practice trials a second set of instructions were provided, to tell the child that the following words would be “more difficult words” that were in the stories previously presented in class. The child was instructed to guess the right answer if he/she did not remember the meaning of the words.

In experimental trials, each trial was separated from the subsequent one by a slide that asked the child to press a button when ready to continue, thus allowing the participant to take a short break between trials. An example of the timeframe of a complete trial is shown in Figure 2.3. Each step was composed of 5 slides that appeared in a fixed sequence. During the first 2 seconds of each step (the first slide) the child was presented both the written and the oral form of the target word (the word was spoken once). After 2 seconds the first response option appeared in the top left of the screen, presented both written and orally, then the second option appeared, both written and orally, and so on till the appearance of the last response option. By the appearance of the last option the child was able to see both the target word at the top of the screen, and the four alternative response options. A new alternative was presented every 3 seconds for the first and second step, and every 5 seconds for the third step. The child was given the possibility of providing an answer at any point after the presentation of the first alternative in the first step; in this step the alternatives were very different from each other, making it unlikely that they would be mistaken for one other. For the remaining two steps, the participant had to wait for all the alternatives to be presented before answering, to ensure that he/she would choose the best possible alternative, and not simply the one presented first. All the children were administered the first step of each trial, but they were presented the following steps only if they had given the right answer in the previous one. For this task, both order of presentation of the words and the position of the four alternatives were randomised, and we measured children’s accuracy in each step. The task took on average 15 minutes to be completed

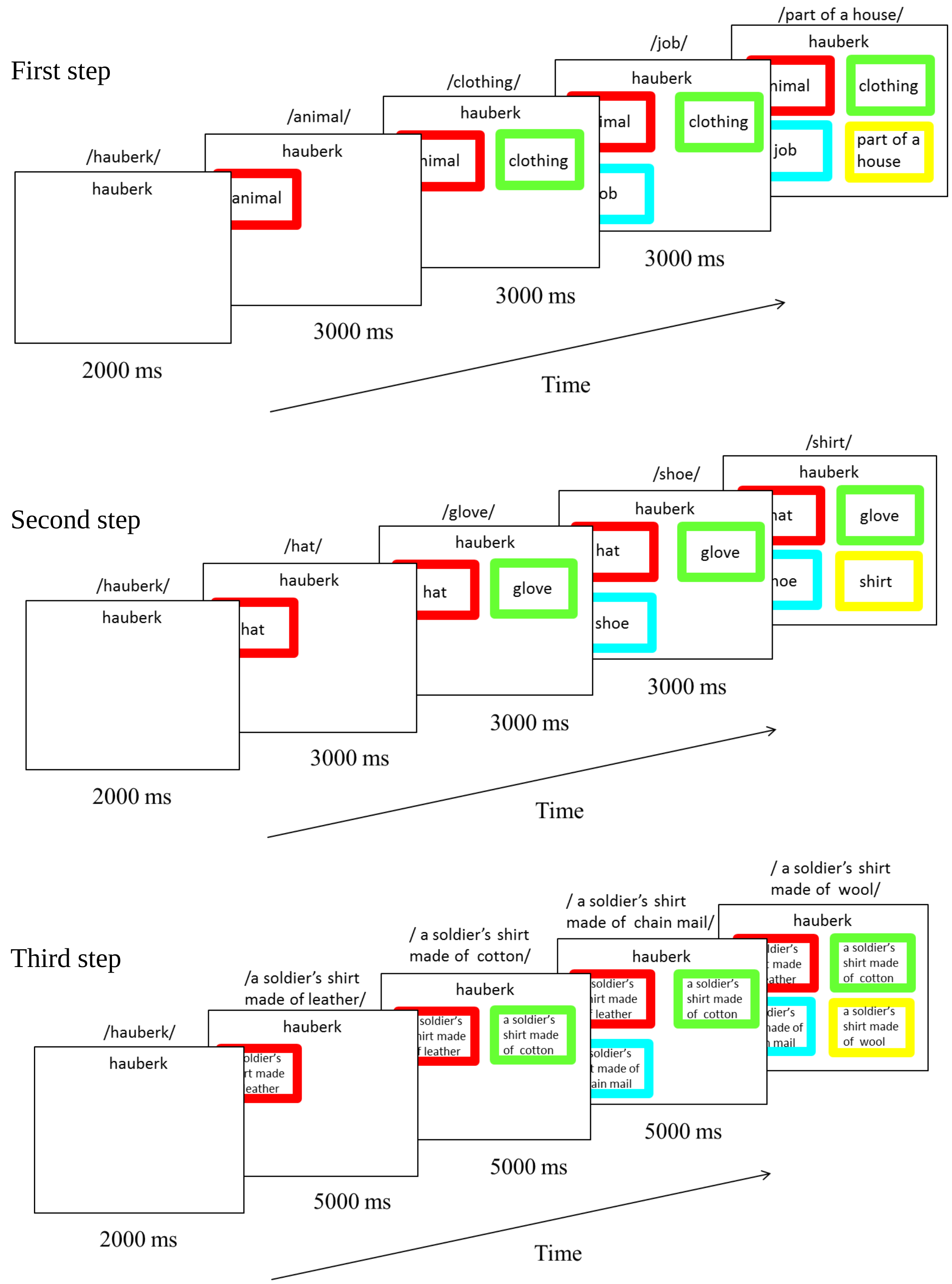


Figure 2.3. Event sequence on a given trial in the semantic task.

The one-step task, used at pre-test and for the control words post-test, corresponded to the first step of the three-step task. Similarly to the three-step task, the child was introduced to the answering procedure through a first set of instructions, which lasted approximately 1 minute, and was followed by four practice trials, where the child had to choose the right category, between four alternatives, for the four practice words (HOUSE, SHIRT, DOG and CAR). Finally a second set of instructions informed the child that the following words would be "more difficult words". The instructions also suggested to guess the answer, in case of uncertainty. The experimental task was composed of 40 trials at pre-test (16 target items, 8 control items and 16 filler words), and 8 during the control words post-test. Each word was tested once during this task. Presentation modality and timing of presentation modality were the same as the first step on the three steps task. Both order of presentation of the words and the position of the four alternatives were randomised and children's accuracy was measured.

2.2.5 - Background Measures

Standardised assessments of reading, oral language and non-verbal abilities were carried out for all the participating children, to explore the effects of individual abilities in these areas on children's ability to learn new words in the different modalities. All were administered according to test manual instructions.

To explore children's cognitive non-verbal abilities the Raven's Coloured Progressive Matrices (CPM; Rust, 2008) were used. The CPM are composed of three sets, Set A, Set AB and Set B, consisting of 12 items each. The items are diagrammatic puzzles that can be solved using analogies and inferences. During the testing session each child was shown an incomplete pattern, and he/she was asked to choose the piece that completed it best, among six alternatives. The raw score was calculated (total number of patterns completed correctly out of 36) and used to derive a standard score (M = 100, SD = 15) for each child. The normative data for the UK used in the 2008 version is based on 608 children from 4 to 11 years of age. The split-half reliability, which is a measure of the consistency of a set of measurements from a given test, is .97.

To explore children's oral language comprehension two tests were administered: the British Picture Vocabulary Scale (BPVS – 3; Dunn, Dunn, & NFER., 2009), for word-level understanding, and the Understanding of Spoken Paragraphs subtest of the Clinical Evaluation of Language Fundamentals (CELF- 4; Semel, Wiig, & Secord, 2006), for sentence and paragraph level understanding.

The BPVS - 3 is an individually administered measure of receptive vocabulary, originally based on the Peabody Picture Vocabulary Test (PPVT; Dunn, 1959). The test is composed of 168 items, divided into 14 sets of 12 items each. During the assessment the child was presented with a set of four pictures from which he/she was asked to choose the one that best represented the stimulus word spoken by the administrator. The raw score (total correct) was calculated and used to derive a standard score (M = 100, SD = 15) for each child.

The Understanding of Spoken Paragraphs subtest of the CELF - 4 examines children's ability to sustain attention while listening, understand oral texts and answer questions about these texts. During the testing session three paragraphs were presented orally to the child, one at a time, each followed by five questions, giving a total of 15 questions. The final raw score was the number of questions answered correctly. This was used to derive a scaled score (M = 10, SD = 3) for each child. The Understanding of Spoken Paragraphs subtest has good test-retest reliability (test-retest reliability: .80).

Two tests to explore word reading abilities were administered: the Test of Word and Reading Efficiency (TOWRE – Second edition; Torgesen, Wagner, & Rashotte, 1999), as a measure of word and nonword reading speed, and the Single Word Reading Test (SWRT 6-16; Foster, 2007), part of the York Assessment of Reading for Comprehension (YARC; Snowling et al., 2009), to assess word reading accuracy independently from speed. In addition, the YARC was used as an index of passage reading abilities.

The TOWRE is a measure of fluency and accuracy of word reading, formed of two subtests. The Sight Word Efficiency subtest (SWE) measures the ability to read real words. To complete this subtest the child was asked to read a list of words as fast as possible, and the number of words correctly identified within 45 seconds formed the raw score for this subtest. The Phonemic Decoding Efficiency subtest (PDE) assessed the ability to decode nonwords. In this subtest children were asked to pronounce a list of nonwords as fast as possible, and the raw score for this subtest was computed as the number of nonwords pronounced correctly in 45 seconds. For both subtests, raw scores were used to derive a standard score (M = 100, SD = 15) for each child. This test has good reliability (correlation between different forms of the test: SWE r = .91, PDE r = .92; test-retest reliability: SWE r = .91, PDE r = .90; inter-scorer reliability: SWE r = .99, PDE r = .99).

The SWRT is a test of word reading accuracy, in which children are asked to read aloud lists of words of increasing difficulty, with no time limit. The final raw score of the SWRT corresponds to the number of words read correctly, up to 60. The raw score (total correct) was calculated and used to derive a standard score (M = 100, SD = 15) for each child. This test was used as further measure of word reading ability, and to select the first passage of the YARC for each child.

The YARC is an individually administered assessment tool to evaluate accuracy, rate and comprehension of reading in primary school children from 5 to 11 years of age. Each child read two passages from Form A, each followed by 8 comprehension questions to tap literal and inferential comprehension skills. For each passage administered, three raw scores are calculated: one for reading accuracy, one for reading rate, and one for reading comprehension. For each of these three measures, an ability score was then calculated, and the average of the ability scores for the two passages in each of the three abilities measures were then transformed into a standardized score (M = 100, SD = 15). Only reading accuracy and reading comprehension were considered in the present research, given that reading rate is influenced by the number of reading errors the child makes. The two measures have good reliability (parallel form reliability for reading accuracy: all rs > .70; Cronbach's alpha for reading comprehension scores from two passages: all αs > .70).

2.2.4 - Procedure

Before the beginning of the project the study was discussed with the headteacher of the school and the teachers of the participating classes, and consent was obtained. Afterwards pupils were informed about the project and they received information leaflets for themselves and information sheets for their parents, with opt-out consent forms attached, to allow parents to opt out of the research. Information about the project was also published in the school weekly newsletter. No one chose to opt out (see Appendices F, G and H for information sheets and consent forms). The procedure devised for this experiment is described in Figure 2.4, and involved six steps.

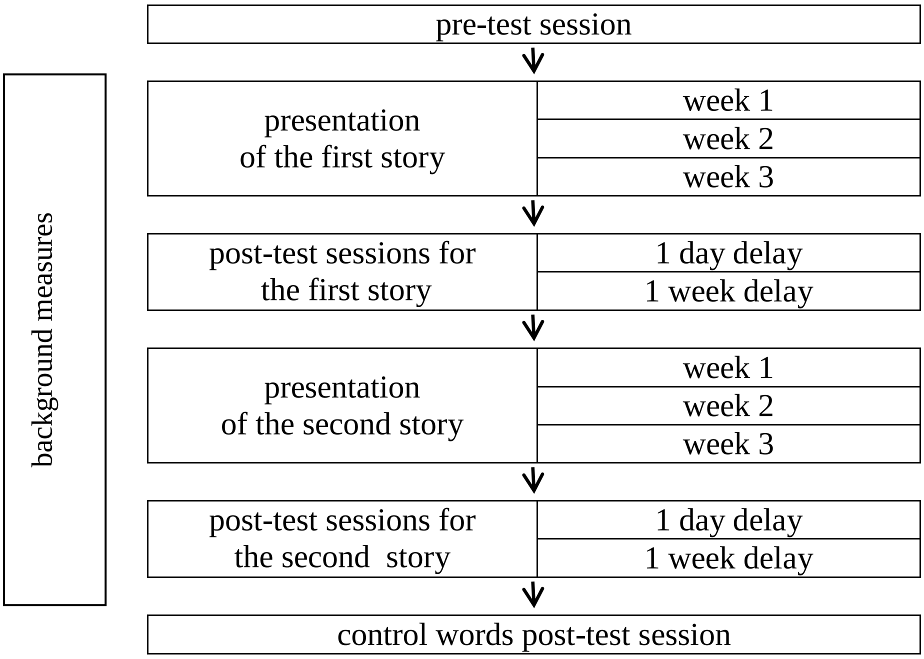


Figure 2.4. Study schedule.

As a pre-test, children were administered the one-step semantic task with 40 items. This task was administered by means of PCs placed in a quiet room in the school. Each child worked independently on the task on a separate PC, and listened to the stimuli via headphones, and 7 to 8 children completed their task independently at the same time. The pre-test session lasted between 15 and 20 minutes.

When all children had completed the pre-test session, the first story was presented. Each story was presented by the teacher in class, once a week for three consecutive weeks. The teachers presented the stories to the class as a class activity; each teacher was provided a set of instructions that described how the story should be presented, and discussed the presentation modality with the researcher before the first reading. The instructions specified how to answer children's questions at the end of the story presentation; teachers were instructed to give details about any word the children might not be sure about, apart from the target words, for which only the given definitions could be provided as an explanation. The teachers were also provided with a guide on how to pronounce the target vocabulary, by means of an audio recording of the correct pronunciation. Before presenting each story to the children the teachers were also asked to briefly present the activity to the class, and point out how this was part of the project they were taking part in. Each story presentation occurred during the same day of the week for three consecutive weeks, approximately at the same time of the day, usually in the afternoon, and the presentation session took about 20 minutes. Teachers recorded pupils' absence and the questions asked by the children on sheets provided.

For the story presented in the listening condition children were asked to sit on the carpet and to listen attentively to the teacher; they had no access to the written form of the story, since a single copy was provided for the teacher's use. For the story presented in the combined condition children were asked to sit at their desks, and they were provided with a copy of the story, in form of a 6-page booklet written in Calibri 14-point font. Children were then asked to follow the reading of the story by the teacher using their booklets. The booklets were collected at the end of each story presentation and stored in a cupboard, so that children could not access them outside of the story presentation sessions. In neither condition were children shown any pictures during the story readings.

The day after the final presentation of each story, children took part in the first post-test session. This included three tasks: the phonological test, the orthographic task, and the three-step semantic task. In all the post-test tasks, children were tested only on the 8 target words linked to the story they had been presented, to avoid exposing the children to the other words (target words for the other condition and control words) repeatedly. During the post-test sessions half of the children completed the phonological task first, and half the orthographic task fist. The semantic task was always administered last. While the orthographic and phonological tasks did not provide information about semantics, the semantic task, where target words were presented both in the oral and written modalities, could have affected children's performance on the other two tasks, if presented first. The order of presentation of the phonological and orthographic tasks was counterbalanced to avoid any confounding order effects. The session lasted around 30 minutes in total, and 7 to 8 children completed the tasks independently at the same time. A week after the first post-test session, on the same day of the week, children took part in the second post-test session, completing the tasks in the same order as in the first session. This session was carried out to assess children's retention of the words.

The second story presentation started 1 week after the last testing session for the first story for Class A, and 2 weeks after the last testing session of the first story presentation for Class B. The second story was presented in the modality children had not yet experienced. The post-test sessions for the second story were carried out following the same method as the post-test sessions for the first story.

The control words post-test session took place when all the children had completed all the story presentation sessions and all the other testing sessions. This session was carried out exactly as the story post-test sessions: the children completed three tasks, a phonological task, an orthographic task, and a semantic task (one-step task). Half of the children completed the orthographic task first, and half completed the phonological task first. These tasks assessed children’s knowledge of the control words. The one-step semantic task assessed children’s semantic knowledge of both control words (8 words) and filler easy words (16 words). This session lasted between 20 and 30 minutes.

Oral language abilities, reading skills and non-verbal abilities of the children were assessed in one session that lasted around one hour, administered individually with each child. The timing of the collection of background measures varied between children, but occurred no earlier than six weeks before the first story presentation, and no later than six weeks after the final post-test with control words (i.e. between April and July 2014).

2.3 - Results

2.3.1 - Sample Characteristics

Children's performance on the background measures is described in Table 2.4. Children's performance in all the measures transformed into Standard Scores follows a normal distribution (all *p*s > .05), except for the Scaled Score of the Understanding of Spoken Paragraphs subtest of the CELF (USP CELF) (*D*(54) = .129, p = .025) and the Standard Score of the CPM (*D*(54) = .123, p = .041). The distribution of the raw scores for all the measures also follows a normal distribution (all *p*s > .05), except for the Sight Word Efficiency subtest of the TOWRE (TOWRE SWE) (*D*(54) = .135, p = .016), the Understanding of Spoken Paragraphs subtest of the CELF (USP CELF) (*D*(54) = .134, p = .016) and the CPM (*D*(54) = .123, p = .041). Wherever normality assumptions were not met, the appropriate non-parametric analyses were conducted. One-sample t-tests and one-sample Wilcoxon signed-rank tests were carried out on all the standardised measures to compare our sample distribution to the distribution of the population. The sample appears significantly better than average in listening (USP CELF: *Median* = 11, *W* = 786.00, *p* = .040) and reading comprehension (YARC comprehension: *t*(54) = 2.09, *p* = .041), and non-verbal abilities (CPM: *Median* = 102.5, *W* = 608.50, *p* = .020), but worse in vocabulary (BPVS: *t*(54) = -4.40, *p* < .001). Nevertheless their mean score was still within 1 SD of the normal distribution for all the measures. Differences between the two classes were also examined by means of two-sample t-tests and Mann-Whitney tests. Children in the two classes did not differ on the background measures (all *p*s > .05).

Table 2.4

Children's performance on background measures

	Class A N=28				Class B N=26				Difference between classes		Overall N=54	
	Min	**Max**	**Mean**	***SD***	**Min**	**Max**	**Mean**	***SD***	***t***	***p***	**Mean**	***SD***
TOWRE SWE												
raw score[1]	35	82	64.93	9.45	41	80	66.92	8.45	308.00	.332	65.89	8.95
standard score	71	121	98.04	11.47	70	117	99.92	10.26	-.64	.528	98.94	10.85
TOWRE PDE												
raw score	11	49	31.46	9.99	14	55	34.50	9.77	-1.13	.265	32.93	9.91
standard score	75	119	97.46	12.57	75	127	101.12	12.41	-1.07	.288	99.22	12.51
SWRT												
raw score	18	54	42.68	9.18	22	55	43.54	7.24	-.38	.705	43.09	8.24
standard score	71	124	101.14	13.87	77	127	103.46	11.99	-.90	.929	103.30	12.88
YARC accuracy												
ability score	39	68	53.50	8.69	42	71	54.85	6.90	-.63	.533	54.15	7.84
standard score	82	122	100.21	11.41	86	126	101.96	9.30	-.61	.542	101.06	10.39
YARC comprehension												
ability score	48	70	58.61	6.82	43	70	58.27	7.17	.18	.860	58.44	6.93
standard score	91	119	102.61	8.74	85	118	102.38	8.96	.09	.927	102.50	8.77
BPVS												
raw score	96	152	122.29	14.04	88	140	120.15	14.94	.54	.591	121.26	14.38
standard score	72	121	94.04	12.62	69	107	92.04	11.03	.63	.531	93.07	11.57
USP CELF												
raw score[1]	8	15	11.57	1.99	4	15	11.15	3.08	360.50	.951	11.37	2.56
scaled score[1]	7	14	10.68	1.87	4	14	10.35	2.80	464.00	> .999	10.52	2.34
CPM												
raw score[1]	25	35	30.29	2.83	16	35	29.23	4.19	325.50	.503	29.78	3.55
standard score[1]	85	125	105.71	11.20	65	125	101.73	13.99	307.50	.323	103.80	12.66

Note. TOWRE SWE = Sight Word Efficiency subtest of the TOWRE (Test of Word and Reading Efficiency); TOWRE PDE = Phonemic Decoding Efficiency subtest of the TOWRE; SWRT = Single Word Reading Test included in the YARC protocol; YARC accuracy = reading accuracy of passages collected as part of the YARC (York Assessment of Reading for Comprehension); YARC comprehension = score associated with the reading comprehension of passages collected as part of the YARC; BPVS = score for the British Picture Vocabulary Scale; USP CELF = score for the Understanding of Spoken Paragraphs subtest of the Clinical Evaluation of Language Fundamentals; CPM = score obtained in the Raven's Coloured Progressive Matrices.
[1]Mann-Whitney U has been reported in place of the t-test, given that this variable is not normally distributed.

Spearman's rank correlation coefficients between the background measures are reported in Table 2.5. A Bonferroni correction was applied (corrected critical p-value = .007)[2]. Measures of reading accuracy and speed for words (TOWRE SWE and SWRT), nonwords (TOWRE PDE) and passages (YARC accuracy) had high correlations with each other. Vocabulary correlated highly with reading comprehension, and moderately with listening comprehension. Our measure of non-verbal cognitive abilities (CPM) correlated moderately with the reading comprehension and oral vocabulary skills.

Table 2.5

Spearman's rank correlation coefficients between the background measures

	TOWRE SWE[a]	TOWRE PDE[a]	SWRT[a]	YARC accuracy[b]	YARC comprehension[b]	BPVS[a]	USP CELF[c]	CPM[a]
TOWRE SWE[a]	-							
TOWRE PDE[a]	.71*	-						
SWRT[a]	.61*	.74*	-					
YARC accuracy[b]	.64*	.72*	.86*	-				
YARC comp[b]	.53*	.23	.54*	.55*	-			
BPVS[a]	.42*	.11	.40*	.36*	.64*	-		
USP CELF[c]	.19	.14	.09	.13	.28	.36*	-	
CPM[a]	.38*	.15	.27	.32	.40*	.41*	.19	-

Note. TOWRE SWE = Sight Word Efficiency subtest of the TOWRE (Test of Word and Reading Efficiency); TOWRE PDE = Phonemic Decoding Efficiency subtest of the TOWRE; SWRT = Single Word Reading Test included in the YARC protocol; YARC accuracy = reading accuracy of passages collected as part of the YARC (York Assessment of Reading for Comprehension); YARC comprehension = score associated with the reading comprehension of passages collected as part of the YARC; BPVS = score for the British Picture Vocabulary Scale; USP CELF = Understanding of Spoken Paragraphs subtest of the Clinical Evaluation of Language Fundamentals; CPM = score obtained in the Raven's Coloured Progressive Matrices.
[a]raw score. [b]ability score. [c]scaled score.
*correlation is significant (p < .006).

[2] Raw scores have been used for all the measures, apart from the YARC measures, for which the ability scores are reported, and the USP CELF, for which the scaled scores are used. The choice to use the scaled scores of the USP CELF, and not the simple raw scores was led by the fact that two different versions of the USP CELF were used depending on the age of the child (see: Method), therefore the use of a more standardised measure that would take into account the use of different versions was deemed necessary.

2.3.2 - Children's performance at pre-test

Details of children's performance at pre-test are provided in Table 2.6. For each child the score in this task was computed as the total number of words for which the category was correctly identified. The maximum scores for each child were 24 for target and control words and 16 for easy filler words. Children's performance for target words and easy filler words were compared to chance. Wherever normality assumptions were not met, the appropriate non-parametric analyses were conducted. Children in our sample showed a performance not different from chance (computed as 6 out of 24: for each item there was a 1:4 probability of choosing the correct answer by chance) for control and target words ($t(53) = 1.83$, $p = .074$). On the other hand, they showed a performance which was significantly better than chance (computed as 4 out of 16) for easy filler words ($W = 1485.00$, $p < .001$). The same results were obtained when children were grouped by class (control and target words: $t(27)_{\text{class A}} = 1.01$, $p = .323$, $t(25)_{\text{class B}} = 1.55$, $p = .135$; easy filler words: $W_{\text{class A}} = 406.00$, $p < .001$; $W_{\text{class B}} = 351.00$, $p < .001$). Children, therefore, did not know the chosen target and control words at pre-test, and their low performance was not dependent on the nature of the task, since they performed well in filler words. The performance of the two classes was also compared using an independent samples t-test and a Mann-Whitney test, and no significant differences between the two was highlighted ($t(52)_{\textit{target and control words}} = -.48$, $p = .636$; $U_{\textit{filler words}} = 341.50$, $p = .684$). Children in the two classes did not differ on their initial knowledge of our stimuli.

Table 2.6

Number of words correctly recognised in the one-step semantic task at pre-test

	Class A N=28				**Class B N=26**				**Overall N=54**			
	Mdn	**M**	**SD**	**%**	**Mdn**	**M**	**SD**	**%**	**Mdn**	**M**	**SD**	**%**
target and control words (out of 24)	6	6.43	2.25	26.79	7	6.73	2.41	28.04	6	6.57	2.31	27.38
easy filler (out of 16)	15	14.29	1.30	88.31	15	13.69	2.45	85.56	15	14.00	1.94	87.50

Figure 2.5 shows the proportion of target and control words known by the children separately. Children knew similar proportions of target and control words ($T =$

441.50; p = 531). Target word performance by condition is reported in Table 2.10 in Section 2.3.5, for a clearer comparison with post-test performance.

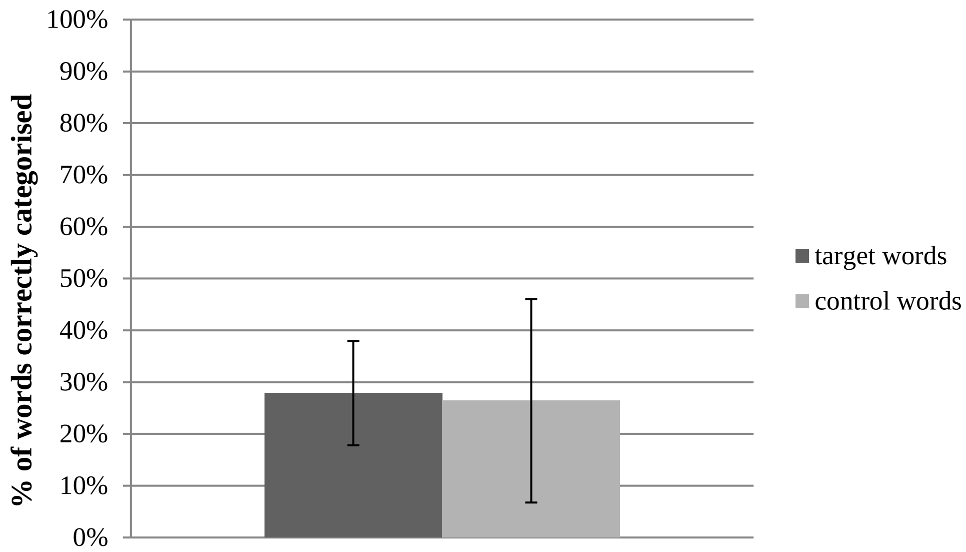


Figure 2.5. Number of control and target words correctly categorised at pre-test.

2.3.3 - Phonological task

In the phonological task children were asked to choose the right phonological form of the words presented in the story between two alternatives. Children were assigned a score of 1 each time they recognised the correct form of the word, and 0 each time they chose the incorrect one; the total number of words correctly recognised was then computed, and formed the score for the task. A score of four out of eight was expected by chance (50% probability of choosing the correct form by chance). Each child provided two sets of scores for the listening condition and the combined condition (post-test 1 and post-test 2), and one score for the control words set.

The mean number of correct words recognised in each condition is described in Table 2.7. Children correctly identified the phonological form for approximately seven words at both post-test 1 and post-test 2 in both conditions, and they identified the right form for five words for the control words set.

Table 2.7

Number of words correctly recognised in the phonological task

		Mean	*SE*	*SD*	Chance comparison	
					W	***p***
listening condition	post-test 1	6.85	.20	1.45	1818.00	< .001
	post-test 2	7.11	.18	1.36	1646.00	< .001
	learning at post-test 1	1.74	.24	1.79		
	learning at post-test 2	2.00	.22	1.63		
combined condition	post-test 1	7.09	.15	1.07	1770.00	< .001
	post-test 2	7.06	.17	1.28	1590.00	< .001
	learning at post-test 1	1.98	.25	1.86		
	learning at post-test 2	1.94	.24	1.78		
control words set	post-test	5.11	.18	1.36	924.00	< .001

Note. Chance comparison was computed by comparing the scores against the expected median of four (50% probability of choosing the correct alternative by chance for the 8 items).

First of all, to test whether children's performance was at chance, scores were compared to a chance performance using one-sample Wilcoxon signed-rank tests (scores were not normally distributed: all *p*s < .05). Chance level was set at four out of eight (50% chance level). As shown in Table 2.7 performance in both conditions and for the control words set were higher than chance. To test whether children performed better with target than control words, performance in the listening and the combined condition was compared to that in the control words set, by means of Wilcoxon signed-rank tests, and found significantly higher than it, at both times of testing (listening condition: $T_{post\text{-}test\ 1} = 105.00$, $p < .001$; $T_{post\text{-}test\ 2} = 28.00$, $p < .001$; combined condition: $T_{post\text{-}test\ 1} = 64.50$, $p < .001$; $T_{post\text{-}test\ 2} = 78.00$, $p < .001$). These analyses suggest that children performed better in this task for words presented in the stories than words they were never presented, thus suggesting that story presentation had a positive effect on a task that assesses phonological knowledge.[3]

A 2 x 2 repeated measures analysis of variance (ANOVA) was then conducted with learning scores as the dependent variable, condition (listening condition *vs.*

[3] This conclusion is speculative, since there were different words in the different sets. Nevertheless the results from the pilot studies (see Method) suggest that there is no reason to expect a different performance in the three sets in the phonological or orthographic tasks.

combined condition) and time (post-test 1 *vs.* post-test 2) as predictors, to test whether the two conditions yielded different amount of phonological learning, and whether learning was retained over time. Learning scores were used, instead of the raw number of words correctly recognised in each condition, because performance in the control words set was higher than chance (see Table 2.7). The control words set was therefore used as a baseline measure of performance in the phonological task for unknown words, and learning scores were computed as phonological score – baseline.[4] The analysis yielded no significant main effect of either condition ($F(1, 53) = .27$, $p = .608$, $\eta^2_p < .01$) or time ($F(1, 53) = 1.08$, $p = .305$, $\eta^2_p = .02$), and no significant interaction between these measures: $F(1, 53) = 1.36$, $p = .249$, $\eta^2_p = .03$. The performance of the children therefore did not differ between the two conditions and between the two post-test sessions, and the effect of time was the same in the two conditions. Children learned the same amount of phonological information about words in the listening and the combined condition and retained their knowledge from post-test 1 to post-test 2. Since our data do not meet the assumption of normal distribution for parametric analysis, non-parametric Wilcoxon signed-rank tests were used to confirm our results. No differences between conditions or time of testing were found (all *p*s > .05).

As a further check that the combined condition does not yield any advantage in phonological knowledge over the listening condition, an orthographic facilitation measure was computed as mean phonological learning in the combined condition minus mean phonological learning in the listening condition, and was compared to 0, by means of one-sample Wilcoxon signed-rank tests. The test was not significant ($Mdn = 0$, $W = 513.00$, $p = .626$).

These analyses therefore show that children learnt the phonological forms of the words equally well when these were presented in the oral and dual presentation modality, performing much better than chance and better than in the set of control words in both conditions.

[4] Prior phonological and orthographic knowledge of the words was not directly tested, to avoid pointing the attention of the children to the presence of new words, and their properties. For this reason it was necessary to use their performance in the phonological and orthographic tasks for control words as baseline, with the assumption that, if children had not learned the form of the words, they would perform similarly for words they have encountered in the stories, and words not encountered in the stories.

2.3.4 - Orthographic task

In the orthographic task children were asked to choose the right orthographic form of the words presented in the story between two alternatives. Scores for the orthographic task were computed similarly to those of the phonological task (i.e. depending on whether children correctly recognised the correct alternative or not). A score of four out of eight was expected by chance (50% probability of choosing the correct form by chance). Scores in the task are described in Table 2.8. Children correctly identified the orthographic form for around six words at both post-test sessions in the combined condition, while they identified the orthographic form of fewer than five words at post-test 1 and of around 5 words at post-test 2 for the listening condition. They identified the right form of five control words.

Table 2.8

Number of words correctly recognised in the orthographic task

		Mean	***SE***	***SD***	**Chance comparison**	
					W	***p***
listening condition	post-test 1	4.61	.22	1.64	653.50	.010
	post-test 2	5.37	.19	1.42	985.00	< .001
	learning at post-test 1	-.35	.29	2.16		
	learning at post-test 2	.41	.26	1.92		
combined condition	post-test 1	6.06	.22	1.62	1363.00	< .001
	post-test 2	6.04	.25	1.81	914.50	< .001
	learning at post-test 1	1.09	.31	2.26		
	learning at post-test 2	1.07	.30	2.24		
control words set	post-test	4.96	.21	1.52	1013.00	< .001

Note. Chance comparison was computed by comparing the scores against the expected median of 4 (50% probability of choosing the correct alternative by chance for the 8 items).

To test whether children's performance was at chance in the orthographic task, scores were compared to a chance performance using one-sample Wilcoxon signed-rank tests. Chance level was set at four (50% chance level). As shown in Table 2.8 performance in both conditions and for the control words set were higher than chance. Performance in the listening and the combined condition was then compared to that in the control words set: performance in the listening condition was not significantly different from the performance in the control words set at either time of testing (post-test 1: T = 369.50, p = .300; post-test 2: T = 339.50, p = .101), while performance in the combined condition was significantly higher than that in the control words set at both post-tests (post-test 1: $T = 287.50$, $p = .001$; post-test 2: $T = 206.00$, $p = .001$). Children performed better in this task for words presented in the combined condition than words they were never presented, while the same effect was not found when the stories were presented only orally.

A 2 x 2 repeated measures analysis of variance (ANOVA) was then conducted with learning scores as the dependent variable, condition (listening condition *vs.* combined condition) and time (post-test 1 vs. post-test 2) as predictors, to test whether the two conditions yielded significantly different amount of orthographic learning, and whether learning was retained over time. Learning scores were used, since performance in the control words set was higher than chance. As for phonological scores, control words' performance was used as a baseline measure, and learning scores were computed as orthographic score – baseline. The analysis yielded significant main effects of both condition ($F(1, 53)= 27.65$, $p < .001$, $\eta^2_p= .34$) and time ($F(1, 53)= 7.41$, $p= .017$, $\eta^2_p= .10$), and a significant interaction between these measures: $F(1, 53)= 8.17$, $p = .024$, $\eta^2_p= .09$. The performance of the children thus differed between the two conditions and between the two post-test sessions, and time had a different effect in the two conditions. To explore the results further, we compared learning scores at post-test 1 and at post-test 2 for the two conditions by means of Wilcoxon signed-rank tests, given that the data did not meet normality assumptions. Children performed equally well at post-test 1 and post-test 2 in the combined condition ($Mdn_{\text{post-test 1}} = 1.5$; $Mdn_{\text{post-test 2}}= 1$, $T = 255.50$, $p = .872$), but they performed significantly better at post-test 2 than at post-test 1 in the listening condition ($Mdn_{\text{post-test 1}} = 0$; $Mdn_{\text{post-test 2}}= 0$, $T = 285.50$, $p = .004$). We can explain this finding by considering that children in the listening condition had not been exposed to the written form of the words during the story presentation sessions, but they were presented with the orthographic forms of these words during the first testing session, by means of the orthographic task itself and the semantic task. Therefore, children may have acquired some information about the orthographic form of the words during the first testing session, and used it during the second testing session. Learning that occurred during the testing session itself did not improve children's performance so much as to make their performance significantly better than their performance with control words, but it was big enough to make their performance in the second session significantly better than their performance in the first session. As expected scores in the two conditions were different at both post-test sessions ($T_{\text{post-test 1}} = 130.00$, $p < .001$; $T_{\text{post-test 2}} = 347.50$, $p = .020$).

To further confirm that the combined condition yielded an advantage in orthographic knowledge over the listening condition, an orthographic facilitation measure was computed as mean orthographic learning in the combined condition minus mean orthographic learning in the listening condition, and was compared to zero, by means of one-sample Wilcoxon signed-rank tests. The test was significant (*Mdn* = 1, *W* = 964.00, $p < .001$).

These analyses therefore show that children learned the written form of the words only when presented with both the written and the oral form of words in the stories (combined condition), and that the oral presentation in the listening condition was not sufficient to elicit orthographic learning.

2.3.5 – Three-step semantic task

The semantic task used at post-test was formed of three steps. In each step children were required to choose the right category or definition for the target word presented, between four alternatives. Each step represented a choice between more similar alternatives than the previous one, and children were presented the subsequent step only if they had chosen the right answer in the previous one. Data from this task were analysed in different ways, first considering all the steps, and then each step separately. To compute the total score to consider all steps, each child was assigned a categorical score from zero to three for each item, depending on the step the child reached for each specific word. A score of zero was assigned if the wrong alternative was chosen in the first step, a score of one was assigned when the right answer was chosen for the first step, but not in the subsequent one, a score of two was assigned when the right answer had been chosen in both the first and the second step, but not in the last step, and a score of three meant that the child chose the right answer in all three steps. Each child received two scores for each word encountered in the listening and combined conditions, one for the 1 day post-test and one for the 1 week delay post-test.

Figure 2.6 describes children's scores computed in this way; the total number of trials was 864 for both conditions (each child provided a score for all eight words in each condition at both time points). As it is possible to see from the figure, for the majority of trials children either obtained a score of zero or three (71% of trials in the listening condition and 78% in the combined condition), while fewer than 20% of the trials are associated with a score of one and fewer than 15% are associated with a score of two. It can also be seen that a higher percentage of trials received a score of three in the combined condition compared to the listening condition. Since we used a repeated measure design, which produces data not suitable for chi-square analyses, the data will be analysed using *by subject* analyses in the following sections.

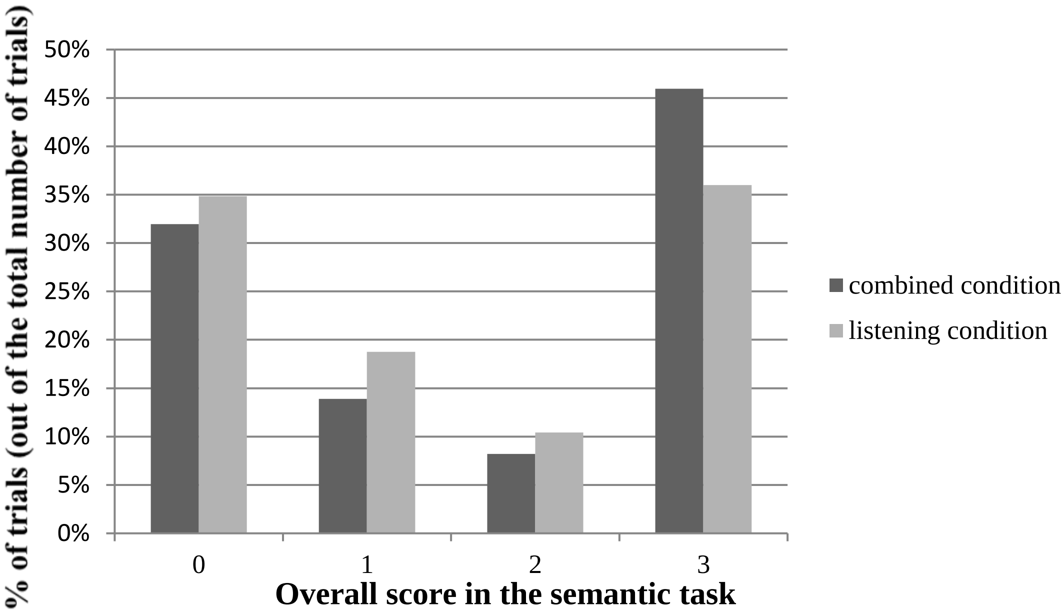


Figure 2.6. Percentages of trials where children obtained a score of 0, 1, 2 or 3 in the semantic task (scores correspond to the Step reached in three steps task). The trials are divided by condition.

Another way to compute scores for the semantic task was to consider only whether children successfully completed all the steps of the semantic task, or failed to do so. For this measure each child performed 16 trials per condition, one for each post-test session. This binomial measure was used to compare the two conditions using a generalised linear mixed-effect model, where both participant and item were considered as random effects. The model is reported in Section 2.3.7.3.

For each child, the Mean Step reached in the semantic task was also computed. For each item a child could obtain a score from zero to three, and these scores were averaged for each child to compute a new score, called Mean Step. The average Mean Step obtained in the two conditions in the two post-test sessions is presented in Table 2.9. The Mean Step scores for the combined condition are higher than the Mean Step for the listening condition.

An orthographic facilitation score was also computed for use in subsequent analyses as the difference between Mean Steps reached in the combined and listening conditions; orthographic facilitation scores for the two post-test sessions were averaged to obtain one orthographic facilitation score for semantic learning.

Table 2.9

Mean Step reached in the three-step semantic task for each condition

		Mean	***SE***	***SD***
listening condition	post-test 1	1.49	.11	.83
	post-test 2	1.46	.10	.77
combined condition	post-test 1	1.68	.11	.78
	post-test 2	1.69	.11	.78
orthographic facilitation		.21	.08	.57

A 2 x 2 repeated measure analysis of variance (ANOVA) was conducted, with Mean Step as the dependent variable, and condition (listening vs. combined) and time (post-test 1 vs. post-test 2) as predictors, to test whether the Mean Step differed depending on time or condition. This analysis yielded a significant main effect of condition ($F(1, 53) = 6.97$, $p = .011$, $\eta^2_p = .12$), but no significant main effect of time ($F(1, 53) = .07$, $p = .790$, $\eta^2_p < .01$) and no interaction between the two ($F(1, 53) = .12$, $p = .734$, $\eta^2_p < .01$). From this analysis it can therefore be concluded that children performed significantly better in the combined condition than in the listening condition, reaching higher steps in the semantic task more often when words were presented in the oral and written modalities simultaneously.

Another way of looking at performance in the semantic task is to compute, for each child, the total number of words for which the correct alternative was chosen at each step. Table 2.10 presents the mean number of words recognised by the children for each step of the semantic task. On average children chose the correct alternative five words out of eight in the first step, for three or four words in the second step, and for two or three words in the third step. The difference between the two conditions increased from the first to the second and third step. Children also correctly recognised around two words at pre-test for all sets, including control words, and they similarly recognised around two words of the control words set at post-test: children did not perform significantly better in the control words set at post-test than at pre-test.

Table 2.10

Children's performance in the three steps of the semantic task –number of words correctly categorised in each step

		Step 1 – Category Recognition		Step 2 – Sub-category Recognition		Step 3 – Definition Recognition	
		Mean	**SD**	**Mean**	**SD**	**Mean**	***SD***
listening condition	pre-test	2.26	1.23	--	--	--	--
	post-test 1	5.33	2.41	3.69	2.39	2.93	2.19
	post-test 2	5.13	2.15	3.74	2.22	2.83	2.06
combined condition	pre-test	2.20	1.35	--	--	--	--
	post-test 1	5.46	2.03	4.31	2.24	3.67	2.16
	post-test 2	5.44	1.88	4.35	2.28	3.69	2.35
control words set	pre-test	2.11	1.57	--	--	--	--
	post-test	2.09	1.01	--	--	--	--

To avoid assuming an increased difficulty in each step and take into account the fact that children were not presented all the words at Step 2 and Step 3, we computed the proportion of words for which the correct alternative was chosen, for each step separately. Proportions were computed as the total number of correct alternatives identified divided by the total number of words presented in each step, which corresponded to eight in Step 1, but varied in Step 2 and 3, depending on performance in the previous steps. For example if a child correctly categorised four words at Step 1, he obtained a proportion of .50 (four words out of eight). If the same child again chose the correct sub-category for all the four words in Step 2 he obtained a proportion of 1 (four words correctly categorised out of the four presented at Step 2), and by choosing the correct definition for 3 of these words, he would receive a proportion of .75 at Step 3 (three words out of four). Children's performance in the semantic task when considered in terms of proportion correct is described in Table 2.11. The table also reports a learning measure for Step 1, computed as the change in the performance in category recognition from baseline to the two post-intervention tests. A learning measure for the control words set was also computed as post-test score – baseline, with baseline being the score obtained at pre-test. On average, children correctly identified around 25 % of the words at pre-test and 25 % of the words for the control set at post-test, but identified more than 60 % of the words presented in all steps at both post-test sessions in both the combined and the listening conditions.

Table 2.11

Children's performance in the three steps of the semantic task - proportions of words correctly categorised in each step

		Step 1 – Category Recognition		Step 2 – Sub-category Recognition		Step 3 – Definition Recogniti on	
		M	**SD**	**M**	**SD**	**M**	**SD**
listening condition	pre-test	.28	.15	--	--	--	--
	post-test 1	.67	.30	.62	.30	.73	.31
	post-test 2	.64	.27	.69	.27	.71	.32
	learning post-test 1	.38	.33	--	--	--	--
	learning post-test 2	.36	.29	--	--	--	--
combined condition	pre-test	.28	.17	--	--	--	--
	post-test 1	.68	.25	.72	.29	.82	.23
	post-test 2	.68	.24	.76	.27	.79	.29
	learning post-test 1	.41	.29	--	--	--	--
	learning post-test 2	.41	.29	--	--	--	--
control words set	pre-test	.26	.20	--	--	--	--
	post-test	.26	.13	--	--	--	--
	learning	.00	.22	--	--	--	--

Children's performance in the three steps was compared to a chance performance by means of one-sample Wilcoxon signed-rank tests (scores were not normally distributed: all *p*s < .05). Chance level was set at 25 %, since children had to choose one correct alternative from four. Performance was compared to chance only for post-test proportions, since chance performance at pre-test had been previously established (see Section 2.3.2). All the measures for target words are significantly higher than chance for both conditions (all *p*s < .001). On the other hand, the performance on control words at post-test was not better than chance (W = 285.00, p = .444). These results led us to conclude that children learned a significant proportion of the words in each step of the semantic task. Conversely, their knowledge of the control words at the end of the study was at chance.

For each Step, a 2 x 2 repeated measures analysis of variance (ANOVA) was then conducted with proportion of words correctly recognised as the dependent variable, condition (listening *vs.* combined) and time (post-test 1 *vs.* post-test 2) as predictors. The analysis for Step 1 yielded no significant main effect of time ($F(1, 53) = .85$, p = .361, $\eta^2_p = .02$) or condition ($F(1, 53) = .51$, $p = .478$, $\eta^2_p = .01$) and no significant interaction between these two ($F(1, 53) = .41$, $p = .527$, $\eta^2_p = .01$). On the other hand the analysis for Step 2 yielded no significant main effect of time ($F(1, 53) = 3.64$, $p = .062$, $\eta^2_p = .07$), but a significant main effect of condition ($F(1, 53) = 9.61$, $p = .003$, $\eta^2_p = .16$) and no significant interaction between these two measures ($F(1, 53) = .75$, $p = .389$, $\eta^2_p = .02$). Similarly the analysis for Step 3 highlighted no effect of time ($F(1, 43) < 0.01$, $p = .965$, $\eta^2_p < .01$), a significant main effect of condition ($F(1, 43) = 5.90$, $p = .020$, $\eta^2_p = .13$), but no interaction ($F(1, 43) = .56$, $p = .460$, $\eta^2_p = .01$). Similar results were obtained using mixed effect models (section 2.3.7.3), leading to the conclusion that children performed equally well in category recognition in the two conditions, but they were able to recognise the right description of the words in Step 2 and Step 3 (sub-category and definition respectively) more easily in the combined than in the listening condition. On the other hand children's performance did not differ between the two post-tests session, thus confirming that knowledge was stable for at least a week.

Since our data did not meet the assumption for parametric analysis, non-parametric Wilcoxon signed-rank tests were used to confirm differences between conditions. The performance in the two conditions did not differ from each other for Step 1 either at post-test 1 ($Mdn_{\text{listening condition}} = .75$; $Mdn_{\text{combined condition}} = .75$; $T = 441.00$, $p = .894$) or at post-test 2 ($Mdn_{\text{listening condition}} = .63$; $Mdn_{\text{combined condition}} = .75$; $T = 319.50$, $p = .143$). The difference between condition was significant for Step 2 for the first post-test ($Mdn_{\text{listening condition}} = .67$; $Mdn_{\text{combined condition}} = .80$; $T = 254.50$, $p = .005$), but not the second ($Mdn_{\text{listening condition}} = .80$; $Mdn_{\text{combined condition}} = .83$; $T = 335.50$, $p = .067$). The difference between condition was not significant for Step 3 at post-test 1 ($Mdn_{\text{listening condition}} = .82$; $Mdn_{\text{combined condition}} = .88$; $T = 180.50$, $p = .118$), while it was significant at post-test 2 ($Mdn_{\text{listening condition}} = .78$; $Mdn_{\text{combined condition}} = .83$; $T = 194.00$, $p = .017$). These results suggest that there was a trend for the combined condition to lead to greater word learning than the listening condition as the task required increasing levels of detailed semantic knowledge. Particularly, the two conditions yielded similar results in the first step of the semantic task, where children were asked to choose the right categories for the new words. Children recognised the right sub-category in Step 2 more often in the combined condition than in the listening condition in the first post-test session, and showed a non-significant trend in the same direction in the second session. Similarly, they showed a non-significant trend in recognising the correct definition more often in the combined than the listening condition. These results are in line with the results of the Mean Step analysis, i.e. children obtained a higher Mean Step for words presented in the combined condition, than for words presented in the listening condition.

2.3.6 – Correlation between the experimental tasks

The relationship between the performance in the three experimental tasks was explored by means of Spearman's rank correlation coefficients (Table 2.12). Only correlations between measures at post-test 1 are reported: given the strong evidence of retention, with no decrement in performance shown at post-test 2, the presence of significant learning already at post-test 1, and similar pattern of results between the two post-test sessions, performance in the first post-test session was considered the purest measure of learning from the story. A Bonferroni correction was applied (corrected critical p-value = .008).

Table 2.12

Spearman's rank correlation coefficients between the scores obtained at post-test 1 in the phonological and the orthographic tasks (number of words correctly recognised), and the semantic task (Mean Step) in each condition

		listening condition			combined condition		
		Phon	Orth	Sem	Phon	Orth	Sem
listening condition	Phon	-					
	Orth	.19	-				
	Sem	.65*	.32	-			
combined condition	Phon	.17	.34	.24	-		
	Orth	.40*	.29	.54*	.17	-	
	Sem	.54*	.29	.63*	.29	.39*	-

Note. Phon = number of words correctly recognised in the phonological task; Orth = number of words correctly recognised in the orthographic task; Sem = Mean Step reached in the semantic task.
*correlation is significant ($p < .008$) after Bonferroni correction.

In the listening condition performance in the semantic and phonological tasks showed a significant association: children who successfully recognised the correct phonological form of the words were also more likely to obtain higher scores in the semantic task. On the other hand, performance in the orthographic task in this condition was not associated with performance in the other two tasks. In the combined condition, scores in the semantic task were significantly associated only with orthographic scores, while scores in the orthographic and phonological tasks were not associated to each other. In summary, semantic knowledge correlated with knowledge of phonological form when the stories were presented only orally, while it was associated with knowledge of orthographic forms and not phonological forms, when the stories were presented in both the written and oral form.

When considering scores in the two conditions, there was an association between semantic learning in the two conditions, but the same was not true for phonological and orthographic learning between the two conditions. On the other hand orthographic learning in the combined condition was associated with phonological learning and semantic learning in the listening condition, and semantic learning in the combined condition was associated with phonological learning in the listening condition.

2.3.7 - Influence of individual abilities on tasks' performance

This section explores the relationship between individual differences, and performance in the word learning tasks. To avoid multicollinearity issues in the regression analyses presented in the following sections, all the measures of reading accuracy, including measures of non-word decoding, word reading and passage reading accuracy, which showed high correlation with each other (see Section 2.3.1), were merged to form a single factor. A principal component analysis (PCA) was conducted on all the background measures with orthogonal rotation (varimax). Two components had eigenvalues greater than 1. In combination these components explained 72.19% of the variance. Table 2.13 shows the factor loadings after rotation. The measures that clustered in component 1 clearly represented a measure of reading accuracy which was highly reliable (Cronbach's alpha = .92), while the interpretation of component 2 was more ambiguous, possibly linked to a general ability factor, and showed comparatively low reliability (Cronbach's alpha = .65). Since the first PCA clearly illustrated that the measures of reading accuracy tended to cluster together to form a reliable factor, a second PCA was carried out, considering only the relevant measures (TOWRE SWE, TOWRE PDE, SWRT and YARC accuracy), to compute more precise factor scores. Factor scores were obtained using the regression method. As expected, only one component had eigenvalues greater than Kaiser's criterion of 1, and this explained 82.03% of the variance (See Appendix I for further information on PCA analyses). Table 2.13 shows the factor loadings for this second analysis. The reading accuracy factor created had a high reliability (Cronbach's alpha=0.93), and was therefore retained for further analysis. Scores in the reading accuracy factor were normally distributed (M = .00, SD = 1.00, $D(63)=0.067$, p=.200).

Table 2.13
Factor loadings for a PCA computed using all the measures (PCA 1), and factor loadings for PCA using only the 4 measures that form the reading accuracy factor (PCA 2)

	PCA 1		PCA 2
	Component 1: reading accuracy	Component 2	Component 1: reading accuracy
TOWRE SWE[a]	**.794**	.346	**.863**
TOWRE PDE[a]	**.948**	.014	**.919**
SWRT[a]	**.890**	.273	**.928**
YARC accuracy[b]	**.865**	.296	**.911**
YARC comprehension[b]	.322	**.775**	
BPVS[a]	.186	**.842**	
USP CELF[c]	.109	**.663**	
CPM[a]	.151	**.707**	

Note. Factor loadings > .40 are in boldface. TOWRE SWE = Sight Word Efficiency subtest of the TOWRE (Test of Word and Reading Efficiency); TOWRE PDE = Phonemic Decoding Efficiency subtest of the TOWRE; SWRT = Single Word Reading Test included in the YARC; YARC accuracy = reading accuracy of passages collected as part of the YARC (York Assessment of Reading for Comprehension); YARC comprehension = score associated with the reading comprehension of passages of the YARC; BPVS = score for the British Picture Vocabulary Scale; USP CELF = Understanding of Spoken Paragraphs subtest of the Clinical Evaluation of Language Fundamentals; CPM = score obtained in the Raven's Coloured Progressive Matrices.
[a]raw score. [b]ability score. [c]scaled score.

Spearman's rank correlation coefficients between the background measures that do not contribute to the new reading accuracy factor, and reading accuracy were computed. A Bonferroni correction was applied to correct for multiple comparison (corrected critical p-value = .01). The reading accuracy factor showed a moderate to strong correlation with the reading comprehension ability score of the YARC (r_s= .52, p < .001), and with the oral vocabulary score of the BPVS (r_s= .39, p = .004), and no significant correlation with either the USP CELF (r_s= .22, p = .117) or the CPM (r_s= .31, p = .021).

2.3.7.1 - Relationship between the phonological task and the background measures. The relationship between the performance in the phonological task in the first post-test session, and the background measures was explored by means of Spearman's rank correlation coefficients, which are reported in Table 2.14. A Bonferroni correction was applied (corrected critical p-value = .01). Reading accuracy significantly correlated with the scores in the phonological task in the combined condition, but not in the listening condition. Phonological scores in the listening condition correlated with listening comprehension (USP CELF) and non-verbal abilities (CPM).

Table 2.14
Spearman's rank correlation coefficients between the scores in the phonological task at post-test 1 (number of words correctly recognised) and background measures

	reading accuracy[d]	YARC comprehension[b]	BPVS[a]	USP CELF[c]	CPM[a]
listening condition	.17	.29	.24	.38*	.41*
combined condition	.36*	.20	.21	.10	.28

Note. Reading accuracy = score on the factor reading accuracy computed in the second PCA, using the regression method (it takes into account the scores for TOWRE SWE, TOWRE PDE, SWRT and YARC accuracy); YARC comprehension = score associated with the reading comprehension of passages collected as part of the YARC; BPVS = score for the British Picture Vocabulary Scale; USP CELF = Understanding of Spoken Paragraphs subtest of the Clinical Evaluation of Language Fundamentals; CPM = score obtained in the Raven's Coloured Progressive Matrices.
[a]raw score. [b]ability score. [c]scaled score. [d]factor score.
*correlation is significant after Bonferroni correction ($p < .01$).

The relationship between the scores in the phonological task and in the background measures was then explored using linear regression models, where all the scores in the two conditions of the phonological task were entered one at a time as the dependent variables, and all the background measures that showed a significant correlation with these were entered as the independent variables, in two backward regression models, in order to find the best fitting explanatory model for each of the phonological scores. When reading accuracy was entered as the sole predictor of scores in the phonological task for the combined condition, the regression explained a significant amount of variance in the phonological scores (F (1,52) = 7.68, p = .008 R^2 = .13), and reading accuracy significantly predicted phonological scores (β = .36, t(53) = 2.77, p = .008). When listening comprehension and non-verbal abilities were entered as predictors of scores in the phonological task for the listening condition, listening comprehension was not retained in the final model. The regression that included non-verbal abilities as the sole predictor explained a significant amount of variance in the phonological scores (F (1,52) = 4.55, p = .038 R^2 = .06), and non-verbal abilities significantly predicted phonological scores in this condition (β = .28, t(53) = 2.13, p = .038). In conclusion the ability to decode words and nonwords predicted the ability to identify the correct phonological form of the new words, but only when both the orthographic and phonological forms of the words were presented in the story (i.e. in the combined condition), while more general non-verbal abilities predicted phonological scores in the listening condition.

Another way to analyse the data from the phonological task is to consider each child's score in each trial as a binomial score (0 if they failed to recognise the correct phonological alternative, and 1 if they recognise it). We used generalised linear mixed models for binomial data (Jaeger, 2008), specifically the function "glmer" from the package "lme4" (Bates, Maechler, Bolker, &Walker, 2014), with the software R (R Core Team, 2013) to analyse the data. Each child produced 16 trials per condition, one score per word for each post-test session; this was used as the dependent variable in each analysis. For each dependent variable, an initial model included a maximal random effects structure that captured our experimental design (Barr, Levy, Scheepers, & Tily, 2013). This entailed the random intercepts terms for both participants and items and the random slopes terms for participants and items that relate to our repeated measures manipulation: condition. However, models including random slopes were prone to non-convergence; therefore, given that results of these models and simpler models that only included the intercepts for item and subject did not differ, the simpler and convergent models are reported (Bates, Kliegl, Vasishth, & Baayen, 2015). We then compared these 'empty' models (using pair-wise Likelihood Ratio Test comparisons; Barr, et al., 2013) with models that additionally included the hypothesised fixed effects: condition (listening *vs.* combined), session (post-test 1 *vs.* post-test 2) and background measures. All continuous factors were centred around the mean for analysis. Hypothesised interactions were included one at a time in the model with all fixed effects, and were retained only if significant. The interactions between condition and the background measures were separately introduced to test whether any background measure had a differential effect on performance depending on presentation modality. Estimates of fixed effects and interactions for the final models are reported in Table 2.15.

Table 2.15

Generalized Linear Mixed Model for performance on the phonological task (scores compute as binomial depending on whether the child had correctly recognised the correct alternative)

Factor	Estimate	St. Error	z values	
			z value	*p*
intercept	2.43	.29	8.39	< .001
condition (listening *vs.* combined)	.08	.17	.48	.629
session (post-test 2 *vs.* post-test 1)	.17	.16	1.03	.305
reading accuracy	.47	.19	2.45	.014
BPVS	.00	.21	-.01	.994
YARC comp	-.16	.23	-.72	.473
USP CELF	.10	.18	.56	.578
CPM	.43	.18	2.30	.021
condition*USP CELF	.39	.15	2.60	.009

The fixed factor model for the phonological task significantly improved fit compared to the empty model (χ^2 (8) = 25.86, p = .001). The final model indicated reading accuracy, non-verbal abilities and the interaction between condition and listening comprehension as significant predictors of performance on the phonological task, with superior learning in children with higher abilities in reading and non-verbal abilities and children with lower scores in the listening comprehension task being more likely to show a higher difference between scores in the two conditions (see Figure 2.7). These results confirmed the lack of a condition effect on phonological learning, and highlighted reading accuracy and non-verbal abilities as good predictors of scores in this task. Furthermore listening comprehension seemed to affect children in the two conditions differently, with listening comprehension having a positive effect on phonological learning in the listening condition, but no similar effect in the combined condition, thus also suggesting that children with lower listening comprehension abilities might experience highest facilitation from a dual modality presentation, or that children who struggle to perceive phonological differences show difficulties in listening comprehension.

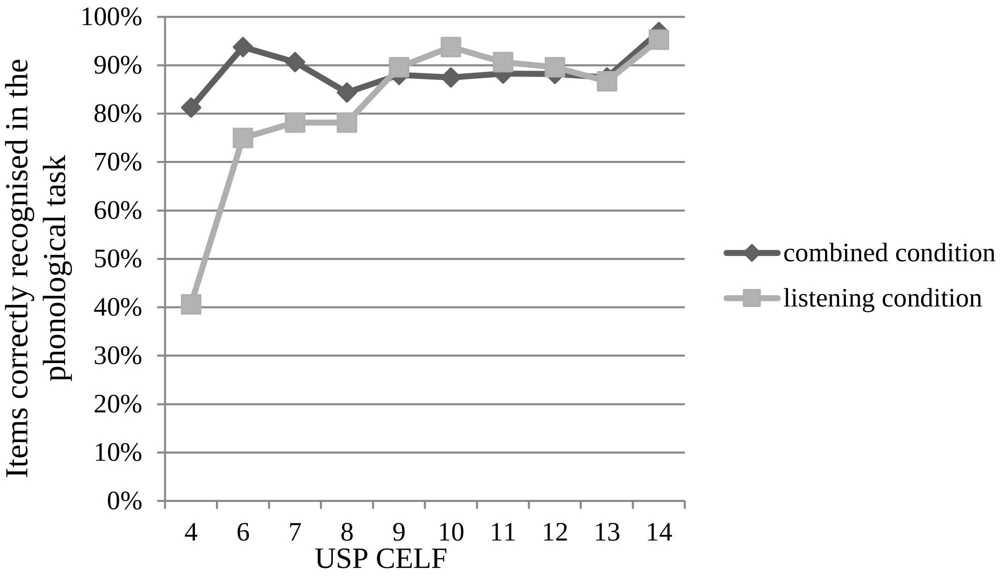


Figure 2.7. Relationship between the number of items correctly recognised in the phonological task (on average in percentage scores) and listening comprehension scores (children were grouped by listening comprehension scores)

2.3.7.2 - Relationship between the orthographic task and the background measures. This section explores whether individual differences affect performance in the orthographic task. Spearman's rank correlation coefficients between orthographic scores in the first post-test session and background measures are reported in Table 2.16. A Bonferroni correction was applied (corrected critical p-value = .01). Reading accuracy significantly correlates with the scores in the orthographic task in both conditions, and no other correlation is significant.

Table 2.16

Spearman's rank correlation coefficients between the scores in the orthographic task at post-test 1 (number of words correctly recognised) and background measures

	reading accuracy[d]	YARC comprehension[b]	BPVS[a]	USP CELF[c]	CPM[a]
listening condition	.35*	.28	.26	.14	.27
combined condition	.35*	.22	.21	.19	.24

Note. Reading accuracy = score on the factor reading accuracy computed in the second PCA, using the regression method (it takes into account the scores for TOWRE SWE, TOWRE PDE, SWRT and YARC accuracy); YARC comprehension = score associated with the reading comprehension of passages collected as part of the YARC; BPVS = score for the British Picture Vocabulary Scale; USP CELF = Understanding of Spoken Paragraphs subtest of the Clinical Evaluation of Language Fundamentals; CPM = score obtained in the Raven's Coloured Progressive Matrices.

[a]raw score. [b]ability score. [c]scaled score. [d]factor score.

*correlation is significant after Bonferroni correction ($p < .01$).

The relationship between the scores in the orthographic task and in the background measures was also explored by means of linear regression models, built as described in the previous section. With reading accuracy as the sole predictor of scores in the orthographic task in the combined condition, the regression explained a significant amount of variance in orthographic scores (F (1,52) = 7.38, p = .009, R^2 = .12), and reading accuracy significantly predicted orthographic scores (β = .35, t(53)= 2.72, p = .009). Similarly the regression computed for the listening condition explained a significant amount of variance in the task (F (1,52) = 7.13, p = .010, R^2 = .12), and reading accuracy significantly predicted orthographic scores (β = .35, t(53)= 2.67, p = .010). In summary the ability to decode words and nonwords correlated with the ability to correctly identify the correct orthographic form of the new words and predicted performance in the orthographic task.

Mixed effects models were conducted for the scores in each trial of the orthographic task (see previous section for a description of the analysis method). In the case of orthographic scores the interaction between session and condition was also considered, given that the analyses in Section 2.3.4 suggested a different effect of session on the two conditions. Estimates of fixed effects and interactions for the final models are reported in Table 2.17.

Table 2.17
Generalized Linear Mixed Model for performance on the orthographic task (scores compute as binomial depending on whether the child had correctly recognised the correct alternative)

Factor	**Estimate**	**St. Error**	**z values**	
			z value	***p***
intercept	1.8	.18	6.99	< .001
condition (listening *vs*. combined)	-.95	.16	-5.96	< .001
session (post-test 2 *vs*. post-test 1)	-.01	.17	-.08	.932
reading accuracy	.39	.11	3.51	< .001
BPVS	.11	.12	.92	.360
YARC comp	.20	.15	1.33	.183
USP CELF	.05	.10	.53	.593
CPM	.09	.11	.71	.478
condition*session	.47	.23	2.09	.036
condition* YARC comp	-.25	.12	-2.12	.034

The fixed factor model for the orthographic task significantly improved fit compared to the empty model (χ^2 (9) = 81.35, p < .001). The final model indicated condition, reading accuracy and the interactions between condition and session and between condition and reading comprehension as significant predictors of performance on the task, with superior learning in the combined condition, and in children with higher reading abilities. As previously explored there was a different effect of session in the two conditions (see Section 2.3.4). Furthermore there seemed to be a different effect of reading comprehension on orthographic learning in the two conditions, with children with medium and high reading comprehension showing more orthographic learning in the combined condition than in the listening condition, while those with lower comprehension skills show little difference in performance across conditions (see Figure 2.8). In summary, this analysis confirms previously established effects of condition, and the different effect of session on the two conditions. It also confirms the results of the previous regression analyses in highlighting reading comprehension as an important predictor of orthographic knowledge for both conditions. Furthermore this analysis seems to suggest that reading comprehension may determine the magnitude of the orthographic facilitation effect for orthographic learning, with better comprehenders experiencing more orthographic facilitation for orthographic learning than less skilled comprehenders.

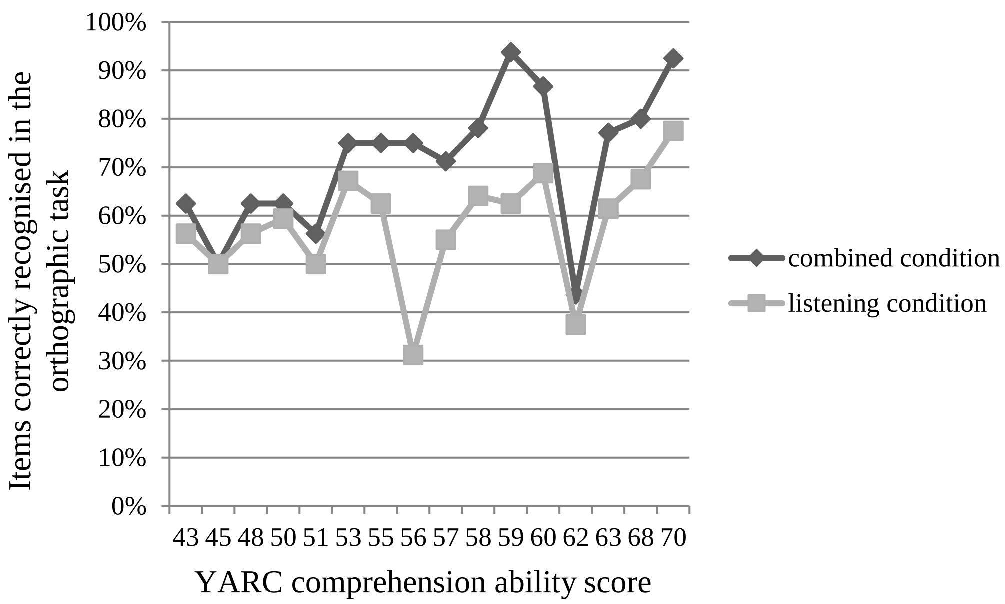


Figure 2.8. Relationship between the number of items correctly recognised in the orthographic task (on average in percentage scores) and reading comprehension scores (children were grouped by reading comprehension)

2.3.7.3 - Relationship between the semantic task and the background measures. The relationship between the performance in the semantic task, specifically Mean Step obtained in the first post-test session, and the background measures was explored first by means of Spearman's rank correlation coefficients, which are reported in Table 2.18. A Bonferroni correction was applied (corrected critical p-value = .01). As it is possible to see from the table all background measures were significantly associated with the Mean Step obtained in both conditions, and the correlations were all either moderate or strong, therefore all measures were retained for further analyses.

Table 2.18

Spearman's rank correlation coefficients between the scores in the semantic task (mean step at post-test 1) and background measures

	reading accuracy[d]	YARC comprehension[b]	BPVS[a]	USP CELF[c]	CPM[a]
listening condition	.42*	.41*	.43*	.39*	.36*
combined condition	.39*	.47*	.57*	.49*	.42*

Note. reading accuracy = factor scores computed in the second PCA, using the regression method (it takes into account the scores for TOWRE SWE, TOWRE PDE, SWRT and YARC accuracy); YARC comprehension = score associated with the reading comprehension of passages collected as part of the YARC; BPVS = score for the British Picture Vocabulary Scale; USP CELF = Understanding of Spoken Paragraphs subtest of the Clinical Evaluation of Language Fundamentals; CPM = score obtained in the Raven’s Coloured Progressive Matrices.

[a]raw score. [b]ability score. [c]scaled score. [d]factor score.

*correlation is significant ($p < .01$).

Two backwards regressions, one for each condition, were then computed.[5] The results of these regression models are reported in Table 2.19. Both models explained variance in the data fairly well (27% for the listening condition and 41% for the combined condition), with listening comprehension (USP CELF) being a significant predictor in both models, and non-verbal abilities (CPM) being retained as a predictor, albeit not reaching significance. The two models differed in the predicting power of reading accuracy, which significantly predicted Mean Step reached in the listening, but not the combined condition, and vocabulary (BPVS), which significantly predicted Mean Step in the combined, but not the listening condition.

[5] scores in the two conditions and in the Mean Step of the semantic task at post-test 1 in the two conditions were entered one by one as the dependent variables, and all the background measures were entered as the independent variables, in two backwards regression. Factors were removed depending on the p-value associated with their Beta, starting from the factors with highest p-value, up until only factors with p-values lower than .100 were retained.

Table 2.19

Final fitted backwards linear regression models to explain Mean Step in the semantic task

	Retained factors for the final model	$\boldsymbol{\beta}$	$\boldsymbol{p}$	**R**	$\mathbf{R^2}$
listening condition	reading accuracy[d]	.29	.027	.56	.27
	USP CELF[c]	.29	.020		
	CPM[a]	.23	.076		
combined condition	BPVS[a]	.36	.007	.67	.41
	USP CELF[c]	.31	.010		
	CPM[a]	.21	.077		

Note. reading accuracy = factor scores computed in the second PCA, using the regression method (it takes into account the scores for TOWRE SWE, TOWRE PDE, SWRT and YARC accuracy); YARC comprehension = score associated with the reading comprehension of passages collected as part of the YARC; BPVS = score for the British Picture Vocabulary Scale; USP CELF = Understanding of Spoken Paragraphs subtest of the Clinical Evaluation of Language Fundamentals; CPM = score obtained in the Raven's Coloured Progressive Matrices.
[a]raw score. [c]scaled score. [d]factor score.

Another way to compute scores for the semantic task was to compute a score that considered only whether children successfully completed either the first step of the semantic task, or all the three steps of the semantic task. While the first measure (i.e. correctly recognising the category for the word) could be considered a measure of more general semantic knowledge, the second measure could more clearly be associated with a measure of detailed semantic knowledge. Each child produced 16 trials per condition, one for each post-test session. These two binomial measures were used to compare the two conditions using generalised linear mixed-effects models for binomial data (Jaeger, 2008), using the function "glmer" from the package "lme4" (Bates et al., 2014), computed with the software R (R Core Team, 2013). Children provided 32 responses to target words (eight for each post-test session, sixteen for each condition); this was used as the dependent variable in each analysis. Models were built and evaluated as described in previous sections for the phonological and orthographic tasks. Estimates of fixed effects and interactions for the final models are reported in Table 2.20.

Table 2.20

Generalized Linear Mixed Model for performance on the semantic task (scores compute as binomial depending on whether the child had correctly recognised the category of the new word (step 1), or all the steps of the task (all steps))

Factor	**Estimate**	**St. Error**	**z values**	
			z value	***p***
Step 1 – category recognition				
intercept	.96	.26	3.75	< .001
condition				
(listening *vs.* combined)	-.19	.12	-1.61	.108
session				
(post-test 2 *vs.* post-test 1)	-.17	.12	-.55	.585
reading accuracy	.32	.16	1.96	.050
BPVS	.32	.18	1.77	.077
YARC comp	.04	.19	.22	.829
USP CELF	.35	.14	2.43	.015
CPM	.20	.17	1.90	.234
All steps – detailed knowledge				
intercept	-.44	.26	-1.67	.095
condition				
(listening *vs.* combined)	-.64	.12	-5.30	< .001
session				
(post-test 2 *vs.* post-test 1)	-.03	.12	-.24	.812
reading accuracy	.33	.19	1.80	.071
BPVS	.43	.21	2.08	.038
YARC comp	.08	.21	.70	.699
USP CELF	.31	.17	1.86	.063
CPM	.48	.20	2.43	.015

Both fixed factor models for the semantic task significantly improved fit compared to the empty models (Step 1: χ^2 (7) = 33.20, p < .001; All steps: χ^2 (7) = 66.03, p < .001). Interaction terms did not significantly improve model fit. The final model for category recognition indicated reading accuracy and listening comprehension as significant predictors, while the model that considers all the steps indicated condition, vocabulary and non-verbal abilities as significant predictors of performance on the semantic task. Therefore, while the model that considers only Step 1 did not highlight a facilitation effect of the combined condition, the model that considers all three steps of the semantic task showed superior semantic learning in the combined condition compared to the listening condition. This confirms the presence of a condition effect only when considering all the steps of the semantic task (see Section 2.3.5). When considering individual differences, the model for category recognition (Step 1) showed better performance in children with higher abilities in reading accuracy and listening comprehension. The model that considers all the steps also highlighted the importance of vocabulary and general non-verbal abilities, although reading accuracy and listening comprehension were nearing significance. The results are also similar to the results of the backwards regression, in highlighting non-verbal abilities as a good predictor of semantic learning when all the steps are considered. Differently from the previous analyses, these also highlighted an effect of vocabulary, independent of presentation modality. Interestingly, a different set of abilities seems to predict performance in the semantic task, depending on the measure considered, for example differentiating category recognition and recognition of the correct alternative in all steps.

2.4 - Discussion

The aim of the study was to assess whether presenting stories simultaneously in an oral and written form would be more beneficial to vocabulary acquisition than an oral presentation (i.e. the presence of an orthographic facilitation effect). It was specifically predicted that new phonological and orthographic forms would be acquired more easily when children were exposed to both phonology and orthography (cf. Ricketts et al., 2009; Rosenthal & Ehri, 2008; Vadasy & Sanders, 2015). Furthermore it was expected that the combined condition would also facilitate semantic learning compared to the listening condition (cf. Ricketts et al., 2009; Rosenthal & Ehri, 2008). These hypotheses were addressed in the first set of analyses (sections 2.3.3 to 2.3.5), and confirmed by the mixed models (sections 2.3.7.1 to 2.3.7.3) and will be explored in the first part of the discussion. The relationship between the experimental tasks (section 2.3.6) will be also discussed in this section, to reach a more comprehensive account of children's vocabulary acquisition.

A second aim of the study was to explore which individual abilities would affect learning. It was particularly expected that children with bigger vocabularies would acquire word meanings better (Robbins & Ehri, 1994; Rosenthal & Ehri, 2008), and that reading abilities would predict learning of word forms (Ricketts et al., 2011; Rosenthal & Ehri, 2008; 2011). These effects were explored in Section 2.3.7, and will be explored in the second part of the discussion. The final part of this second section will be devoted to the discussion of whether any individual ability showed an effect on the magnitude of the orthographic facilitation effect (mixed models in Sections 2.3.7.1 to 2.3.7.3). In this case it was expected that better readers may show a larger facilitation effect that less skilled readers (Ricketts et al., 2009; Rosenthal & Ehri, 2008). A conclusion will then be drawn to summarise the results from this first experiment.

2.4.1 - Story presentation modality effects

Children learnt the phonological forms of the words very proficiently, averaging seven words recognised out of eight in both conditions, but the expected advantage for the combined condition for phonological learning (Miles et al., 2016; Ricketts et al., 2009; Rosenthal & Ehri, 2008) was not found. It must be noted that the means for the phonological recognition task were very high, suggesting that this task may have been too easy to show a difference between conditions. In any case, it is not possible to conclude from the results that the additional presence of orthography yielded any advantage for learning of the phonological forms of the new words. Children were able to use the phonological information provided equally well in the two conditions to learn phonological forms.

The results from the orthographic task are quite different. In fact, as expected, children showed significant learning of the orthographic forms of the words only in the condition where orthography was directly presented. The provision of the phonological form alone did not prompt children to create an orthographic form for the new item. This result also seems to suggest that children were paying attention to the written text provided in the combined condition, using it to extract orthographic information regarding the new words. On the other hand, in the listening condition, they did not seem to form a representation of the written form of the words. This result supports previous studies in highlighting an orthographic facilitation effect for orthographic learning (Miles et al., 2016; Vadasy & Sanders, 2015). On the other hand, the poor performance in the orthographic task when words were presented orally seems to contradict the idea proposed by Wegener et al. (2017), who suggest that school-aged children form an orthographic skeleton for orally presented words. Children, in fact performed similarly for control words and words presented in the listening condition, thus failing to show any positive effect of knowledge of words' oral forms on orthographic learning. It is possible that the difference between our result and that of previous research might stem from a difference in how orthographic knowledge was probed. Wegener et al., in fact, assessed familiarity with the orthographic form of the words through eye-movements, while our study tested orthographic knowledge more directly. It is possible that children had formed implicit expectations regarding the orthographic form of the new words in our study, but these expectations would not have been specific enough to change their behaviour and lead to a correct selection of the spelling of our new words. Furthermore, it is important to note that children's ability to create an orthographic skeleton for new words depends on the consistency of the words' spelling pattern (Wegener et al., 2017). The fact that we used real words, without directly controlling for the consistency of their spelling, might therefore have had a negative impact on children's ability to form a reliable orthographic skeleton from the oral presentation.

In this study, semantic learning was assessed using a task in three steps (step 1 category recognition, followed by step 2 sub-category recognition, and then step 3, corresponding to definition recognition). Several measures of word learning were therefore considered, particularly Mean Step (i.e. the mean step reached by each child in each condition), proportion of words correctly recognised in each step between those assessed in that specific step, and two binary measures, corresponding to whether or not they recognised the correct category in step 1, and whether they recognised the correct alternative in all steps, respectively (see Section 2.7.3). The results clearly showed that children learned the categories of words equally well from the listening condition and the combined condition, both when considering proportions correct on step 1 and the binary recognition measure for step 1. In addition, they particularly showed good category learning, by identifying the right category for around five words out of eight in both conditions. This result contrasts with previous studies that showed better word learning in conditions where both phonology and orthography were presented, compared to conditions where only phonology was presented (Ricketts et al., 2009; Rosenthal & Ehri, 2008; Vadasy and Sanders, 2015). Despite this lack of difference in the first step of the semantic task, children showed better word learning in the combined condition as increasing levels of details were required: they were better at recognising the right definition of the words when they had been presented with these both orally and written, compared to when they were presented only orally. This result was replicated using Mean Step, proportion scores and the binary measure that considered all the steps. Therefore we could conclude that children were equally good in learning about the categories that words belong to, in the two conditions, but they seem to be better able to acquire detailed knowledge of the words in the combined condition compared to the listening condition, despite the fact that the same stories were used in the two conditions, and that all the words were accompanied by a definition. Whether or not the ability to recognise sub-categories and definitions indicates a deeper semantic knowledge is a debatable issue, which will be explored in detail below. First we will consider some possible reasons behind the better performance in the combined condition.

Different explanations for these results could be proposed, as discussed in Chapter 1. One possibility is that the combined condition enhanced children's performance by allowing children to form a more complete representation of the given words. Following the Lexical Quality Hypothesis (Perfetti & Hart, 2002), for example, we could hypothesize that, when children were presented with both the orthographic and phonological forms of the words, they were more able to link the orthographic, phonological and semantic representation of them in memory, than when they were presented with the words only orally. This hypothesis seems supported by the data, given that children did not acquire orthographic information in the listening condition, thus showing a lower quality representation for these words. In relation to semantic learning, the quality of the representation may be more evident in steps that require the retrieval of more detailed information, thus affecting step 3 (definitions) more than step 1 (categories), although there is no evidence for this specific assumption in the literature. Another possible explanation is that the redundancy of the information freed attentional resources for children to attend to the meaning of the text online, therefore resulting in deeper semantic learning (Mayer, 2014; Mayer et al., 1999; Paas et al., 2003). The difference between these two ideas will be explored more in depth in Chapters 3 and 4.

The previous hypotheses point to the idea that the advantage in the combined condition was obtained thanks to the combined presence of the oral and written modalities, but we cannot determine from this study whether the presence of reading alone could have produced a similar advantage. If this was the case we might expect an advantage of reading stories over listening to them; in contrast this advantage would not be found if the facilitation was given by the combination of two modalities of presentation. The results of the study by Suggate et al. (2013), who found an advantage of the oral presentation over the written presentation of stories, seem to exclude the possibility that the advantage for the combined condition might be due to the presence of the written presentation only. On the other hand, the results of studies of second language learners (Brown et al., 2008), where students tend to learn as much from reading as they do from listening and reading combined, but less from listening only, seem to suggest that the simple presence of written text might be the cause of the orthographic facilitation effect found for semantic learning. Unfortunately, our data cannot exclude any of these hypotheses. The absence of a reading only condition in the final design makes it impossible to explore whether the orthographic facilitation effect might be due to the presence of the written modality only, as opposed to the presence of two modalities (written and oral) combined.

A related possibility is that, in the listening condition, children were forced to follow the reading pace of the teacher, and they were not able to revisit specific parts of the story, while the provision of the written text gave them more freedom to revisit parts of the story, even after these were read. Children's better performance in step 3 of the semantic task in the combined condition might reflect this possibility: when presented with the written text, children may have spent more time reading and elaborating the definitions of the new words, perhaps not reading the story at the same pace at which it was read to them. The time spent elaborating the definition could have led to an advantage in the second and third step of the semantic task, when a deeper understanding of the words was required, but not in the first step, which assessed more superficial learning. Chapter 4 will explore how children attend to the story in a dual presentation modality.

It might be important to note that items were presented both orally and in the written form in the semantic tasks. This presentation modality corresponds to the dual presentation modality of the combined condition, while it differs from the unimodal oral presentation of the listening condition. The similarity between presentation modality and testing modality might have therefore favoured the combined condition. However, we consider it unlikely. For instance, if this was the cause of the better performance in the combined condition, we would expect to see this in all three steps of the semantic task, which we did not. It might be possible that presentation modality would affect the sub-category and definition more, because both of these were directly presented in the story, in the same modality as testing in the combined condition, but only orally in the listening condition. Categories, on the other hand, were not directly presented in the stories. Nevertheless this explanation is rendered less likely by the fact that even sub-categories and definitions were not exactly the same as those presented in the story, but paraphrased. We recognise, however, that a different testing modality might have solved these issues. Nevertheless the combined testing modality was considered the most appropriate, despite possible disadvantages (see conclusion for a more detailed discussion).

Some consideration regarding characteristics of the semantic task should be addressed. The study was originally designed to consider step 1, step 2 and step 3 of the semantic task as measures of increasing detail of semantic knowledge regarding the target words. Step 1 required a more general ability to recognise the correct category for the words, step 2 required more specific knowledge, such as the ability to choose the correct sub-category of the word, while step 3 required children to recognise the correct definition of the word out of four very similar ones. It was originally reasoned that recognising a general category required less detailed knowledge of the words than recognising their definitions amongst very similar ones. Step 1 was expected to be the easiest for the children, since the literature shows that, by the age of 8, children tend to use basic-level categories to organise their mental lexicon (Hashimoto, McGregor, & Graham, 2007; Lin & Murphy, 2001). For this reason the task was organised to allow children to proceed to later steps only if they correctly completed earlier ones.

However, it could be argued that recognising the correct category (step 1) may have been in fact the most difficult task for the children, for two different reasons. One reason is that the category recognition step requires abstraction. To recognise the correct category children need to abstract category-level knowledge about the target words from the specific information provided in the story, while recognising the correct definition (step 3) might be easier, since it required children to recognise a definition similar to that provided in the story. A second possibility is that step 3 might be easier than step 1 because between the alternatives in the definition recognition step, only one was a definition for a word presented in the story. Conversely, all the alternatives in the category recognition step correctly described one of the new words presented in the story. Therefore, it might be easier to recognise that *destrier* is *a horse used for fighting*, amongst other kind of horses, because only this type of horse was presented in the story. On the other hand, to recognise that *destrier* is an *animal*, and not an *object* or a *job* or *clothing* the child had to remember the specific link between the new word and its meaning, distinguishing it from the meaning for other new words.

For these reasons, it is possible that, although recognising the definition (step 3) necessitated a higher level of detail, recognising the correct category might in fact have been harder. However, the set-up of the semantic task ensured that children only progressed to step 3 if they had chosen the right category for the word in step 1 and 2. In summary, given the ideas mentioned in this section, it is not possible to reach a definite conclusion regarding what led to the advantage of the combined condition over the listening condition, or of the difference between category recognition and other measures of semantic knowledge. It is acknowledged that, in future studies, the assumption that knowledge assessed in step 1 is more easily learned than knowledge assessed in step 2 and step 3 should be avoided.

2.4.2 - Relationship between semantic, phonological and orthographic learning

Children's performance in the semantic task in the listening condition correlated with the performance in the phonological task, while children's performance in the semantic task in the combined condition correlated mostly with performance in the orthographic tasks. It is therefore possible that, in the listening condition, either the ability to acquire phonological forms influenced the ability to acquire semantic forms (or vice versa), or similar factors were influencing both phonological and semantic learning. On the other hand, in the combined condition, the acquisition of word meanings seemed more associated with the acquisition of written than oral forms. Interestingly, children's ability to acquire phonological and orthographic forms in the combined condition are not associated, thus suggesting that different children tend to learn different properties of the words from a dual presentation.

When considering performance in both conditions, children who learnt phonology best in the listening condition also learnt orthography best in the combined condition, and were more likely to acquire more semantic information in both conditions. This result suggests that paying attention to the oral form of the word in the listening condition and the written form in the combined condition were the best strategies to learn new words, or that the same underlying abilities that drove phonological acquisition during oral presentation, and orthographic acquisition during a combined presentation also best explained semantic learning in both conditions.

2.4.3 - Individual differences effects

Reading accuracy, measured as children's ability to read words and non-words aloud quickly and accurately, and to read a text accurately, was one of the measures that best predicted children's learning of new words. This measure predicted performance in the phonological task, particularly in the combined condition, and orthographic learning in both conditions. In the combined condition children with higher reading abilities were probably able to link the phonological and orthographic forms of the new words more effectively during story presentation, and use this knowledge to succeed in both phonological and orthographic tasks.[6] The importance of reading abilities in learning

[6] This conclusion is tempered by the lack of correlation between the phonological and orthographic task in the combined condition, but this lack of correlation may reflect the different distribution of the scores in the two tasks, and the relatively higher scores in the phonological task. See Appendix J.

word forms is in line with previous results (Bowey & Miller, 2007; Ricketts et al., 2009; Rosenthal & Ehri, 2008; 2011). When considering orthographic learning in the listening condition, we might hypothesize that children with higher reading accuracy ability were able to use their knowledge of orthographic-phonological mapping to make a more informed guess about the orthographic forms of the words they had previously learnt only orally, while less skilled decoders were less able to do so: better readers might have created an orthographic skeleton from oral presentation more readily than less skilled ones (Wegener et al., 2017).[7]

Reading accuracy also predicted performance in the semantic task in the listening condition. It could be hypothesized that better readers were able to form some orthographic representations of the new words despite lack of input, and were more likely to form good quality representation of the words, therefore being able to encode and/or retrieve their meanings more easily (Perfetti & Hart, 2002). This result, alongside the similar result in the orthographic task, and the orthographic facilitation effect found for the semantic task, supports the hypothesis that a good mapping between the oral and written forms of the words enhances the acquisition of word meanings. When considering both conditions and both post-test sessions, reading accuracy is highlighted as a significant predictor of category recognition. Being able to easily form a link between orthographic and phonological forms of new words might be important for semantic acquisition in the combined condition as well.

The orthographic facilitation effect anticipated for semantic, orthographic and phonological learning was expected to interact with reading accuracy, with better readers showing a larger facilitation effect that less skilled readers (Ricketts et al., 2009; Rosenthal & Ehri, 2008). However, no interaction between condition and word and nonword reading was significant. In our sample good and less efficient decoders experienced similar benefit from the dual modality condition. The effect of reading accuracy skills on orthographic facilitation might be more pronounced when words are presented in isolation, or when children's attention is focussed specifically on the word learning task, while the dual presentation of both words and context might be useful for less skilled readers, as well as more skilled ones, in an incidental learning context.

In conclusion, reading accuracy, which included children's ability to read nonwords, words and texts, played a role not only in predicting children's orthographic

[7] Although it is important to remember that overall performance in the orthographic task in the listening condition was not better than performance on control words trials.

learning, but also their phonological and semantic learning. Children's understanding of the mapping between oral and written forms of the words, as expressed by higher scores in this measure, helps when learning both the forms, and, more interestingly, the meanings of new words. Children who were able to form some orthographic representations of the new words, even when lacking orthographic input, were perhaps more able to form good quality representations of the words, and were therefore better able to encode and/or retrieve meanings of the words more easily. The presence of a written presentation in addition to the oral presentation did not particularly advantaged good or less skilled decoders, suggesting that both groups benefited similarly from a dual presentation,

Differently from previous research (Cain et al. 2003; Jenkins et al., 1984; Swanborn & de Glopper, 1999), reading comprehension was not a very good predictor of children's performance. Reading comprehension skills only showed a significant interaction with condition in explaining orthographic learning. Better comprehenders showed a bigger difference between scores in the two conditions, in favour of the combined condition, compared to less skilled readers (Figure 2.8). The presentation of the written text seemed therefore to facilitate good readers more than less skilled ones in acquiring orthographic knowledge. Good comprehenders, thus, experienced an advantage for orthographic learning when meeting the words embedded in a meaningful text. Understanding the story may have been easier and less effortful for good comprehenders, thus freeing resources for them to attend to the written representation of the words. Nevertheless this result was not confirmed in all analyses, suggesting that this effect was probably not robust.

Listening comprehension predicted both phonological and semantic learning (Mean Step) in the listening condition. The fact that children who were good at understanding spoken texts fared better in the listening condition seems also quite natural. Children with good oral comprehension skills were probably better at understanding the stories, and therefore more likely to learn the words embedded in them, by understanding the definition and the context of these words best. Even for the combined condition, performance in the semantic task was explained by children's oral language abilities, including passage comprehension and vocabulary knowledge: when provided with both the oral and the written input, what made a difference in children's performance was their ability to attend and gain as much as possible from the oral modality. It could also be hypothesized that children with higher scores in oral vocabulary and oral passage comprehension, being very proficient listeners, were able to attend to both modalities more easily, since, for them, attending to oral language input was a less demanding task than for other children. Less proficient listeners, on the other hand, were probably less able to take advantage of the presence of the oral presentation, acquiring less information from it, and with more effort. When considering both conditions and both post-test sessions, listening comprehension is highlighted as a significant predictor of category recognition, while vocabulary knowledge better explains the ability to complete all three steps of the semantic task. Our measure of vocabulary seems to predict more detailed semantic knowledge. It is possible that children of all abilities could form a representation of the words, but those with bigger vocabularies included the new words in a richer semantic network, which made it easier for them to acquire specific details of the words (for example by differentiating them from words with similar meanings). The mixed models also show an interaction between condition and listening comprehension when considering performance in the phonological task. This suggests that listening comprehension might affect the different amount of phonological learning in the two conditions (see Figure 2.7). Children with poorer listening comprehension abilities might be facilitated in a dual presentation over an oral presentation for phonological learning. These children, having either low oral abilities or a lower oral attention, might benefit from the presence of the written text, to focus their attention on the words, and encode them better.

It is perhaps surprising that oral vocabulary did not play a bigger role in word learning, especially in the listening condition, contrary to the hypothesis based on previous research (Robbins & Ehri, 1994; Sénéchal et al., 1995). It must be noted that our oral vocabulary measure and our oral passage comprehension measure were correlated, and it is therefore very likely that part of the variability explained by oral vocabulary was also explained by oral passage comprehension. This explanation is in line with the results from second language learners (Lin, 2014), with children with the highest oral proficiency obtaining greater vocabulary gains. In general it could be observed that the passage comprehension test used as a background measure assessed oral attention as well as oral comprehension. It could therefore be hypothesized that the children who scored higher on this test were those who had a tendency to pay more careful attention to the story presentation, both in the listening and the combined condition, and for this reason this measure was prominent in predicting performance in both conditions. This consideration highlights the lack of more direct measures of attention to and comprehension of our stories, which should be addressed in future studies.

Non-verbal abilities predicted phonological scores, mainly in the listening condition, and semantic scores when all steps were considered. It is possible that children's attention to detail and ability to consider pattern similarities from a varied input may have helped children in succeeding in a phonological task that required them to distinguish very similar sound patterns and choose the one they had previously heard. Attention to detail might particularly help in the listening condition, where no other clue to phonology during story presentation was present (i.e. they could not rely on their knowledge of orthography and orthography-phonology rules). Conversely, other skills, such as knowledge of the link between phonology and orthography, better predicted phonological scores when the orthographic form was presented. It might also be considered that non-verbal ability played a role in learning new words at increasing levels of demand (i.e. all steps *vs.* category recognition only): children with higher non-verbal abilities tended to progress further on the semantic task, than less skilled children. Since in this research we used the Coloured Progressive Matrices as a measure of non-verbal abilities, and this test is highly correlated with other measures of I.Q. and fluid intelligence (Banks & Sinha, 1951; Martin & Wiechers, 1954), we could hypothesize that children with higher non-verbal abilities measured by this test were able to distinguish similar alternatives more easily (able to make meaning out of confusion) in Steps 2 and 3 of the semantic task. Furthermore it is possible that such children were better able to use the information in the first post-test session (for example, that their answer was wrong in step 2, because they did not proceed to step 3, or the orthographic information provided in the semantic task), and use it in the second post-test session (by choosing a different alternative) more effectively than less skilled children.

It is noteworthy that individual differences did not interact with condition in explaining performance in the semantic task, thus suggesting that the orthographic facilitation shown in the semantic task (i.e. the better performance in the combined condition compared to the listening condition in the third step of the semantic task) was not influenced by individual differences. This was interpreted as evidence that all the children benefited similarly from the presence of orthography. In conclusion, although reading accuracy might have been expected to have an effect on the extent of orthographic facilitation (Ricketts et al., 2009; Rosenthal & Ehri, 2008), this does not seem to be the case for either learning of word forms or meaning. It is possible that reading accuracy might play a more prominent role in the orthographic facilitation effect when new words are presented in isolation, but, when presented in context, other type of abilities, especially the ease with which this context is processed, as measured by listening or reading comprehension, might play a more important role in the extent to which the presence of a written presentation facilitates vocabulary acquisition.

2.4.4 - Conclusion

The prediction that children would acquire orthographic forms more proficiently when they were exposed to the written form of the words, compared to a condition where the orthographical form was not presented was clearly confirmed by the results. On the other hand, the similar prediction for phonological learning was not confirmed: children learnt phonological forms equally well in the two conditions. This result seems to confirm that children can acquire a good oral representation of the words irrespective of a written presentation, as they efficiently do in the earlier years of their life before they learn to read. Nevertheless, the difference between this result and previous research (Rosenthal & Ehri, 2008) may be also attributed to the ease of the recognition task used in the present study, compared to the more complex production tasks used in previous studies.

The expectation that the combined condition would lead to more semantic learning than the listening condition (Ricketts et al., 2009; Rosenthal & Ehri, 2008) was only in part confirmed from our results. While category recognition was similar in the two conditions, children were more likely to recognise more detailed features of the new words in the combined condition than in the listening condition. These results suggest that a dual presentation modality facilitates deeper semantic learning. However, this result is deeply linked to the nature of the task used, and the different testing procedures used in other research, which include word-picture matching tasks (Ricketts et al., 2009), and sentence completion (Rosenthal & Ehri, 2008), might explain the different results reported.

Regarding children’s abilities, it was expected that children with higher vocabulary knowledge would learn more word meanings than children with smaller vocabularies, in both presentation modalities (Robbins & Ehri, 1994; Rosenthal & Ehri, 2008). While vocabulary scores did indeed correlate with scores in the semantic task, so did all the other background measures. Nevertheless, our measure of listening comprehension predicted semantic learning better than vocabulary knowledge itself. The results observed seem to bear similarities with the results from second language learners (Lin, 2014), with children with highest oral proficiency obtaining greater vocabulary gains. This result suggests that general oral language proficiency in some cases explains vocabulary learning better than vocabulary itself, at least when vocabulary is presented in context and not in isolation. In any case, since the two measures were significantly correlated, it is possible that part of the variability explained by oral vocabulary was also explained by oral passage comprehension.

It was also hypothesised that decoding skills would predict phonological, and orthographic learning in both conditions (Ricketts et al., 2011; Rosenthal & Ehri, 2008; Rosenthal & Ehri, 2011), and the orthographic facilitation effect expected in the semantic and orthographic tasks would interact with decoding abilities, with better readers showing a larger facilitation effect than less skilled readers (Ricketts et al., 2009; Rosenthal & Ehri, 2008). The importance of decoding skills in determining performance was confirmed in all tasks. On the other hand, decoding skills failed to reliably predict the extent of orthographic facilitation. Listening comprehension and reading comprehension had a large impact on the difference between the two conditions in the phonological and orthographic tasks respectively, possibly suggesting that more general comprehension skills allow children to encode word forms more easily when new words are presented in a story context. Nevertheless, these results were not confirmed in all analyses.

2.4.5 - Limitations and future steps

One of the main limitations of this study is that it is not possible to clearly discern the origin of the advantage of the combined condition over the listening condition in semantic learning. Our data, in fact, do not permit conclusions regarding whether the advantage was obtained thanks to the combined presence of the oral and written modalities, or to the written presentation only, or whether the advantage was mostly a by-product of the combined testing modality used.

As seen in Chapter 1 there may be different possible explanations for an advantage of the combined condition over a single modality condition. A dual presentation modality could in fact provide a representation of higher quality, by presenting both oral and written forms of the words (Perfetti & Hart, 2002). This facilitation could stem from the direct presentation of two word forms, or from the indirect process of freeing attentional resources for the purpose of word learning, by making the task of attending to the story easier. The results of this study do not permit us to distinguish between the two possibilities, although we can assume that a combined oral and written presentation certainly produces a representation of higher quality, since children did not learn the written form of the words in the listening condition. Assessing children's comprehension of the stories would have allowed us to assess whether a dual presentation modality may be facilitating vocabulary acquisition through a more general facilitation of the comprehension process.

Another open question is whether the advantage of the combined condition stems from the dual presentation modality, or the presence of the written form only. A relatively easy way to address this question would be to add a reading condition to the study, and compare the learning resulting from this presentation modality to learning acquired in the two other presentation modalities. As seen in Chapter 1, if we consider studies with first language learners, such as the research carried out by Suggate et al. (2013), who found an advantage of oral presentation of stories over written presentation, we would expect that the advantage for the combined condition might be due to the combined presentation of the two modalities. Therefore, we would expect a better performance in the combined condition compared to the reading condition. Instead, if we consider the results of studies on second language learners (Brown et al., 2008), where students tend to learn as much from reading as from listening and reading combined, but less from listening only, we would expect that the mere presence of written text would be the cause of the orthographic facilitation effect found for semantic learning. This would predict similar results from a combined condition and a reading condition, with both leading to better performance than the listening condition. The addition of a reading condition to our design would, consequently, help to answer this theoretical question.

In the present research the presentation of the stories was carried out in the class context by the teachers. This story presentation modality, which was chosen for its ecological validity, leaves open a series of questions. For example it is not possible to ascertain whether all the children were paying attention to the presentation of the story, and if attention during story presentation had a prominent, underlying role in children's performance that could explain the difference between the conditions. Furthermore, this story presentation modality leaves open the issue of whether children were attending mostly the written or the oral presentation, or both equally, in the combined condition. A few ways to control for these effects could be proposed: firstly, a more controlled story presentation procedure, with children being presented the stories in smaller groups, might ensure more equal attention to the story in different modalities. The introduction of a few comprehension questions after story presentation would also promote children's attention to the story, and provide a measure of children's comprehension of the stories themselves.

The question of whether the advantage of the combined condition in the semantic task was an artefact of the testing modality is more complex to address. Two features of the testing modality may have influenced performance in the semantic task: presentation modality at testing and the three steps of the semantic task.

The first issue arises from the fact that both words and the alternative options of the semantic task were presented in both oral and written modalities at test. This testing method was chosen to allow children in the combined condition to access semantics via all available sources. This, however, might have advantaged children in the combined condition. In the listening condition, in fact, participants were presented with new information regarding the words (written forms) during testing, and could have been distracted by the presence of this new information. Unfortunately, it is not possible to conclude that children would have performed the same if they had been presented the words in only one modality at testing. This surprise effect is made less likely by the presentation of written forms in the orthographic task and at pre-test, both of which were completed before the semantic post-test. For this reason, the combined presence of orthography and phonology might not have been completely unfamiliar to children in the listening condition during testing. Furthermore, all possible solutions to this issue carry some degree of difficulty. Alternative ways of testing semantic learning would have been to test it in one modality, the same for both conditions, or different modalities depending on condition. Testing in the written modality would have disadvantaged children in the listening condition. By testing the words only orally, on the other hand, we would have forced the children in the combined condition to rely solely on phonology to access meaning. This might have helped us assess whether the strength of the link between phonology and meaning was affected by the presence of orthography. Nevertheless, this testing modality would had forced children in the combined condition to rely on one link (phonology-meaning), when, in fact, they could have used a different one (orthography-meaning), or an indirect one (orthography-phonology-meaning). This might therefore have prevented them from showing the full extent of their knowledge. It could be argued, in fact, that this would not have assessed children's entire knowledge of the words, but only the children's knowledge accessible via phonology. Another solution could have been to match testing modality with story presentation modality, thus testing words only orally in the listening condition and both orally and written in the combined condition. This testing modality would have resulted in a more difficult comparison between conditions, since the testing modality would have been an added source of variation in the data. For these reasons, the combined oral and written modality, albeit not free from difficulties in interpretation, was deemed the best to tap semantic knowledge acquired from the stories.

A further issue related to the design of the semantic task was that children were not allowed to perform the definition recognition task if they had failed the category recognition task. This design decision was originally built on the assumption that recognising the category of the words would be easier than recognising detailed definitions, but, as discussed in Section 2.4.1, this may not be the case. Nevertheless the most obvious solution – to allow children to complete all three steps regardless of performance in the previous step – would have allowed children to learn new information about the words in the first testing session, and apply this learning at the second test session. Nevertheless, separating the three tasks might be considered an option in future studies.

Given the features of the story used in this study, the origin of the semantic knowledge acquired by the children in our story presentation is not clear. The stories in fact provided both definitions and contextual information, and the words were an integral part of the stories themselves. It is possible that different features of the word were acquired from different features of the stories. For example, more detailed features from a definition, and more general categories from the context (see Section 1.6.1). It might, for example, be the case that children were able to acquire sufficient information to choose the correct alternative in step 1 and step 2 from context, while specific attention to definitions might have influenced performance in step 3. It is also possible that children in the two conditions used the two sources of information differently. In future studies, the differential roles played by contextual information and definitions might be explored more in depth, by providing definitions for some target words, and contextual information for others. This would allow to distinguish which semantics aspects can be learnt from context, and which from definitions, and whether these aspects are differentially acquired in different presentation conditions.

Chapter 3: Study 2

3.1 - Introduction

Building on the results of Study 1 (Chapter 2) this study explored how children learn new words from stories presented in three different modalities (listening, reading, and a combined listening and reading modality). While in Study 1 all the children were presented stories in both story presentation modalities, in this study different groups of children were presented the same story, in different modalities, making this a between subjects design, with the groups of children matched for oral, reading and non-verbal abilities.

This study also had two specific additional aims compared to the previous experiment. The first aim was to elucidate the source of the advantage of the combined condition over the listening condition for semantic learning, by adding a reading condition. This allowed us to distinguish the effects of the presence of the written text per se, and the effect of the combined presence of the oral and written texts (see Section 2.4.1). A further aim was to distinguish the source of semantic learning within the stories: we particularly decided to present words in the stories either accompanied or not accompanied by a definition, to explore whether children use different sources of information in different conditions to learn the meaning of new words.

As with Study 1, learning of the phonological and orthographic forms of the words was explored, as well as learning of the semantic features of these words. In contrast with the set-up of the semantic task in Study 1, where children progressed to further steps only if they had correctly identified the right answer in the previous ones, the semantic task in this study was divided into its three components, so that all the children were assessed on all the words in all three semantic tasks. With this testing modality it was possible to avoid the assumption, previously postulated, that the easier task for the children would be to recognise the categories of the new words, and the most difficult task the recognition of the definitions. In addition to the tasks used in the previous study, a story comprehension task and a measure of executive control were added.

3.1.1 - The effect of presenting words in different conditions

As summarised in Chapter 1 several studies have compared the effects of presenting stories in different conditions for word learning. Presenting words through reading seems less beneficial than presenting them through listening (Suggate et al., 2013) or through a combination of listening and reading (Rosenthal & Ehri, 2011), at least for monolingual children. On the other hand, studies with learners of English as a second language generally find an advantage for the presentation of material in the written form, compared to an oral only presentation (Brown et al., 2008; Neuman & Koskinen, 1992; Sydorenko, 2010).

The addition of a reading condition in the design of the second study had a dual purpose. First adding a reading condition allowed us to investigate whether the advantage of the combined listening and reading condition was due to the presence of the written text alone (Brown et al., 2008; Neuman & Koskinen, 1992; Sydorenko, 2010) or if it was due to the combined presence of both texts (Ricketts et al., 2009; Rosenthal & Ehri, 2008; 2011). In the first case, we would expect a similar performance in word learning from children in the reading and the combined groups, while, in the second, we would expect higher level of semantic learning from the combined group, compared to the reading group. In addition, the comparison between word learning in the reading and the listening groups would allow us to explore whether children in Year 4 tend to learn more semantic information from listening or from reading.

3.1.2 - Learning from context and learning from a definition

Another difference between Study 1 and the present study is the additional contrast between words presented accompanied by a definition, and words whose meaning needed to be inferred from their context alone. Children can extract meaning of new words even from a context that is not particularly informative (Nagy et al., 1987; Shatz & Baldwin, 1986), but several studies have shown a beneficial effect of presenting a definition for the new words to be learnt (Coyne et al., 2004; Dickinson, 1984; Justice et al., 2005; Wilkinson & Houston-Price, 2013). Nevertheless, it is unclear whether children extract the same kind of information from context and definitions, with some theories suggesting this is not the case (Reichle & Perfetti, 2003). For instance, if definitions are particularly useful to acquire more specific information regarding word meanings, we could expect the presence of a definition to benefit definition recognition more than category recognition. Furthermore, while the provision of more specific information facilitates semantic learning from both oral (Wilkinson & Houston-Price, 2013) and written presentations (Ricketts et al., 2011), it is unclear whether definitions have the same effect in different conditions. For instance, when presented with a written definition children might spend more time reading it, something they could not do with an oral presentation. Thus the presence of a definition might be particularly helpful in the written, and dual presentation modalities, compared to the oral modality. We could therefore expect the presence or absence of a definition to interact with the orthographic facilitation effect found for definition recognition in Study 1.

Regarding individual differences, it is known that children with different abilities may be more or less able to make effective use of context and definitions to derive word meanings (see Section 1.6). For example, children with low reading comprehension abilities seem to have difficulties in integrating information in the text, having more difficulties when information is presented further apart in the text (Cain et al., 2003; Cain et al., 2004; Ehrlich, & Remond, 1997), and this difficulty may be mediated by working memory (Yuill, Oakhill, & Parkin, 1989). For this reason, children with lower reading comprehension skills may be more negatively affected by the lack of a definition in our study, than children with higher reading comprehension abilities.

3.1.3 - The effect of story comprehension

In this study a measure of story comprehension was included, mainly to control for children's attention to the story presentation and ability to comprehend it (see Section 2.4.3.2). Even though the measure was quite simple, it can also be used to elucidate whether story comprehension mediates the effects of story presentation modality. In a study that directly measured comprehension of the given story (Cain et al., 2004) this factor did not significantly predict word learning from stories, but this effect might have been masked by the relatively high comprehension scores achieved by all the children, even less skilled comprehenders. If children show the same orthographic and phonological facilitation effects in both vocabulary acquisition and story comprehension, this would support the idea that a dual modality of presentation facilitates children's acquisition of information from the story, by freeing resources from lower level tasks, such as decoding and attention to the speech stream.

3.1.4 - The effect of executive control

A test of children's ability to shift attention between two different series was added as a further background measure in the present study. In Chapter 1, it was hypothesized that a possible source of facilitation for the combined condition was the redundancy of the information, since the text was provided both orally and written. Executive control is crucial for tasks that demand the ability to divide attention between two sources. The central executive hypothesized by Baddeley and Hitch (1974) is a system that controls attention allocation between other slave systems: the visuospatial sketch pad and the phonological loop. Participants are usually able to carry on a visual task and an oral task without experiencing interference, as long as the two tasks are simple (Cocchini, Logie, Della Sala, MacPherson, & Baddeley, 2002), while interference tends to happen with increasing processing demands from the two tasks (Brown, 1997).

In the combined condition the children need to pay attention to two sources of information at the same time, similarly to dual task situations, but the information is the same in both modalities, thus redundant. It is therefore unclear from the literature whether children are more likely to be facilitated in this condition, compared to the others, or if the task requires more processing demands (Kalyuga & Sweller, 2014; Mayer, 2014). The results of Study 1 point towards a facilitation effect for the combined condition. It is also hypothesized that the role of the executive control would be crucial in helping children to split their attention between the oral and the written presentation, and therefore that a measure of executive control will explain children's performance in the combined condition more than in other conditions.

3.1.5 – Aims and Hypotheses

As for Study 1, the principal aim of Study 2 was to explore the benefit of a dual presentation modality, compared to a single modality, either written or oral, for phonological, orthographic and semantic learning. The effect of presentation modality on story comprehension was also explored. In relation to dual versus single modality presentations, it was expected that children would acquire similar phonological knowledge in the listening and the combined groups. This prediction was in line with results from Study 1, but in contrast with results of previous research (Rosenthal & Ehri, 2008). Given the same tasks were used in Study 1 and Study 2 we assumed results would be similar between the two. It was expected that the reading group would also acquire phonological knowledge regarding the new words through recoding (Share, 1995), but this knowledge was expected to be less well developed than that obtained via direct oral presentation. It was also expected that children in the reading group would acquire orthography better than children in the listening group, who were not directly exposed to written forms. Regarding orthographic learning, we expected better performance in the combined condition than in the other two conditions, in line with previous research (Ricketts et al., 2009; Rosenthal & Ehri, 2011). It was also hypothesised that the combined condition would elicit superior semantic learning (Ricketts et al., 2009; Rosenthal & Ehri, 2008) compared to the other two conditions. On the other hand, given the conflicting evidence, it was difficult to predict whether the listening group would out-perform the reading group in learning word meanings, as suggested by the first language literature (Suggate et al., 2013), or whether the reading group would be advantaged, as suggested by the second language and the memory literature (Brown et al., 2008; Menne & Menne, 1972).

A second aim of this study was to explore whether the presence of a definition embedded in a story would facilitate children's vocabulary learning, compared to a condition where the meaning had to be inferred from context alone. It was expected that the presence of an accompanying definition would foster greater semantic learning, in all groups (Coyne et al., 2004; Dickinson, 1984; Wilkinson & Houston-Price, 2013), and that the effect of the definition would be more marked when asking children to recognise specific features of the words (i.e. sub-category and definition recognition), which are directly provided by a definition, compared when asking them to recognise the category of the words, which might be more easily extracted from the general context. Building from the results of Study 1, it was also expected that the presence of the definition would have a bigger impact on the combined group and the reading group, where the definition was provided in written form, and children were at liberty to spend more time focussing on it, than in the listening group, where the definition, given only orally, might have been given less attention.

Similarly to Study 1, a further aim of this research was to determine whether and how children's oral and written language abilities and non-verbal cognitive abilities, and additionally children's executive control abilities, impact on vocabulary learning within each condition. Finally, with this research, it was possible to explore whether children's oral and written language abilities and non-verbal abilities affected their ability to make proficient use of the definition provided to learn the meaning of new words. Regarding individual abilities, it was expected that children with higher levels of vocabulary knowledge would learn more word meanings than children with smaller vocabularies, in all presentation modalities (Robbins & Ehri, 1994; Rosenthal & Ehri, 2008). Furthermore, decoding skills were expected to predict phonological, and orthographic learning in all groups (Ricketts et al., 2011; Rosenthal & Ehri, 2008; 2011).

In addition, given the results of Study 1 (section 2.3.7), it was expected that the orthographic facilitation effect (i.e. the better performance in the combined group compared to the listening group) for orthographic learning would be greater for children with good reading comprehension abilities, while the effect for phonological learning would be greater for children with lower listening comprehension skills. Additionally, it was expected that the phonological facilitation effect (i.e. the better performance in the combined group than in the reading group) would be greater for less skilled decoders than for more skilled ones (Rosenthal & Ehri, 2011). Finally, it was expected that executive control abilities would impact on performance of the combined group more strongly than on the performance of the other groups, since executive control should help children to maintain and divide their attention between the two different presentation modalities. To our knowledge, this is the first investigation of whether monolingual children show greater word learning from listening to, reading, or both listening to and reading stories, and simultaneously investigate the effects of individual abilities on word learning in these conditions.

3.2 - Method

3.2.1 - Participants

Participants for this study were recruited amongst five Year 4 classes in four different primary schools. Seventy-one children aged 8 to 9 years participated in the study (*M* age = 9.03 years; *SD* = 0.31 years; 28 boys). All children had normal or corrected to normal vision and teachers confirmed an absence of learning or neurological disabilities. All children were native English speakers. Children were assigned to either the listening, reading or combined condition in order to form three comparable groups matched on key background measures: the three groups did not differ for any of the background measures collected at pre-test, gender or age (see Results for details). Children from each school were equally distributed across conditions. Twenty-four children were assigned to the combined condition, 24 to the listening condition and 23 to the reading condition.

3.2.2 - Design

To explore our hypotheses we used a between subject design in which all the participants were exposed to one story. The story was presented in one of three presentation modalities: 1) listening, where the children listened to the story via headphones, 2) combined modality, where the children listened to the story via headphones, and they were invited to follow it on their booklets, and 3) reading, where children read the story on their booklets silently at their own pace. Thus, as in Study 1, in one condition (listening) children were exposed only to the phonology of the new vocabulary and in another condition (combined) they were exposed to both the words' orthography and phonology. Additionally, in a third condition (reading), children were exposed to orthography only.

Children were therefore divided into three groups, each corresponding to a different presentation modality. The story was presented twice, once a week for two consecutive weeks. Each child participated in three sessions: an initial one-to-one session, where background measures were collected, and children's knowledge of our stimuli was assessed, and two sessions where children were exposed to the stories and asked to complete other tasks. The study schedule is described in Figure 3.1.

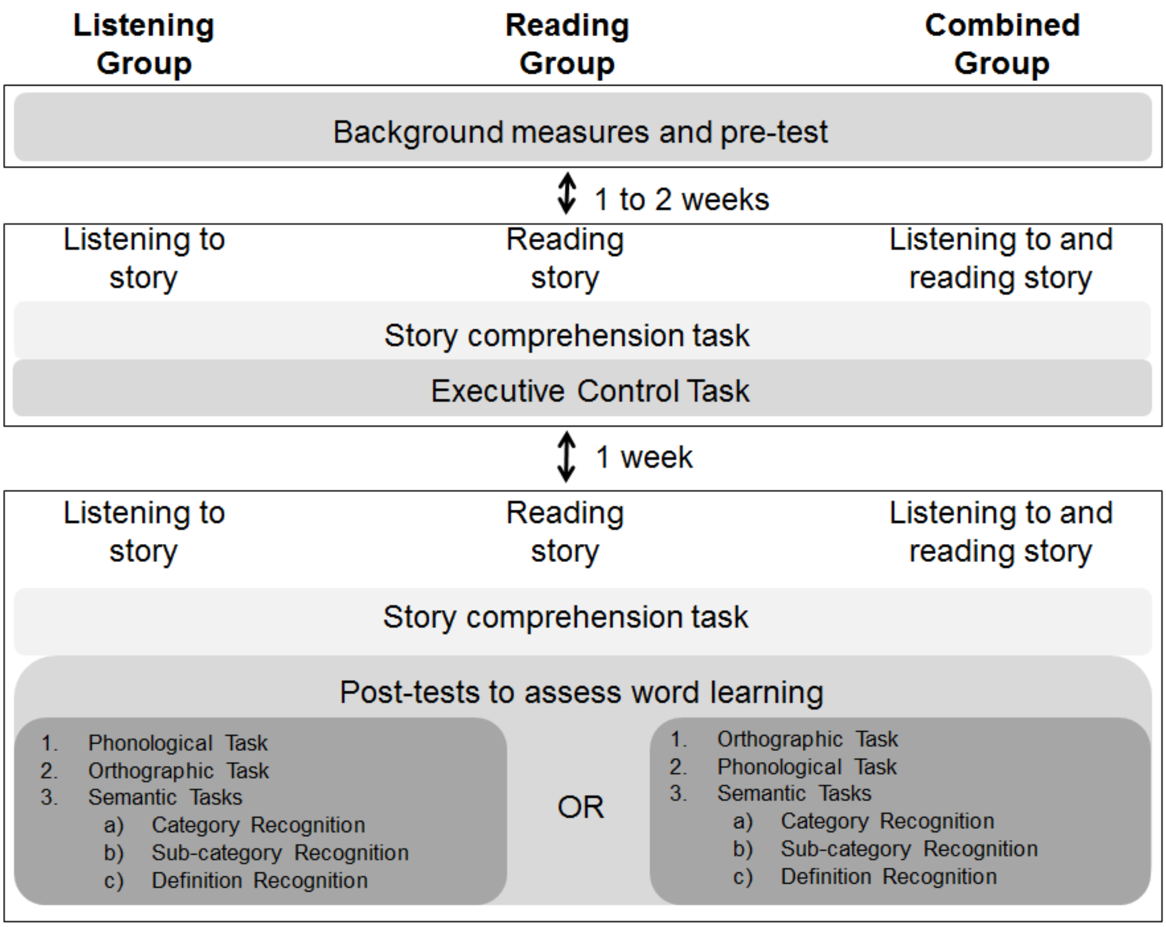


Figure 3.1. Study schedule.

3.2.3 – Experimental Stimuli

Target items were the same as those used in the Normans story in Study 1. These were eight low frequency English concrete nouns (e.g., destrier, hauberk), each belonging to one of eight categories (e.g., animal, object). Eight control words were also selected from the word lists used in Study 1. Control words were selected to match the category of each target word and to be matched on the following properties: word length, frequency (Van Heuven et al., 2014), and previous study results (adults' spelling and pronunciations and children's performance at pre-test in Study 1). Control and target words were matched in pairs, and a series of chi-square tests and Fisher Exact Probability Tests were computed on these measures to compare each target word with its control word counterpart. The results of these analyses are reported in Appendix K. Only one significant difference was highlighted (pre-test performance on *pottage* and *toddy*: $\chi^2 = 16.81$, $p < .001$, where *toddy* was less likely to be correctly categorised). To further test whether the two lists of words (target and control words) were matched a series of Mann-Whitney tests were performed on the same measures, none of which highlighted a significant difference between the two lists of words (all *p*s > .500).

To test our hypothesis regarding the additional benefit of the presence of a definition, children were presented with the definitions for only half of the words within the stories. The 8 target words were therefore divided into two lists of four items (word lists A and B) and each child was presented a story containing either definitions for list A or list B (see story versions in the next section). A list of contextual references and definitions is provided in Appendix L. The words in the two lists were paired for length, frequency (Van Heuven et al., 2014), the aforementioned results from Study 1, the length of the definition and the distance between the first and the second mention of the word in the story (Appendix M). Mann-Whitney tests revealed no difference between the lists on these measures (all *p*s > .200).

The template story used for this study was a modified version of the Normans Story used in the previous study (Appendix N). The template story was 1382 words in length, had 5% passive sentences, a Flesh Reading Ease of 84.1 and a Flesh-Kinkaid grade level of 5.3. Contextual references to the meaning of each word were included to ensure that the children could learn the meaning of the words from context alone. The following excerpts provide an example of how definitions and contextual references for the target word *pottage* were embedded in the stories (target word and definitions in bold, contextual reference underlined).

"I can offer you my famous ***pottage****. It is* ***a thick liquid food that farmers usually eat****!" replied the old man, happy that Fred and I could help him out."*

"We worked all day, and we fixed the wooden palisade for the old man. We did a very good job, but the <u>farmer's</u> ***pottage*** *was a poor reward for our work, and we woke up the next morning extremely <u>hungry</u>!"*

"The servants saw how very tired we were, and offered us a place to stay, and some <u>watery</u> ***pottage****. When the knight found out we were staying in his castle, he invited us to visit him in his chamber."*

Two different versions of the story were created, so that either the definitions of word list A or word list B were included in the story as part of the text following the first mention of the word. The versions were matched for length (1346 and 1347 words respectively). For story presentation, recordings of a female native English speaker reading the stories were created: background noise elimination and other minor editing were performed using the software Audacity (Audacity Team, 2012). Both recordings lasted 8 minutes and 35 seconds. Booklets for the presentation of the written stories were also prepared: these only included the written text of the stories, presented in seven pages, written in Calibri 14. Pilot studies established how long, on average, children took to read the stories independently, and the recordings of the adult storyteller were controlled to match this. When all the children in our study were considered, children who read independently were exposed to the story for a similar amount of time to children who listened to the story (reading group: M = 548 s; SD = 172 s); listening and combined groups: M = 507 s; SD = 2 s).

3.2.4 - Pre-test word knowledge assessment

Children's initial knowledge of target and control words was assessed using a definition production task. Children were told that they would be shown a series of words on cards, and that they would need to say everything they knew about the meaning of these words. It was also pointed out that these words were very rare words, and that it was likely that they would not know them. The children were then introduced to the task by being shown the easy word *sandwich* and by being asked a question regarding its meaning. After this practice trial the children were shown, one by one, the 8 target words and the 8 control words in a fixed, mixed order. Three additional easy words were introduced during the task to ensure children's continuous compliance with the task, and attention. Each word was read by the researcher, as well as being shown on a card.

3.2.5 - Story comprehension task

The story comprehension task assessed children's comprehension of the story. In this task children were asked to choose the right answer for 5 comprehension questions.

The task was programmed in E-Prime 2.0 (Schneider et al., 2002), and was designed as follows: at the beginning of the task the child was presented a first set of instructions that introduced the answering procedure. To choose the right answer for the question the child had to press one of four buttons, corresponding to the right answer on the screen: four buttons were highlighted in different colours on the keyboard, each corresponding to the position and colour of one of four rectangles on screen (red at top left, green at top right, blue at bottom left and yellow at bottom right). This first set of instructions lasted approximately 1 minute, and was followed by three practice trials, where the child had to choose the right answer for three easy questions on general knowledge (*What is the colour of the sun? - yellow, blue, black, green, What's the animal that makes a meow sound?-cat, swan, squirrel, horse, Where are you now? - school, coffee shop, hospital, playground,* order of alternatives randomised). The questions were presented both orally and written at the top of the screen, while the answers were presented, one at a time, orally and visually: drawings that depicted the alternatives were presented inside the 4 rectangles. Drawings were chosen over written words to minimise cognitive load. Finally a second set of instructions were provided, to tell the child that the following questions would be "more difficult questions regarding the story". The child was also instructed to guess the right answer if he/she did not know the answer to the question. The experimental task was composed of 5 trials, each corresponding to one question. An example of a trial is depicted in Figure 3.2. The questions were presented in a fixed order, but the position of the four possible answers was randomised. Children's accuracy in each question was measured. Questions and answers are described in Appendix O.

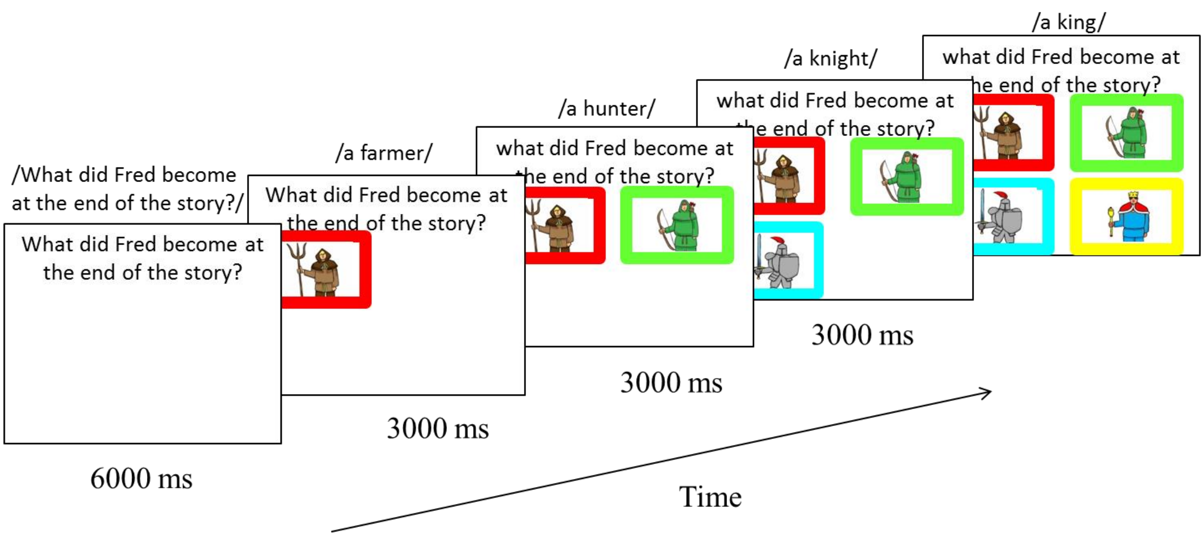


Figure 3.2. Event sequence of one of the questions in the story comprehension task.

A control group of 12 Year 4 children completed the story comprehension task to ensure answers were not easy to guess without exposure to the story; these children were recruited from 2 of the participating schools. The performance of this control sample is described in Appendix P. The performance of these children in each of the questions was compared to chance by means of a Fisher Exact Probability Tests (chance performance for each question was set at 25%, since the children could choose between four possible alternatives). Our control sample showed a performance that was significantly better than chance on Question 2 (*Where di Fred need to go to become a knight?*, p = .003), but not on other questions (all ps > .500). This question was therefore discarded from further analysis, while the remaining questions were considered a reliable measure of children's attention to and comprehension of the story.

3.2.6 - Vocabulary acquisition tasks.

For this study the same phonological and orthographic tasks used in Study 1 were used (see Chapter 2 for details). In the phonological task the children heard two possible oral forms of target and control words, while in the orthographic task the children saw two possible written forms of target and control words. In both tasks they were asked to choose the correct form between the two presented, by choosing the corresponding dinosaur. Each task was composed of 16 trials, 8 for target words, 8 for control words. Target and control words were mixed, and both the order of presentation of the words and the link between the two dinosaurs and the correct answer were randomised as in Study 1. Accuracy for each item was recorded. Each task took around 3-5 minutes to complete.

For this study the same semantic task used in Study 1 was used (see Chapter 2), with the following modification: the three steps that composed the original semantic task were separated into three separate sub-tasks, all of which were completed by all children, regardless of their performance at each step. The first semantic task (category recognition) assessed children's ability to recognise the right category of the words, between four alternatives, the second semantic task (sub-category recognition) assessed their ability to recognize the correct sub-category, between four, while the third semantic task (definition recognition) assessed recognition of the definitions of the words, between four. The three tasks were completed by the children in this fixed order. The tasks were designed in a way that ensured that no information provided from the earlier steps could help answer later steps. The children, therefore, were asked to choose the correct category for all the target and control words, before being asked to choose the correct sub-category for all the words, and then the correct definition for all the words.

As for the semantic task used in the previous study, each of the three tasks in the present study was preceded by an initial set of instructions, some practice trials, and a second set of instructions. The instructions were similar to those used for Study 1. For the first semantic task the instructions were more detailed, and children completed four practice trials, while for the second and third tasks the instructions were more concise and the children were asked to complete only 3 practice trials. See Section 2.2.4.3 for more details. During all three semantic tasks, knowledge of both target and control words was assessed. Both order of presentation of the words and the position of the four alternatives were randomised, and we measured children's accuracy in each step. Each task took on average 10 minutes to be completed.

3.2.7 - Background measures

In addition to the experimental tasks specifically designed for this study, assessments of reading, oral language, non-verbal abilities and executive control were carried out for all the participating children. As for Study 1, measures of non-verbal abilities (Raven's Coloured Progressive Matrices; Rust, 2008), vocabulary (British Picture Vocabulary Scale – 3; Dunn, Dunn, & NFER., 2009), oral language comprehension (the Understanding of Spoken Paragraphs subtest of the Clinical Evaluation of Language Fundamentals - 4; Semel, Wiig, & Secord, 2006), and reading comprehension (The York Assessment of Reading for Comprehension; Snowling et al., 2009) were collected. For reading abilities three measures were retained from the four used in Study 1: the Single Word Reading Test (Foster, 2007), the Phonemic Decoding Efficiency subtest of the Test of Word and Reading Efficiency – Second edition (PDE; Torgesen et al., 1999), and the reading accuracy measure derived from the YARC (Snowling et al., 2009).[8]

To test executive control, a paper and pencil version of the Trail Making Test was used. The Trail Making Test is widely used to discriminate brain-damaged individuals

[8] To reduce the testing battery it was decided to retain only one measure of word reading, the SWRT (Foster, 2007), dropping the Sight Word Efficiency subtest of the TOWRE (Torgesen et al., 1999).

from unimpaired individuals (Lezak, 1995; Stuss, Stethem, Hugenholtz, & Richard, 1989), and to assess executive functions in children (Ardila, Pineda, Rosselli, 2000; Lehto, Juujärvi, Kooistra, & Pulkkinen, 2003).

The original test is given in two parts: Trail Making Test, Part A (TMT-A) involves drawing a line connecting consecutive numbers from 1 to 25, while Part B (TMT-B) involves drawing a similar line, connecting alternating numbers and letters in sequence (i.e., 1-A-2-B and so on). For both parts the time taken to complete the task is recorded. In the standardized procedure the examiner should point out errors as they occur, and the time taken to correct the errors impacts on the total time taken to complete the test (Lezak, 1995; Reitan & Wolfson, 1993). While both parts require visuo-perceptual and motor abilities, Part B additionally requires the ability to shift between two series, more careful planning, and control of attention and inhibition (Wodka et al., 2008). Several authors have highlighted that an additional source of difficulty of Part B compared to Part A for children is the need to maintain and follow the alphabetical series (Närhi & Ahonen, 1995; Närhi, Rasanen, Metsapelto, & Ahonen, 1997). For this reason, before administering Part B of the Trail Making Test, we administered Part A and an additional Part, similar to Part A, but requiring children to connect with a line all the letters of the alphabet in the right order (Närhi et al. 1997). Furthermore, differently from the original administration procedure, the Trail Making Task was administered as a group task, at the end of the second session. Children were asked to be as fast as possible, being careful in not making any mistakes, and the experimenter supervised the children while they completed the task, pointing out mistakes, allowing the children to correct them. Each Part of the TMT was preceded with a practice trial, to familiarize the children with the procedure. The total time in seconds needed by each child to complete each task was recorded. A ratio score (TMT Ratio) was then obtained, by dividing the time taken to complete Part B by the mean time taken to complete the other two Parts. The Ratio score has been identified as a more accurate measure of switching cost than time taken to complete Part B on its own (Arbuthnott & Frank, 2000).

3.2.8 - Procedure

Before testing began, the study was discussed with the headteacher of the school and the teachers of the participating classes, and the consent of the headteacher was obtained. Afterwards pupils were informed about the project and they received information leaflets for themselves and information sheets for their parents, with opt-in consent forms attached (see Appendix Q for information sheets and consent forms). The first one-to-one session with the children in each school started when a reasonable number of consent forms were returned to the teachers of that specific school. Each session was carried out in a quiet room within the school.

The first session was a one-to-one session carried out by the experimenter with each participating child individually of approximately one hour. During the first session all the background measures were collected. Additionally children's pre-existing knowledge of target and control words was assessed.

Second sessions started only when all the children in each school had completed the first session. The second session was carried out in small groups of 3 to 5 children. Each group was composed of children of mixed abilities, all assigned to the same story presentation condition (combined, listening or reading).

For this session children were seated comfortably around a big table. Before starting the story presentation children were introduced to the procedure; the experimenter particularly told them that they would be asked to listen, read or listen and read a story (whichever presentation modality they were assigned to), and they would then play some games and quizzes on the computer regarding the story. Children were instructed to pay particular attention to the story, since they would be asked some questions about it. They were also asked to read or listen quietly, to avoid disturbing the other children in their group.

Children in the listening group were asked to listen carefully to the story that they would hear through the headphones (HP 530 Headset headphones: Frequency range: 20Hz-20,000Hz; Sensitivity: 105dB S.P.L at 1KHz; Rated power: 100mW) connected to a laptop, whose screen was made blank to avoid distraction. Children in the combined group also listened to the story through headphones connected to a laptop, but were also given a booklet in which to read the story at the same time. Children in the reading condition were asked to silently read the story in their booklets. They were particularly asked to read the story carefully to try to understand it, reading at their own pace. Children's reading times were recorded during the reading condition, to compare exposure time to the story between conditions, but children were not made aware of this. Children were also instructed to give back the booklets to the experimenter when the story had finished.

When all the children in the group had completed the story presentation procedure, they were asked to complete the story comprehension task. This task took around five minutes. Once this was completed, the laptops were moved to make space for the pencil and paper TMT. This task was presented as a timed game. Children were given a pencil, and each given a sheet containing the practice for the TMT Part A. For this task the children were asked to connect the numbers in the right order as fast as possible, when the experimenter said “Go!”. All the children completed the practice trial first, all starting at the same time, and then they were given the trial itself, for which the same procedure was followed. The time taken to complete the trial for each child was recorded. After the completion of the TMT Part A children were asked to complete the TMT letter part, by connecting the letters of the alphabet as fast as possible. The children were asked to say the alphabet out loud before this task, to help them remember the right sequence. As for the TMT Part A the real task was preceded by a practice trial. Finally all the children completed the TMT Part B. As with the other tasks, children were instructed to be as fast as possible, trying not to make mistakes. If children made a mistake they were asked to correct it. As for the other Parts the real task was preceded by a practice trial, all the children started the task at the same time, and the time taken to complete it was recorded. At the end of the session the children were told that they would be presented the story a second time after a week, and that they would be asked more difficult questions and details regarding it. The second session lasted around 20 minutes.

The third session was carried out a week after the second session. This session was carried out in the same groups as the second session, and children in each group were presented the story in the same modality. First, children were introduced to the procedure: they were told that they would be asked to read or listen or listen and read the same story as the previous week, and they were asked to pay particular attention to the story, because they would be asked more difficult questions about the story. The children in each group were then presented the story for the second time. After the story presentation each child completed the story comprehension task for the second time. Each child was then asked to complete the phonological and the orthographic tasks (half of the children completed the phonological task first, and half of the children the orthographic task first), then all the children were asked to complete the three steps of the semantic task in their fixed order. The tasks were completed on a laptop with headphones, and the children were allowed breaks between the tasks. At the end of this session children were given stickers as a reward. The third session lasted between 50 and 75 minutes. Since this session was quite long, the children were offered a supervised break after the completion of the first semantic task.

3.3 - Results

3.3.1 - Sample Characteristics

Groups' features are described in Table 3.1, while children's performance on the background measures is described in Table 3.2. Children's performance on most measures followed a normal distribution (exceptions were: age: $D_{combined}(24) = .188$, $p = .028$; SWRT standard score: $D_{combined}(24) = .189$, $p = .026$; YARC accuracy: $D_{combined}(24) = .190$, $p = .025$; USP CELF score: $D_{listening}(24) = .200$, $p = .014$; USP CELF Scaled Scores: $D_{combined}(24) = 0.218$, $p = .004$; $D_{listening}(24) = .220$, $p = .004$; CPM: $D_{combined}(24) = .180$, $p = .043$; $D_{listening}(24) = .234$, $p = .001$, ; TMT Ratio: $D_{reading}(23) = .304$, $p < .001$).

Differences between the three groups were examined by means of Chi-squares for gender and school, one-way ANOVAs for all the normally distributed measures and Kruskal-Wallis tests for the non-normally distributed ones. The groups had similar numbers of boys and girls, and children from the different schools (Table 3.1), and children in the three groups did not differ on any of the background measures collected at pre-test (Table 3.2). On the other hand, the groups differed in the Trail Making Test Ratio, with children in the reading group obtaining lower scores than children in the combined group ($U = 21.65$, $p_{adjusted} = .001$). Nevertheless, there were no differences between children in the listening group and those in the reading group ($U = 12.69$, $p_{adjusted} = .105$), or those in the combined group ($U = 8.96$, $p_{adjusted} = .398$). Children in the reading group were therefore faster in switching between alphabetical and numerical sequences, when their time to complete the two sequences was taken into account, compared to children in the combined group[9].

[9] Since the TMT test was administered during the second session, it was not possible to control for children's differences in this task. It must be noted that, due to the group administration of the task, it is possible that results of this task may not be highly reliable, thus we cannot exclude that the difference between groups may reflect random variance, rather than a genuine difference between children.

Table 3.1

Gender and number of children from each school by story presentation modality group

	listening group (N=24)				**reading group (N=23)**				**combined group (N=24)**				**Difference between groups**	
Schools	**A**	**B**	**C**	**D**	**A**	**B**	**C**	**D**	**A**	**B**	**C**	**D**	χ^2	**p**
Number of children	7	2	9	6	6	2	7	8	6	3	9	6	1.09	.982
Gender	**Boys**		**Girls**		**Boys**		**Girls**		**Boys**		**Girls**		χ^2	**p**
Number of children	9		15		9		14		10		14		.09	.957

Table 3.2

Performance on the background measures by story presentation modality group

Variable	listening group (N=24, 9 boys)		reading group (N=23, 9 boys)		combined group (N=24, 10 boys)		Difference between groups	
	Mean (*SD*)	**Range**	**Mean (*SD*)**	**Range**	**Mean (*SD*)**	**Range**	***F***	***p***
Age[1]	9.06 (.28)	8.58 - 9.50	8.96 (.31)	8.50 - 9.50	9.07 (.34)	8.41 - 9.50	2.05	.359
TOWRE PDE								
Raw score	34.38 (10.19)	9 - 51	36.04 (8.05)	23 - 53	35.13 (11.81)	13 - 59	0.16	.854
Standard score	101.46 (12.34)	71 - 121	104.96 (10.32)	86 - 126	103.17 (15.66)	73 - 145	0.43	.655
SWRT								
Raw score	43.63 (9.85)	17 - 57	43.87 (8.29)	23 - 55	42.83 (7.91)	27 - 54	0.09	.913
Standard score[1]	105.96 (15.22)	74 - 130	106.61 (13.89)	75 - 130	104.21 (13.69)	82 - 127	0.47	.792
YARC accuracy								
Ability score[1]	53.13 (9.13)	33 - 71	54.00 (8.47)	40 - 74	54.00 (7.81)	41 - 69	0.13	.938
Standard Score	101.29 (11.06)	82 - 123	103.00 (11.44)	82 - 128	102.17 (10.28)	87 - 127	0.14	.867
YARC comprehension								
Ability score	58.63 (6.79)	37 - 68	59.78 (6.41)	46 - 77	59.33 (5.62)	48 - 68	0.20	.817
Standard Score	104.08 (8.03)	83 - 118	105.78 (8.58)	87 - 128	104.63 (7.09)	90 - 118	0.28	.756
BPVS								
Raw score	121.96 (12.66)	94 - 144	120.87 (11.91)	94 - 146	122.58 (12.16)	94 - 143	0.11	.889
Standard score	96.71 (13.47)	79 - 119	97.74 (13.00)	70 - 122	97.46 (12.76)	70 - 115	0.04	.962
USP CELF								
Raw Score[1]	11.33 (1.63)	7 - 14	11.43 (1.95)	8 - 15	11.38 (2.34)	6 - 15	0.14	.931
Scaled Score[1]	10.25 (1.62)	7 - 13	10.09 (2.15)	7 - 14	10.38 (2.29)	6 - 14	0.48	.788
CPM								
Raw score	30.04 (3.46)	23 - 35	28.35 (3.83)	20 - 34	28.78 (3.41)	22 - 33	1.43	.246
Standard Score[1]	105.00 (13.99)	80 - 125	100.65 (15.47)	75 - 130	101.25 (12.70)	80 - 125	1.33	.515
TMT Ratio[1]	3.12 (.87)	1.55 – 5.17	2.76 (1.03)	1.55 – 6.73	3.40 (.68)	1.55 – 5.17	13.02	.001

Note. TOWRE PDE = Phonemic Decoding Efficiency subtest of the TOWRE; SWRT = Single Word Reading Test included in the YARC protocol; YARC accuracy = reading accuracy of passages collected as part of the YARC; YARC comprehension = score associated with the reading comprehension of passages collected as part of the YARC; BPVS = score for the British Picture Vocabulary Scale; USP CELF = score for the Understanding of Spoken Paragraphs subtest of the CELF; CPM = score obtained in the Raven's Coloured Progressive Matrices.

[1]The Kruskal-Wallis test is reported, due to the measure being non-normally distributed in at least one of the groups.

Spearman's rank correlation coefficients with Bonferroni correction (corrected p-value = .007) between the background measures are reported in Table 3.3. The table shows that measures of decoding skills and reading accuracy for reading both words (SWRT), nonwords (TOWRE PDE) and passages (YARC accuracy) have high correlations with each other (all *r*s > .80). Vocabulary (BPVS) showed moderate correlations with reading comprehension, non-verbal scores and single word reading while the correlation with listening comprehension did not reach significance. Scores in tests that assess comprehension of both oral and written passages also correlate moderately. Our measure of non-verbal cognitive abilities (CPM) also correlated moderately with two out of three measures of reading accuracy (specifically word reading and passage reading accuracy). On the other hand the measure of executive control did not correlate with any of the other measures.

Table 3.3

Spearman's rank correlation coefficients between the background measures

	TOWRE PDE[a]	SWRT[a]	YARC accuracy[b]	YARC comprehension[b]	BPVS[a]	USP CELF[c]	CPM[a]	TMT Ratio
TOWRE PDE[a]	-							
SWRT[a]	.84*	-						
YARC accuracy[b]	.81*	.84*	-					
YARC comp[b]	.17	.24	.29	-				
BPVS[a]	.15	.33*	.25	.46*	-			
USP CELF[c]	.00	.01	.04	.39*	.31	-		
CPM[a]	.30	.49*	.35*	.21	.39*	-.01	-	
TMT Ratio	.08	.13	.13	.26	.18	-.02	.11	-

Note. TOWRE PDE = Phonemic Decoding Efficiency subtest of the TOWRE; SWRT = Single Word Reading Test included in the YARC protocol; YARC accuracy = reading accuracy of passages collected as part of the YARC (York Assessment of Reading for Comprehension); YARC comprehension = score associated with the reading comprehension of passages collected as part of the YARC; BPVS = score for the British Picture Vocabulary Scale; USP CELF = Understanding of Spoken Paragraphs subtest of the Clinical Evaluation of Language Fundamentals; CPM = score obtained in the Raven's Coloured Progressive Matrices.

[a]raw score. [b]ability score. [c]scaled score.

*correlation is significant after Bonferroni correction.

As in Study 1 (Section 2.3.7), all the measures of reading accuracy (TOWRE PDE, SWRT and YARC accuracy) were merged to form a single factor. This seemed appropriate in view of the high correlation between the three measures. A principal component analysis (PCA) was conducted on these three measures[10]. Scores for each participant were computed using the regression method. These scores have a mean of 0.00 and a Standard Deviation of 1.00, and the range of the scores varied from a minimum score of -2.89 to a maximum score of 2.24 (combined group: M = -.07, SD = .99; listening group: M = .10, SD = 1.07; reading group: M = .00, SD = .90). This measure was normally distributed for the listening and the reading groups ($D_{Listening}(24)$ = .138, p = .121; $D_{Reading}(23)$ = .138, p = .200), but not for the combined group ($D(24)$ = .177, p = .050). The groups did not differ significantly on this measure (H (2) = .02, p = .992).

Spearman's rank correlation coefficients with Bonferroni correction between this factor and other background measures were computed. Reading accuracy showed a moderate correlation only with the CPM (r_s= .38, p = .001), and no significant correlation with other scores (YARC: r_s= .12, p = .336; BPVS: r_s= .16, p = .177; USP CELF r_s= -.02, p = .846; TMT Ratio: r_s= .03, p = .798).

3.3.2 - Children’s pre-test knowledge

Only three participants showed pre-existing knowledge of one target word (two children knew *pottage* and one knew *motte*), thus it was judged that prior knowledge of target words was negligible. Three further children showed pre-existing knowledge of one control word (two children knew *catacomb* and one knew *verandah*). Analyses that excluded these participants yielded the same pattern of results as those reported. Thus, all participants were retained in the analysis.

[10] The Kaiser-Meyer-Olkin measure verified the sampling adequacy for the analysis, KMO = .77 (“good” according to Field, 2009), and the Bartlett’s test of sphericity ($\chi^2(3)$ = 197.10 , p < .001) confirmed that this factor reduction was appropriate for our data. The factor created explained 87.77% of the variance and shows a high reliability (Cronbach’s alpha = .92).

3.3.3 - Story comprehension task

In the story comprehension task children were asked to answer 4 comprehension questions regarding the story by choosing the right answer between four alternatives[11]. The story comprehension task was administered twice to participants, once after each story exposure, giving a maximum score of 8, and a chance level performance of 2 (there was a 25% probability that children would choose the right alternative by chance, in each trial). Scores in this task are reported in Table 3.4. Children in the combined group correctly answered around 7 questions, children in the listening group correctly answered around 6 questions, and children in the reading group around 5 questions. The combined group performed best in this task: 7 children correctly answered all the questions, and the lowest score in this group was 4. On the other hand the reading group performed less well, with only 3 children correctly answering all the questions, and two correctly answering only 2 questions. Scores in this task were compared to chance performance (Table 3.4). Wherever normality assumptions were not met, the appropriate non-parametric analyses were conducted. Children in all groups performed significantly better than chance.

Table 3.4

Children's performance in the story comprehension task

	Minimum (N)	**Maximum (N)**	**Mean**	**SD**	**Chance Comparison**	
					W	***p***
listening group	2 (1)	8 (7)	6.25	1.81	276.00	< .001
reading group[1]	2 (2)	8 (3)	5.35	1.84	8.80	< .001
combined group	4 (1)	8 (7)	6.83	1.04	300.00	< .001

Note. *W* = one-sample Wilcoxon signed-rank test.
[1]one-sample t-test are reported in place of one-sample Wilcoxon signed-rank test – the distribution of the sample is normal

A Kruskal-Wallis test was computed to compare the performance of the three groups, and showed a significant difference ($H(2) = 8.36$, $p = .015$). Post-hoc analyses yielded a significant difference between the combined group and the reading group ($U = 16.63$, $p_{adjusted} = .014$), but not between the combined group and the listening group

[11] Only 4 questions were used for the analyses (see Method).

(U = 5.02, $p_{adjusted}$ > .999) or the listening group and the reading group (U = 11.61, $p_{adjusted}$ = .146). Children were therefore paying attention to the story in all three conditions and were able to comprehend it, since their performance was better than chance. On the other hand the three conditions did not yield the same levels of comprehension of the story, as children in the combined group significantly outperformed children in the reading group in answering the comprehension questions.

Spearman's rank correlation coefficients with Bonferroni correction (corrected p-value = .008) were computed between scores on the story comprehension task and scores in other tasks. Significant correlations emerged with scores in the phonological task (r_s= .34, p = .004), and with scores in the semantic task (category recognition; r_s= .35, p = .002; sub-category recognition: r_s= .38, p = .001; definition recognition: r_s= .45, p < .001), but not with scores in the orthographic task (r_s= .04, p = .714). These correlations suggest that children's comprehension of the story was linked to children's phonological and semantic learning of the words in the stories. When correlations were performed for each group separately only the correlation between comprehension and phonological knowledge was significant, for the reading group (r_s= .58, p = .004).

Spearman's rank correlation coefficients with Bonferroni correction (corrected p-value = .008) were also computed between the story comprehension task and background measures. Significant correlations emerged between story comprehension and vocabulary knowledge (r_s= .39, p = .001), and story comprehension and listening comprehension (r_s= .37, p = .001). When performance was divided by group, story comprehension significantly correlated with comprehension of spoken passages in the combined (r_s= .53, p = .008) and reading groups (r_s= .64, p = .001), while story comprehension scores in the listening group correlated significantly only with oral vocabulary (r_s= .61, p = .001) and non-verbal abilities (r_s= .53, p = .008). When the background measures that showed significant correlations with the story comprehension measure were entered in a backwards linear regression as predictors of story comprehension scores, the results were very similar: in the models computed for the combined and the reading groups only oral comprehension was retained in the final model (combined group: β = .55, p = .005; reading group: β = .65, p = .001), while in the model for the listening group only oral vocabulary scores (β = .57, p = .004) was retained as a significant predictor. These results therefore suggest that different abilities influenced story comprehension in the three groups, particularly oral comprehension for both the combined and the reading groups and oral vocabulary for the listening group.

3.3.4 - Approach to Data Analysis

For each post-test task, two sets of analyses were carried out. The first set compared the recognition of target and control words, and compared both of these to chance (50% for phonological and orthographic tasks, 25% for the semantic tasks). The second set of analyses used a mixed-effects modelling approach to explore our hypotheses relating to (1) presentation modality (listening *vs.* reading *vs.* combined groups), (2) definitions, and (3) individual differences, within each of the three experimental tasks in turn. All background measures were included. For the semantic task further analyses were performed to explore the effect of presence of a definition more in depth (section 3.3.6.3).

Since the data collected on each trial were binomial (a child could either choose a correct or incorrect alternative, obtaining a score of 1 or 0), mixed-effects models were conducted using generalised linear mixed models for binomial data (Jaeger, 2008), using the function “glmer” from the package “lme4” (Bates et al., 2014), computed with the software R (R Core Team, 2013). Each of the 71 children provided eight responses to target words on each task; this was used as the dependent variable in each analysis. For each dependent variable, an initial model included a maximal random effects structure that captured our experimental design (Barr et al., 2013). This entailed the random intercepts terms for both participants and items and the random slopes terms for participants and items that relate to our repeated measures manipulation: presence of a definition. However, models including random slopes are prone to non-convergence; therefore, given that results of these models and simpler models that only included the intercepts for item and subject did not differ, the simpler and convergent models are reported (Bates et al., 2015). We then compared these ‘empty’ models (using pair-wise Likelihood Ratio Test comparisons; Barr, et al., 2013) with models that additionally included performance on control words as a control variable and the hypothesised fixed effects: group (combined *vs.* listening *vs.* reading), presence of definition (definition present *vs.* definition absent) and all background measures. All continuous factors were centred around the mean for analysis. Hypothesised interactions were included one at a time in the model with all fixed effects, and were retained only if significant. The interactions between group and the background measures were separately introduced to test whether any background measure had a differential effect on performance in each task, depending on presentation modality.

3.3.5 - Phonological task

In the phonological task, children were asked to choose the right phonological form of the words presented in the story between two alternatives. Children were assigned a score of 1 each time they recognised the correct form of the word, and 0 each time they chose the incorrect one. A total number of 4 words recognised out of 8 was expected by chance (there was 50% probability of choosing the correct form by chance). Each child provided scores for the target words and for the control words set.

The mean number of correct words recognised in each condition in the phonological task is presented in Table 3.5. Children in both the combined and the listening groups correctly identified the phonological form for more than 6 target words, while children in the reading group identified the correct form of more than 5 target words. Children in the combined and in the listening groups correctly identified around 5 control words, while children in the reading group correctly identified around 4 control words. Kolmogorov-Smirnov tests of normality were performed, and whenever the assumption of normality was not met, the appropriate non-parametric tests were conducted. Target word performance was above chance for all the groups, while the performance for control words was above chance only for the combined group. Tests that compared performance for target and control words also showed that children in all three groups recognised phonological forms more often for target words than control words. It is also noteworthy that the combined group tended to perform better than chance for both target words and control words.

Table 3.5

Performance in the phonological post-test

Group	**Items**	**Mean (*SD*)**	**Range**	**Chance Comparison**		**Target *vs.* Control**	
				W	***p***	***T***	***p***
listening group	target words	6.42 (1.10)	4 - 8	253.00	< .001	0.00	< .001
	control words	4.58 (1.50)	1 - 7	124.00	.087		
reading group[2]	target words[1]	5.43 (1.53)	1 - 8	4.49	< .001	3.74	.001
	control words[1]	3.91 (1.41)	1 - 7	-0.30	.770		
combined group	target words	6.67 (1.69)	1 - 8	244.00	< .001	29.00	.002
	control words	5.29 (1.30)	3 - 8	165.00	< .001		

[1]one-sample t-tests are reported in place of one-sample Wilcoxon signed-rank test – the distribution of the Sample is normal (see footnote)
[2]paired t-test is reported in place of Wilcoxon signed-rank test – the distribution of both control and target words is normal (see footnote)

Estimates of fixed effects and interactions for the final models are reported in Table 3.6. The fixed factor model for the phonological task significantly improved fit compared to the empty model (χ^2 (12) = 61.16, p < .001). The final model (see Table 3.6) indicates reading accuracy and control word scores as significant predictors of performance on the phonological task, with superior learning in better readers and those better at correctly identify the phonological forms of control words. Importantly, phonological learning did not differ across groups. The interaction between reading accuracy and group was also significant, and shown in Figure 3.3. The figure suggests that reading accuracy has a similar effect on the combined and the reading group, with children with high reading accuracy learning more phonological forms, compared to children with low reading accuracy. On the other hand reading accuracy seems to have a negative effect on the performance of the listening group, with children with high reading accuracy performing less well than children with medium or low reading accuracy.

Table 3.6
Generalized Linear Mixed Model for performance on the phonological task (scores compute as binomial depending on whether the child had correctly recognised the correct alternative)

Factor	**Estimate**	**St. Error**	**z values**		χ^2	
			z value	***p***	χ^2	***p***
intercept	.88	.23	3.89	< .001		
group					2.04	.360
listening *vs*. combined	-.10	.26	-.38	.708		
reading vs. combined	-.37	.19	-1.39	.165		
reading vs. listening	-.28	.25	-1.12	.263		
definition						
def. *vs*. no def.	.10	.19	.55	.584		
reading accuracy	.46	.19	2.48	.013		
YARC comp	-.01	.12	-.06	.950		
BPVS	.03	.12	.21	.831		
USP CELF	.21	.11	1.87	.061		
CPM	.22	.11	1.85	.065		
TMT Ratio	-.10	.12	-.89	.373		
control words	.56	.11	5.17	< .001		
group*reading accuracy					6.85	.033
combined vs listening	-.61	.25	-2.48	.013		
combined vs reading	-.21	.27	-.78	.431		
listening vs reading	.40	.24	1.67	.095		

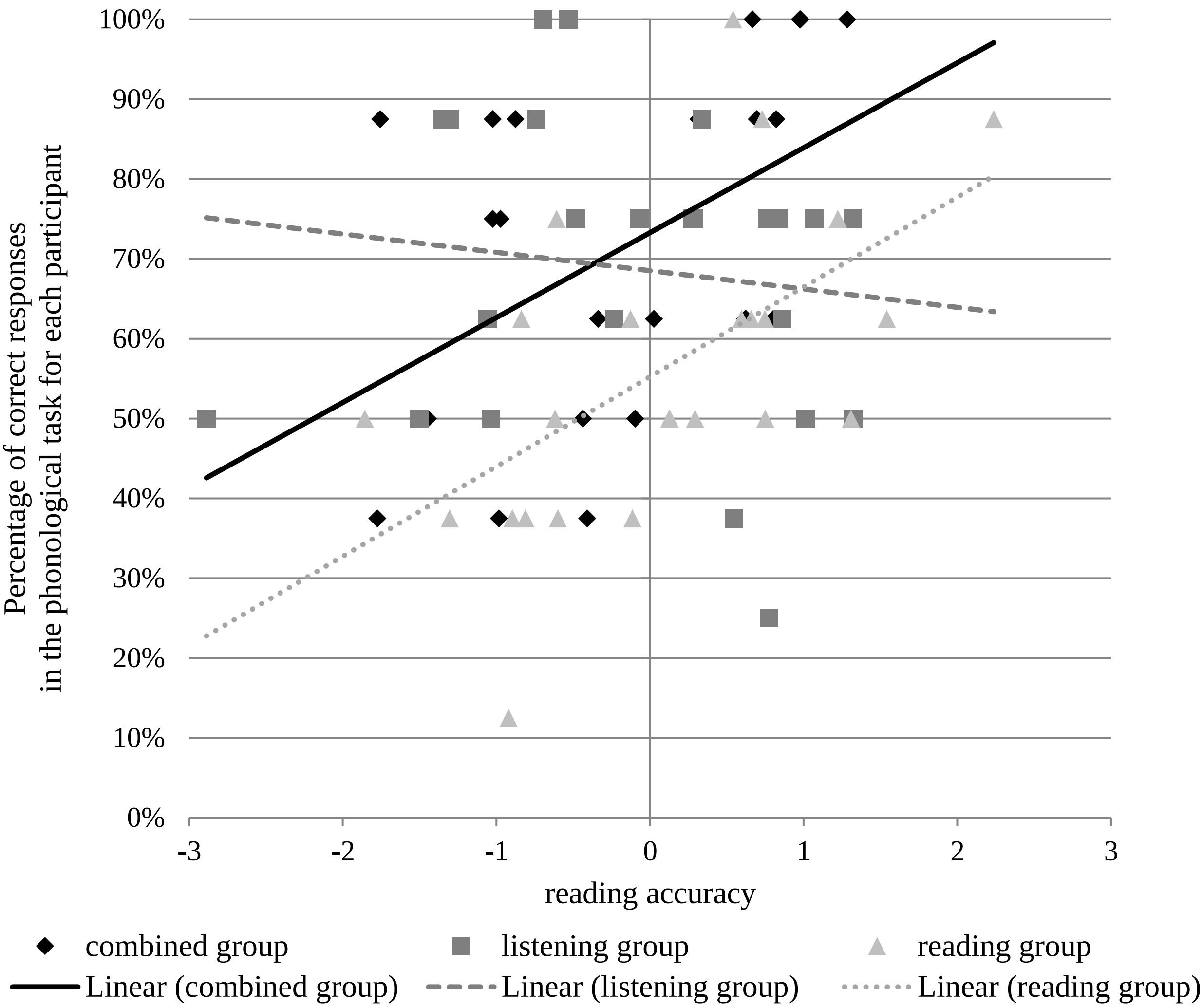


Figure 3.3. Relationship between the number of items correctly recognised in the phonological task (on average in percentage scores) and reading accuracy scores

To further explore the interaction between condition and decoding skills in predicting performance in the phonological task, two further models were run, parallel to the main model. One model included scores in the reading and the combined groups, to assess whether the performance of these groups was similarly affected by decoding abilities, and one model was computed for the listening group only, to check whether decoding skills had an effect on the performance of this group. For the combined and reading groups no significant interaction between decoding skills and condition was found (*Estimate* = -.21, z = -.75, p = .451), but the main effect of decoding skills was significant (*Estimate* = .49, z = 2.61, p = .009), showing that decoding skills determined performance in the phonological task for both the reading and the combined groups, and for both it determined performance similarly. For the listening group, decoding skills also predicted performance, but negatively (*Estimate* = -.47, z = -2.03, p = .043).

These results confirm the lack of a main effect of condition on phonological learning, and highlight reading accuracy and performance on control words as good predictors of scores in this task. However, the impact of reading accuracy depended on condition, with good readers in the reading and combined groups learning more phonological forms than their less skilled peers. Good readers, on the other hand, performed less well than their peers in phonological learning in the listening group. These results are similar to those of Study 1 in highlighting an effect of reading accuracy on performance, although, in this study, this result seems mostly driven by the reading and the combined condition, with the listening condition showing a reverse of the effect. Unlike in Study 1, listening comprehension did not explain phonological learning in the present study.

3.3.6 - Orthographic task

In the orthographic task children were asked to choose the right orthographic form of the words presented in the story between two alternatives. Children were assigned a score of 1 each time they recognised the correct form of the word, and 0 each time they chose the incorrect one. A total number of 4 words out of 8 was expected by chance (50% chance level). Each child provided scores for the target words and for the control words.

The mean number of correct words recognised in each condition in the orthographic task is presented in Table 3.7. Children in the combined and the reading groups correctly identified the orthographic form of around 6 target words, while children in the listening group identified the correct form of more than 4 target words. Children in all three groups identified the correct orthographic form of around 4 control words. Kolmogorov-Smirnov tests of normality were performed, and whenever the assumption of normality was not met, the appropriate non-parametric tests were conducted. Target word performance was higher than chance in the reading and the combined groups, but not in the listening group, while the performance for target words was higher than chance for none of the groups. Tests that compared performance for target and control words also showed that children in the combined and reading groups learned the orthographic forms of at least some of the words in the story. On the other hand children in the listening group performed similarly for target and control words.

Table 3.7

Performance in the orthographic post-test

Group	Items	Mean (*SD*)	Range	Chance Comparison		Target *vs.* Control	
				W	***p***	***T***	***p***
listening group	target words	4.54 (1.44)	2 - 8	137.50	.076	100.50	.598
	control words[1]	4.37 (1.31)	2 - 7	1.40	.175		
reading group	target words[1]	6.00 (1.38)	3 - 8	6.94	< .001	6.00	< .001
	control words	4.39 (1.16)	2 - 6	85.00	.143		
combined group	target words[1]	6.21 (1.35)	3 - 8	8.01	< .001	3.00	< .001
	control words	4.13 (1.33)	1 - 6	97.00	.603		

[1]one-sample t-tests are reported in place of one-sample Wilcoxon signed-rank test – the distribution of the Sample is normal

Estimates of fixed effects and interactions for the final models are reported in Table 3.8. The final model for the orthographic task significantly improved fit compared to the empty model (χ^2 (12) = 40.23, p < .001). The final model indicates group and reading accuracy as significant predictors of orthographic performance, with children in the combined and reading groups showing better performance than children in the listening group (who did not show significant orthographic learning), and greater reading accuracy associated with higher performance. The model also highlighted performance in the executive function task as a predictor of orthographic learning, especially in interaction with group. From Figure 3.4 it would seem that poorer executive functions (higher scores) are associated with a better performance in the combined group, but not in the other two groups.

Table 3.8

Generalized Linear Mixed Model for performance on the orthographic task (scores compute as binomial depending on whether the child had correctly recognised the correct alternative)

Factor	Estimate	St. Error	z values		χ^2	
			z value	***p***	χ^2	***p***
intercept	1.24	.30	4.20	< .001		
group					16.98	< .001
listening *vs*. combined	-.81	.25	-3.28	.001		
reading vs. combined	-.08	.29	-.28	.780		
reading vs. listening	.89	.27	3.27	.001		
definition						
def. *vs*. no def.	-.07	.20	-.40	.687		
reading accuracy	.40	.12	3.38	.001		
YARC comp	.04	.13	.31	.754		
BPVS	.06	.13	.44	.660		
USP CELF	-.01	.12	-.11	.911		
CPM	-.12	.12	-1.00	.319		
TMT Ratio	.61	.26	2.34	.019		
control words	-.20	.11	-1.88	.061		
group*TMT Ratio					6.30	.043
combined vs listening	-.77	.32	-2.46	.014		
combined vs reading	-.39	.35	-1.10	.271		
listening vs reading	.38	.30	1.30	.193		

To further explore the relationship between executive functions and performance in the orthographic task, two further models were run, parallel to the main model. One model included scores in the reading and the listening groups, to confirm that the performance of these groups was not affected by executive functions, and one model was computed for the combined group only, to confirm the effect of executive functions in this group. For the listening and reading groups no significant interaction between executive functions and condition was found (β = .40, z = 1.30, p = .191), and the main effect of this factor was not significant (β = -.14, z = -.74, p = .459), showing that the ability to switch between sets and maintain the sets in working memory did not affect performance in the listening and reading groups. For the combined group, on the other hand this ability predicted children's performance in the orthographic task (β = .70, z = 2.51, p = .012), but in the opposite direction than expected (higher ratio scores imply children were hindered in completing the dual task compared to the numbers or letters only tasks).

Similarly to the results of Study 1, the present results confirm that only children who were exposed to orthography (combined and reading groups) learnt the orthographic forms of the words, and in all groups children with better reading accuracy skills performed better in this task than less skilled readers and decoders. In the present study the ability to switch between sets (alphabetical and numerical sets) and maintain both sets in memory predicted performance in the combined group, although not in the expected direction: the more children struggled with the executive functions task, the better they recognised the new words' orthography. On the other hand, differently from Study 1, reading comprehension did not predict the performance of the combined group.

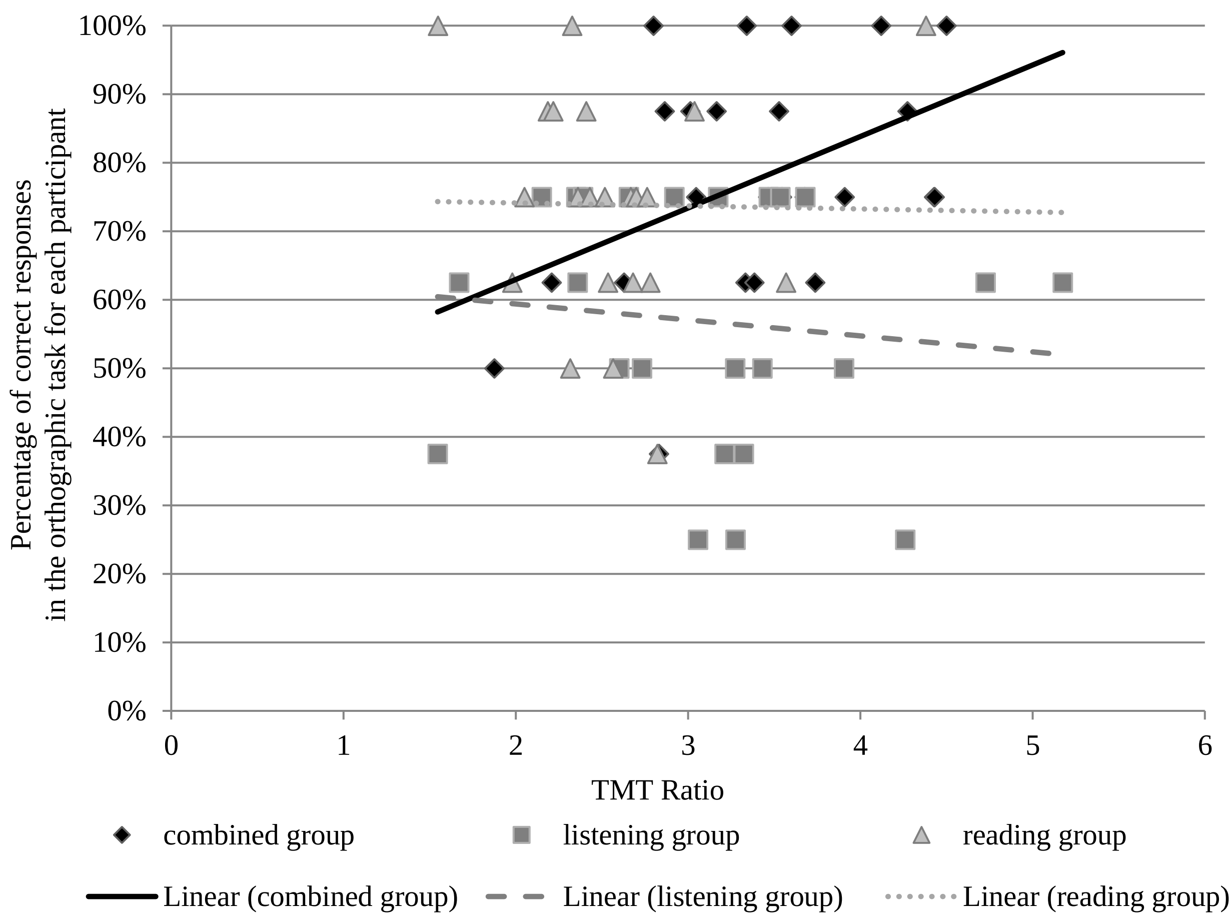


Figure 3.4. Relationship between the number of items correctly recognised in the orthographic task (on average in percentage scores) and ratio scores in the TMT

3.3.7 - Semantic tasks

The semantic task was composed of three parts, administered to the children in a fixed order. In the first task children had to choose the right category for the words between four alternatives, in the second task children had to choose the right sub-category for the words between four alternatives, and in the third task children were asked to choose the right definition for the words, between four very similar ones. The mean number of words correctly recognised by each group in each semantic task is presented in Table 3.9, along with the results of analyses comparing target and control word performance to each other and to chance. Children in the combined group recognised the right category for nearly 5 target words, while children in the other two groups recognised the right category of nearly 4 target words. All the groups recognised the right sub-category for more than 4 target words. The combined group recognised the right definition for 5 target words, and the listening and the reading groups recognised the right definition for more than 4 words. In all three tasks all the groups correctly categorised around 2 control words. Kolmogorov-Smirnov tests of normality were performed, and whenever the assumption of normality was not met, the appropriate non-parametric tests were conducted. All groups performed significantly above chance for target words and significantly better for target words than control words in all semantic tasks. Control word performance was not significantly above chance in any of the tasks for any of the groups.

Table 3.9

Performance in the semantic post-tests – number of words correctly recognised by the three groups

		Category Recognition	**Chance Comparison**		**Target *vs.* Control**		**Sub-category Recognition**	**Chance Comparison**		**Target *vs.* Control**		**Definition Recognition**	**Chance Comparison**		**Target *vs.* Control**	
		Mean (SD)	***W***	***p***	***T***	***p***	**Mean (*SD*)**	***W***	***p***	***T***	***p***	**Mean (*SD*)**	***W***	***p***	***T***	***p***
listening group	target words	3.67 (1.69)	4.85[1]	< .001	25.50	.003	4.63 (1.56)	210.00	< .001	10.50	< .001	4.50 (1.77)	267.00	< .001	10.00	< .001
	control words	2.42 (1.25)	130.00	.129			2.54 (1.22)	116.00	.054			1.75 (1.26)	58.50	.379		
reading group	target words	3.65 (1.37)	203.00	< .001	13.50	< .001	4.35 (1.43)	7.85[1]	< .001	0.00	< .001	4.13 (1.89)	206.50	< .001	0.00	< .001
	control words	1.70 (1.22)	55.50	.302			2.13 (1.10)	98.00	.559			1.96 (0.93)	42.50	.552		
combined group	target words	4.71 (1.60)	8.29[1]	< .001	7.50	< .001	4.96 (1.57)	9.21[1]	< .001	15.00	< .001	5.00 (1.47)	9.97[1]	< .001	0.00	< .001
	control words	1.92 (1.21)	67.00	.635			2.42 (1.10)	89.00	.084			1.83 (1.09)	55.00	.479		

[1]one-sample t-test are reported in place of Wilcoxon signed-rank test

Estimates of fixed effects and interactions for the final models are reported in Table 3.10. When considering category recognition, the fixed factor model significantly improved fit compared to the empty model ($\chi^2(10) = 57.28$, $p < .001$). Interaction terms did not significantly improve model fit. Group, presence of definition, reading accuracy and vocabulary were significant predictors in the final model. The combined group performed better than the other two groups, and greater reading accuracy and oral vocabulary knowledge was associated with better performance. Category recognition was better for words presented without a definition than with a definition.

When considering sub-category recognition, the fixed factor model significantly improved fit compared to the empty model (χ^2 $(10) = 31.26$, $p < .001$). Interaction terms did not significantly improve model fit. Presence of definition and non-verbal abilities significantly predicted performance, with the presence of a definition and greater non-verbal abilities associated with better performance. No group effect was highlighted in the results: the three groups performed similarly on this task.

When considering definition recognition, the fixed factor model significantly improved fit compared to the empty model (χ^2 $(10) = 34.63$, $p < .001$). The presence of a definition and oral vocabulary knowledge were significant predictors but group was not, on its own. However, addition of the group by reading comprehension interaction also improved model fit. The final model significantly improved fit compared to the empty model (χ^2 $(12) = 44.15$, $p < .001$).

Table 3.10

Generalized Linear Mixed Model for performance in the semantic post-tests

Factor	Estimate	St. Error	z values		χ^2	
			z value	*p*	χ^2	*p*
Category Recognition						
intercept	.68	.28	2.45	.014		
group					9.97	.004
listening *vs.* combined	-.73	.23	-3.13	.002		
reading vs. combined	-.58	.24	-2.37	.018		
reading vs. listening	.15	.24	.63	.527		
definition						
def. *vs.* no def.	-.42	.18	-2.29	.022		
reading accuracy	.29	.11	2.74	.006		
YARC comp	.06	.12	.50	.619		
BPVS	.35	.12	2.98	.003		
USP CELF	.06	.11	.57	.572		
CPM	.12	.11	1.17	.241		
TMT Ratio	-.05	.11	-.05	.608		
control words	.18	.09	1.98	.054		
Sub-category Recognition						
intercept	.41	.36	1.14	.253		
group					2.21	.332
listening *vs.* combined	-.32	.24	-1.36	.173		
reading vs. combined	-.29	.25	-1.20	.231		
reading vs. listening	.03	.24	.11	.913		
definition						
def. *vs.* no def.	.46	.19	2.42	.015		
reading accuracy	.01	.11	.14	.886		
YARC comp	.05	.12	.42	.676		
BPVS	.13	.12	1.05	.293		
USP CELF	.12	.11	1.08	.279		
CPM	.37	.12	3.15	.002		
TMT Ratio	.01	.11	.05	.956		
control words	.01	.11	.14	.890		
Definition Recognition						
intercept	.34	.26	1.33	.185		
group					4.11	.128
listening *vs.* combined	-.21	.23	-.88	.377		
reading vs. combined	-.45	.24	-1.83	.067		
reading vs. listening	-.24	.24	-.99	.324		
definition						
def. *vs.* no def.	.45	.18	2.44	.015		
reading accuracy	.13	.11	1.20	.228		
YARC comp	-.03	.21	-.14	.891		
BPVS	.32	.12	2.66	.008		
USP CELF	.19	.11	1.75	.079		
CPM	-.03	.11	-.26	.794		
TMT Ratio	-.01	.12	-.11	.910		
control words	.17	.10	1.66	.097		
group*YARC comprehension					9.52	.009
listening vs combined	.47	.26	1.81	.071		
reading vs combined	-.28	.26	-1.06	.290		
reading vs listening	-.75	.25	-3.00	.002		

Figure 3.5 illustrates the relationship between reading comprehension and group in predicting definition recognition performance. To explore the group by reading comprehension interaction further, a separate model was computed for each group. When considering the full models for each group, the presence of a definition predicted performance of the combined group (β = .99, z = 3.05, p = .002). The performance of the listening group was positively influenced both by vocabulary (β = .60, z = 2.58, p = .010) and reading comprehension (β = .51, z = 2.27, p = .023), while the performance of the reading group was influenced negatively by reading comprehension (β = -.63, z = -2.40, p = .016), and positively by listening comprehension (β = .65, z = 2.69, p = .007), and performance on control words (β = .96, z = 3.81, p < .001).These supplementary analyses suggest that the group by reading comprehension interaction reflects a positive association between reading comprehension and definition recognition for the listening group, but a negative association for the reading group. In the reading group, children with better reading comprehension performed less well on the definition recognition task than children with lower reading comprehension scores. In addition, the presence of definitions may have supported later definition recognition more for children in the combined group, and existing vocabulary knowledge particularly influenced performance in the listening group. However, the group by vocabulary interaction was not significant, so this finding will not be interpreted further here, but may warrant further attention in future studies.

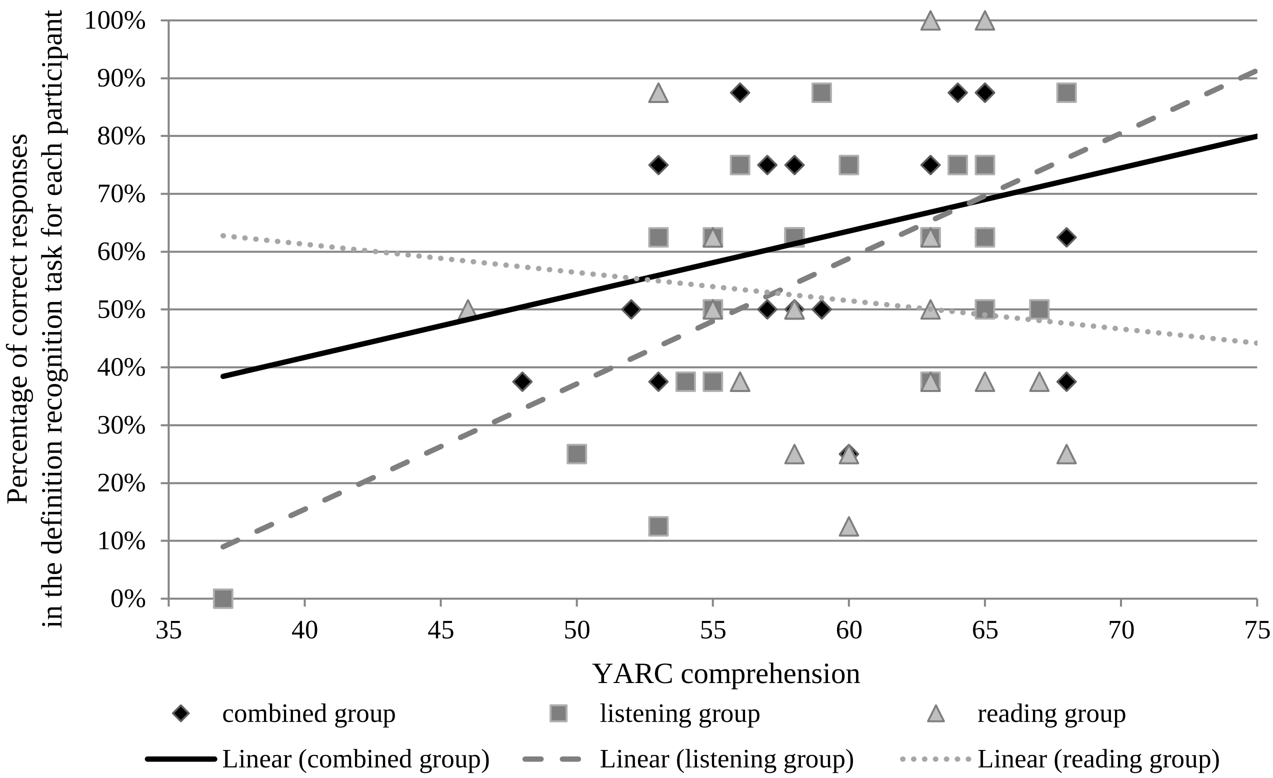


Figure 3.5. Relationship between the number of items correctly recognised in the definition recognition task (on average in percentage scores) and reading comprehension.

3.3.7.1 - Effect of definitions. The previous results suggest that the presence of a definition hindered category recognition, but facilitated the recognition of sub-categories and definitions. Further analyses were carried out to explore the relationship between presence of definition and other variables. In each version of the template story 4 words were presented accompanied by a definition and 4 words were presented without a definition. The number of words correctly categorised by the children in each semantic task depending on presence of definition is presented in Table 3.11. Wilcoxon signed-rank tests were carried out, since measures were not normally distributed. Children choose the correct category more often for words presented without definition than with definition (this difference was significant for the reading group, but not for the other groups), while children chose the correct sub-category and definition for words presented with definitions more often than for words presented without definitions, although only significantly so in the definition recognition task for the combined group.

To further explore these results an interaction effect between group and presence of definition was included in the fixed factor models for the three semantic tasks, but none of these interactions were significant (category recognition: χ^2 *(2)* = 2.75, *p* = .253; sub-category recognition: χ^2 *(2)* = .18, *p* = .913; definition recognition: χ^2 *(2)* = 4.73, *p* = .094), suggesting that the overall negative effects of definitions on category recognition, and positive effects on sub-category and definition recognition were similar between the three groups.

The effect of presence of a definition was further explored to analyse whether children of different abilities were influenced differently by the presence of a definition. To explore this effect, interactions between presence of a definition and all background measures were considered in turn within the models for the three semantic tasks. No interaction was significant (see Appendix R), suggesting that all children were similarly affected by the presence or absence of a definition in the semantic tasks.

Table 3.11

Children's performance in the semantic task divided by presence of definition

		Category Recognition	**Target *vs.* Control**		**Sub-category Recognition**	**Target *vs.* Control**		**Definition Recognition**	**Target *vs.* Control**	
		Mean (SD)	***T***	***p***	**Mean (SD)**	***T***	***p***	**Mean (SD)**	***T***	***p***
listening group	definition	1.67 (1.05)	53.50	.263	2.54 (1.02)	43.00	.104	2.33 (.96)	52.50	.384
	no definition	2.00 (1.14)			2.08 (1.06)			2.17 (1.05)		
reading group	definition	1.48 (.85)	18.00	.007	2.35 (1.03)	43.50	.187	2.33 (1.10)	41.50	.471
	no definition	2.17 (.89)			2.00 (.85)			2.17 (1.13)		
combined group	definition	2.33 (1.01)	75.50	.960	2.63 (1.10)	57.00	.199	2.92 (.97)	28.00	.005
	no definition	2.38 (.97)			2.33 (.92)			2.08 (.93)		

3.3.8 - Relationship between the tasks

Since scores in the tasks were not normally distributed (see Sections 3.3.4.1, 3.3.5.1 and 3.3.6.1), the relationship between performance in the three experimental tasks was explored by means of Spearman's rank correlation coefficients, (Table 3.12). A Bonferroni correction was applied (corrected critical p-value = .01).

Table 3.12

Spearman's rank correlation coefficients between the scores obtained in the phonological, the orthographic and the semantic tasks (number of words correctly recognised) in each group

		Phonology	Orthography	Category	Sub-category	Definition
listening group	Phonology	-				
	Orthography	.23	-			
	Category	-.16	.11	-		
	Sub-category	.41	.17	.38	-	
	Definition	.23	-.14	.41	.54*	-
reading group	Phonology	-				
	Orthography	-.05	-			
	Category	-.20	.23	-		
	Sub-category	.05	.47	.33	-	
	Definition	.25	.22	.38	.39	-
combined group	Phonology	-				
	Orthography	.38	-			
	Category	.55*	.28	-		
	Sub-category	.69*	.20	.40	-	
	Definition	.39	.03	.60*	.46	-

*correlation is significant ($p < .01$) after Bonferroni correction.

In the listening group the only significant association was between performance in the sub-category and definition recognition tasks. Similarly to the results of Study 1, the association between performance in the phonological and semantic tasks was higher than association between the orthographic and semantic tasks, but no correlation reached significance. For the reading group, the measures showed no significant association. For the combined group, the association between category and sub-category recognition and performance on the phonological task reached significance, and performance in the category recognition and definition recognition tasks were significantly correlated. Overall, it could be concluded that performance in the semantic tasks showed a trend towards a significant positive association in all groups. Furthermore, children with better semantic learning in the combined group showed higher phonological learning, and this relationship was significant.

3.4 - Discussion

3.4.1 - Story presentation modality effects

3.4.1.1 - Phonological and Orthographic learning. For the phonological task it was predicted that children would acquire new phonological forms better when exposed to the oral form of the words (i.e. in the combined and listening condition), compared to the reading condition. Furthermore, following the results of Study 1 (see Chapter 2), it was expected that children would acquire new phonological forms equally well in the listening and combined groups, in contrast with previous research (Ricketts et al., 2009; Rosenthal & Ehri, 2008). Children learnt the phonological forms of the words quite proficiently, and, similarly to the results of Study 1, no advantage for the combined group over the listening group was found. In contrast with the hypothesis of a phonological advantage for phonological learning, the results of the mixed models also failed to highlight any difference between the three conditions: once performance for the control words was taken into account, the phonological facilitation effect was not significant, and the three groups showed similar performances. In conclusion, no phonological advantage for phonological learning can be inferred by our results: children who had only been exposed to the written form of the words (reading group) were as likely as children who had been exposed to the oral form of the words to recognise the right phonological form of the words in the story. This result demonstrates that children in the reading group were able to use orthography-phonology conversion rules to learn the phonological form of words they only saw written, thus supporting the self-teaching hypothesis proposed by Share (1995), who considers phonological recoding an essential part of word learning through reading.

Can we be certain that the phonological task provided a robust measure of phonological learning in this study? Three issues are worthy of mention. Performance on the task involved distinguishing between targets and plausible foils. Foils were generated, where possible, from adult mispronunciations of the written forms (see Section 2.2). Consequently, it is possible that the task probed abilities other than children's in-task learning, such as their general sensitivity to word-likeness. This idea is supported by the finding that control word scores significantly predicted performance on this task. To account for such general effects, scores on control word trials were included in all models. Thus, we are confident that our results demonstrate phonological learning in all conditions. A second consideration is that the use of plausible foils may have made the task particularly challenging for the reading group. Children in this group were not directly provided with the phonological forms of the new words, and may therefore have generated an alternative, which aligned more closely with the foil than the target, making it harder for them to reject the incorrect alternative. However, this does not seem to be the case. Instead, they performed as well as the other groups. Finally, it is possible that our task was not sufficiently sensitive to detect subtle differences in phonological representations between the groups. We used a recognition task. Had we used a production task (cf. Rosenthal & Ehri, 2008) it is possible that we might have detected such group differences. Future research could build on this study by including alternative measures of phonological learning.

Similarly to the prediction for the phonological task, for the orthographic task it was predicted that children would learn orthographic forms only when exposed to them i.e. in the combined and the reading groups. This hypothesis was supported by the results of the previous study, where children only learned orthographic forms in the combined condition, and further confirmed by the present results, where only children in the combined and reading groups learnt the orthographic forms of the words. The children also learned orthographic forms equally well in these two conditions. This result, therefore, suggests that the orthographic form of a word was not easily and automatically extrapolated from oral presentation, and only the direct presentation of written text prompted orthographic learning. Furthermore, the presentation of the oral form alongside the written form did not enhance orthographic learning compared to a written only presentation, in contrast to the results of previous research (Rosenthal & Ehri, 2011).

Differently from the phonological form of the words, which is actively learnt even when not presented, children did not use their knowledge of phonology-orthography conversion rules to learn the orthographic form of the words when not presented with them. Extrapolating the written form of words from an oral text is not as automatic as extrapolating the phonological form from a written text, at least for children. This suggests two linked conclusions. The first is that learning the phonological form of a new word is more important than learning its orthographic form for the creation of a stable representation of the new word in memory. This would explain why children automatically recode written forms to phonological forms to learn them, but not the reverse. The first learning mechanism for children is through oral language: when they start learning words in the first year of their life they learn oral information not accompanied by written information. They first build an oral vocabulary, and only subsequently a written vocabulary. Likewise, in the evolution of language, oral communication came before its written counterpart. It therefore seems reasonable that children might rely more on phonological representations of new words, than orthographic representations. Indeed, if a phonological representation is an essential component of a word's representation, word learning may not be possible unless a phonological representation is constructed when this is not provided directly. Second, since children tend to acquire both forms of the words when exposed to the written form only, but they only acquire the phonological form when exposed to the phonological form only, it appears that learning new words from reading leads to a more complete representation of the words, that includes both phonological and orthographic forms. Conversely, learning new words from listening only expands oral vocabulary knowledge, with little impact on written vocabulary. What is important for a stable representation of words in memory is, in fact, how representations are stored, and not how they were presented in the first place. Children that are successful in deriving phonological form from written forms will therefore be advantaged by a written presentation, despite the primacy of oral language over written language. This, however, might be dependent on the specific features of the language explored: it is, for example, possible that an oral presentation might prompt orthographic learning in a language with more transparent phono-orthographic correspondencies than English.

3.4.1.2 - Semantic learning. For the semantic task, it was expected that children would acquire more semantic information when presented with the oral and written forms of the words combined, than when presented with only one form of the words (Ricketts et al., 2009; Rosenthal & Ehri, 2008; 2011). Specifically, following the results of Study 1, it was expected that the combined condition would elicit deeper semantic learning. Children in the combined group were expected to learn more information about the word meanings than children in the listening group. Furthermore, we expected that semantic learning might be greater in the reading than listening group (Brown, et al., 2008) or vice versa (Suggate et al., 2013).

As expected, a difference between the combined and the listening groups emerged from this study, with children in the combined group outperforming children in the listening group. Nevertheless, unlike in Study 1, this difference was found for category recognition, rather than definition recognition. Similarly, the combined group outperformed the reading group in category recognition. No difference in semantic learning was found between the reading and the listening groups. Our study, therefore, confirmed the presence of an orthographic facilitation effect (i.e. better performance of the combined group over the listening group in semantic learning), albeit not on the expected measures, and the presence of a phonological facilitation effect (i.e. a difference between the combined group and the reading group), and no superiority of the oral medium over the written medium for vocabulary learning. This final result leads us to suggest that, by Year 4, children can learn as much semantic information from reading as from listening.

3.4.1.3 – Explaining the advantage of the combined condition. Several possible explanations for the advantage of a combined presentation over a single presentation have been proposed in the previous chapters. The hypothesis that the advantage of a dual presentation over the oral presentation was due to the presence of the written text only (Brown et al., 2008), can be now discarded, as children in the combined condition outperformed children in both single modality condition, and the reading group did not outperform the listening group.

It was also hypothesised that children would be able to form a representation of higher quality when presented with orthographic and phonological information, and the lack of one representational form of the words in memory would negatively impact either encoding or retrieval of semantic information. The Lexical Quality Hypothesis (Perfetti & Hart, 2002) suggests that words with higher quality representations are more easily retrieved from memory. This theoretical perspective focuses on existing lexical representations rather than their acquisition. Nonetheless, it is consistent with the proposal that the combined condition promotes building of better-specified representations, facilitating access to stored word knowledge at test. This framework would appear to predict consistent differences in the quality of the phonological, orthographic, and semantic representations formed in the combined and the single modality presentations, as previously reported in studies finding orthographic (Ricketts et al., 2009; Rosenthal & Ehri, 2008) and phonological facilitation effects (Rosenthal & Ehri, 2011). Such consistent effects across tasks were not found in the current study. As we have seen, unexpectedly, both the combined and the reading groups built a double representation for the new words, one group because it was directly provided both forms, and the other because children derived one form (phonology) from the given one (orthography). We could therefore claim that both children in the combined and the reading group had a more complete representation of the words than the listening group. Considering the framework of the Lexical Quality Hypothesis, therefore, both groups should outperform the listening group, whose representation was less well developed. Nevertheless, only the combined group performed significantly better than the listening group in the semantic task. On the other hand, it is possible that, even though both the combined and the reading groups formed phonological and orthographic representations of the words, the phonological representations of the reading group were less well developed, since they were built from orthography, and not directly provided. Although the current study provides no evidence that this was the case, the relative insensitivity of our phonological and orthographic measures might have masked subtle differences between the groups.

An alternative possibility for the advantage of a combined presentation over a single presentation is that the combined condition reduced cognitive load during word learning, freeing resources for comprehension and word meaning extraction (Mayer et al., 1999). Compared to the combined group, the reading and listening groups were charged with additional processing demands at the point of encountering the new words – the reading group in the form of spontaneous phonological recoding (as evidenced by their performance on the phonological task) and the listening group due to the attentional demands associated with continuously monitoring the oral story presentation (without any 'back-up' support from the written text). Children in the combined group may, therefore, have had more resources available to allocate to processing the contextual support (including definitions) that immediately followed a new word, thus encoding the word meanings better. This perspective would predict word form representations to be of similar quality across conditions, but representations of word meanings to be better in the combined condition. This hypothesis is consistent with the pattern found in our data: children in the combined group performed better than children in the listening group in word learning, and they performed better than children in the reading group in both story comprehension and in word learning. It could be argued that children presented with only one form of the words had to allocate their resources to low level processes, such as phonological recoding in the reading condition, and comprehension in the listening condition, so that fewer resources were available for the children to integrate the new words in memory, and abstract information about their meaning.

Children in the combined group seem therefore to have had more resources available to allocate to comprehending the story and learning the new word meanings. It is less clear whether this is due to the general lower demands of the task, or because children formed a better representation of the words in this task, and therefore had more resources left for more complex learning. The fact that a different performance between presentation modalities emerges both in comprehension (reading *vs.* combined), and vocabulary acquisition (reading and listening *vs.* combined) might point towards the general account of resource availability. Nevertheless it cannot be excluded that children had better representation in the combined condition, not revealed by the phonological and orthographic tasks used.

Although both of the previous approaches could, in theory, explain the greater semantic learning of the combined group, it remains to be discussed why this effect was seen only in one of our semantic tasks, the category recognition task. This is in direct contrast with the result of Study 1, where the difference was seen when all three steps were combined, but not when considering category recognition alone. Two potential explanations for the results occur to us. First, only the category recognition task required children to abstract category-level knowledge about the target words from the information provided in the story; the category label for each new word was never directly provided in the narrative. In contrast, recognising the correct sub-category or definition in the other semantic tasks required participants to choose a sub-category or definition that was very similar to those provided in the story. Second, choosing the correct response on the sub-category and definition recognition sub-tests is likely to have been easier than choosing the correct form in the category recognition task. In the former tasks, only one of the four alternative response options had been encountered in the story (e.g., among the 4 clothing options offered in the definition task for the word 'hauberk', only *a soldier's shirt made of chain mail* had been mentioned in the story, making the other alternatives less likely, regardless of any learning of the label for this item). In contrast, all four of the alternatives in the category recognition task correctly described one of the new words presented in the story (i.e., to recognise *hauberk* as a piece of *clothing* in the category recognition task, children had to know that this new word did not identify a new *animal*, *job*, or *part of a house*, each of which had been encountered in the story). These factors could have made the category recognition task the most challenging, and therefore the most sensitive measure of children's semantic learning. By freeing resources, the combined condition might have especially facilitated this task due to the level of abstraction and/or precision of the mapping required, a hypothesis that warrants further investigation.

Differences between the results of Study 1 and Study 2 may be explained by children's lower proficiency in this second study, as shown by their lower learning scores. This could be attributable to the number of exposures to the story (three in the previous study, and two in this study) and/or the lack of a consolidation period (sleep) between second exposure and testing (Henderson, Devine, Weighall, & Gaskell, 2015). It is possible that, in the previous study, children acquired information regarding the category of the target words more readily in the combined condition than in the listening condition, but the additional presentation, and the opportunity to consolidate their learning through sleep, was enough to enhance category learning in the listening condition, therefore eliminating this difference. Furthermore, it is possible that, in Study 1, the presence of a definition particularly enhanced performance in the combined condition, facilitating definition recognition in this condition compared to the listening condition. Some analyses, in fact, suggest that presence of a definition may interact with condition (section 3.3.7.1), although the interaction was not significant in the mixed models. This enhancement, and thus the condition effect on definition recognition in the present research, would have been less marked, since not all words were presented accompanied by a definition.

3.4.2 - Story comprehension effects

As previously described, the story comprehension task used in this study was designed to be quite simple, to control for children's attention and comprehension of the story; nevertheless a difference between the three groups was found in this task. Specifically, the combined group performed better than the reading group, with the listening group's performance falling between that of the other two groups. Comprehending the story, therefore, was a harder task when the story was presented only written. Nevertheless, since this task was built to assess children's memory of facts explicitly provided in the story, and asked the children to only remember more significant features of the story, a more detailed assessment would need to be used to draw clearer conclusions about whether presenting the story both orally and written improves comprehension, compared to presenting the story in written format alone.

Previous research showed that, while there was an effect of comprehension abilities on semantic word learning from stories, the memory for the literal content of the stories did not affect new word learning (Cain et al., 2004). Nevertheless, the number of comprehension questions correctly answered in our study was significantly associated with both phonological learning and semantic learning, especially learning of definitions of the words; we might therefore hypothesize that similar underlying abilities influenced performance on these three tasks, or that comprehending the story positively influenced performance in both of these measures of learning.

3.4.3 - The effect of the presence or absence of a definition

It was hypothesised that the presence of an accompanying definition alongside each new word would foster semantic learning in all groups (Coyne et al., 2004; Dickinson, 1984; Wilkinson & Houston-Price, 2013). The presence or absence of a definition did not have an impact on phonological or orthographic learning. On the other hand, as expected, the presence of a definition had a positive impact on learning of sub-categories and definitions of the words, while, unexpectedly, it had a negative effect on category learning. The positive effect of definition on definition recognition seems to be mostly driven by the combined group, as predicted, although the interaction effect between definition and group was not significant. This suggests that, although the combined group might have shown a bigger difference in definition recognition for words presented with and without a definition, a similar effect was shown by the other groups. Similarly, the negative effect on category recognition might be mostly attributed to the performance of the reading group, although all groups performed similarly.

The positive effect of the presence of a definition on learning definitions and sub-categories could be explained by considering that the definition gives children all the information needed to succeed in the definition recognition task directly. Although the definitions used within the story and in the task were phrased differently, both conveyed the same meaning. Furthermore, within a definition, children were provided with either the sub-category or a synonym of the sub-category, so that the level of abstraction required to succeed in the sub-category recognition task was necessarily lower than the level of abstraction required to succeed in the category recognition task. In the absence of a definition, children need to extract this information from the text, synthesising various cues to construct a coherent representation of the meaning of the word. On the other hand, the abstraction of the information needed to identify the words' categories would have been similar, whether or not a definition is provided, as this information was not directly supplied within definitions. It is perhaps not surprising then that definitions did not support performance on all three semantic tasks. In terms of the negative effect that definitions had on category learning, it is possible that, when provided with a definition, children may have tried to learn more specific information regarding the meaning of the word, rather than trying to incorporate it in their lexicon. Focussing on learning specific features of the meanings of the words may have discouraged children from trying to abstract a more general representation for the meanings of the words. The finding that the negative impact was strongest in the reading group supports this idea, since in this group children could re-read the definition multiple times, and thus had time to focus on the details of this definition, while in the listening group this would have been more difficult, due to the nature of the presentation. Interestingly, children in the combined group were most influenced by the presence of the definition, but interactions between group and definition were not significant, thus no strong conclusion can be drawn regarding this relationship.

3.4.4 - Individual differences effects

3.4.4.1 - Reading accuracy. As in Study 1, these results highlight the importance of reading accuracy not only for learning of the orthographic and phonological forms of new words, but also for learning of words meanings. Reading accuracy, in fact, predicted both phonological and orthographic learning. Specifically, it had a positive effect on phonological learning for the combined group (as in Study 1), and the reading group, although it showed a negative effect for children in the listening group. Similarly to previous research (Ricketts, et al., 2009; Rosenthal & Ehri, 2008), but differently from the results of Study 1, reading accuracy showed a significant interaction with group for phonological learning, thus suggesting that the extent of the orthographic facilitation effect (i.e. the difference between the combined and the listening group), depends on reading accuracy skills, with children with better reading skills showing higher facilitation (see Figure 3.3). Thus, while better readers acquire phonology better from a dual modality presentation, this does not seem to be the case for less skilled readers, who, on the other hand, seem to acquire phonology best from an oral only presentation. We can assume that, while skilled readers can use all the cues provided in a dual modality presentation, acquiring phonology directly, and easily linking it with orthography (similar to the redundant phonological representation proposed by Perfetti, 1992), less skilled readers may struggle with the reading process itself, which might detract attention from the acquisition of new words' forms. The negative effect on phonological learning in the listening group is more difficult to account for: it is possible that better readers might be used to rely on orthography for their precise phonological mapping, and the lack of an orthographic presentation might prompt them to create less specified phonological codes.

Reading accuracy also predicted orthographic learning for all children. Category learning was also influenced by reading accuracy (as in Study 1), and in line with previous research (Ricketts et al., 2011; Swanborn & de Glopper, 1999). Higher reading accuracy skills may have had an effect of reducing cognitive load similar to that of the dual presentation modality, by improving the ability to learn word forms more easily, thus facilitating learning of more complex semantic information (Ehri, 2014).

3.4.4.2 - Vocabulary, listening comprehension and reading comprehension. Both oral vocabulary and oral passage comprehension played a role in the comprehension of our specific story, and oral vocabulary was a predictor of category and definition recognition. The association between oral vocabulary, semantic learning (Shefelbine, 1990; Cain et al., 2004; Oakhill, 1983), and reading and listening comprehension (Beck & McKeown, 1991; Nagy, 2007) is in line with previous research. It is perhaps less obvious why oral passage comprehension, and not reading comprehension, would affect story comprehension, especially in the reading and in the combined groups. It is perhaps relevant that oral passage comprehension resembled the story comprehension task more than the reading comprehension task. Both the oral passage comprehension task and our story comprehension task assessed comprehension as well as memory and attention for the material presented: children could not rely on the presented material during testing (something they were allowed to do in the reading comprehension measure). For these reasons we might conclude that general language comprehension abilities, memory and attention for the material played a bigger role in performance in the story comprehension task for the combined and the reading groups, than more basic reading abilities.

Written text comprehension, on the other hand, was positively related to definition recognition, but only for the listening group. This finding is surprising, since children in the listening group were not presented with the story in written form. It is possible that a third factor may play a role in both reading comprehension and word learning from listening to stories, such as the ability to build a representation of the incoming story or other discourse, without the need to revisit it (e.g., by re-reading). The negative relationship between reading comprehension and definition recognition in the reading group is also surprising. Nevertheless, it is possible that better comprehenders in the reading group might have paid attention to the general meaning of the story, understanding the story, without acquiring the more specific details regarding the words. Although the interaction between reading comprehension and group was significant for definition recognition, the combined group did not differ from any of the other two groups, therefore reading comprehension did not seem to have an effect on the orthographic or phonological facilitation effect.

3.4.4.3 - Non-verbal abilities and executive control. Non-verbal abilities predicted performance in the sub-category recognition task. As for Study 1, non-verbal abilities might play a role in learning new words at increasing levels of demand. Nevertheless this hypothesis does not appear supported by the fact that non-verbal abilities fail to predict category recognition, the task that requires the highest level of abstraction.

A test of children's ability to shift attention between two different series was added as a further background measure in this study, and it was hypothesized that executive control abilities might impact on performance of the combined group more than the performance of the other groups. This hypothesis was confirmed for the orthographic task, where executive control predicted performance in the combined group, but not in the other groups. Nevertheless, the direction of the effect was opposite to what was expected, with children slower to complete the dual task obtaining higher scores in the orthographic task. This might be explained by a speed-accuracy trade-off: children who prioritise rapidity over accuracy, for example, might have obtained lower speed scores in the executive functions task, thus appearing more proficient than children who prioritised accuracy. Prioritising accuracy, on the other hand, could have been a successful strategy to obtain higher scores in both the phonological and the orthographic task, where children were allowed to read and listen to the stimuli multiple times. Accurate children might have chosen to listen to or read the alternatives if they were unsure of the correct answer, where children who maximised speed might have preferred to make a hasty choice. This effect might particularly influence the orthographic task, due to its comparative difficulty, thus advantaging the slower, but more accurate children, who would pay more attention during the orthographic task. Nevertheless, given that the results are unexpected, and difficult to justify, and the fact that the groups differed on executive control, measuring executive functions in a different way, particularly exploring the ability to pay attention to different stimuli at the same time, might clarify this effect in future research.

3.4.5 - Conclusion

The results of this study support two main hypotheses regarding word learning, and the nature of the advantage of a combined phonological and orthographic presentation over a single modality presentation. On one hand, the results that show an advantage of a combined presentation for semantic learning, and the results that show the importance of reading accuracy skills for learning phonological, orthographic and semantic information of the words support the Lexical Quality Hypothesis (Perfetti & Hart, 2002). These results show, in fact, that children with a representation of higher quality, either given by a dual presentation, or created thanks to effective reading skills, fared better when asked to recognise phonological, orthographic and semantic information regarding the words. On the other hand, the difference between a combined presentation and a single form presentation might be conceptualised as a difference in the availability and allocation of cognitive resources. We could in fact hypothesize that children in the combined group needed fewer resources to encode the form of the words, since they were given both phonological and orthographic information at once, and did not need to create a phonological representation from the orthographic representation, as the reading group did. Thus, it is possible that they were able to encode the words more easily and had spare attention and cognitive resources for higher level text processing, such as story comprehension and word meaning comprehension. This idea is supported by the fact that children in the combined group are those who benefit the most from the positive effect of the presence of a definition: they can allocate their spare resources to link definitions and new words. Unfortunately, none of these two hypotheses account for the presence of a difference between the combined and the reading group, but not the listening group, in story comprehension. In any case, the two accounts are not mutually exclusive, and it is possible that both the quality of the representation, and the allocation of resources impact on children's performance.

A problematic result, especially for the attention resources availability, is the fact that our executive task did not predict many of the measures, and, in predicting orthographic learning, its effect was in the opposite direction than expected. A reason for this unexpected result could be the low reliability of the task, compared to the other background measures, due to group administration. The effect of executive functions, therefore, warrants further investigation.

In conclusion, it is difficult to distinguish between the Lexical Quality Hypothesis (Perfetti & Hart, 2002) and the resource availability hypothesis (Mayer et al., 1999) and it is possible that both processes intervene in enhancing the combined group performance. A possible way to test the Lexical Quality Hypothesis more directly would be to test the nature of the representations of the words more thoroughly, especially in the combined and the reading groups, for example by assessing the link between the phonological and the orthographic representations in the two groups. This could be achieved by using more demanding tasks, such as production tasks, or by comparing performance in lexical configuration and lexical engagement (James et al., 2017; Tamura et al., 2017). An alternative, more sensitive measure of the allocation of attention and resources hypothesis, would be the comparison of eye movements, which enable the exploration of the allocation of attention, between the combined and the reading conditions. Such a tool would enable us to explore whether children have more free resources in the combined condition compared to the reading condition. For instance, if children learn words spending less time reading them in the combined condition, then we could suggest that this condition frees attentional resources.

Chapter 4: Study 3

4.1 - Introduction

Chapters 2 and 3 explored word learning from stories when these are presented orally (listening), written (reading), and in the two modalities simultaneously (combined). Both studies showed that children tend to learn word meanings better when they are presented with stories in a dual modality (combined condition). Children learned words' phonological forms equally well, regardless of modality of presentation, but were more likely to learn orthography when the written form was directly provided (i.e., in the reading and combined conditions).

The present study explored possible reasons for the facilitation effect of dual modality over one modality of presentation for semantic learning. This facilitation could be associated to the Lexical Quality Hypothesis (Perfetti & Hart, 2002), suggesting that exposure to both the written and the oral forms of new words creates a better representation of these words in memory, and allows easier access to these representations. On the other hand it could also be proposed that the redundancy of the information in dual modality conditions could free attentional resources online, even before a representation of the new word has formed, and therefore enabling children to attend to the meaning of the text more thoroughly (Mayer et al., 1999). Both of these hypotheses may be traced back to a reduction in required attentional resources for encoding word meanings and/or comprehending the story in the combined condition. The Lexical Quality Hypothesis would suggest that, due to the creation of a better quality representation of words presented in a dual modality condition, fewer resources would be required to process these words, after a first presentation. Attentional resources could then be allocated to story comprehension, and possibly encoding of specific semantic details, especially on further reading. The second hypothesis, on the other hand, would suggest a reduction in attentional resources online, even at first presentation of a new word, before a representation of the word has been created. It is important to note that these two ideas are not mutually exclusive, but they might both facilitate vocabulary acquisition in the combined condition.

The present study, therefore, aimed to explore the strategies children use to learn words in a dual presentation modality, compared to a reading condition. To identify why the combined condition supported greater learning we examined children's reading behaviour in the two conditions. We considered that different resource allocation to reading and semantic processing might be reflected in their attention to the words, definitions and context surrounding these. The ideal methodology to explore children's strategies in exploring the text is eye-tracking, which can elucidate different attentional strategies online. Eye movements on the text, specifically on the words of interest and their definitions, were therefore explored to compare reading time on the target words and on relevant information in the two conditions.

A further aim of this study was to explore children's knowledge of new words in more depth, to establish which aspects of the lexical representation are superior in the combined condition. Studies 1 and 2 tested phonological and orthographic knowledge separately, and failed to show a difference in phonological and orthographic representations between the reading and the combined condition. In this study, rather than probing children's knowledge of the forms of the words, we tested their knowledge of the link between orthographic and phonological forms. The use of this task helped us ascertain whether the combined condition prompted children to form a stronger link between word forms than a single modality presentation, given that previous results suggest that children show similar knowledge of the two forms in the reading and combined condition, when these forms are assessed separately (Chapter 3). Regarding semantic knowledge, category learning was assessed with a recognition task, similarly to our previous studies, while learning of words' definitions was assessed using a definition production task. This helped us explore whether the effects found in Chapter 3 for the category learning task, but not the definition recognition task (section 3.3.7) were specifically linked to the different requirements of the two tasks (i.e. recognising the category *vs.* recognising the definition), or whether the effect was evident in the category recognition task because this was the most difficult of the semantic tasks.

The following sections will describe the results of eye-movement research relevant to the present study, while the specific aims and hypotheses of this study will be described in Section 4.1.5.

4.1.1 - What eye movements reveal about how word pronunciations and meanings are accessed during reading

While reading, the eyes of the adult reader remain fixed on the words of the text for periods ranging from 50 to 500 ms, called fixations. The eyes move from one fixation to another through movements called saccades. The eyes of the reader generally move forward in the text, but backward saccades (regressions), are not uncommon, especially when comprehension of the text has not been achieved, and readers need to revisit previous areas of the text (Rayner et al., 2012). Other eye-movement measures generally utilised in eye-movement research are first fixation duration, the duration of the first fixation on a specific word, gaze duration, which is the initial time spent on a word, before moving further in the text (the sum of all first-pass fixations), re-reading time, which is the time spent on a region of the text after having moved away from it, and total reading time, the sum of gaze duration and re-reading time. Gaze durations are shorter for more frequent, familiar, shorter and more predictable words. On the other hand, second pass time measures, such as re-reading time and total reading time, are influenced by the ease with which a word's meaning is integrated in the context: readers are more likely to make regression to a given word if the processing of its meaning is hard or the initial selection of the meaning was incorrect (Clifton, Staub, & Rayner, 2007). Therefore, while first pass measures mostly indicate processing of the new item itself, second pass measures tend to reflect integration of the word with following context.

Eye movements during reading differ between children and adults. Specifically, as children's reading skills increase, the duration and the number of fixations tend to decrease, the spatial length of the saccades increases, allowing children to move their eyes further in the text, and the number of regressions also decreases (Rayner, 1998; Blythe et al., 2006). These changes reflect the transition from developing to expert reader. In becoming expert readers, children become able to obtain more information from the text in a shorter amount of time, and process words that are further away in the text, i.e. their perceptual span increases (Rayner, 1986).

The amount of time readers spend on a specific region of the text reflects the ease or difficulty of processing that region of text. The time readers spend on specific words seems to give information regarding ease of access to the words' phonological features (Frost, 1998; Rastle & Brysbaert, 2006) and meanings (Clifton et al., 2007; Rayner, 1998). Phonological information is extracted very early during reading: information about a word’s phonological form is extracted even before the eyes land on the specific word, especially during the reader’s previous fixation (Pollatsek, Lesch, Morris, & Rayner, 1992; Henderson, Dixon, Petersen, Twilley, & Ferreira, 1995). Studies that explore the effect of preview benefit often use the boundary technique, a technique in which the identity of a target word in a sentence changes before it is fixated. This creates a *preview* of a word which is different from the word the reader fixates. In these studies reading time on the target word is reduced if the preview word is a homophone of the target, providing accurate phonological but not orthographic information (Pollatsek et al., 1992). Children as young as seven show preview effects, suggesting that their phonological processing while reading is similar to that of adults in this task (Blythe, Pagán, Dodd, 2015). As seen from the preview benefit effect, ease of access to the phonological form of a word reduces the time spent reading the word itself. New words, being new in both phonology and orthography, will result in longer reading time than known words, and this additional time spent reading the new word would be associated, amongst other things, with phonological processing. We would therefore expect that reading time, especially at first pass (i.e. gaze duration), would predict subsequent phonological recognition. However, it must be noted that spending less time reading a word could indicate either that a poorer phonological representation has been constructed, or that less phonological processing is needed to build an adequate representation, for example due to previous experience of the word.

The ease with which readers extract the meaning of the word also influences time spent reading it. This is supported by the observation that several variables related to semantic access influence reading time, such as word frequency, word familiarity, age-of-acquisition, number of meanings of the word, plausibility and predictability (Clifton et al., 2007). Specifically, first fixations and gaze durations are longer on lower frequency words (Inhoff & Rayner, 1986; Just & Carpenter, 1980; Kennison & Clifton, 1995; Raney & Rayner, 1995) and words that are more predictable from the preceding text are skipped more often and have shorter reading times than less predictable words (Ehrlich & Rayner, 1981; Rayner & Well, 1996). These results confirm that the easier the access to a word’s meaning, the less time the reader needs to spend on that word. Children’s eye movements are affected by the same features, although not always to the same extent (Reichle et al., 2013; Blythe & Joseph, 2011). Children also show an interaction between length and frequency effects, needing more time to process long low-frequency words than long high-frequency ones (Tiffin-Richards & Schroeder, 2015). This result has been interpreted as indicating that children tend to rely on effortful sublexical decoding when reading unfamiliar words, instead adopting a more efficient lexical-level approach when reading higher frequency words.

Word-level variables are not the only variables that influence word reading times. Indeed, text features also affect the ease with which the meanings of words are extracted. The time spent reading a word not only reflects how long the reader needs to make sense of the meaning of the word itself, but it represents how easy it is to access its meaning *within* its context. Specifically, gaze durations and first fixations are longer when the meaning of the fixated word is not easily extracted (Binder & Rayner, 1998; Sereno, O'Donnell, & Rayner, 2006), for example when there is some conflict between preceding context and the target word. In these instances adult readers also spend more time on the disambiguating parts of texts and make more regressions to relevant context, when they had originally selected the wrong meaning for the word (Duffy, Morris & Rayner, 1988). Readers therefore use previous context to predict the meaning of the words, and this impacts on and interacts with word identification time.

4.1.2 – Eye movements during reading: processing new words

Recent studies have investigated how readers attend to new words and extract their meaning by considering readers' attention to the text (Blythe et al, 2012; Brusnighan & Folk, 2012; Brusnighan, Morris, Folk, & Lowell, 2014; Chaffin, Morris, & Seely, 2001; Godfroid, Boers, & Housen, 2013; Pollatsek, Slattery, & Juhasz, 2008; Williams & Morris, 2004). Eye-movement research generally considers the time spent on a word as indicative of its status in the lexicon and whether it has been successfully encoded. Given the research on the link between gaze duration, frequency and predictability summarised above, it would be expected that new words and non-words, of which the reader has no prior knowledge, will be fixated longer than known words. In fact, these new words will have both lower frequency and less predictability than known words. Several studies support this prediction (Brusnighan & Folk, 2012; Chaffin et al., 2001; Godfroid et al., 2013; Pollatsek et al., 2008), suggesting increased processing demands for new words. With multiple exposures, new words become easier to process, and reading time decreases at subsequent encounters (Joseph, Wonnacott, Forbes, & Nation, 2014). In some studies (Brusnighan & Folk, 2012; Chaffin et al., 2001) it was found that skilled readers can build a partial but useful representation of the meaning of a new word even after just one encounter. For example, adult participants in Chaffin et al.'s study were exposed to two sentences. The first sentence contained target items alongside a context with high or low informativity. Items were either new words, words of low familiarity or words of high familiarity, all with similar meanings, congruent to the sentence in which they were embedded. The second sentence, on the other hand, contained a synonym or category of the previous word (for example for words indicating musical instruments, the second sentence contained the word *instrument*). As expected, readers spent more initial processing time, and more regression out of context and towards the word, for words with low familiarity and new words, compared to more familiar words. This indicated their efforts to extract the new item's meaning. However, in the second sentence, they did not spend more time reading the synonym of a new word compared to that of highly familiar words. Access to the meaning of the synonyms was similarly easy for new or known words. Interestingly, this was only the case when the context was informative: given non-informative context readers spent more time on the synonym, in the second sentence, following a new word than a known one. This indicates that, when provided with the meaning of a word within context (informative context condition) both known and new words had been retrieved or stored sufficiently well to allow the easy recognition of a synonym in the following sentence. Similar effects can be found when new words are encountered by beginner readers. For example, in a recent study, Joseph and Nation (2017), showed that, similarly to adults, 9- and 10-year-old children spend less time on new words at later, compared to earlier exposures, both in terms of gaze duration and total reading time, and that they spend more time on the new words (total reading time) when the context is not informative.

Studies on new word learning from texts have also investigated how learners use context to infer word meanings. Specifically, research has shown that adult readers tend to spend more time reading the context in which new words are embedded when it is informative than when it is not. They also spend more time reading the context of new words than known words, and make more regressions out of context for new words than known ones (Brusnighan & Folk, 2012; Chaffin et al., 2001; Williams & Morris 2004). Most eye-movement research tends to explore the word learning process while it develops online, but a few studies have attempted to include both eye-movement measures and off-line vocabulary learning measures in the same study, to combine information regarding how the word is acquired with the extent of this acquisition.

These studies have shown that adults spend more time reading sentences containing new words that are learned, compared to new words that are not (Brusnighan & Folk, 2012). Furthermore, adult readers show shorter gaze duration on newly learnt words, compared to unlearnt words, but they spend more time re-reading learnt than unlearnt items (Williams & Morris, 2004). Conflicting findings come from L2 literature, which shows an increase in total reading time for learnt compared to unlearnt words (Godfroid et al., 2013), but no decrease in gaze duration. These results seem to indicate that readers who spend more time on a word tend to learn it more efficiently. These studies not only differ in terms of the sample (monolingual *vs.* bilingual adults), but also in the stimuli used (single *vs.* multiple sentences). Nevertheless, they both suggest that later fixation time measures (either re-reading time or total reading time), might be stronger predictors than gaze duration, given that a difference in gaze duration between learnt and unlearnt words was evident in only one of the studies. Given that neither of these studies involved children, it is difficult to predict what children would do in the presence of new words, but it might be hypothesized that children would also show longer re-reading and total reading time on the learnt words, and possibly shorter gaze duration on these.

4.1.3 – Eye movements during oral presentations: reading aloud and reading while listening

When reading aloud, fixations tend to be longer in duration (Inhoff, Solomon, Radach, & Seymour, 2011; Laubrock & Kliegl, 2015). This is generally interpreted as a tendency of the eyes to surpass the voice, and then remain in place to allow the oral presentation to catch up with the eyes' position (Levy-Schoen, 1981). Participants, therefore, tend to integrate the two streams of information; to do so adult readers slow their reading rate to wait for the voice to catch up. Similar results have been found with adults reading captions from videos (Ross & Kowler, 2013).

It has also been shown, in research with videos or images and captions, that adults tend to attend to the text carefully (i.e. read it) (d'Ydewalle, Praet, Verfaillie, & Van Rensbergen, 1991; Rayner, Rotello, Stewart, Keir, & Duffy, 2001), even when the text is redundant to the oral information or is not useful (Ross & Kowler, 2013; Wang & Pomplun, 2012). On the other hand, in shared picture-book reading contexts, children do not always attend the text as thoroughly as adults in similar conditions. Specifically, pre-readers tend to spend very little time on the print (Evans & Saint-Aubin, 2005), and time spent on the written text increases with age and reading ability, and depends on the nature of the text itself (Roy-Charland, Saint-Aubin, & Evans, 2007): older children (9-10 year-olds) read alongside the oral presentation more often than younger ones, and younger children read along mostly when the text is appropriate for their reading ability. The time spent on the written presentation also varies as a function of reading ability, with more able readers spending more time on the text. Nevertheless, even older children do not spend all of the allotted time on the written text (reaching 66% of the time at 9 – 10 years of age), if provided with a picture. Overall, there is a developmental shift where readers tend to pay attention to the text more and more as their reading skills grow, until the text becomes impossible to ignore, even when not useful, as in subtitles or visual scenes (Ross & Kowler, 2013; Wang & Pomplun, 2012) or when performing unrelated tasks (see the Stroop Test as an example: Comalli, Wapner, & Werner, 1962).

Since the strategy used to read a text while listening depends on experience, it is possible that, when compared to better readers, less able readers might skim through the text, rather than reading it. Nevertheless, 9 year-olds would probably read along most of the time, if presented with a text adequate for their reading level (Roy-Charland et al., 2007). Research that includes videos and subtitles also shows that, although including subtitles does not always improve comprehension performance (Danan, 2004; Garza, 1991), participants who tend to read the subtitles perform better than participants who do not use subtitles as efficiently (Kruger & Steyn, 2014). It could therefore be hypothesized that children's attention to the written text while listening would influence their comprehension of the text, and their vocabulary acquisition from it.

Given the shortage of literature in the area of attentional allocation while reading and listening to text containing new words, it is unclear what reading strategy would enhance word learning the most in a dual presentation modality. Previous research does not clarify whether increased attention to the written text over time provides an advantage to the reader, or it is a by-product of his or her increasing reading abilities. Therefore it is difficult to predict whether children would be advantaged by dual presentation in learning new word meanings by attending to the written text, if they learn new words mostly by attending to the oral text, or by integrating the two. Regarding simultaneous attention to different modalities, research on shared story-book reading in pre-school children (Evans & Saint-Aubin, 2013) and younger readers

(Duckett, 2003) shows that children are able to use pictures presented with the text meaningfully to understand the text more clearly. For example, children who looked at the relevant parts of pictures (i.e. those parts that give them information regarding the meaning of the words) while a word was being spoken, or soon after the word was spoken, were more likely to learn it. Thus children can use the link between two presentation modalities (i.e. oral and visual), at least in specific situations, to learn new words. This ability might generalise to the oral and written presentation of a text, although, given the difference between the processing of visual information (pictures) and written information (written words), this remains to be seen.

4.1.5 – Aims and Hypotheses

This study aims to investigate how attention is allocated when children learn new words while reading and listening to stories at the same time, compared to when they only read the stories. The study investigates both the processes children use to acquire new words when exposed to stories in two different modalities, and the products of this process, i.e. how well children learn the link between novel orthographic and phonological forms and meanings of new words. Thirty-four Year 4 children (9 year-olds) were exposed to two stories in two conditions, a reading only condition where they were presented with stories through the written modality only, and a combined condition where they both listened to and read stories simultaneously. Stories were divided into passages, each containing one new word repeated three times. Eye-movement data were collected while the children were exposed to the stories and off-line measures of word learning were obtained following story reading. This study aimed to explore several hypotheses regarding how presentation modality affects learning and text processing, how words are processed differently when learnt or not, and whether the processing of learnt and unlearnt words differs across the two presentation modalities.

Similarly to the previous two studies, children's vocabulary learning in the two conditions was explored. Since no difference was highlighted between the phonological and orthographic representation of the words of the combined and the reading group, in this study we measured the link between the two representations, rather than the two representations separately. Nevertheless, given the similarity between this measure and our previous measures it was expected that the learning of the link between phonological and orthographic forms would be similar whether items were presented in written or both oral and written texts. Conversely, it was expected that semantic learning would be better in the combined than in the reading condition, especially for category learning (see Chapter 3).

Secondly, it was expected that children would process the text differently in the two presentation modalities. To successfully process the text in the reading condition, children would have to read the text carefully. However, it is more difficult to predict how children approach the task of comprehending a text when this is also orally presented. The studies previously mentioned on multimodal presentation (Ross & Kowler, 2013; Wang & Pomplun, 2012) suggest that adults tend to read the text even when exposed to the oral presentation. Studies with children indicate that, as they get older, young readers tend to increase the time they spend on the written text while reading and listening to stories at the same time (Roy-Charland et al., 2007). It is possible that 9-year-olds' eye movements (specifically the length and direction of fixations and saccades) in the two conditions would be similar: children in this age range might have reached the stage where they are compelled to read written text, regardless of presentation modality (as adults do). On the other hand, since, in picture book presentation, not even 9 and 10 year-olds spend all of their time reading along (Roy-Charland et al., 2007), and since children's experience and ability to rely solely on the written text may be less developed than adults', it is possible that they would rely more on the oral presentation. This strategy might lead the children to move their eyes on the text without reading it. This would result in fewer reading-like eye movements (mainly rightwards and leftwards fixations) in the combined presentation modality, where they might show more upwards and downwards movements over the text.

When considering the effect of reading time on the learning of word forms and meanings, it was expected that, in the reading condition, first pass reading times (i.e. gaze duration) would predict learning of the link between phonological and orthographic forms, with learnt words fixated longer than unlearnt words. This prediction was based on the assumption that the creation of a phonological representation starts as soon as the word is encountered, and in some cases even earlier (Pollatsek et al., 1992; Henderson et al., 1995). The effect of form learning would therefore be evident at first pass, as suggested by the literature on phonological effects (section 4.1.1). Considering the learning of word meanings, research with adults indicates that readers show longer re-reading and total reading times on learnt compared to unlearnt words, and, in some research, shorter gaze duration (Godfroid et al., 2013; Williams & Morris, 2004): the same pattern was therefore expected in the present research, at least for the reading condition.

Hypotheses for the combined condition were more tentative, given the lack of literature on this presentation modality. It is possible that the pattern of differences in eye-movements between learnt and unlearnt words might be the same for both modality conditions, across the different reading measures. This idea is supported by findings that show that children in the age range of our participants are able to process oral and written information simultaneously, reading alongside the spoken presentation (Roy-Charland et al., 2007). Nevertheless, an alternative idea can be proposed, specifically, that learning words in the combined condition might be best predicted by the time spent on words and definitions while simultaneously hearing words and definitions. If this is the case, we might expect that children's learning of the link between the phonological and orthographic forms of the words would be predicted by the time children spend looking at the target words while these were spoken (coincident time), rather than the total time spent looking at the words. On the other hand, learning of word meaning in this condition could be linked to the time spent looking at the word while the definition was being spoken or the time spent looking at the definition while the word was being spoken (cross-coincident time). The prediction that longer cross-coincident looking times might predict learning of the meanings of the words was based on findings relating to younger children's exploration of pictures in picture-book shared reading (Evans & Saint-Aubin, 2013). Children presented with picture-books learn words best when they pay attention to the image and hear the word it depicts at the same time. This can be translated into the idea of more proficient word learning when processing oral and written texts simultaneously.

From a theoretical perspective, considering the difference between the two conditions, a further hypothesis regarding processing time can be proposed. Specifically, if the combined condition frees attentional resources, children would need less processing time on the new words in this condition. If children spend less time on the new words in the combined condition even at first presentation, but nevertheless show equivalent or better learning of these words, this would suggest that a dual modality presentation facilitates learning by diminishing online processing demands (Mayer et al., 1999). On the other hand, if the time spent to process the new words is similar in the two conditions, or if new words are processed more rapidly at second and third presentation, but not at first presentation in the combined condition, this would suggest that the combined condition frees attentional resources only after the creation of a more stable representation of the word in memory, thus supporting the idea that attentional resources are freed due to better representations in memory (Perfetti & Hart, 2002).

Individual differences were also expected to influence children's ability to learn new words, as in Studies 1 and 2. It was expected that reading accuracy would predict learning of the link between orthographic and phonological forms (Ricketts et al., 2011), and vocabulary knowledge would predict semantic learning (Cain et al., 2004; Penno et al., 2002).

4.2 - Method

4.2.1 - Participants

Thirty-four children aged 8 to 9 years participated in the study (M_{age}= 8.93 years; *SD* = .29 years; 15 boys). Participants were recruited from three primary schools in South-East England. Informed parental consent was received for all participants (see Appendix S for information sheets and consent forms). All children had normal or corrected to normal vision and teachers confirmed an absence of learning or neurological disabilities. All children spoke English as their first language.

4.2.2 – Design

Story presentation modality was manipulated within subjects and all children were presented one story in the reading condition, and one story in the combined condition. Order of condition, story and list of target non-words were counterbalanced. The eight resulting counterbalancing groups are described in Table 4.1.

Table 4.1

Counterbalancing procedure

Counterbalancing group	N	First Condition	First Story	Target word list used in first story
1	5	Reading	Pirate	a
2	4	Reading	Knight	a
3	4	Combined	Pirate	a
4	4	Combined	Knight	a
5	4	Reading	Pirate	b
6	4	Reading	Knight	b
7	4	Combined	Pirate	b
8	5	Combined	Knight	b

4.2.3 - Stimuli

Twelve non-words to be used as target items were chosen from existing datasets of non-words: set B, C and D of the TOWRE – Second edition (Torgesen et al., 1999 - set A was used as background measure), DTWRP (Forum for Research in Literacy and Language, 2012), WIAT-II (Wechsler, 2005), and Chaffin (1997) (see Appendix T). Six were bisyllabic non-words while six were composed of three syllables. To select the target items pronunciation data from a sample of 45 non-words were collected from a pilot sample of 13 adults: each of the non-word chosen was correctly pronounced by at least 69% of the pilot sample. The pilot sample also supplied data on how plausible it seemed that the non-words were real words (word-likeness), rating the items on a likert-scale from 1 - very word-like, to 3 – not word-like at all. They were then asked to judge the ease of pronunciation on a similar scale (1 – very easy to pronounce, 3 – difficult to pronounce). Target non-words were divided into two lists of 6 items (list a and list b). Items in the two lists were paired, each pair associated with a category (e.g. animal, job – Appendix T). The two lists were matched for length, bigram frequency (Medler & Binder, 2005) and phonotactic probability (Vitevitch & Luce, 2004) (all *p*s > .100). The words were also matched for measures from the pilot study: number of adults correctly pronouncing the word, word-likeness and ease of pronunciation (all *p*s > .400).

To control for task effects in the phono-orthographic task, 12 control non-words were also selected from the list used in the pilot study, but not presented in the story (Appendix T). Target and control non-words were matched on length, number of adults correctly pronouncing the word, word-likeness and ease of pronunciation, bigram frequency (Medler & Binder, 2005) and phonotactic probability (Vitevitch & Luce, 2004) (all *p*s > .300). Real words were also selected to be included in the phono-orthographic task (Appendix T); the list of real words was also matched with the list of target non-words in length, bigram frequency (Medler & Binder, 2005) and phonotactic probability (Vitevitch & Luce, 2004) (all *p*s > . 100). Control non-words and real words were also arbitrarily assigned to word list a or word list b for presentation in the phono-orthographic task. Alternative pronunciations for the non-words were also created for the phono-orthographic task, either by selecting the most frequent incorrect pronunciation produced by same pilot sample of 13 adults who were asked to read the non-words aloud, or by selecting the errors produced by the children in reading aloud the non-words, as reported for the standardised DTWRP (Forum for Research in Literacy and Language, 2012) (see Appendix T).

Two stories were written for this study, henceforth called the Pirate story and the Knight story (see Appendix U). Each story was divided into eight passages. The first two passages were introductory passages of around 50 words in length, while the following six passages (101 to 133 words in length) each introduced one target non-word, repeated three times, accompanied by clues to its meaning. The order in which new item categories were presented was matched across the two stories (job, clothing, food, building, object, and then animal). The stories were similar in length (821 and 848 words respectively). The two stories had a Flesch reading ease and Flesch-Kinkaid Grade Level appropriate for the age of the children (Flesch reading ease: M_{Knight} = 84.14; M_{Pirate} = 82.93; Flesch-Kinkaid Grade Level: M_{Knight}= 4.61; M_{Pirate}= 4.29), and passages in the two stories did not differ on these measures or length (all ps > .300).

Meanings of the non-words were conveyed in the story by including a definition for the item the first time it was mentioned, and clues to its meaning the second and the third time it was mentioned. Definitions were four words in length, and comprised information about the word's sub-category and a further phrase to specify it (see Appendix V). For example, for the item for the category *clothing* in the pirate story, the definition provided was '*dress worn by men'*, which comprises both the sub-category information *'dress'*, and the specific characteristic *'worn by men'*. Before the creation of the stories, 24 adult speakers of English were given the first three words of each definition and asked to supply the fourth, to assess predictability of the last word contained in each definition, while 15 further adults assessed the internal plausibility of the definitions as a whole using a 5-point likert scale (two examples were provided to prompt the adults to assess internal plausibility: specifically "a castle made of stone" was used as an example of plausible definition, while "a castle made of water" was used as an example of implausible definition). The definitions for each pair of items linked to a specific category were paired for length and matched for plausibility and predictability, as well as word frequency and number of orthographic and phonological neighbours for each word in the definition (Masterson, Stuart, Dixon, & Lovejoy, 2003), and the definitions in the two stories did not differ significantly on these measures (all ps > .100). Definitions and clue positions were controlled to be at similar distance and position respective to the non-words in each target passage. The first time target non-words were presented in a passage these were preceded by an adjective, to minimize the probability of skipping the previous word and control for preview benefit: inserting a preceding adjective ensured that the time spent on the target non-word was time for the analysis of the target, and did not include the time to process a previous article or other short word (which are often not fixated).

As described in Table 4.1, two different versions of each story were created, so that either non-word list a or non-word list b was included in the story. Recordings of the stories were read by a female native English speaker. Results from previous studies were used to predict typical reading time for Year 4 children for each passage, and the recordings of the adult storyteller were controlled to match these predicted reading times. Mean reading time of the storyteller was 46 s (*SD* = .62 s), and this did not differ from the time children took to read the passages on their own, which was 49 s (*SD* = 11.56 s, *T* = 231.00, *p* = .256).

4.2.4 – Vocabulary acquisition tasks

Knowledge of the link between orthographic and phonological forms of the target non-words was assessed by asking children to decide whether the word they could see written on the screen was read aloud correctly by a recorded voice. This task was deemed to assess children's knowledge of the link between orthographic and phonological forms of the new items, where children could employ similar skills required by read aloud tasks, but in comprehension, rather than production. This phono-orthographic task was delivered through a laptop using E-prime (Schneider et al., 2002). In this task a word was presented visually in the middle of the screen, and an associated spoken form was simultaneously presented. The written word was presented 350 ms before the presentation of the oral form, to allow children to familiarise themselves with the written word. Target non-words presented in the story, and control non-words and real words from the same word list were assessed after each story. All items were presented twice in random order, once accompanied by the correct pronunciation and once accompanied by an alternative incorrect pronunciation. Participants' task on each trial was to decide whether the word was pronounced correctly or not, as fast as possible, by pressing one of two buttons. On-screen instructions followed by eight practice trials (4 familiar words repeated twice, once correctly, once incorrectly) ensured that children understood the demands of the task. To ensure comprehension of the task, children could proceed to the target items only if they made fewer than 3 mistakes during the practice trials, otherwise they had to repeat all practice trials. Both accuracy and reaction times were recorded in this task.

Two semantic tasks were administered on an individual basis by the experimenter following the phono-orthographic task. Only knowledge of target non-words was assessed in the semantic tasks. The first task was a *category recognition task*. For this task the experimenter told the child that he or she would be asked to select the right category for some words that had been presented in the story. A written list of 8 categories was presented, and read aloud by the experimenter. Then the child was presented with each target non-word in turn in a random order: the experimenter read the item aloud, and showed it written on a card, then asked the child to select the correct category for this item. The same list was used for all the non-words, and it comprised the 6 correct categories, one for each non-word, plus two other categories (PLANT and VEHICLE), to minimise the probability of choosing the correct category by guessing. For this task accuracy was recorded for each item.

After the category recognition task, children completed a *definition production task* where they were asked to produce a definition for each target non-word. This task was designed to elicit production of all the information children remembered regarding each item. First children were asked “X was mentioned in the story, do you remember what X means?”. Children were also invited to say “everything they remembered”. If children were unable to produce a full definition for the non-word, they were given prompts. The first prompt was the correct category of the word: for *dress worn by men*, for example, children were told “X was an *item of clothing* in the story. Do you remember something more about it? What *item of clothing* was X in the story?”. If the child still failed to produce the entire definition for the item, the first part of the definition was provided, for example for *dress worn by men* children were told that the item was a *dress* and asked if they remembered anything further for this *dress*. This task was scored on a 0-4 scale, with children obtaining a 4 when able to produce a complete definition without prompt, 3 if able to produce part of the definition without prompt, 2 if able to produce the entire definition after the category prompt, 1 if they either produced only part of the definition after the category prompt, or they correctly produced the second part of the definition after the second prompt, or 0 if they failed to produce any part of the definition, even after prompts.

4.2.5 - Background measures

Children completed background measures in one session either before or after the eye-tracking session. All were standardized assessments and were administered according to test manual instructions. Measures of non-verbal abilities (CPM; Rust, 2008), vocabulary (BPVS – 3; Dunn, Dunn, & NFER., 2009), word and non-word reading abilities (Set A of the TOWRE – Second edition; Torgesen et al., 1999), and reading comprehension and written passage accuracy (YARC; Snowling et al., 2009) were used to assess individual differences (see Chapter 2 for details of the measures). Differently from Study 1 & 2, for this study, all children completed the same age-appropriate passage of the YARC (level 4).

4.2.6 – Procedure

Participants were tested in two different sessions, the main session when both stories were read and both learning tasks took place, and a background measures session, either before or after the main session, when measures of children's oral and reading skills were collected. Figure 4.1 summarises the time course of the main session. The main session was completed in a quiet room within the school and lasted around 1 hour. Stories were presented on a computer screen, while participants' eye movements were recorded. Before story presentation, calibration took place: the child sat in front of the eye-tracker, with forehead resting on the head rest, and was asked to look at 9 dots, one at a time, in 9 different positions on the screen. The same procedure was repeated immediately for validation. Calibration and validation were repeated twice during the story reading.

Each story was divided in passages, and each passage was presented on the computer screen, one passage at a time. Before each passage, a contingency box (a square on screen) appeared. The child had to look at the contingency box for 500 ms before the passage was presented: the contingency box was placed at the beginning of the first line of the passage, to ensure that the child would be ready to start to read as soon as the passage appeared. When the participants finished each passage, the computer displayed a comprehension question that required participants to answer either YES or NO by pressing different buttons on a response device. These questions served to assess basic comprehension of the passages and maintain children's attention to the story. After the presentation of the first story, children moved to a different computer to complete the experimental tasks that assessed learning of the items presented in the story. Tasks were presented in a fixed order, with the task assessing the link between orthography and phonology first, followed by the semantic tasks (category choice and prompted definition production). This order minimised any impact of previous tasks on later tasks (see Chapter 2 for a discussion of how this order minimises order effects due to learning during testing).

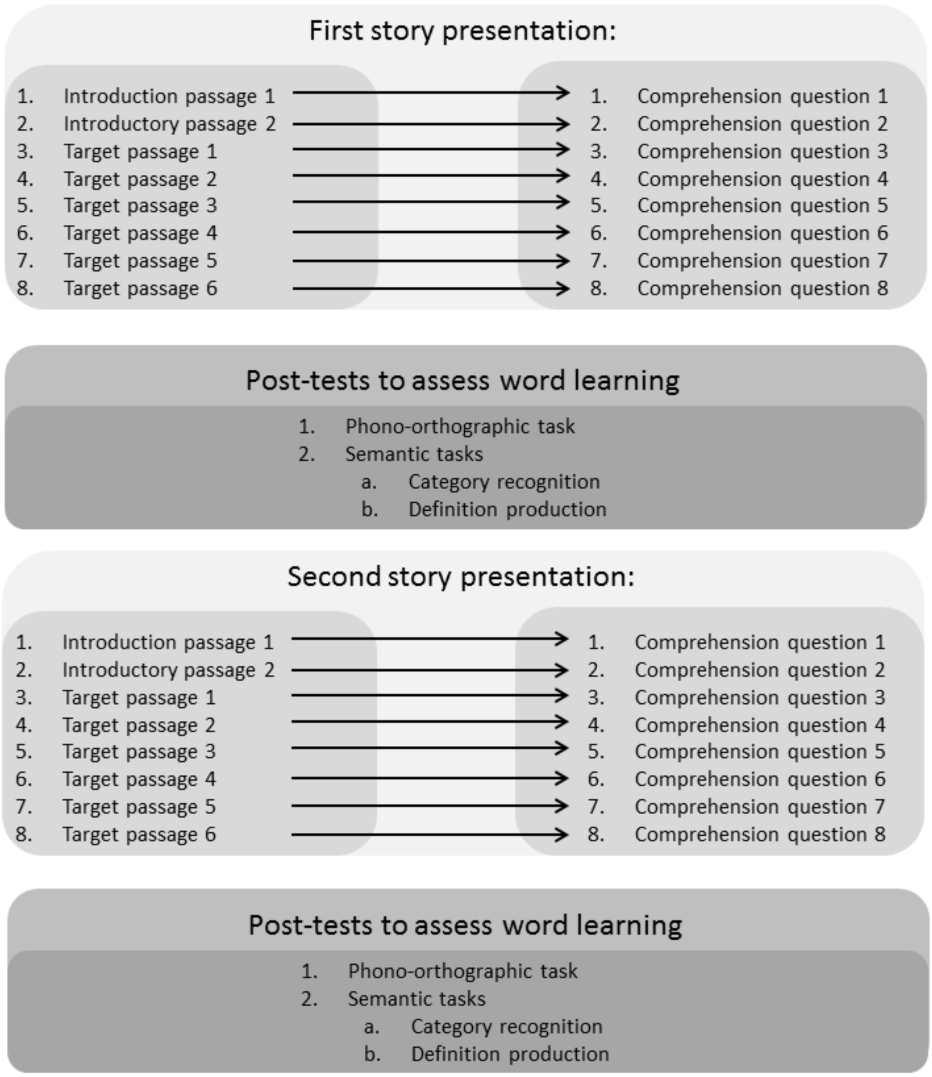


Figure 4.1. Procedure for the main session

During story exposure participants were asked to read at their own pace (reading condition) or listen to and read (combined condition) the two stories. The order in which children completed the two conditions was counter-balanced (see Table 4.1). They were told to try to comprehend each passage as well as possible and to answer the comprehension questions after each passage. They were further told that they would participate in a couple of tasks at the end of the story presentation. The presence of unknown words within the stories was mentioned, to reduce surprise and possible movements away from the head rest while recording eye movements. Children were told they might encounter words they did not know in the story, but their main focus was story comprehension. In the combined condition, children were instructed to listen to the story they could hear via the headphones, and read along on the screen. In the reading condition children only had to read the story, but they still wore headphones, to isolate them from external noise, and make the two presentation modalities as similar as possible, except for the lack of oral presentation. In the reading condition children had to press a button to indicate they had finished reading each passage, while in the combined condition the passages were timed: children were asked to press the button when they had finished reading and listening to the story, but the presentation of each passage ended automatically 500 ms after the end of the oral passage. This ensured that, in the combined condition, children could not re-read the passage after having listened to it. To ensure children read the passage only once in the reading condition, they were directly asked to do so, and their eye movements were monitored closely. After each story, children completed the word learning tasks, where their knowledge of the target non-words presented in the specific story was assessed.

4.2.7 – Eye-tracking apparatus and eye-movement data collection procedure

Right eye movements were recorded by an Eyelink 1000+ eye-tracker with a refresh rate of 1000hz. The eye-tracker was interfaced with a computer that controlled stimulus display and data storage, and the computer screen where the passages were presented (screen resolution: 1920 by 1080, refresh rate 59hz, length: 33.8 cm, height: 26.6 cm). The display was 60 cm from the reader. Participants viewed the screen with their heads positioned in a deep chin rest and a forehead rest to minimize movements. A nine-point calibration and validation procedure was performed three times during story presentation. For the condition that included listening children heard the stories through HP 530 Headset headphones (Frequency range: 20Hz-20,000Hz; Sensitivity: 105dB S.P.L at 1KHz; Rated power: 100mW).

For the eye-movement analyses, interest areas and interest periods were created. Interest areas are defined as specific parts of the text, either single words or a collection of words. For each passage six interest areas were created, corresponding to the three repetitions of the target non-word, the area including the definition and the areas including each of the two clues. Figure 4.2 illustrates an example of a passage and the relevant areas of interest. Time spent on these areas was then assessed: measures of gaze duration, re-reading time, total reading time and regressions were considered. Gaze duration was defined as the sum of the initial fixations within an interest area, prior to the eyes moving outside the area, and re-reading time was the sum of all the fixations to the interest area after the eyes re-entered the area the second time. Total reading time was defined as the total time spent on the area of interest (thus the sum of gaze duration and re-reading time). To measure regressions, an either/or measure that considered whether regressions were or were not made to the interest area was computed. To analyse eye-movement data, interest periods were also created to explore how the children in the combined condition attended to the written text in reference to the oral text. Interest periods were defined as the timeframe in the oral passage when a specific word or set of words was spoken aloud, including half of the pause between the previous word and the word of interest, and half of the pause between the word of interest and the following word. Specifically six interest periods for each passage were created, an interest period for each of the three times the target item was pronounced aloud (M = 698 ms; SD = 137 ms), an interest period that included the definition (M = 1650 ms; SD = 309 ms), and an interest period for each of the two clues (M = 896 ms; SD = 366 ms).

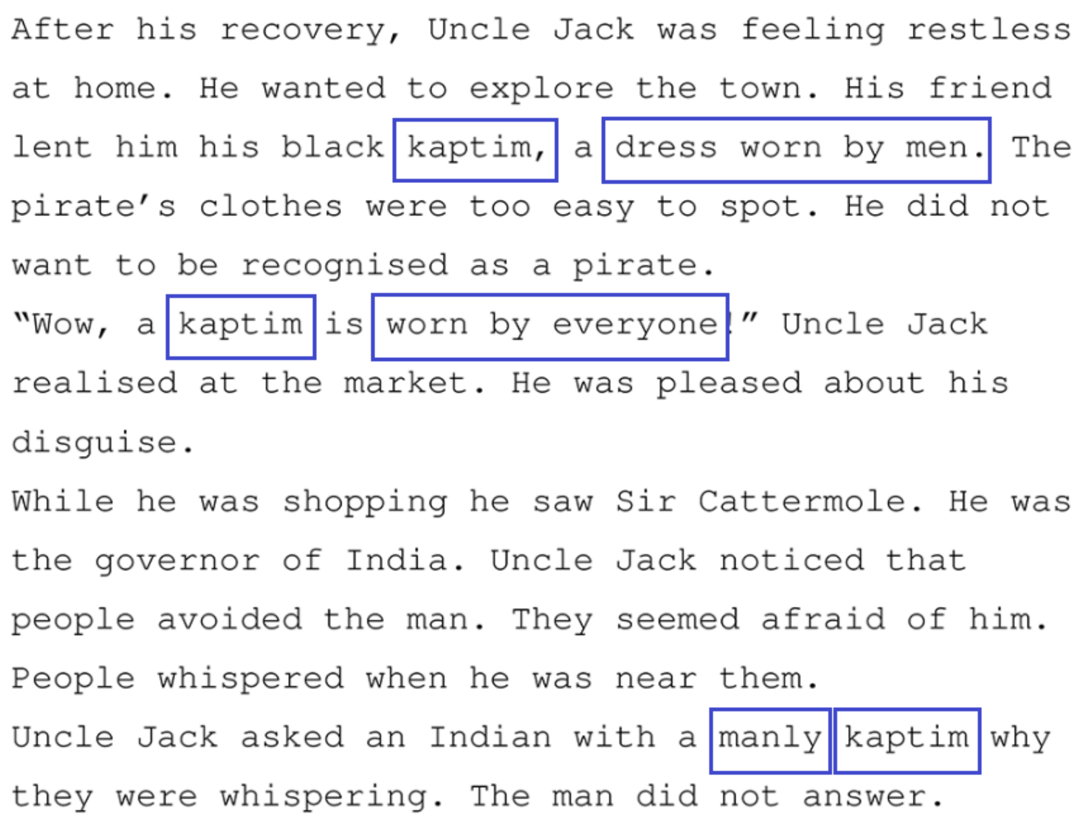


Figure 4.2. Example of passage including interest areas

4.3 - Results

4.3.1 - Sample Characteristics

Children's performance on the background measures is reported in Table 4.2. Children performed in the normal range on all the measures considered. As in Chapter 2 and 3, a reading accuracy composite was formed by merging TOWRE and YARC text reading accuracy scores into a factor, using the regression method (*M* = 0.00; *SD* = 1.00; range: -2.65 – 2.09); the creation of this factor was supported by moderate to high correlations between its constituent measures (TOWRE SWE (word reading) and TOWRE PDE (non-word reading): $r = .78$, $p < .001$; TOWRE SWE and YARC errors: $r = -.54$, $p = .001$;TOWRE PDE and YARC errors: $r = -.72$, $p < .001$).

Table 4.2

Performance on the background measures

	Mean (SD)	**Range**
TOWRE SWE		
Raw Scores	67.82 (7.58)	56 - 87
Standardised Scores	104.94 (10.98)	85-134
TOWRE PDE		
Raw Scores	38.94 (10.72)	16 – 58
Standardised Scores	108.47 (13.70)	77 - 135
BPVS		
Raw Scores	117.76 (13.81)	96 - 152
Standardised Scores	95.06 (14.09)	72 - 126
CPM		
Raw Scores	27.88 (4.07)	19 - 36
Standardised Scores	99.26 (16.43)	70 - 135
YARC[12]		
Number of errors	7.24 (5.02)	1 - 25
Comprehension Questions	5.24 (1.60)	2 - 8

4.3.2 - Story comprehension task

After each paragraph of each story children were asked to answer a yes/no question. This task was included to maintain children's attention on the story.

[12] Standardised scores could not be derived for this measure, since only one passage was administered.

Nonetheless, performance was analysed to compare the two presentation modalities. In both conditions accuracy was significantly better than chance, computed as a score of 4 out of a possible 8 (combined condition: *Median* = 6.00; $W = 508.50$, $p < .001$; reading condition: *Median* = 6.00; $W = 373.00$, $p < .001$). Children correctly answered 6 questions in both conditions, and no difference between the conditions was highlighted ($T = 146.50$, $p = .455$), suggesting that children were able to follow stories in both conditions reasonably well.

4.3.3 – Vocabulary acquisition tasks results

4.3.3.1 - Phono-orthographic task. In this task children were asked to decide whether the given pronunciation of a written word or non-word was correct. All written items were presented twice, once with the correct pronunciation, and once with an incorrect but plausible one. Table 4.3 reports the mean number of items for which the correct form was recognised (hits), and the incorrect form was correctly rejected (correct rejection), as well as a measure that considers both, by considering only the items for which the correct form was recognised and the incorrect one rejected (we will refer to this measure as 'tot'). An A' score (Stanislaw & Todorov, 1999) for each participant was also computed, to take response bias into account, and means are reported in the table. A' is a sensitivity index which takes misses and false alarms into account, whose scores range between 0 and 1, where 1 corresponds to perfect performance, and .5 corresponds to chance. Scores for target items, real words and control words were compared in the two story conditions. Scores in the two conditions were very similar for real words and control non-words, and were highest for real words, and lowest for control non-words, with the performance on target words in between the two. Scores for target items were higher in the combined condition than in the reading condition.

Table 4.3

Results of the phono-orthographic task by condition

	reading condition		combined condition	
	***M* (*SD*)**	**Range**	***M* (*SD*)**	**Range**
target non-words				
hits	4.29 (1.19)	2 – 6	4.82 (1.24)	1 – 6
correct rejection	3.65 (1.18)	1 – 6	3.85 (1.42)	1 – 6
tot	2.65 (1.20)	0 – 6	3.32 (1.45)	1 – 6
A’	.72 (.18)	.33 – 1.00	.78 (.21)	.16 – 1.00
words				
hits	5.91 (.29)	5 – 6	5.85 (.36)	5 – 6
correct rejection	5.18 (.87)	3 – 6	5.24 (.89)	3 – 6
tot	5.18 (.76)	3 – 6	5.09 (.97)	3 – 6
A’	.96 (.04)	.84 – 1.00	.96 (.04)	.88 – 1.00
control non-words				
hits	4.09 (1.42)	1 - 6	4.12 (1.34)	1 – 6
correct rejection	3.56 (1.21)	1 - 5	3.47 (1.38)	1 – 6
tot	2.62 (1.42)	0 - 5	2.56 (1.38)	0 - 5
A’	.68 (.23)	.16 – .96	.69 (.21)	.23 – .96

Note: hits = items for which the correct phonological form was correctly identified; correct rejections = number of items for which the incorrect phonological form was correctly rejected; tot = number of items for which the correct phonological form was identified and which the incorrect phonological form was rejected; A’ = A’ score.

A repeated measures ANOVA was carried out to confirm the observed trends. This ANOVA used the number of correct responses (both hits and correct rejections) as its dependent variable, and condition (reading *vs.* combined), item type (real words *vs.* target non-words *vs.* control non-words), and type of correct response (accept *vs.* reject) as the independent factors. The analysis highlighted a main effect of item type (F (2, 66) = 95.63, p < . 001), with real words (M = 5.54) recognised more often than target non-words (M = 4.15, p < . 001) and control non-words (M = 3.81, p < . 001), and target non-words recognised more often than control non-words (p = . 043). Type of correct response was also significant (F (1, 33) = 23.76, p < .001), with number of hits (M = 4.85) higher than number of correct rejections (M = 4.16). The effect of condition was not significant, nor were there any interactions (all ps > .200). Thus, this analysis suggested that it was easier for children to correctly accept the correct phonological form of items than to reject the wrong one. The performance for target non-words was better than that for control non-words, confirming that some learning of the link between phonological and orthographic forms had occurred during story reading. Performance for target non-words was also worse than that for words. Thus, the link between the orthographic and phonological forms of target non-words was not yet as well established as the same link for known words. Parallel patterns of results were obtained when ANOVAs were conducted using total number of items correctly responded to (tot) and A’ scores as the dependent variables and condition (reading *vs.* combined) and item type (words *vs.* target non-words *vs.* control non-words), as the independent factors.

No difference between the two conditions was highlighted in this main analysis, but this result is not surprising, given that we did not expect an effect of condition for words or control non-words. More surprising is the lack of interaction between condition and item type (F (2, 66) = 1.51, p = .228), given that we would have expected a difference for target non-words, but not other items. We therefore decided to perform analysis including only the performance on target non-words, to ascertain whether a difference might be present, but not strong enough to be significant in the main analysis. Non-parametric Wilcoxon signed-rank tests were carried for all measures with non-normally distributed differences between scores (hits: D(34) = .172, p = .012; correct rejections: D(34) = .159, p = .030; tot: D(34) = .158, p = .031), while parametric paired sample t-tests were carried out for normally distributed measured (A’: D(34) = .141, p = .084). Children correctly accepted the correct phonological form of the target non-words more readily in the combined than the reading condition (T = 76.00, p = .015), but there was no difference in their ability to reject wrong phonological forms (T = 127.00, p = .525). The difference was also significant when considering both hits and correct rejections (tot) (T = 72.00, p = .007), while the analysis with A’ scores was not significant (t(33) = 1.65, p = .109).

When considering all the analyses, the results therefore suggested that there was a trend for a difference in the performance in the two conditions in some analyses, but, this difference was not evident in the main analyses, where control items (both real words and control non-words) were considered.

4.3.3.2 - Semantic tasks. Table 4.4 reports the results of the semantic tasks for the two conditions. The following measures were computed: for the category recognition task, the total number of correct categories recognised per child per condition, while for the definition production task, the total number of full definitions produced per child per condition, corresponding to the number of target non-words for which a score of 4 was obtained, the total number of target non-words for which at least a correct feature was produced (i.e. at least a score of 1 out of 4 was obtained), and the mean overall score (where each item was scored between 0 and 4). Both tasks were difficult for the children, and means for all the measures were very low. We computed the number of words for which at least a feature was produced in the definition task, since half or more of the words received a score of 0 for each child.

In the category recognition task children recognised on average the correct category for 1 of the items in the reading condition, and 2 of the items in the combined condition, out of 6. Chance performance in this task was set at .75 (the probability of selecting the correct answer from 8 alternatives on 6 trials). Wherever normality assumptions were not met, the appropriate non-parametric analyses were conducted. Performance was significantly better than chance in the combined condition (W = 559.00, p < .001), but only approached significance in the reading condition (W = 409.00, p = .054). Children performed significantly better in the combined than in the reading condition in the category recognition task (Table 4.4).

Table 4.4

Scores in the semantic tasks in the two conditions

	reading condition		**combined condition**		**difference between conditions**	
	***M* (*SD*)**	**Range**	***M* (*SD*)**	**Range**	***T***	***p***
category recognition	1.32 (1.41)	0 – 6	2.09 (1.31)	0 – 5	78.50	.022
definition production						
n. correct full definitions	.53 (.96)	0 – 3	.82 (1.17)	0 – 4	58.00	.123
n. feature production	2.12 (1.75)	0 – 6	2.59 (1.88)	0 – 6	132.00	.161
mean feature production[1]	.78 (.84)	0 – 2.67	.98 (.91)	0 – 3.33	1.53	.136

[1]paired-sample t-test is reported in place of Wilcoxon signed-rank test – the distribution of the differences is normal

Similarly, in the definition production task, not all the children were able to produce one full definition for the target non-words without prompt, and even when prompted the scores remained low. For definition production there was no difference between the two conditions for any of the measures. Nevertheless Figure 4.3 suggests that there was a trend for children to obtain higher scores in the definition production task in the combined condition, although this trend is not significant.

In conclusion, children performed better in the combined condition than in the reading condition in the category recognition task, but not in the definition production task. Nevertheless, given the very low performance in the definition production task, a lack of significant difference is not surprising, and may be due to an overall floor performance in this task, rather than the absence of a presentation effect.

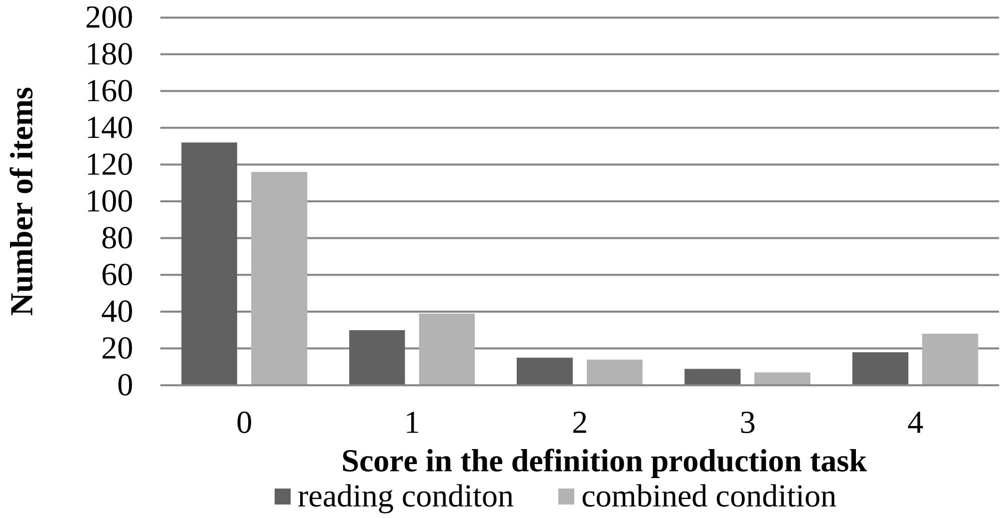


Figure 4.3. Number of target non-words in each score category for the definition production task by condition

4.3.4 – Approach to analysis of eye-movement data

Before analysing how children dealt with the presence of new word forms within the stories, a more general comparison of eye-movement data between the two conditions was carried out, to ascertain whether children were approaching the task differently when reading a story, and when reading and listening to a story at the same time. The first passage was used as a practice trial for the children to become acquainted with the procedure, and eye-movement results from this trial were excluded from all further analysis. The data from all other passages, including those for the second passage, which did not contain any target words, were considered for these preliminary analyses. For both conditions, fixations that were shorter than 80 ms were excluded from the analysis, since such short fixations as unlikely to reflect meaningful processing (see Inhoff & Radach, 1998 for a discussion). This led us to exclude 1.9 % of the fixations in both conditions. No cut-off was applied to fixations longer than usual (fixations longer than 1200 ms formed 0.08 % of fixations in the reading condition and 0.27 % in the combined condition).

For an initial comparison between conditions, mean number of fixations per passage per condition, mean fixation durations per passage per condition, in ms, and saccade amplitudes, i.e. the mean spatial distance between two fixation points, per passage, per condition, in degrees of visual angle, were considered (see Section 4.3.5). We averaged the number of fixations by passage first, then by condition, while we averaged fixation durations by passage for each child separately, then averaging passages by condition. Repeated measures t-tests and Wilcoxon signed-rank tests were used to compare the two conditions for the aforementioned measures.

For the main analyses, on the other hand, we considered eye-movement measures on the specific areas of interest only, specifically the three repetitions of the target non-words, and the definitions. We considered gaze duration, i.e. the time spent in the area before moving to another area, in ms, re-reading time, i.e. the time spent in the area at second-pass, after having left, it in ms, and total reading time, in ms, on the three presentations of the target non-word and the definitions. We averaged measures by child first, and then by condition to obtain these measures. It is important to note that the term “gaze duration” was used more loosely than usual in the analyses for definitions, since we considered time spent on larger areas than just one word, usually not defined as “gaze duration”. We implemented a looser definition of the term to aid comparison with previous studies (Williams & Morris, 2004). In Section 4.3.6 we considered eye movements on the target non-words and definitions without considering children’s performance on the word learning tasks, to analyse whether children attended to the interest areas differently in the two conditions. We also considered eye movements on the three repetitions of the words separately. We used repeated measures ANOVAs to compare the two conditions for the three repetitions of the target non-words, and t-tests to compare the two conditions on looking times on definitions.

In Sections 4.3.7 and 4.3.8 we analysed whether the pattern of eye movements influenced performance on the phono-orthographic task, the category recognition task and the definition recognition task. To ascertain whether word learning in the two conditions was associated with specific strategies for exploring the text, the analyses considered performance on one measure of learning, as the dependent variable. Performance in each task was a dichotomous variable, where items correctly recognised or defined (learnt items) were assigned a score of 1, and items not correctly recognised or defined (unlearnt items) were assigned a score of 0. All dependent measures were binomial variables, and the analyses were conducted using generalised linear mixed models for binomial data (Jaeger, 2008), using the function “glmer” from the package “lme4” (Bates et al. 2014), computed with the software R (R Core Team, 2013). This approach was deemed the most appropriate to consider all possible influencing factors in one analysis. In these models we entered eye-movement measures as predictors of learning. Condition and the interaction between condition and the eye-movement variables were also entered as predictors in these analyses, as were the background measures that were expected to influence children's performance. For these analyses gaze duration, re-reading time, total reading time and regressions to specific areas of interest in the text were considered. We considered gaze duration, re-reading time and total reading time to all three repetitions of the target non-words, and whether regressions were made to the interest area that included the definition. Regressions to the definition were considered in these analyses due to the hypothesis that children would learn the items' meanings by linking definitions and new words. Since no hypotheses were considered for more general analyses for this measure, regressions to the definition were not considered in previous analyses. Number of regressions was not considered, since the occurrence of more than one regression was very rare. Four different models, one for each of the eye-movement measures were considered. We did not enter all the measures in one analysis, since these measures were not independent of each other. For each dependent variable, an initial model included random intercept terms for both participants and items. We then compared these 'empty' models (using pair-wise Likelihood Ratio Test comparisons; Barr, et al., 2013) with models that additionally included the hypothesised fixed effects: condition (combined *vs.* reading), a single eye-movement measure, and background measures that were hypothesized to have an effect on learning. The interactions between background measures and condition were also included, one at a time, in the main fixed effect model and retained only if significant. All continuous factors were centred around the mean. Each of the 34 children provided six responses to the six target non-words in each condition; this was used as the dependent variable in each analysis. When a difference in looking times was highlighted in the models, means and standard deviations were presented to explore these differences further.

4.3.5 - Eye-movement data by story and passage: comparisons between conditions

Table 4.5 reports the mean number of fixations, fixation duration (ms), and saccade amplitude (degrees of visual angle), per passage, per condition. Paired-sample tests were carried out to compare the two conditions. Wherever normality assumptions were not met, the appropriate non-parametric analyses were conducted. As shown in

Table 4.5, children made significantly fewer fixations in the combined condition, but these were significantly longer. The mean distance between two fixations (saccade length) was also significantly longer in the combined condition than in the reading condition. Furthermore, children made significantly more downwards and upwards, but fewer leftward and rightward movements on the text in the combined condition than in the reading condition. Overall, when the oral text was available, children's approach to the written text was different to when they were reading without this support.

Table 4.5

Eye movement differences in the two conditions

	reading condition		**combined condition**		**comparison between conditions**	
	Mean (SD)	**Range**	**Mean (SD)**	**Range**	***T***	***p***
Number of fixations						
Total (per trial)	183 (39.05)	110 - 259	162 (12.15)	138 - 186	131.00	.004
Rightward (per story)[1]	769 (142.34)	468 - 1099	668 (74.60)	531-816	3.90	< .001
Leftward (per story)	359 (103.51)	217 - 590	323 (55.96)	225 - 435	177.00	.039
Upward (per story)	13 (10.24)	4 - 52	18 (9.66)	6 - 46	90.00	< .001
Downward (per story)	12 (9.09)	2 - 43	20 (8.65)	6 - 35	75.00	< .001
Fixation duration[1]	240 (23.44)	196-298	256 (20.72)	222 - 300	-6.14	< .001
Saccade amplitude[1]	3.30 (.48)	2.35 – 4.37	3.46 (.35)	2.88 – 4.27	-2.63	.013

[1]t-test was computed since the differences between measures were normally distributed

4.3.6 - Eye-movement data by interest areas: comparisons between conditions

Table 4.6 presents eye movement measures for interest areas. For these comparisons we considered gaze duration, i.e. the time spent in the area before moving to another area, in ms, re-reading time, i.e. the time spent in the area at second-pass, after having left it in ms, and total reading time, in ms, on the three presentations of the target non-word and the definitions. Children spent more time on the target non-words in the reading condition than in the combined condition, especially considering gaze duration and total reading time. This difference was particularly noticeable on the first presentation of the target non-word. Furthermore, time spent on the target non-word seemed to diminish with exposure, with the first presentation of the target being fixated longer than the second and the third presentation, in all three measures considered. No difference is seen between the conditions in reading time for definitions.

Table 4.6

Means (and Standard Deviations) for gaze duration, re-reading times and total reading times (in ms) on the three presentations of the target non-words and the definitions

		reading condition *M (SD)*	**combined condition** *M (SD)*
Gaze Duration	target 1	927.53 (407.24)	681.5 (224.83)
	target 2	580.03 (219.21)	555.35(175.97)
	target 3	546.05 (165.19)	550.36 (168.07)
	definition	1194.64 (414.37)	1162.28 (352.90)
Re-reading Time	target 1	491.55 (423.42)	480.40 (346.66)
	target 2	372.15 (318.25)	288.19 (239.97)
	target 3	169.09 (233.88)	233.97 (231.63)
	definition	630.94 (590.86)	615.24 (500.93)
Total Reading Time	target 1	1397.11 (570.55)	1138.70 (313.73)
	target 2	937.25 (387.72)	809.51 (302.36)
	target 3	678.80 (253.71)	759.42 (254.77)
	definition	1796.00 (439.92)	1777.51(432.01)

Repeated measures ANOVAs were carried out to confirm the trends observed in relation to time spent on the target words. Three ANOVAs used gaze duration, re-reading times and total reading times respectively as their dependent variable, and condition (reading *vs.* combined) and presentation of the target (first *vs.* second *vs.* third) as the independent factors. The analysis for gaze duration highlighted a main effect of condition (F (1, 33) = 8.10, p = . 008), and presentation of the target (F (2, 66) = 28.43, p < . 001), and an interaction between these (F (2, 66) = 5.80, p = . 005), which confirmed that gaze durations were shorter on target non-words in the combined than reading condition, and that there was a decrease in gaze duration from first to second ($p_{reading} < .001$; $p_{combined} = .048$) and third ($p_{reading} < .001$; $p_{combined} = .035$) presentations of the target non-word in both conditions, but no difference between second and third presentation ($p_{reading} > .999$; $p_{combined} > .999$). It also highlighted a steeper decrease in gaze duration for the reading than the combined condition, with the difference between the two conditions significant only for the first presentation of the words ($p_{target\ 1} = .003$; $p_{target\ 2} > .616$; $p_{target\ 3} = .902$). See Figure 4.4 for a representation of the interaction.

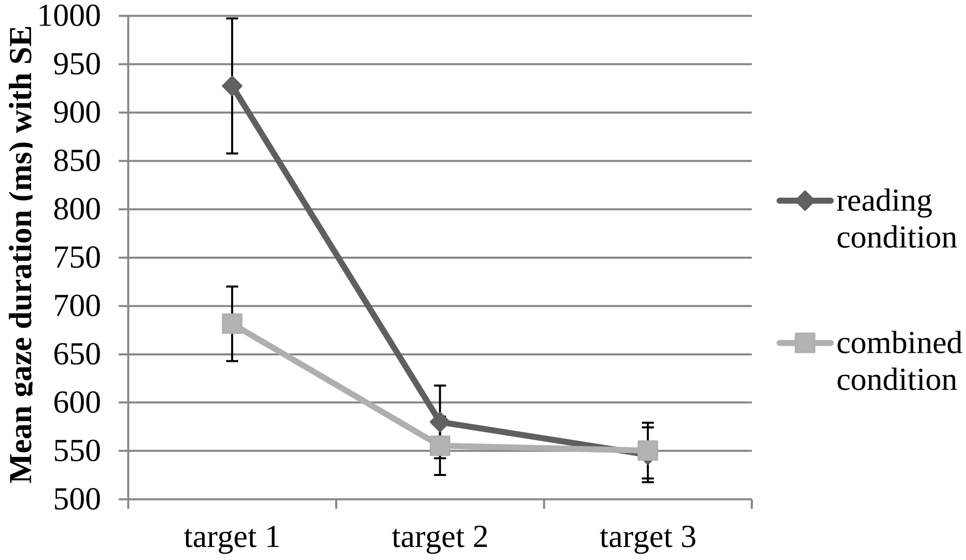


Figure 4.4. Mean gaze duration on the three repetitions of the target non-words by condition

The analysis for re-reading time only highlighted a main effect of presentation of the target (F (2, 66) = 25.38, p < . 001), with re-reading time decreasing from first to second to third presentation (target 1 *vs*. target 2: F (1, 33) = 11.99, p = . 001; target 1 *vs*. target 3: F (1, 33) = 30.72, p < . 001; target 2 *vs*. target 3: F (1, 33) = 24.26, p < . 001), and no effect of condition (F (1, 33) = .31, p = . 861) or interaction (F (2, 66) = 1.38, p = . 260).

The analysis for total reading time also highlighted a main effect of presentation of the target (F (2, 66) = 70.91, p < .001), and an interaction (F (2, 66) = 6.07, p = .004), but no main effect of condition (F (1, 33) = 2.60, p = .117). The interaction showed a significant difference between the conditions at the first presentation of the target (p = .010), but not at the second (p = .183) or third (p = .208). The interaction also highlighted a marked decrease in reading time throughout the presentations for the reading condition (all *p*s < .001). On the other hand, in the combined condition, total reading time was higher in the first presentation compared to subsequent presentations (both *p*s < .001), but there was no decrease between the second and the third presentation (p = .986). See Figure 4.5 for a depiction of the interaction.

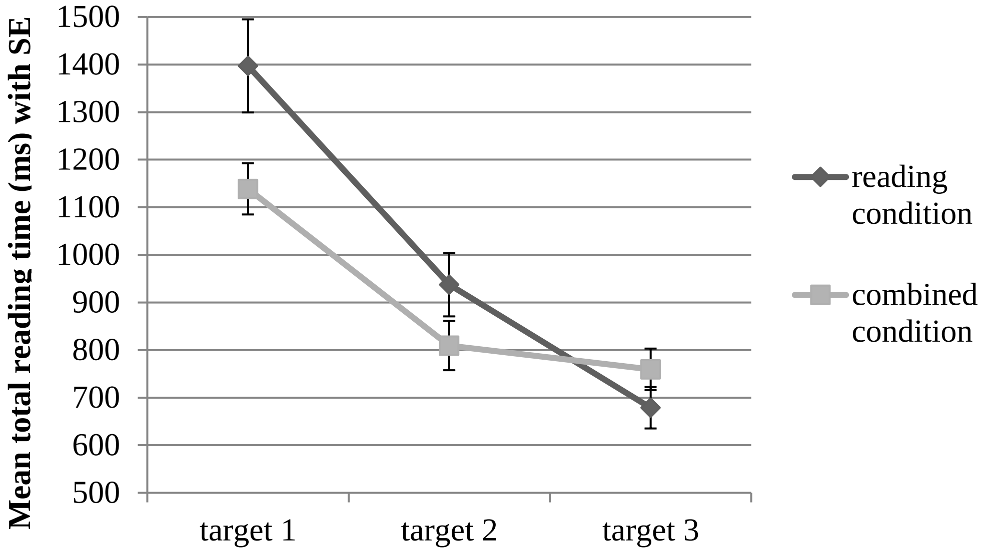


Figure 4.5. Mean total reading time on the three repetitions of the target non-words by condition

We can therefore conclude that, overall, children spent less time on the target non-words in the combined condition, significantly so in all three presentations when considering gaze duration, but only significantly so at the first presentation of the word when considering total reading time, and overall they spent less time on the items the second and third time they were presented.

When considering definitions, t-tests were carried out to compare reading time in the two conditions. Differently from the analyses for target non-words, the analyses for looking time at definitions did not highlight any difference between the two conditions ($t_{gaze\ duration}$ (33) = .45, p = .651; $t_{re\text{-}reading\ time}$ (33) = .14, p = .890; $t_{total\ reading\ time}$ (33) = .18, p = .861).

4.3.7 - Eye-movement data by interest area: predictors of word learning in the phono-orthographic task

The dependent measure used in analyses of performance on the phono-orthographic task was the number of items for which participants responded correctly to both phonological forms (the 'tot' measure, see section 4.3.3.1). For this task the models of interest included: condition (combined *vs.* reading), an eye-movement measure, vocabulary score, the composite reading accuracy score and non-verbal ability. Since it was hypothesised that time spent reading the target non-words would predict learning of the link between phonology and orthography, three main models were considered: a model that included gaze duration to all three repetitions of the target non-words within a passage, a model that included re-reading time to all three repetitions of the target words, and a model that considered total reading time. We further considered models that only included the first repetition of the target non-words, to ascertain whether time spent looking specifically at the first repetition of the item significantly affected learning of the link between orthography and phonology. These models showed the same patterns of results of the main models, and are therefore not reported. Models are reported in Table 4.7. Two further factors were included in a third set of analyses, namely tot scores for both control non-words and real words. These factors were included to control for participants' baseline ability to match phonological and orthographic forms, in light of the results of previous analyses for the phono-orthographic task (section 4.3.3.1). It was expected that controlling for performance in control non-words and real words would eliminate any condition effect in this task, as this was the case for the ANOVAs in Section 4.3.3.1. The patterns of the two sets of analyses were similar, and performance in control non-words and real words were not significant predictors (see Appendix W).

Table 4.7

Generalized linear mixed models for accuracy in the phono-orthographic task exploring the role of looking times towards the three repetition of the target non-words

Factors	**Model 1: gaze duration**		**Model 2: re-reading time**		**Model 3: total reading time**	
	Estimate	**p-value**	**Estimate**	**p-value**	**Estimate**	**p-value**
(intercept)	-.28	.178	-.28	.178	-.28	.161
gaze duration	.01	.846				
re-reading time			.04	.660		
total reading time					.05	.578
condition	.54	< .001	.54	< .001	.55	< .001
reading accuracy	.50	< .001	.50	< .001	.49	< .001
vocabulary	-.05	.661	-.06	.658	-.07	.616
non-verbal abilities	.20	.124	.21	.116	.23	.089
gaze duration* condition	-.05	.724				
re-reading time* condition			-.03	.838		
total reading time* condition					-.05	.727
Final Model *vs.* Empty Model	$\chi^2(6) = 82.20$, $p = .001$		$\chi^2(6) = 82.28$, $p < .001$		$\chi^2(6) = 36.68$, $p < .001$	

As it is possible to see from Table 4.7 all models highlighted a significant effect of condition. These results differed from the results of previous ANOVAs (section 4.3.3.1), and suggested that children were more likely to acquire a stable link between phonology and orthography in the combined than in the reading condition. The models also highlighted a positive effect of reading accuracy. The inclusion of performance on control non-words in the models reduced the effect of reading accuracy, but did not eliminate it completely (see Appendix Q): it is possible that performance on control non-words and real-words in this task was also associated with higher reading abilities, and this could explain the shared variance between these factors. Gaze duration, re-reading time and total reading time did not predict performance on this task.

4.3.8 - Eye-movement data by interest area: predictors of word learning in the semantic tasks

4.3.8.1 - Category recognition task. For the models for the category recognition task the following effects were included: condition (combined *vs.* reading), an eye-movement measure, and background measures which were hypothesized to have an effect on semantic learning (vocabulary, reading accuracy, reading comprehension and non-verbal abilities). The interaction between background measures and condition were also included one at a time in the main fixed effect model and retained only if significant. Three main models were computed: a model than included gaze duration to all three repetitions of the target non-words within a passage, a model that included re-reading time to all three repetitions of the target non-words, and a model that included total reading time. The three models for recognition of the correct category for target non-words are reported in Table 4.8.

Table 4.8

Generalized linear mixed models for accuracy in the category recognition task considering gaze duration and re-reading time on the three repetitions of the target non-words

Factors	**Model 1: gaze duration**		**Model 2: re-reading time**		**Model 3: total reading time**	
	Estimate	**p-value**	**Estimate**	**p-value**	**Estimate**	**p-value**
(intercept)	-1.61	< .001	-1.57	< .001	-1.59	< .001
gaze duration	.22	.011				
re-reading time			-.13	.255		
total reading time					.12	.194
condition	.85	< .001	.82	< .001	.84	< .001
reading accuracy	.14	.404	.13	.455	.13	.443
vocabulary	.62	.001	.63	.001	.62	.001
reading comprehension	.19	.326	.18	.352	.19	.324
non-verbal abilities	-.14	.377	-.14	.365	-.13	.416
gaze duration* condition	-.35	.024				
re-reading time* condition			.24	.118		
total reading time* condition					-.08	.583
vocabulary* condition	-.40	.009	-.40	.008	-.40	.007
Final Model *vs.* Empty Model	$\chi^2(7) = 56.56$, $p < .001$		$\chi^2(7) = 51.48$, $p < .001$		$\chi^2(8) = 50.70$, $p < .001$	

All three models for category recognition highlighted a significant effect of condition. This result was in line with the results in Section 4.3.3.2, and suggested that children were more likely to recognise the correct category for target non-words in the combined than in the reading condition. All models also highlighted a significant effect of vocabulary and an interaction between vocabulary and condition, with vocabulary specifically affecting semantic learning in the reading condition, and the combined presentation particularly supporting word learning in children with low vocabulary (see Figure 4.6). Gaze duration also predicted category learning, while re-reading time and total reading time did not predict learning. Furthermore, the interaction between condition and gaze duration was significant, showing that first pass reading time predicted learning differently in the two conditions. Children spent more time looking at learnt non-words than unlearnt non-words in the reading condition, but this difference was less evident in the combined condition (see Figure 4.7). This interaction suggests that looking longer at target non-words was particularly beneficial for vocabulary learning in the reading condition, while learning in the combined condition was less associated with looking time.

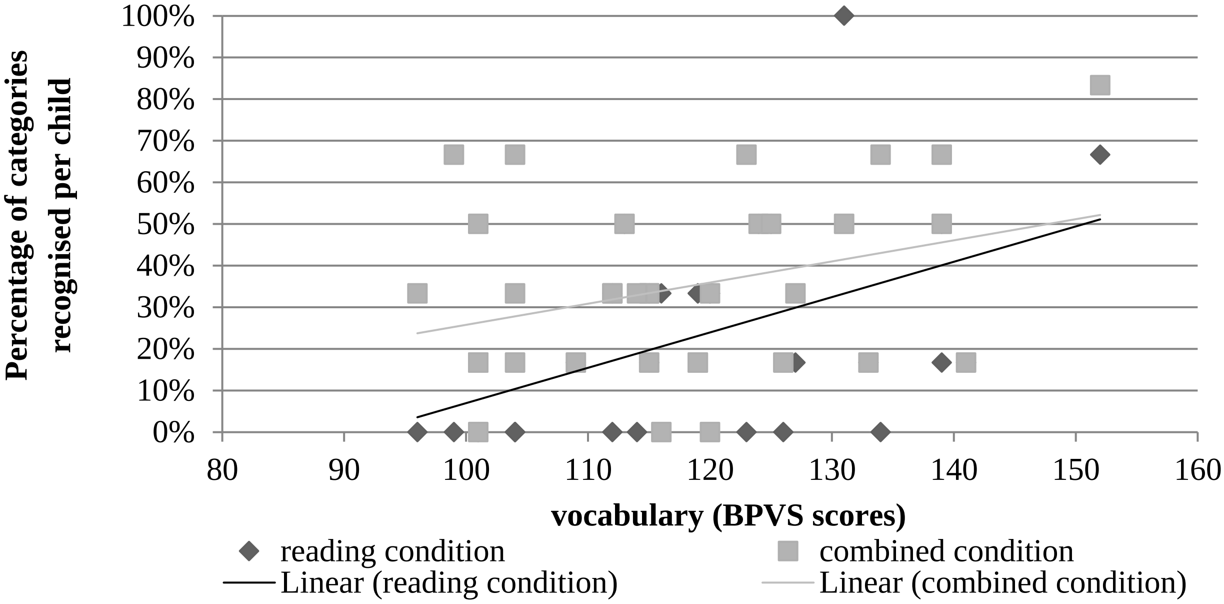


Figure 4.6. Percentage of correct category recognitions by condition and vocabulary.

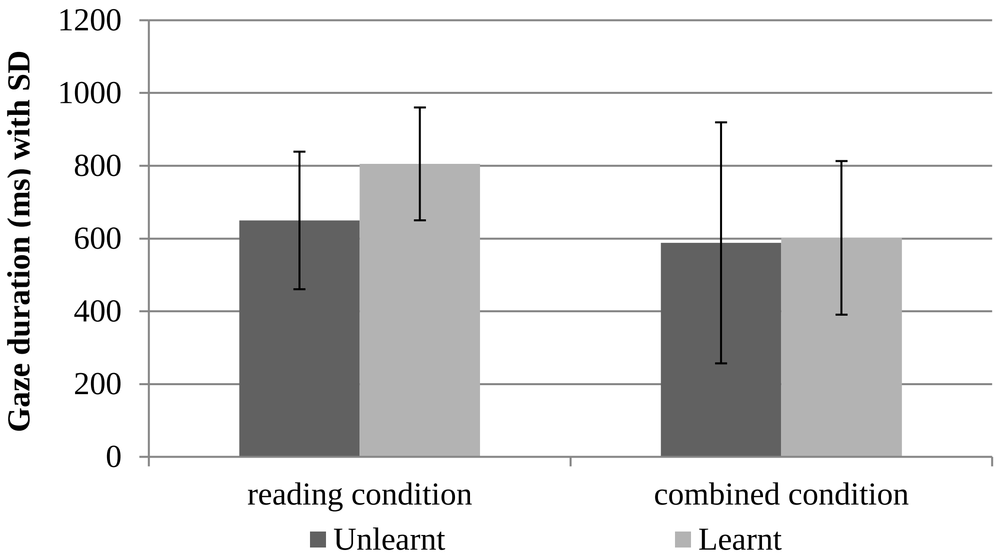


Figure 4.7. Mean gaze duration by condition and semantic learning (category recognition task).

A second hypothesis regarding semantic learning was tested, namely that children might learn a word's meaning by making a regression from the target non-word to the definition. For this reason we computed a model including a measure of regression between words and definitions (the measure considered whether or not children made a regression to the interest area of the definition). This analysis is reported in Table 4.9, and showed that probability of making a regression to the definition did not predict category learning.

Table 4.9

Generalized linear mixed model for accuracy in the category recognition task exploring the effect of regression to the definition

Factors	**Estimate**	**p-value**
(intercept)	-1.47	< .001
regression	1.49	.141
condition	.77	.001
reading accuracy	.15	.355
vocabulary	.38	.018
reading comprehension	.18	.352
non-verbal abilities	-.09	.538
regression*condition	-1.87	.161
Final Model *vs.* Empty Model	$\chi^2(7) = 25.59$, $p = .001$	

Overall, these analyses show that children learned more categories for items in the combined than in the reading condition. Gaze duration on target items significantly predicted semantic learning: specifically, children spent more time looking at the learnt items than the unlearnt items in the reading condition, while this difference was less evident in the combined condition. Vocabulary also predicted category recognition, with children with larger vocabularies learning more new item meanings, especially in the reading condition. Regressions to the definitions, on the other hand, did not predict semantic learning.

4.3.8.2 - Definition production task. For the definition production task items were scored on a scale from 0 to 4 depending on how much information children were able to produce regarding the meaning of the word, and how many clues they needed to produce a definition. Nevertheless, since half or more of the items received a score of 0, for the following analyses a binomial scoring system was used, dividing items for which children did not produce any information (items that originally had a score of 0), from items for which at least some information was produced (items with a score higher than 0). Analyses were conducted using generalised linear mixed models for binomial data. The same factors used for the analyses for category recognition were entered in the analyses for definition production and the same procedure was followed. Analyses including gaze duration, re-reading time and total reading time to all three repetitions of the target non-words, as well as regressions to the definitions were considered. These analyses are reported in Tables 4.10 and 4.11.

Table 4.10

Generalized linear mixed models for accuracy in the definition production task considering gaze duration, re-reading time and total reading time on the three repetition of the target non-words

Factors	**Model 1: gaze duration**		**Model 2: re-reading time**		**Model 3: total reading time**	
	Estimate	**p-value**	**Estimate**	**p-value**	**Estimate**	**p-value**
(intercept)	-.77	.007	-.76	.008	-.77	.008
gaze duration	.10	.269				
re-reading time			.04	.722		
total reading time					.12	.221
condition	.46	.002	.46	.001	.48	.001
reading accuracy	-.02	.923	-.03	.885	-.03	.886
vocabulary	.82	< .001	.83	< .001	.82	< .001
reading comprehension	.71	.014	.69	.014	.70	.014
non-verbal abilities	-.30	.138	-.30	.139	-.28	.168
gaze duration* condition	-.23	.160				
re-reading time* condition			.19	.208		
total reading time* condition					.04	.780
reading comprehension* condition	-.35	.031	-.35	.029	-.37	.024
Final Model *vs.* Empty Model	$\chi^2(8) = 37.17$, $p < .001$		$\chi^2(8) = 44.30$, $p < .001$		$\chi^2(8) = 43.05$, $p < .001$	

Table 4.11

Generalized linear mixed models for accuracy in the definition production task considering the effect of regression to the definition

Factors	**Estimate**	**p-value**
(intercept)	-.76	.003
regression	-.02	.989
condition	.43	.072
reading accuracy	.02	.920
vocabulary	.80	< .001
reading comprehension	.41	.108
non-verbal abilities	-.28	.080
regression*condition	-.05	.750
Final Model vs. Empty Model	$\chi^2(7) = 30.68$, $p = .001$	

As it is possible to see from Table 4.10 all models for definition production highlighted a significant effect of condition. These results were in line with the results of the analyses carried out for category recognition, where a significant difference between the conditions was highlighted. On the other hand, these results differ from those reported in Section 4.3.3.2, which did not show any significant difference between the conditions for definition production. The lack of difference in the analyses in Section 4.3.3.2 could be attributed to the overall floor performance in the task, which would be expected to have a smaller impact on the mixed models, which consider all items separately. All models also highlighted a significant positive effect of vocabulary and reading comprehension, and an interaction between reading comprehension and condition, with low reading comprehension having a more negative impact in the reading condition than in the combined condition (see Figure 4.8). Neither gaze duration, nor re-reading time nor total reading time on the target non-words predicted definition production. The model that included regressions to the definition (Table 4.11) only highlighted a significant effect of vocabulary knowledge, and only a trend towards a difference between conditions. No effect of regression to the definition was identified. This indicates that looking back at the definition did not significantly predict semantic learning in this study. Furthermore, the lack of an effect of condition in one of the analyses for definition production (Table 4.11) suggested that this result might be unstable, and the difference between conditions might not be reliable.

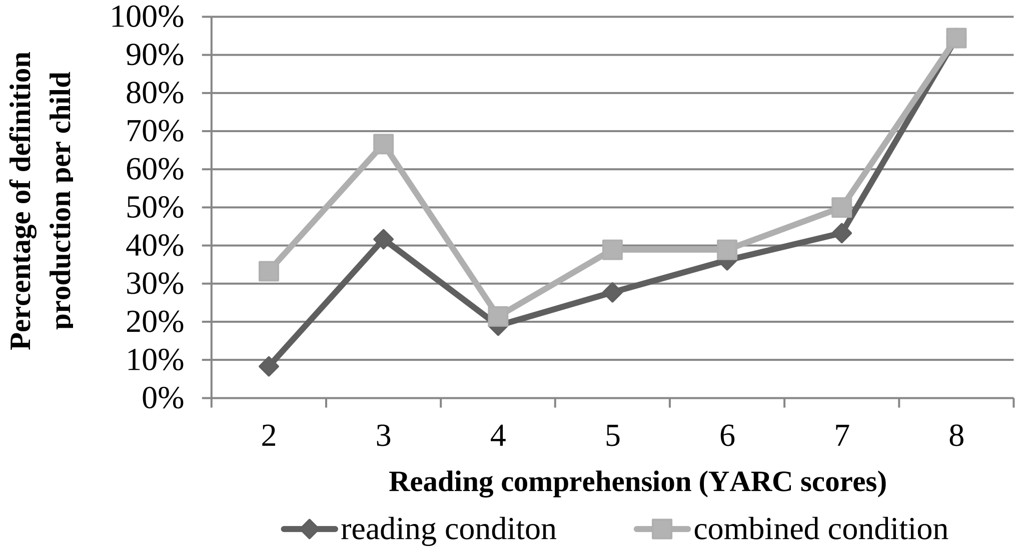


Figure 4.8. Percentage of target non-words for which a semantic feature was produced by condition and reading comprehension.

Overall, these analyses show that children were more likely to produce definitions for items in the combined than in the reading condition, although this difference was not significant in all analyses. Eye-movement measures and time spent on the target items and regressions to the definitions failed to predict semantic learning, when definition production was considered. Vocabulary predicted definition production in all models, with children with larger vocabularies producing more definitions. Reading comprehension also predicted definition production, especially in the reading condition, although this effect was not significant in all models.

4.3.9 - Analyses of coincident and cross-coincident time for the combined condition

During the combined condition children could hear the words and definitions spoken as well as read them. It was therefore hypothesized that looking time at the specific areas of interest in the text at the same time the oral text was heard would explain word learning. It was hypothesized that children would form a stronger link between orthographic and phonological forms of the non-words when looking at the non-word while hearing it spoken (coincident times). It was further hypothesized that they would learn the meaning of the target non-words more easily if they looked at the word while hearing either the word itself (coincident times), or the definition or the clues (cross-coincident time). Similarly they were expected to learn meanings better if they were reading the definition while hearing the word (cross-coincident time).

The analysis of whether coincident time predicted performance on the phono-orthographic task is reported in Table 4.12. The analysis included coincident time on the three repetitions of the words. Total coincident time spent on the target non-words did not predict performance on the phono-orthographic task. As expected from previous analyses, reading accuracy significantly predicted performance in this task. This analysis confirms the previous results showing no relationship between eye movements and the learning of phono-orthographic mappings.

Table 4.12

Generalized linear mixed model for accuracy in the phono-orthographic task in the combined condition considering total coincident time spent on the three repetitions of the target non-words

Factors	**Model**	
	Estimate	**p-value**
(intercept)	.35	.240
total coincident time	-.02	.784
reading accuracy	.78	.001
vocabulary	< .01	.991
non-verbal abilities	.13	.568

The analyses for category recognition and definition production are reported in Table 4.13. These analyses included coincident time on the three repetitions of the target non-words and cross-coincident time between target non-words and semantic clues, including definitions. For this second measure we considered both total time spent on the non-words while hearing the definition or a clue, and time spent on a definition while hearing the target non-words. Total coincident or total cross-coincident time did not have an effect on semantic learning. It seems that children were not building better semantic knowledge by looking at the definition while hearing the words and vice-versa.

Table 4.13

Generalized linear mixed models for accuracy in the semantic tasks in the combined condition considering total coincident time spent on the three repetition of the target non-words

Factors	**Model 1: category recognition coincident**		**Model 2: category recognition cross-coincident time**		**Model 3: definition production coincident time**		**Model 2: definition production cross-coincident time**	
	Estimate	**p**	**Estimate**	**p**	**Estimate**	**p**	**Estimate**	**p**
(intercept)	-1.00	.009	-1.12	.008	-1.00	.009	-1.13	.008
total coincident time	.01	.981			.01	.982		
total cross-coincident time			-.01	.675			-.01	.675
reading accuracy	.12	.703	.14	.721	.12	.703	.14	.721
vocabulary	.23	.451	.22	.548	.23	.451	.22	.548
reading comprehension	.31	.411	.40	.369	.31	.411	.40	.369
non-verbal abilities	-.10	.725	-.10	.780	-.10	.725	-.10	.780

4.4 - Discussion

In this study children were exposed to new word forms (non-words), presented in stories. Stories were presented only once, and target non-words were repeated three times within each story. Several predictions were made. Firstly it, was hypothesized that, similarly to our previous studies (Chapter 2 and 3), children would learn word forms similarly in the combined and reading condition, but they would learn the meaning of the new items more proficiently in the combined than the reading condition. The results of the learning tasks were reported in Section 4.3.3, and will be discussed first. Secondly, it was expected that children would attend to the text differently in the two conditions, given the addition of the oral presentation in the combined condition. This hypothesis was explored in Section 4.3.5 and will be discussed in Section 4.4.2. A third hypothesis concerned attention to the target items in the two conditions, and the pattern of eye movements for learnt and unlearnt words. When considering learning of the link between orthographic and phonological forms, we expected longer gaze duration on learnt compared to unlearnt words (see Section 4.1.2), while, on measures of semantic learning, we expected longer re-reading times and total reading times, and possibly shorter gaze duration on learnt compared to unlearnt words (Godfroid et al., 2013; Williams & Morris, 2004). These hypotheses were explored in Sections 4.3.7 and 4.3.8, and will be considered in Section 4.4.3. A specific pattern of fixation was also explored for the combined condition: it was expected that the time spent reading the word while the word was presented orally (coincident time), and the time spent reading the word or the definition when the word or the definition was presented respectively (cross-coincident time), would affect learning in this condition. This hypothesis was explored in Section 4.3.9, and will be discussed in Section 4.4.4. Section 4.4.5 considers the effect of individual differences on vocabulary acquisition, as explored by the mixed models (Sections 4.3.7 and 4.3.8).

4.4.1 – Vocabulary acquisition results

In line with the findings of studies 1 and 2, it was expected that learning of the link between orthographic and phonological forms would be similar in the reading and the combined condition, while semantic learning would be better in the combined condition, at least on some of the measures. The results broadly confirm the hypotheses.

Children were able to learn the link between word forms quite proficiently in both conditions, even given the small amount of exposure (they were only exposed to the new words three times). A significant difference between the conditions, favouring the combined condition was highlighted in some of the analyses, especially the mixed models (section 4.3.7), and analyses that considered target non-words only, but not in the ANOVAs that also included control words and non-words (section 4.3.3.1). In this study, children demonstrated that they had learned the link between orthography and phonology in both conditions, performing better on target words than control non-words. On the other hand, their performance for target words was worse than for real words, for which they presumably had more stable representations. The results, therefore, confirm that children tend to spontaneously build a phonological representation when encoding new words from the written text (Share, 1995), creating a mental representation that includes the oral form of a word. However, when children hear and read the words at the same time, they are advantaged in creating a link between oral and written representations in memory, and this may result in a representation of higher quality (Perfetti & Hart, 2002).

In the semantic tasks, learning the new words meanings was quite hard for children, and their performance was better than chance only when they were presented with the stories in both the oral and written forms simultaneously (section 4.3.3.2). Performance on the category recognition task was significantly better in the combined than the reading condition in all analyses (section 4.3.3.2 and 4.3.8.1). On the other hand, for definition recognition, the combined condition improved performance only in the mixed models (section 4.3.8.2).These results are in accordance with the results reported in Chapter 3. As discussed in Chapter 3, the fact that this difference is especially evident in the category recognition task could be because recognising the category of a word requires a higher level of abstraction than the other tasks. Alternatively, given that performance in all semantic tasks, but particularly in the production task, was overall quite low, floor performance may have had a negative impact on the ability to detect differences between conditions in the definition production task.

4.4.2 - Allocation of attention in the two conditions

As expected, participants explored the texts differently, depending on the presence or absence of the oral information. Specifically, similarly to adults reading aloud (Inhoff et al., 2011; Laubrock & Kliegl, 2015), in the combined condition children made fewer but longer fixations than in the reading condition, and they also made longer saccades. The majority of fixations moved horizontally (mostly from left to right) in both conditions, showing a pattern similar to normal reading. Children moved their eyes upwards and downwards significantly more often in the combined than the reading condition, resonating with the eye movements of children of similar age in shared story-book reading studies (Roy-Charland et al., 2007). In Roy-Charland et al.'s study, for example, children in grade 3 and 4 made more than 70% of horizontal or reading-like saccades, and 20 to 30% of non-reading like ones. Like adults in similar situations (i.e. reading aloud), children pay attention to the text and tend to read it, even when the text is simultaneously narrated. Nevertheless, in this condition, they also tend to skim through the text more often, making more upwards and downward movements than in the reading condition. Since, to our knowledge no similar research has been carried out with adults, it is difficult to know whether adults would show a similar pattern. It is possible that adults would react to the two modalities similarly to children. When the text is heard as well as read, participants, both children and adults, might feel less bound to the written text, since the oral presentation gives them the same information. On the other hand, since adults might rely more on the written text (Brown et al., 2008; Sydorenko, 2010), they might attend to the text similarly in the two conditions. Some hypothesis on the specificity of this effect in children may be also proposed. Children might pay more attention to the oral presentation if they are more familiar with this presentation modality. Attending to the oral presentation may also be less taxing for children, since it does not require them to rely on their developing reading skills. The fact that younger children spend less time looking at the text in shared story reading situations, and do so more often for stories at their reading level than for more difficult stories (Roy-Charland et al., 2007), supports this hypothesis. On the other hand, children might find linking two streams of information more difficult, due to their developing executive functions (Altemeier, Abbott, & Berninger, 2008), leading them to rely on one modality more than the other. Future research could tease these hypotheses apart. First, these two different ideas might be easily tested by exposing adults to a written and written and oral presentation, monitoring their eye movements and comparing their reading strategies. Second, the presentation of partially degraded information in either the reading or the listening modality, when the two are presented simultaneously, might clarify whether children and/or adults tend to rely on one modality more than the other. Third, children's executive functions might be measured as a factor in their different reading strategies in the two conditions.

Similarly to previous research with adults reading new words in context (Joseph et al., 2014), children in our study showed shorter reading times the second and third time they encountered the new words both in terms of their initial gaze duration on each word, re-reading time and total reading time. Specifically, children showed shorter gaze duration the second and third time they read the target non-word, compared to the first time they encountered it in both conditions, and they showed decrease in total reading time from the first to the second to the third presentation in the reading condition. On the other hand, in the combined condition, total reading time was lower for the second and third presentation compared to the first, but no difference was found between second and third presentation. This suggests that the target non-words were processed more easily at each encounter, in line with previous findings with adults (Joseph et al., 2014). On the other hand, total reading time in the combined condition showed a decrease only from the first to the second presentation, possibly suggesting faster integration of the word in the lexicon in this condition.

4.4.3 – Influences of allocation of attention on word learning in the two conditions

It was expected that gaze duration would predict learning of the link between orthographic and phonological forms, with learnt words fixated longer than unlearnt words. This hypothesis was based on the assumption that, by looking longer at the new words, children would better establish the link between the two forms in memory. This assumption was based on the finding that the ease of access to phonological features influences reading time (Pollatsek et al., 1992): since reading time is influenced by ease of access to the phonological form, it might also predict the formation of a phonological representation. This prediction was not confirmed by the data for either presentation modality. No reading time measure predicted learning of the link between orthographic and phonological forms. This result suggests that children automatically and rapidly create a phonological representation for written words, even in the reading condition.

The process of creating a link between orthography and phonology appears to have been more effortful in the reading condition, with children spending more time on the target non-words in this condition compared to the combined condition. Nevertheless, this did not translate in longer looking times on learnt words. Looking longer at the new words does not facilitate the acquisition of the link between orthography and phonology of new words. The lack of an effect of reading times was consistent across all measures, including first and second pass reading times. Children therefore spent more time on the target non-words in the reading than the combined condition, but looking time at the target non-words did not predict phono-orthographic learning.

When learning is considered in terms of gains in semantic knowledge, adults in previous research have shown shorter gaze duration, but longer re-reading times and total reading times on learnt compared to unlearnt words (Williams & Morris, 2004). This eye-movement pattern might reflect a tendency to try to form a representation of the meaning of the word in context, at the expense of paying attention to the word's form itself. This translated in longer reading time on learnt words only after having left it the first time, to explore the context and its meaning. On the other hand, adult second language learners showed longer total reading time on learnt words, but no opposite effect on gaze duration (Godfroid et al., 2013). We expected children to behave similarly to adults, showing longer second pass times (either re-reading or total reading time), and, possibly, shorter gaze duration on learnt compared to unlearnt words. Contrary to our predictions, children's gaze duration on the target non-words were longer when the items were learnt than not learnt, while re-reading time and total reading time between learnt and unlearnt items were similar: re-reading and total reading time did not predict semantic learning. This pattern was particularly marked for the reading condition. Contrary to our predictions of the importance of linking the new word with its meaning, whether or not children made a regression to the definition did not impact semantic learning. It appears, therefore, that initial time spent on the new words helped children better establish their meanings in memory. This result is particularly striking. Children spent more time on words they were going to learn the meaning of, at first pass, even before moving to other part of the text that gave information regarding the meaning (i.e. longer gaze duration). This result appears to bear some similarities to that found for adults learning new words in their second language (Godfroid et al., 2013): in that study, in fact, the researcher found a trend for longer gaze duration for learnt than unlearnt words. Nevertheless, our results differ from those on adults L2 learners, since we did not find any effect for total reading time.

Two explanations for the difference between adults and children's exploration of the text occur to us: either children and adults differ in the way they explore text, or the difference is attributable to the material used in our study compared to the Williams and Morris' (2004) one. Considering the first hypothesis, compared to adults, children might pay more attention to the form of a new word, rather than its meaning, especially while reading. In the present study new words were fixated multiple times, and for long periods: children were probably trying to encode the new word forms in memory, and encoding a new form may have taken precedence over determining a word's meaning. Heightened initial processing time on learnt words might therefore reflect the child's efforts to encode the word. Children might be more used than adults to encountering new words in the text, and their strategy may be more focussed on decoding the specific items' forms, rather than building a more general representation of the text by exploring word meanings.

The second account posits a methodological confound in the nature of the passages presented in our study versus previous studies with adults where single sentences are presented one at a time (Williams & Morris, 2004). Previous research has found different reading patterns for sentences and paragraphs (Radach, Huestegge, & Reilly, 2008; Wochna & Juhasz, 2013), finding that new or low frequency words had longer gaze durations but shorter re-reading times when encountered in sentences compared to passages (although this second result was not always replicated, c.f. Wochna & Juhasz, 2013). Furthermore, informative context was fixated longer in sentences than paragraphs. Adults, therefore, tend to focus their attention on new words and information regarding their meaning when these are presented in sentences rather than paragraphs, presumably due to the new words' perceived importance for comprehension in sentences compared to longer texts (Freebody & Anderson, 1983).

However, the difference between reading strategies to sentences and paragraphs is unlikely to account for the findings of the current study. Godfroid et al. (2013), for instance, presented multiple sentences to adult readers, but still found effects mainly at second pass (albeit in total reading time rather than re-reading). Furthermore, the research previously presented shows that lengthening the contextual presentation (sentences *vs.* passages) affects gaze duration negatively (shorter gaze duration) and re-reading positively (longer re-reading time), and learning the words (learnt *vs.* unlearnt) has the same effect. Given that our research used paragraphs we should have found even shorter gaze durations and longer re-reading times compared to the sentences in Williams and Morris (2004) study. Since both learning and longer material tends to have the same effects, it should have been even more likely for us to replicate the results of Williams and Morris (2004). Presenting passages, rather than sentences should not have changed the direction of the effect for gaze duration, since the evidence does not support such an effect.

While children did not need to spend more time on the words to learn their forms, they did spend more time on them to learn their meanings, especially in the reading condition. Extracting the meaning from text is an effortful process, and the results show that children learn meanings better if they look at the words longer. It seems that initial effort in attending to new items, before exploring the rest of the text, assists children's semantic encoding. On the other hand, in the combined condition, where the phonological forms of the words were directly provided, children did not need to spend as much extra time on the words to learn their meaning.

4.4.4 – Effects of coincident and cross-coincident looking times on word learning in the combined condition

It was hypothesized that in the combined condition, coincident (oral word – written words) or cross-coincident (oral word – written definition or oral definition – written word) looking times might predict learning. This prediction was based on two previous findings. First, adults, while reading, tend to learn meanings of new words by regressing back from the meaningful content to the word (Williams & Morris 2004). Children might similarly try to connect words and contexts when learning word meanings. Second, as children have a tendency to link oral stories to relevant pictures when learning new vocabulary (Evans & Saint-Aubin, 2013), they might also be able to make meaningful links between oral and written presentations. It was therefore expected that, rather than making regressions from the *written* context to the *written* word, children would link words to their meanings by looking at the *written* word or context while listening to the *oral* context or word. Specifically, it was hypothesized that more time spent looking at the word while listening to its oral form would support learning of the link between orthography and phonology, while more time spent on the word or its context, while the context or the word were spoken aloud, would elicit higher levels of semantic learning. Neither prediction was confirmed: coincident looking times did not predict learning of the link between orthography and phonology, and cross-coincident looking times did not predict semantic learning.

In summary, the process of encoding new word forms and meanings was less effortful in the combined condition than in the reading condition: children did not need to spend as much time on the words, and they learnt them more efficiently. The dual modality presentation clearly freed resources for the children, but how they used these additional resources is less clear. Since children were exposed to the stories in the two conditions for similar lengths of time, and spent more time on the words in the reading condition, in the combined condition they must have had additional time to spend on the rest of the text. The analyses confirm that this additional time was not spent on the definitions of the words: children looked at the definitions for similar lengths of time in the two conditions. It is possible that children used their extra resources in the combined condition to explore those parts of the text that were less clear, thus enhancing comprehension. Although no difference between comprehension of the two stories was highlighted in the story comprehension task (section 4.3.2) this conclusion cannot be completely discarded: a difference in story comprehension between the reading and the combined condition was found in Chapter 3. In addition, our present measure, requiring a forced choice between two alternative, might not have been sensitive enough to highlight a condition effect. Alternatively, it is possible that children were not using their freed resources purposefully. In the combined condition they had more freedom to skim through the text without keeping their attention focussed on the written presentation at all times, without losing track of the story. Indeed, both of these hypotheses are in line with the finding that children made more "non-reading-like" eye movements in the combined condition. Further research might be able to elucidate whether "non-reading-like" eye movements were targeted to enhance comprehension of the text or were linked to the presence of new words, by manipulating presence or absence of new words in the text, or manipulating text difficulty.

4.4.5 - Individual differences

Similarly to Studies 1 and 2, individual differences predicted word learning. Specifically, reading accuracy predicted learning of the link between phonological and orthographic forms, while vocabulary knowledge predicted category recognition. As discussed in Chapter 2 and 3, children with better decoding skills were better able to build links between phonological and orthographic forms of the new words during the story presentation, when both forms were presented, and use this knowledge to succeed in the phono-orthographic task. In contrast to previous studies (Ricketts et al., 2011; Swanborn & de Glopper, 1999), reading accuracy did not have an effect on learning of word meanings, possibly due to the overall difficulty of the semantic tasks or the complexity of the stories, which may have led children to rely more heavily on their previous vocabulary knowledge or on reading comprehension. The association between oral vocabulary and semantic learning has been well established in the literature (Cain et al., 2004; Ricketts et al., 2011; Shefelbine, 1990), and is further confirmed in our results. In the present study, reading comprehension also influenced word learning (c.f. Cain et al., 2003), specifically in the definition production task, and in the reading condition. Children were less negatively affected by low vocabulary and reading comprehension skills in the combined condition, which suggests that this presentation modality especially helps vocabulary acquisition in less able learners.

4.4.6 - Limitations

This study had a dual purpose: on one hand we were interested in exploring the source of the advantages of the combined condition, by analysing the process of word learning using eye-movement measures, and on the other hand, we sought to explore children's representation of the words in more depth, using more demanding tasks. We therefore recorded children's eye movements while they read stories, and asked them to complete more complex tasks to explore their learning. To measure semantic learning, both recognition and production tasks were used. Unfortunately, the increased difficulty of the semantic tasks compared to those of previous studies, coupled with the smaller number of exposures to the stories and the words, made the task very difficult for the children. Learning new words from the reading condition, and producing a definition, appeared especially challenging. This may have reduced the sensitivity of the tasks: it is likely that children might have shown more semantic learning, if easier tasks had been used, or if children were exposed to the stories more than once. The difficulty of the tasks notwithstanding, a difference in how learnt and unlearnt words were processed was found in the reading condition. Nevertheless, the results might have been clearer and more consistent between analyses, if easier tasks had been used. It is also important to consider that the learnt item list was composed of far fewer words than the unlearnt list, especially in the reading condition, thus making the two groups different in size and heterogeneous.

The variability of the stimuli chosen for this research may also have impacted on the sensitivity of the eye-movement measures. We used passages that together made full stories, to engage children's attention and reproduce, at least in part, a real life situation in which words are presented in context. Due to the nature of these stimuli, it was not possible to control for all the variables that might impact on eye movements to the areas of interest. For example target non-words were preceded by words that were different in length and frequency, the position of the words within the text were different, and sentence structures were not controlled. For these reasons, many variables may have influenced the length of time children spent on the target non-words, regardless of learning. The lack of control over those variables was due to the fact that the study was built as a study of word learning in context, rather than an eye-movement study, and eye-movement measures were used as a measure of processing and a predictor of learning, rather than a direct measure of learning. A study using more controlled stimuli might produce clearer results, with more stable patterns across analyses.

4.4.7 - Conclusions

The results of Study 3 confirm that presenting a story in two modalities at the same time (oral and written modalities), is more beneficial for word learning compared to a single modality presentation, specifically a written presentation. These results are in line with those of Studies 1 and 2. As in Study 2 this effect was especially evident in category recognition, although the mixed models analyses highlighted similar effects for definition production. The results of the phono-orthographic task were less clear, as, while a dual presentation may be beneficial for learning of the link between orthography and phonology, this effect was not found in all analyses. Overall, results suggest that the final product of word learning was better in the combined condition, thus supporting the idea that, in the combined condition, children produce representations of better quality for the words (Perfetti & Hart, 2002).

Children's eye-movement patterns also support the hypothesis that the process of encoding new words is less effortful in the combined condition; children spent less time on the words overall, and did not need to spend as much time on words that they learned in this condition as in the reading condition. The finding that children spent less time on new words in the combined condition than in the reading condition, even at the first presentation of the words, suggests that the combined condition frees attentional resources online during text processing, even before a representation of the new words has been formed (Mayer et al., 1999).

The results of the phono-orthographic task further suggest that children spontaneously build a phonological representation of new written words while reading, and that this process does not add additional costs, since phono-orthographic learning was not predicted by time spent on the words. However, having to phonologically recode the words impacted on children's ability to learn words meanings in the reading condition. In the reading condition, children needed more extra processing time on the words to form a semantic representation, and they were less effective in creating such a representation, compared to when they encountered new words in the combined condition.

Chapter 5: Discussion

The studies reported in this thesis explored vocabulary acquisition in Year 4 (8-9 year-old) children. In all three studies children were exposed to new words in the context of meaningful stories specifically written for the present research, and devised to engage children's attention, and incorporate new words in a meaningful context (for example, in Study 1 the India Story contained words that related to the British Empire in India). Words were repeated three times within the stories, either accompanied by a definition, or not. Stories were presented three times in Study 1, twice in Study 2, and only once in Study 3. All three studies explored the effect of presentation modality on vocabulary acquisition. To this end, stories were presented either orally (listening condition – Studies 1 & 2), in written form (reading condition – Studies 2 & 3), or in the oral and written modalities simultaneously (combined condition – Studies 1, 2 & 3). Vocabulary acquisition was explored by assessing learning of knowledge regarding word forms (phonological and orthographic forms), and meanings (semantics). Study 1 was designed to be naturalistic: story presentation took place in the classroom and the class teacher presented the story to the children as a class activity, similar to the approach often employed in a classroom setting. In Study 2 stories were presented in a more controlled way, with each child following the story on his or her assigned laptop; this study further examined the effect of the presence or absence of a definition within the story. Study 3 explored children's visual attention to stories presented in one or two modalities, by recording their eye-movements.

5.1 – Summary of the main results

In Studies 1 and 2, phonological and orthographic learning were explored using recognition tasks, which required children to identify the target spoken or written form from a choice of two. In Study 3, the strength of the link between orthographic and phonological forms was assessed through a phono-orthographic task, which required children to decide whether the given pronunciation of a written word was correct or incorrect. In Studies 1 and 2 children learned similar numbers of phonological forms in all conditions, including reading, listening and combined. This result, which suggests that 9 year-olds are able to create an oral representation of good quality from a written presentation, contrasts with the findings of previous research on single word learning, where children acquire better phonological representations from dual modality presentations (Rosenthal & Ehri, 2008). Our findings support Share's (1995) self-teaching hypothesis, which suggests that able readers automatically recode new written forms into oral forms, storing both forms in memory while reading.

In contrast to the results of the phonological task, in the orthographic task children showed knowledge of the orthographic forms of the new words in the combined and the reading conditions, but not in the listening condition, where performance was at chance. Thus, while phonological knowledge was acquired from all presentation modalities, knowledge of the written forms of the new words was only acquired when the written word was presented (combined condition in Study 1 and combined & reading groups in Study 2). The results of Study 3 were less clear: children performed better in the combined condition on some measures of phono-orthographic learning, but not others (section 4.3.3.1), suggesting that this effect might be small, and dependent on specific words' properties.

Semantic learning was assessed using several measures. Studies 1 and 2 employed three measures: a category recognition task, a sub-category recognition task and a definition recognition task. In all tasks the child was asked to choose the correct alternative between four. While the alternatives of the category recognition task were all categories presented within the story, the alternatives of the other two tasks were similar to the correct answer, but not presented within the story. In Study 1 children could complete subsequent tasks only if they had correctly recognised the alternative for previous ones, while in Study 2 children completed all tasks for all the items. In Study 3 semantic learning was assessed using a category recognition task similar to that used in the previous studies, but allowing a choice between eight alternatives, and a prompted definition production task. All three studies revealed better performance in the combined condition over the other conditions, suggesting that the dual modality presentation enhanced semantic learning. The facilitation associated with a dual modality presentation was particularly highlighted in the category recognition task in Studies 2 and 3, and in analyses of performance on all three steps of the semantic task in Study 1. These findings are considered further in the next section. The presence of a dual modality facilitation effect is congruent with previous findings of orthographic and phonological facilitation in single word learning (Ricketts et al., 2009; Rosenthal &

Ehri, 2008; 2011), but contrasts with reports from adult L2 research, where participants appear to learn as much from reading as they do from dual presentations (Brown et al., 2008; Sydorenko, 2010).

5.2 – Interpretation of the main results

The results of Studies 1 and 2 concerning phonological and orthographic learning suggest that children automatically recoded written words into their phonological forms, but did not recode new oral words into orthographic forms. As explored in Section 3.4.1.1, this suggests that creating a phonological representation for new words is of primary importance for vocabulary acquisition, while acquiring an orthographic representation is not. The creation of a phonological representation appears to be an automatic and necessary process for vocabulary acquisition, which happens irrespective of presentation modality. The creation of a written form, on the other hand, seems secondary: children do not automatically create a written representation in memory from an oral presentation. This result resonates with the orthographic skeleton hypothesis put forward by Wegener et al. (2017), which proposes that children create orthographic representations for orally presented words, but that these are not fully specified. It remains to be seen whether the primacy of phonological representations remains throughout adulthood, or whether there is a shift towards more automatic encoding of written forms (see Section 5.5).

An additional finding of interest in Studies 1 and 2, concerning learning of word forms, is that children in the combined condition did not create better representations of word forms than children in either the listening (in terms of phonological representation) or the reading condition (in terms of either phonological or orthographic representation). This lack of differences between conditions contrasts with previous findings of orthographic (Miles et al., 2016; Ricketts et al., 2009; Rosenthal & Ehri, 2008) and phonological facilitation (Rosenthal & Ehri, 2011) in word learning. Section 5.4 ('Limitations') explores whether this discrepancy might be due to the specific features of the tasks employed. Nevertheless, the difference between the present results and previous research may also be due to the different nature of the learning task. In most of the cited studies, learners were explicitly taught individual words (Miles et al., 2016; Ricketts et al., 2009; Rosenthal & Ehri, 2008), while in the studies presented in this thesis new words were acquired incidentally through stories. It is possible that a dual presentation might have a positive effect on form learning when children are consciously trying to learn new words, and are presented with single items only, while it might be less helpful in more naturalistic presentation contexts, and when the whole story is presented in both modalities. For example, the presentation of only target words in both oral and written forms, as in Rosenthal and Ehri's study (2011), where only target words were pronounced, might focus children's attention on the words' forms more effectively than the presentation of words embedded in context presented in the same way. Thus, the facilitation of dual presentation for form learning may be counteracted by the presentation of both words and context in dual modality.

The results of Study 3 were less clear in regards to form learning: children performed better in the combined condition on some measures and analyses of the phono-orthographic task, but not others (Sections 4.3.3.1 & 4.3.7). The phono-orthographic task in Study 3 assessed the link between orthographic and phonological forms, rather than the strength of the two forms separately. The results thus suggest that the combined condition might foster a stronger link between orthographic and phonological representations. Presenting words in two modalities simultaneously, therefore, does not seem to improve the actual representations of phonology or orthography, but it might strengthen the link between these. This result is particularly important in view of possible explanations of the orthographic and phonological facilitation of semantic learning considered later. One of the proposed accounts, in fact, considers that a dual presentation might create a better representation of the words in memory, which might involve stronger links between phonology and orthography. However, this result should be confirmed by further studies, given that it was not replicated in all analyses.

The results of all three studies suggest that a dual modality presentation enhances semantic learning. The main theoretical question behind this result relates to the source of this advantage. The account that has been given most consideration in this thesis is that combined oral and written presentation frees attentional resources, which can then be allocated to word learning and general comprehension processes. As discussed in Section 1.5.3, there are two main hypotheses regarding the source of this facilitation effect. The main difference between the two proposed frameworks is whether a dual presentation frees resources online during story presentation, in line with multimedia learning accounts (Mayer et al., 1999), or creates a better representation in memory, freeing resources when the word is later encountered, in line with ideas proposed by the Lexical Quality Hypothesis (Perfetti & Hart, 2002). Our data provide some insights into the validity of these two non-mutually exclusive accounts.

Considering the Lexical Quality Hypothesis (Perfetti & Hart, 2002), we would expect word representations to be of better quality in the combined condition. Specifically, we would expect children to be facilitated in accessing information regarding words for which they have stored more information in memory. If the combined condition enhances representations of the words, children should access phonological and orthographic information more easily in this condition, showing a facilitation effect not only for semantic learning, but for phonological and orthographic learning as well. Results of the phonological and orthographic tasks from our studies largely fail to indicate facilitation for phonological learning, and show no facilitation for orthographic learning when the combined condition is compared to the reading condition. These results undermine the idea that a dual presentation modality might facilitate vocabulary acquisition by creating better representations in memory. Nevertheless, it is possible that more sensitive tasks might highlight more subtle differences between the representations of the words in the different conditions (section 5.4). We cannot therefore completely exclude this account.

The idea that a dual presentation might facilitate vocabulary acquisition by freeing resources online, however, is particularly supported by the results of Study 3. When seeing a new word for the first time, children needed to spend more time on the item in the reading condition than in the combined condition, as shown by measures of gaze duration. One possible interpretation of this result is that children needed to spend more time on the words in the reading condition to phonologically recode them and memorize them, while the combined presentation spared them the need to spend resources on recoding. This would have left children more resources to encode word meanings in the combined condition. These results suggest that, regardless of the strength of the representations created, dual presentation frees resources online: children needed to spend less time on the words in the combined condition, despite learning their meanings better. The results of Study 2, which showed that children had better overall comprehension scores in the combined condition than the reading condition, seem to support this idea. If dual presentation enhances the acquisition of word meanings by freeing resources online, during story presentation, these resources may also be useful for story comprehension, leading to better story comprehension in the combined condition. Nevertheless, the lack of a similar effect for story comprehension in Study 3, and given that it is unclear from children's eye movements how they employed the freed resources, leaves questions about the nature of this online facilitation process, and whether the resources were employed for semantic learning and/or more general comprehension processes.

In conclusion, a dual modality presentation seems to free resources online, facilitating word processing even the first time a word is presented. Results regarding words' representations (i.e. phonological and orthographic learning in Studies 1 & 2, and the phono-orthographic task in Study 3), however, were not conclusive, and it is also possible that the dual modality facilitation effect might be due, in part, to the creation of a better representation. This is especially the case when comparing the combined and the listening conditions: these two conditions, in fact, differ in the type of representation created, with the combined condition promoting orthographic learning and the listening condition failing to do so.

Another important result to be considered is that the three studies showed the benefit of the dual presentation in different tasks. In Studies 2 and 3, differences were evident for category recognition and in Study 1 for definition recognition, or in measures that demanded more detailed knowledge. Study 3 also found a facilitation effect in some analyses of children's definition production. Different hypotheses were suggested to account for the different results in the different tasks and in the different studies.

First, recognising a category might require a higher level of abstraction than recognising or producing a definition, since it requires abstraction from the specific information, directly provided within the story, to more general knowledge, not directly provided. We could, thus, hypothesise that recognising the correct category might be a measure that requires abstraction from the specific information, lexical integration, as well as lexical access. Lexical integration evolves more slowly over time (Tamura et al., 2017), and this might explain why category recognition might be a more abstract task than definition recognition. It is possible that dual presentation particularly supports this process of integrating new vocabulary into the lexicon, rather than the process of acquiring specific information about the word. From a theoretical perspective, it is possible that a dual presentation might facilitate the integration of new information within the neocortical memory system, where these representations become more integrate with existing knowledge (McClelland et al., 1995), a facilitation effect similar to that proposed for sleep consolidation (James et al., 2017).

Another possible explanation is that the category recognition task is more complex than the other tasks because it is the only one where children need to recognise the precise link between label and meaning; in the other recognition tasks, in fact, children could succeed by recognising a meaning for a new word, without recalling its specific label, because only one of the alternative answers was present in the story. This second account could explain the results of Study 2, and it could also be applied to the results of Study 3, which also found a trend towards better performance in the combined condition in the definition production task: producing a definition also depends on having a direct link between label and meaning. A dual presentation modality, in this framework, might particularly facilitate the creation of this link.

Differences between Studies 2 & 3, where the difference between conditions was highlighted in category recognition, and Study 1, where the difference was highlighted in definition recognition, might also arise from the differing numbers of presentation of the words and the story. As the three studies presented the stories a different number of times, and, in Study 2, only half of the words were presented accompanied by a definition, the three studies might reveal different stages in the process of orthographic and phonological facilitation. The results of Study 3 would then represent the type of facilitation after one presentation, the results of Study 2 the type of facilitation after two presentations, and the results of Study 1 the type of facilitation effect obtained after three presentations of the story. It is possible that the facilitatory effect of a dual presentation modality might interact with presence of a definition and number of presentations. A dual presentation modality might facilitate the acquisition of more general information regarding the words (category recognition) when the story is presented only once or twice, while, by the third presentation, children might be able to abstract the necessary information in all presentation modalities (explaining why, in Study 1, children acquired category knowledge in all conditions). Conversely, the positive effect on definition recognition in Study 1 might be due to the fact that the presence of a definition has a stronger effect on learning in the combined condition. This possibility will be explored in the next section.

In conclusion, the differing results of the semantic tasks might be due to differences in the level of abstraction needed to succeed in the tasks, or to whether or not the task could be successfully completed without forming a direct link between the word's form and its meaning. Thus, dual modality might facilitate performance in tasks that require more abstraction, or tasks that require 1:1 mapping between word forms and meanings. Furthermore, differences in the dual modality facilitation found in the three studies might be accounted for by the numbers of story presentations and the type of context presented, such that the three studies show the development of facilitation over time.

5.3 – Contribution to theoretical models

The results of the present research are relevant to existing theories of organisation of the lexicon. For instance, in the Introduction chapter (Section 1.1.2) it was suggested that the simultaneous acquisition of orthographic, phonological and semantic information of the same words would support the idea of distributed and interconnected orthographic, phonological and semantic representations within the lexicon, in line with connectionist models (Gaskell & Marslen-Wilson, 1997; Seidenberg & McClelland, 1989), rather than fundamentally separate lexicons (Levelt, 1989; Morton, 1979). The correlations between semantic learning and phonological learning in Study 1 for the listening group and in Study 2 for the combined group seem to suggest that phonological and semantic information are heavily interconnected. These results suggest that, when learning new words, participants create a representation which includes both oral form and meaning, at least when the oral form of the words is directly provided; thus conceptualising two fundamentally separate stores for word forms (such as the lexeme) and meaning (such as the lemma) might not represent how learners acquire lexical items. Conversely, models where information regarding phonology and meaning are fundamentally interconnected (e.g. Gaskell & Marslen-Wilson, 1997), might be more easily applied to the present data. However, in our analyses we only explored correlations between semantic and form tasks with data organised by participant; to draw stronger conclusions on this issue, data should be organised by participant and by item. This would confirm whether the correlations are driven by the acquisition of phonological and semantic information of the same items, rather than participant effects (i.e. children with a good performance in the phonological task also perform well in the semantic task, irrespective of items). At present our results confirm that children's abilities to acquire phonological and semantic information of new words are correlated, at least in some cases, and especially when the phonological form was directly provided. Future research employing different correlation analyses and more items might confirm this hypothesis.

As explored in Section 1.1.2, models of reading speak to vocabulary acquisition, and are particularly relevant to the present research, given their focus on the connection between different representations of the words, as well as their meaning. The Triangle Model (Plaut et al., 1996; Seidenberg & McClelland, 1989) fits the findings from the present research particularly well, given its emphasis on the link between orthographic, phonological and semantic representations. In fact our results confirm that the provision of phonological and orthographic information for new words supports the acquisition of semantic information, which is in line with the Triangle Model's view of the importance of the connections between orthography, phonology and semantics. Furthermore, the positive effect of a dual presentation modality for semantic learning and the lack of orthographic learning from oral presentation point to differences in the processing of words acquired through oral, written or combined oral and written presentation. In their recent article, Monaghan, Chang and Welbourne (2017) suggest that words acquired before or after the onset of literacy are processed differently. More specifically, words acquired prior to the onset of literacy tend to have stronger connections between phonology and semantics, which are used even when reading the words aloud, even if the retrieval of semantic information is not directly required, while words acquired after the onset of literacy have stronger connections between orthography and phonology, and this route facilitates not only reading aloud, but also written lexical decision. There are parallels between the distinction between words acquired before or after the onset of literacy, and the present research, which compares words acquired through phonology (oral presentation), through orthography (written presentation) or both (dual presentation). In our semantic task children were presented with words both orally and written, thus they were allowed to access semantic information through either phonology or orthography. From our data it is unclear whether children in the listening condition and in the reading condition accessed semantic information through different routes (such as directly from orthography to semantics or via phonology); nevertheless it is clear that children in the combined condition, who had both routes available, had an advantage in accessing semantic information. This might suggest that the simultaneous use of phonology-semantics and orthography-semantics paths facilitates retrieval of semantic information, supporting the importance of the connection between orthography, phonology and semantics.

The results of the semantic tasks in Study 2 could also support the principles of the Hierarchical Network Model (Collins & Quillian, 1969) which suggests that conceptual information is organised hierarchically within the semantic network, and superordinate categories, such as those used in our semantic category task, are less easy to retrieve than subordinate categories, such as those used in our sub-category task. Although our results seem to confirm this conclusion, with children obtaining higher scores in the sub-category task than in the category task (see Table 3.9), this difference should be interpreted with caution, given the differences between the two tasks (see Section 5.5).

5.4 – Secondary effects

5.4.1 – Definitions

Study 2 compared vocabulary acquisition from stories when words were presented accompanied or not accompanied by a definition, with the hypothesis that semantic learning would be higher for words accompanied by a definition (Coyne et al., 2004; Dickinson, 1984; Wilkinson & Houston-Price, 2013). This hypothesis was confirmed for learning of sub-categories and definitions. However, the opposite effect was found for category recognition, where children recognised more categories for words not presented with a definition. As reasoned in Chapter 3, to recognise the correct category, children needed to abstract information from the text, and the presence of a definition might have led children to focus on acquiring specific details, disadvantaging more general abstraction processes. Indeed, the presence of a definition might have discouraged the process of abstraction: since children were provided with all the necessary information, perhaps they did not need to integrate the information about words with other information from the text. This idea is similar to Vlach and Sandhofer's (2010) proposal that more difficult word learning tasks promote better long-term retention. Vlach and Sandhofer particularly argued that the effort needed to integrate the word into the lexicon facilitates word acquisition in the long term, but not immediately after the learning task, especially if generalization and abstraction is needed. In Study 2, therefore, the effort to integrate information given throughout the text might have made semantic acquisition harder the first time that the story was presented, but it might have prompted the creation of links between the words and other words in their semantic lexicon long-term.

It is possible that the presence of the definition had a particularly positive effect in the combined condition. This might help to explain the better performance in the combined condition over the listening condition in definition recognition in Study 1. Children could have made more proficient use of definitions in the combined condition, perhaps because they could re-read the definition in the combined condition, leading to better learning of definitions. A similar effect might have been partially masked in Study 2 by the relatively low number of items presented with a definition. These considerations are only speculative, however, since there was only a trend for an interaction between definitions and presentation modality (Section 3.3.7).

In conclusion, the presence of a definition facilitated learning of specific details regarding the words, but not learning of the words' categories.

5.4.2 – Story comprehension

Measures of story comprehension were included in Studies 2 and 3 to ensure that children remained engaged with the task. Nonetheless, we did explore the effect of modality on comprehension finding better performance in the combined than reading condition in Study 2, when comprehension was assessed through four general forced-choice questions between four alternatives. This suggests that a dual modality presentation might support story comprehension. From a theoretical point of view it is possible that the dual presentation modality, by freeing resources during story presentation, might facilitate not only vocabulary acquisition, but also story comprehension. This idea is in line with research on multimodal presentation, which shows that students acquire more information when presented with study material in different modalities, such as orally and graphically (Mayer, et al. 1999). However, this difference between conditions was not observed in Study 3, weakening this conclusion. Furthermore research with second language learners show that dual presentation modality might, in fact, have a negative effect on comprehension, compared to a single modality presentation (Diao & Sweller, 2007). However the lack of an effect in Study 3 might be explained by the low sensitivity of the task used in that study. In Study 3, in fact, story comprehension was assessed through yes/no questions presented after each paragraph, and the performance of the children in the story comprehension questions in Study 3 was relatively high: ceiling effects and the ease of the task might have masked differences in story comprehension between the two conditions. Given the different results of the two studies and the listening group performing similarly to the combined group, the link between presentation modality and comprehension of the story warrants further study (section 5.5). Additionally, examining how children explore the text while they read, for example through eye movements, might help us explore whether the process of reading the text is facilitated by the oral presentation (see Section 5.3.3). In conclusion, children in the combined condition might have a better representation of the text as a whole, compared to children in the reading condition, given the results of the story comprehension task in Study 2, but this result needs to be further verified.

5.4.3 – Exploration of the text

In Study 3 we collected measures of children's eye movements, to ascertain whether children explored the text differently in the two conditions. We also wanted to explore whether the dual modality presentation had an effect on children's attention to the new words online, during story presentation, as well as having an effect at later testing (semantic learning). We found that children explored the text differently in the two conditions: they made fewer, but longer fixations, and longer saccades in the combined condition. It seems that the additional oral presentation increased children's perceptual span, since they did not need to fixate the text as often as in the reading condition, or at least the oral presentation allowed them to obtain the same information with a lower number of eye movements during reading. Since the differences in number of eye movements and saccade length parallels reported differences between children's and adults' reading habits (Rayner, 1998; Blythe et al., 2006), these findings suggest that a dual presentation facilitates processing of the text online. Furthermore, in the combined condition, children made more non-reading-like movements on the text than in the reading condition, suggesting that they might not read the text all the time during the story presentation, a finding also reported in previous research (Roy-Charland et al., 2007). Children spent overall less time on the target words in the combined condition. Time spent reading the target words was not related to phono-orthographic learning, but was related to semantic learning, with children spending more initial time (gaze duration) on learnt words compared to unlearnt words. This result was particularly true for the reading condition (Figure 4.7). Regressions to the definitions did not predict word learning, neither did coincident or cross-coincident time (section 4.3.8).

In conclusion, it is clear that the processing of the written text in the two conditions is different, with children in the combined condition needing less attentional resources to encode the new words, as evidenced by the lower amount of time they spent on the new words in this condition, even the first time they meet them. The good performance on the post-tests in the combined condition confirms that children acquired equal or even better representation of the words in the combined condition, despite lower effort. Children were more free to move their eyes around the text in the combined condition, without risk of losing important information. Nevertheless, it is unclear how they used the additional resources, since they did not seem to be purposely seeking out the most useful parts, such as definitions.

5.4.3 – Individual abilities

In all three studies reading accuracy was a good predictor of learning orthographic and phonological forms (Studies 1 & 2), and the link between these forms (Study 3). This result is in line with previous studies (Ricketts et al., 2009; Rosenthal & Ehri, 2008), and clearly suggests that better readers create better representations of word forms in memory than less skilled readers. Furthermore, in Study 2 there was an interaction between reading accuracy and condition (Figure 3.3), which suggests that a dual presentation is more effective than an oral presentation for good readers, but not for less skilled readers. This finding is in line with research showing that orthographic facilitation for word form learning is particularly effective for good readers (Ricketts et al., 2009; Rosenthal & Ehri, 2008). This finding is congruent with Ehri's (1992; 2014) idea of the fundamental role of the connections that are formed between spellings and sounds in memory formed by good readers.

In Studies 1 & 2, reading accuracy predicted category recognition, with better readers faring better in this semantic task. It is possible that better readers, by creating a representation of word forms faster and more easily, would have more resources left for a deeper analysis of words meaning, and use these resources to abstract the category of new words. However, unlike the first two studies, Study 3 does not highlight an effect of reading abilities on semantic learning. This difference might be due to the heightened difficulty of the word learning task in Study 3, and/or to the difference in the number of presentations. Reading abilities might exert a facilitatory effect on semantic learning over time, by facilitating reading of the same story and of the new words when these are re-encountered, while it might have a lesser effect the first time new words are encountered. When encountering new words in Study 3 children had to create a phonological representation for them, and, since this task might have taken up children's resources, even for the more skilled readers, this might not have allowed for the freeing of resources for acquiring the meaning of the items. In Studies 1 and 2, where reading accuracy had an effect on semantic learning, its effect did not interact with condition: the positive effect of a dual presentation was not mediated by reading abilities. Previous research has shownthat orthographic facilitation (i.e. the positive effect of a dual modality presentation) might be particularly effective for good readers (Ricketts et al., 2009). Nevertheless, this heightened orthographic facilitation for good readers was not highlighted in all studies, for example Ricketts et al. (2015) found that children with SLI, who had lower reading skills, showed levels of orthographic facilitation equivalent to those of normally developing children and children with autism. It is possible that, while better readers learn new words more effectively that less skilled decoders both in story and single word presentation, they experience a more marked facilitation advantage for dual presentation modalities in single word presentation (Ricketts et al., 2009), compared to whole story presentation.

Another measure that impacted strongly on semantic learning was vocabulary knowledge: vocabulary predicted overall performance on all steps of the semantic task in Study 1, and performance on categories and definitions in Studies 2 and 3. The prominence of vocabulary knowledge for word learning is in line with previous literature (Cain et al., 2004; Ricketts et al. 2011): children who are better equipped for vocabulary acquisition will have a larger vocabulary, by this age, and will learn words more easily, creating a positive loop. Some authors attribute this facilitation to the easier integration of new vocabulary in larger networks (James et al., 2017; Lewis & Durrant, 2011; Wilhelm et al., 2012), while others suggest it might be ascribed to the presence of more effective word learning strategies (Cain et al., 2004).

Other measures collected in the present research were listening and reading comprehension and non-verbal abilities and divided attention. While all these measures predicted children's performance in some analyses and some studies, with listening comprehension having a more prominent role in the results of Study 1, none of these had such a consistent effect as reading accuracy or vocabulary. We can therefore conclude that reading accuracy and vocabulary knowledge have the most fundamental impact on word learning from stories, with reading accuracy affecting learning of word forms and meanings (to some extent), and vocabulary having a more marked effect on semantic acquisition. Furthermore, although, in some analyses, individual skills affected the extent of the dual modality facilitation (sections 2.3.7, 3.3.5, 3.3.6 & 3.3.7), overall the results indicate that children of all skills and abilities were facilitated by a dual presentation modality.

5.5 – Limitations

Limitations relating to each study have been discussed in detail throughout the thesis (see Sections 2.4.5, 3.4.1 & 4.4.6) but some key limitations are worthy of discussion here. The phonological and orthographic tasks used in the present research, for instance, might be less sensitive than assessments of learning used in other studies (Rosenthal & Ehri, 2008), such as production tasks. A difficulty that arises from the use of recognition tasks is that, although the children in both Studies 1 and 2 showed similar form learning in all conditions, we cannot reject the hypothesis that they obtained representations of higher quality from dual modality presentations, since the tasks used might not be sufficiently sensitive to highlight such a difference. For example, the very high means in the phonological task in both Studies 1 and 2 in the listening and combined groups (Tables 2.7 & 3.5) suggest that a near-ceiling performance might be masking subtle differences between conditions. The slightly different results of the phono-orthographic task in Study 3, which found stronger links between orthographic and phonological representations in the combined condition, suggest that further research is needed to confirm whether children acquire representation of similar quality in different conditions. We particularly interpreted the performance on the phono-orthographic task as probing the creation of the link between orthographic and phonological forms, similarly to a reading aloud task. However, similarly to reading aloud, the phono-orthographic task might have probed abilities other than word form learning, such as more general reading abilities.

A further difficulty arising from the use of recognition tasks to probe phonological and orthographic learning is the choice of alternatives provided. In fact, while in a production task the participant directly provides his alternative pronunciation or spelling for a given item, in our recognition tasks the participants were asked to select the correct pronunciation or spelling for the item (Studies 1 and 2) or to decide whether a given pronunciation was correct or not (Study 3). It was reasoned that real words have correct and incorrect spellings and pronunciations, therefore the categorisation of the choices between correct and incorrect, at least for Studies 1 and 2, should not be problematic. However, the English language is characterised by inconsistent pronunciations (i.e. the same letter cluster produces different sounds, as in the final sound in *go* and *do*) and inconsistent spelling (i.e. the same sound can be written using different combination of letters, as in *see* and *sea*), which results in difficulties when attempting to spell new words, or pronounce new written items. This is consistent with the finding that even adults in the pilot group did not pronounce and spell all the items correctly, producing spelling and pronunciation mistakes that were used as alternatives to the canonical pronunciation and spelling in the phonological and orthographic task. These inconsistencies in the English orthographic system could have affected children's performance in the phonological and orthographic tasks, particularly in the listening and reading condition: without the provision of both the oral and written form, the children in these two conditions could have created an 'incorrect' representation of the form of the items, that would have been, nonetheless, justified (given the adults themselves produced them). It is therefore important to be cautious when considering a given choice 'correct' or 'incorrect' in the phonological and orthographic task: for example, children in the listening group in Study 2 might simply have created an orthographic representation of the word that corresponded to English rules, but not to the correct spelling of the word itself. It is possible that the use of plausible foils made the orthographic tasks particularly challenging for the listening group, who was not provided with an orthographic form. However, the reverse does not seem to have been the case for the reading group, who was not exposed to phonological forms: the reading group performed just as well as the other groups on the phonological task. Nevertheless, the difference between the good performance of the reading group in phonological learning, and the chance performance of the listening group in orthographic learning might be due to a higher consistency in pronunciation than spelling of the given items. Therefore the present results should be confirmed with items with controlled regularity in both spelling and pronunciation or by carrying out a similar study with speakers of a language with transparent orthography, such as Italian.

The same limitation is present in Study 3 when considering the phono-orthographic task, where, by using nonwords, we incur the problem of determining the 'correct' pronunciation of a new item. In Study 3 we used non-words presented in standardised tests. For most of the non-words, the tests provided both the 'correct' or expected pronunciation, and some plausible but 'incorrect' ones. Importantly, these tests considered as 'correct' only the pronunciations that followed phono-orthographic correspondence rules. By using items from standardised tests, we attempted to minimise the problem of determining the 'correct' pronunciation of a new item. Nevertheless, it must be acknowledged that, while in the combined condition the 'correct' pronunciations of the items were directly provided in the story, thus making the alternative pronunciations effectively 'incorrect' in terms of the task, in the reading condition the choice of 'correct' pronunciation was more arbitrary, and might not have reflected the pronunciation inferred by the children.

Overall, it is possible that the orthographic, phonological and phono-orthographic tasks probed abilities other than children's in-task learning, such as their general sensitivity to an oral form's word-likeness. This idea is supported by the finding that control word scores often significantly predicted target word scores on these tasks, and general reading accuracy predicted performance in these tasks in all studies. To account for such general effects, scores on control word trials and reading accuracy were included in all analytical models, enabling us to have confidence that our results reflect children's phonological learning within the task.

An additional difficulty arises from the use of real words in Studies 1 and 2. Specifically in these studies each word form was intrinsically linked to its definition: word forms could not be assigned to a specific definition or a specific story arbitrarily. As a baseline, to assess learning from story presentation, an additional set of items, the control set, was selected. The control set was not presented in the stories and vocabulary acquisition from the story was computed as the difference between the performance on target items and that on control items. The phono-orthographic task in Study 3 also suffers from the difficulty of the use of a control set of items. Given that the focus of the three studies was the different level of vocabulary acquisition between different presentation modalities, we prioritised the comparison between presentation modalities to the comparison between presented and non-presented items. Specifically, in Study 1, the two stories used were presented in both conditions, in a counterbalanced design; similarly in Study 3 both story and item set were counterbalanced. In Study 2 the same story was presented in all conditions. These features of the designs of our studies minimised variability between conditions, thus allowing us to be confident in the reliability of any difference between conditions arising from the studies. However, we did not apply the same level of control to the comparison between presented and non-presented items, where different sets of items (target and control items) were compared to assess learning. The comparison between different sets of items adds a source of variability to the data, where part of the difference between the performance on control and target words might be attributed to the specific features of the items themselves, rather than the fact that some items were previously encountered in a story and some were not. When considering the semantic tasks, in both Studies 1 and 2 this source of variability was controlled by testing children's knowledge of the meaning of both target and control words before story presentation. Nevertheless, due to the nature of the tasks, the same could not be applied to the phonological and orthographic tasks. When comparing performance for target and control words for the phonological and orthographic tasks, therefore, it is important to consider that part of these effects may be due to item differences, rather than presentation. To partly counteract these effects, the control words list was matched with the target words list on several measures, specifically number of letters, frequency and number of adults who pronounced and spelled the words correctly. However, it must be acknowledged that a more detailed analysis of the items and more stringent matching criteria, especially in terms of phonological properties, such as number of syllables and syllable structure, would have decreased variability between item sets, increasing our confidence in the learning data emerging from the phonological and orthographic tasks. Matching criteria for the items in Study 3 were stricter, considering bigram frequency and phonotactic probability. In Studies 1 and 2 matching control and target words was particularly complex, due to the use of real words, nevertheless we acknowledge that, in future, counterbalancing not only condition, but also target and control items, might help reduce variability and improve confidence in the results.

Another limitation of the measures used in this research concerns the acknowledged differences between the requirements of the category recognition task and sub-category and definition tasks: these differences have an impact on the interpretation of the results of these semantic tasks, so that the conclusions regarding the difference between the tasks can only be tentative. The category recognition task, in fact, differed from other tasks both in the level of detail children were required to recognise, but also in that all response options in this task were included in the story. This could have made the sub-category and definition recognition tasks easier than the category recognition task, independent of the level of abstraction needed: children could have chosen the correct definition or sub-category by remembering the item presented in the story without making the connection between the label and its meaning. This suggests that it would be inappropriate to interpret the definition recognition task as assessing deeper semantic learning than the category task. The differences between the results of the category recognition task and the other semantic tasks could therefore be due to the higher level of abstraction required in this task, and because the category recognition task was the only task assessing the creation of a stable link between labels and meanings. Had children been asked to recognise the correct definition between definitions that were present in the story, these two possible sources of differences could have been separately explored.

Another limitation of the present research, specific to Study 3, is that the stories used were not as controlled as materials usually used in eye-tracking research. Since, for each participant, we compared looking behaviour towards different items, embedded in different sentences and passages, it is possible that the eye-movement measures were not as sensitive as those used in other research (i.e. Godfroid et al., 2013; Williams & Morris 2004). This might explain the different effects of learning found between our studies and previous ones: with more controlled material it is possible that the difference between conditions might have been highlighted, not only in gaze duration, but also in later measures (total reading time and re-reading), as it was expected.

5.6 – Future directions

5.6.1 – The role of orthographic and phonological representations in word learning

Is the creation of a phonological representation for new words necessary for the acquisition of new lexical items? The present research suggests that, when acquiring new words from reading, children automatically create a phonological representation for the new items. Our results even suggest that this representation might be as good as the representation stored for orally presented words. On the other hand, children do not seem to recode oral words into written representations as automatically. Some researchers suggest that children do create partial or skeletal written representations for words presented only orally (Wegener et al., 2017), but this representation is not as good as the oral representation created from written form.

We might therefore ask whether the creation of an oral representation is fundamental for vocabulary acquisition, or whether, over time, readers can create representations in memory that contain written and semantic information only. Such a possibility would be supported by the proposal of a semantic route to reading in the dual-route model (Coltheart, 2005). This model postulates direct links between orthographic and semantic representations that do not necessarily involve the activation of the word's phonological form. Some models of reading development also suggest that the latter stage of reading development is characterised by direct orthographic processing, without the need for phonological conversion (Frith, 1985). The importance of orthographic processing is also suggested by the research on second language learners, where adults acquire more words from reading than listening (Brown et al., 2008), and by research that shows that orthographic effects exert their effects at lower prime exposure times than phonological effects, in priming studies (Ferrand & Grainger, 1993). It might be hypothesised, therefore, that, later in development, participants could acquire knowledge of written words without forming a phonological representation, similarly to how they do the reverse (encoding only phonology) earlier in life, before learning to read. Nevertheless, not all research supports such a developmental shift (Ehri, 1992). According to the Lexical Quality Hypothesis (Perfetti & Hart, 2002), words encoded in one modality only might result in a representation of lesser quality: an orthographic-only representation would be expected to negatively affect semantic encoding. To explore this hypothesis, future research might explore orthographic, phonological and semantic representations of new words presented only in written form, in adults, to ascertain whether the creation of a phonological form is as automatic for adults as it is for children, and whether the presence or absence of such a representation affects semantic learning. Second language learners might be included, since they might be more likely to create phonological representations of lower quality. Furthermore, methods to discourage phonological recoding, such as concurrent articulation, might be employed to stop the creation of a phonological representation.

5.6.2 – Developmental changes in orthographic and phonological facilitation

For some measures, the combined condition promoted greater learning than a single modality condition. However, this did not interact with existing reading, oral language or nonverbal abilities. Thus, if children benefitted from the combined condition, they did so to the same extent, irrespective of existing knowledge and skills. This is somewhat surprising: we might in fact have expected that a certain level of reading abilities would be necessary to experience a dual modality facilitation. Some previous research has found that, in fact, reading accuracy has an effect on the extent of the orthographic facilitation effect, with better readers being more facilitated by the presence of orthography than less skilled readers (Ricketts et al., 2009; Rosenthal & Ehri, 2008). In the present research we tested 8- to 9-year-olds: this specific age range was chosen to ensure all children would be able to read independently. The lack of an interaction between reading ability and presentation modality might, therefore, be due to the fact that our participants might have reached a sufficiently proficient level of reading ability, at which the written text supports their vocabulary acquisition. Future research might explore whether similar effects are found with less skilled and younger readers. We might, for instance, expect that younger readers might learn as much from listening as from listening and reading at the same time, or even learn less from a dual modality presentation, which might encourage them to try to decode the text, thereby focusing less on capturing the meaning.

Another possibility would be testing adults in similar modalities, to explore whether the facilitatory effect of a dual modality presentation over a written presentation remains, even at higher levels of reading competency. We would expect adults to learn vocabulary to a similar extent from reading and combined presentations. We might even expect lower performance in the combined condition: the dual presentation might be completely redundant for adults, who can activate phonology directly from orthography automatically, and complete redundancy might affect learning negatively (Kalyuga & Sweller, 2014). In children, the dual modality presentation might, in fact, be considered not completely redundant, since the presence of the oral presentation facilitates the still complex process of phonological recoding of written forms. On the other hand this process could be considered completely automatic in adults. As the expertise reversal principle (Kalyuga, Ayres, Chandler, & Sweller, 2003) suggests, techniques that reduce working-memory load in less expert learners (in this case children), have smaller or even reverse effects in expert learners. Expert learners need less information to acquire the same knowledge, and providing too much information that is not needed might slow down their learning. Similarly, providing an oral presentation might impair adults' acquisition of the words, requiring them to pay attention simultaneously to the oral and written presentations, without any further advantage from the effort, since they could have built a phonological representation of the new words from the written text without any further help.

5.6.3 - Orthographic and phonological facilitation effects on story comprehension

The present research was designed to specifically explore the effects of presentation modality on the acquisition of new vocabulary. In two studies we explored the effect of presentation modality on comprehension of the material, but the findings were unclear, with Study 2 highlighting a positive effect of dual modality on comprehension, not confirmed by Study 3. Research on multimedia learning has shown that learning of academic material is enhanced in multimodal presentations (Mayer, 2014). We might therefore expect that, with more in depth analysis of comprehension of the text, we might find more consistent facilitation of story comprehension, given dual presentation of the material. Future work might determine whether better vocabulary acquisition in the combined condition is a by-product of better understanding of the material.

5.6.4 – Avoiding chance performance and capturing individual differences

In the present research we aimed to explore whether children learned phonological, orthographic and semantic information of new words from a story context, and which individual abilities predict their learning. While in both Study 1 and Study 2 children showed semantic learning in all conditions, in Study 3 children showed significant semantic learning, especially in terms of category recognition, only in the combined condition. The differences between the findings of the three studies might reflect differences in the difficulty of the tasks used. Study 3, in fact, required the recognition of the correct category amongst eight alternatives, compared to the four options of the tasks used in Studies 1 and 2. Alternatively, the difference between the results might be reconducted to the increased difficulty of the learning task in Study 3, when the story was presented only once, compared to the multiple presentation of

Studies 1 and 2. The low performance in Study 3 might therefore have masked the importance of individual differences in predicting learning, giving that the amount of semantic learning was low and mostly restricted to the combined condition. Nevertheless, given that some children showed semantic learning, it seemed appropriate to run models to explore the predictors of this learning, especially considering that mixed effect models take into account random variability between participants and items, thus increasing their likelihood to detect fixed effects. However, floor effects might have had an impact on the models, reducing variability, and masking individual differences effects. To avoid similar difficulties and explore in more depth which individual differences have a greater impact on vocabulary acquisition, and whether individual differences affect attention to the text, in future research simpler semantic tasks, more in line with those used in Studies 1 and 2 could be used, to avoid floor effects.

Exploring attention to the text through eye-movements and vocabulary acquisition as in Study 3, but with a simpler semantic task would enable a more in depth analysis of the effects of individual differences. Specifically, the increased variability arising from simpler tasks, would increase the likelihood of detecting interactions between individual differences and measures of attention to the text, allowing to explore whether individual differences affect how participants explore the text online. For instance, individual difference might have an indirect effect on vocabulary acquisition, by determining how the participant attends to the text itself.

Another possibility would be to explore attention to the text by recording eye-movements while reading stories multiple times, or perform the vocabulary acquisition tasks on subsequent days to allow for sleep consolidation. Previous research has, in fact, shown the beneficial effect of multiple presentation (Swanborn & de Glopper, 1999) and sleep consolidation on vocabulary acquisition (Dumay & Gaskell, 2007; Williams & Horst, 2014). Multiple presentations and delayed testing would also help avoiding possible confounding effects of floor performance, as shown by the results of Studies 1 and 2. Furthermore it is possible that individual differences might determine how children explore the text over time, rather than at first presentation. By recording eye-movements on multiple occasions it would be possible to explore these effects.

5.7 – Conclusions

In conclusion, the work reported in this thesis shows that 9 year-old children learn word meanings better when presented with stories both orally and visually, compared to when they listen to stories or when they read stories themselves. Conversely, they acquire information regarding orthography and phonology similarly from reading as from a dual presentation: we did not find phonological facilitation for orthographic and phonological learning. On the other hand, when listening to stories, children acquire phonological information regarding new words, similarly to the other two conditions, but they do not acquire orthographic information, thus showing orthographic facilitation for orthographic learning. It is unclear whether the facilitation of the dual modality presentation for semantic learning would be shown by younger children or adults, but, at least in our sample, individual differences in reading, vocabulary, oral language and non-verbal abilities, do not seem to consistently reduce or enhance this facilitation effect: all children learn vocabulary best from a dual presentation.

From a theoretical point of view the results of the three studies presented suggest that, regardless of the strength of the representation created, dual presentation frees resources online, in accordance with cognitive load theory in multimedia learning (Mayer, 2014): in Study 3, in fact, children needed to spend less time on the words in the combined condition compared to the reading condition, despite learning their meanings better. Nevertheless we cannot exclude that part of the advantage of the combined condition over the other single modalities of presentation may be due to a representation of better quality in this condition, in accordance with the Lexical Quality Hypothesis (Perfetti & Hart, 2002): in Studies 1 & 2, in fact, words presented orally only had a less stable representation in memory, due to lack of orthographic information, compared to those presented in the written or oral and written presentations.

The practical implications of these findings for the classroom are that children are able to learn information about the phonological forms, orthographic forms and meanings of new words when they are listening to and/or reading stories. Importantly, both listening to and reading stories can support the learning of new phonological forms, while opportunities to see the new words written down are crucial for building representations of their orthographic forms. Finally, allowing children to hear stories while they are reading along may be optimal for learning, and especially for the extraction of semantic information, supporting teachers' practice of reading aloud in the classroom. Furthermore, as shown in previous research (Jenkins et al., 1984; Robbins & Ehri, 1994; Sénéchal et al., 1995; Swanborn & de Glopper, 1999), better readers and children with an extensive vocabulary are more successful at learning new words from stories during incidental presentation; thus, efforts to promote good reading skills and broad vocabulary knowledge will simultaneously boost children's proficiency at learning further words from stories.

References:

Adlof, S., Frishkoff, G., Dandy, J., & Perfetti, C. (2016). Effects of induced orthographic and semantic knowledge on subsequent learning: a test of the partial knowledge hypothesis. *Reading and Writing*, *29*(3), 475-500. doi: 10.1007/s11145-015-9612-x

Aicher, K.A. (2016). The Influence of Word Form on the Acquisition of Meaning: An Adult Visual Word Learning Study (Doctoral dissertation, University of Connecticut, USA). Retrieved from: http://digitalcommons.uconn.edu/dissertations

Aitchison, J. (2012). *Words in the mind: An introduction to the mental lexicon*. Chicester, UK: John Wiley & Sons.

Alario, F. X., Perre, L., Castel, C., & Ziegler, J. C. (2007). The role of orthography in speech production revisited. *Cognition*, *102*(3), 464-475. doi: 10.1016/j.cognition.2006.02.002

Allen, M. L., Hartley, C., & Cain, K. (2015). Do iPads promote symbolic understanding and word learning in children with autism? *Frontiers in Psychology*, *6*, 138. doi: 10.3389/fpsyg.2015.00138

Altemeier, L. E., Abbott, R. D., & Berninger, V. W. (2008). Executive functions for reading and writing in typical literacy development and dyslexia. *Journal of Clinical and Experimental Neuropsychology*, *30*(5), 588-606. doi: 10.1080/13803390701562818

Anderson, R. C., Wilson, P. T., & Fielding, L. G. (1988). Growth in reading and how children spend their time outside of school. *Reading Research Quarterly*, *23*(3), 285-303. doi: 10.1598/rrq.23.3.2

Apel, K. (2010). Kindergarten children’s initial spoken and written word learning in a storybook context. *Scientific Studies of Reading, 14*(5), 440–463. doi: 10.1080/10888431003623496

Apel, K., Brimo, D., Wilson-Fowler, E. B., Vorstius, C., & Radach, R. (2013). Children develop initial orthographic knowledge during storybook reading. *Scientific Studies of Reading*, *17*(4), 286-302. doi:10.1080/10888438.2012.692742

Apel, K., Thomas-Tate, S., Wilson-Fowler, E. B., & Brimo, D. (2012). Acquisition of initial mental graphemic representations by children at risk for literacy development. *Applied Psycholinguistics*, *33*(02), 365-391. doi:10.1017/s0142716411000403

Arbuthnott, K., & Frank, J. (2000). Trail making test, part B as a measure of executive control: validation using a set-switching paradigm. *Journal of Clinical and Experimental Neuropsychology*, 22(4), 518-528. doi:10.1076/1380-3395(200008)22:4;1-0;ft518

Ardila, A., Pineda, D., & Rosselli, M. (2000). Correlation between intelligence test scores and executive function measures. *Archives of Clinical Neuropsychology*, 15(1), 31-36. doi:10.1093/arclin/15.1.31

Aslin, R. N. (1993). Segmentation of fluent speech into words: Learning models and the role of maternal input. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage, & J. Morton (Eds.) *Developmental Neurocognition: Speech and Face Processing in the First Year of Life* (pp. 305-315). Springer Netherlands. doi:10.1007/978-94-015-8234-6_25

Audacity Team (2012) Audacity®. Version 2.0.0. Audio editor and recorder. Available from: http://audacityteam.org/

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database* [CD-ROM]. Philadelphia: Linguistic Data Consortium, University of Pennsylvania. Retrieved from: http://celex.mpi.nl/

Backman, J., Bruck, M., Hebert, M., & Seidenberg, M. S. (1984). Acquisition and use of spelling-sound correspondences in reading. *Journal of Experimental Child Psychology*, *38*(1), 114-133. doi:10.1016/0022-0965(84)90022-5

Baddeley, A. D., Eysenck, M. W., & Anderson, M. C. (2009). *Memory*. New York: Psychology Press.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation*, Vol. 8 (pp. 47–90). New York: Academic Press.

Baldwin, D.A. (1991). Infants' contribution to the achievement of joint reference. *Child Development*, 62, 875–890. doi:10.2307/1131140

Banks, C., & Sinha, U. (1951). An Item analysis of the Progressive Matrices Test. *British Journal of Psychology (Statistics Section)*, 4, 91-94. doi:10.1111/j.2044-8317.1951.tb00311.x

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255-278. doi:10.1016/j.jml.2012.11.001

Barron, R. W. (1986). Word recognition in early reading: A review of the direct and indirect access hypotheses. *Cognition*, *24*(1), 93-119. doi:10.1016/0010-0277(86)90006-5

Bartolotti, J., & Marian, V. (2017). Orthographic knowledge and lexical form influence vocabulary learning. *Applied Psycholinguistics*, *38*(2), 427-456. doi:10.1017/s0142716416000242

Bates, D., Kliegl, R., Vasishth, S., & Baayen, R. H. (2015). Parsimonious mixed models. Retrieved from: https://arxiv.org/pdf/1506.04967.pdf

Bates D., Maechler M., Bolker B., &Walker S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1-7, http://CRAN.R-project.org/package=lme4

Bates, E., Goodman, J.C. (1999). On the emergence of Grammar from the Lexicon. In B. MacWhinney (Ed.), *The Emergence of Language*. Mahwah, NJ: Lawrence Erlbaum.

Baumann, J. F., Edwards, E. C., Font, G., Tereshinski, C. A., Kame'enui, E. J., & Olejnik, S. (2002). Teaching morphemic and contextual analysis to fifth-grade students. *Reading Research Quarterly*, *37*(2), 150-176. doi:10.1598/rrq.37.2.3

Beck, I., & McKeown, M. (1991). Conditions of vocabulary acquisition. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of Reading Research* (Vol. 2, pp. 789–814). Mahwah, NJ: Erlbaum.

Beck, I., McKeown, M., & Kucan, L. (2002). *Bringing words to life: Robust vocabulary instruction*. New York, NY: Guilford.

Beck, I., Perfetti, C. A., & McKeown, M. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology*, *74*(4), 506-521. doi:10.1037/0022-0663.74.4.506

Becker, C. A., & Killion, T. H. (1977). Interaction of visual and cognitive effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *3*(3), 389-401. doi:10.1037/0096-1523.3.3.389

Benedict, H. (1979). Early lexical development: Comprehension and production. *Journal of Child Language*, *6*(02), 183-200. doi:10.1017/s0305000900002245

Bergelson, E., & Swingley, D. (2015). Early word comprehension in infants: Replication and extension. *Language Learning and Development*, *11*(4), 369-380. doi:10.1080/15475441.2014.979387

Berndt, R. S., Mitchum, C. C., & Haendiges, A. N. (1996). Comprehension of reversible sentences in "agrammatism": A meta-analysis. *Cognition*, *58*(3), 289-308. doi:10.1016/0010-0277(95)00682-6

Biemiller, A. (2003). Vocabulary: Needed if more children are to read well. *Reading Psychology*, *24*(3-4), 323-335. doi:10.1080/02702710390227297

Binder, K. S., & Rayner, K. (1998). Contextual strength does not modulate the subordinate bias effect: Evidence from eye fixations and self-paced reading. *Psychonomic Bulletin and Review*, *5*(2), 271-276. doi:10.3758/bf03212950

Blachowicz, C. L.Z., & Fisher, P. (2000). Vocabulary instruction. In M.L. Kamil, P.B. Mosenthal, P.D. Pearson, R. Barr (Eds.) *Handbook of Reading Research, Vol. 3* (pp. 503-523). Mahwah, NJ: Lawrence Erlbaum.

Blythe, H. I., Joseph, H.S. (2011). Children's eye movements during reading. In Liversedge, S.P., Gilchrist, I.D., Everling, S. (Eds.) *Oxford Handbook on Eye Movements* (pp. 643–662). Oxford, UK: Oxford University Press. doi:10.1093/oxfordhb/9780199539789.013.0036

Blythe, H. I., Liang, F., Zang, C., Wang, J., Yan, G., Bai, X., & Liversedge, S. P. (2012). Inserting spaces into Chinese text helps readers to learn new words: An eye movement study. *Journal of Memory and Language*, *67*(2), 241-254. doi:10.1016/j.jml.2012.05.004

Blythe, H. I., Liversedge, S. P., Joseph, H. S., White, S. J., Findlay, J. M., & Rayner, K. (2006). The binocular coordination of eye movements during reading in children and adults. *Vision Research*, *46*(22), 3898-3908. doi:10.1016/j.visres.2006.06.006

Blythe, H. I., Pagán, A., & Dodd, M. (2015). Beyond decoding: Phonological processing during silent reading in beginning readers. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(4), 1244-1253. doi:10.1037/xlm0000080

Bolger, D. J., Balass, M., Landen, E., & Perfetti, C. A. (2008). Context variation and definitions in learning the meanings of words: An instance-based learning approach. *Discourse Processes*, *45*(2), 122-159. doi:10.1080/01638530701792826

Bond, L. (2014). *The Effects of Contextual Reading and Feedback on Orthographic Development* (Master's thesis, Concordia University, Montreal, Canada). Retrieved from: http://spectrum.library.concordia.ca

Booth, A. E., McGregor, K. K., & Rohlfing, K. J. (2008). Socio-pragmatics and attention: Contributions to gesturally guided word learning in toddlers. *Language Learning and Development*, *4*(3), 179-202. doi:10.1080/15475440802143091

Borovsky, A., Ellis, E. M., Evans, J. L., & Elman, J. L. (2016). Semantic structure in vocabulary knowledge interacts with lexical and sentence processing in infancy. *Child Development*, *87*(6), 1893-1908. doi:10.1111/cdev.12554

Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology*, *112*(4), 417-436. doi:10.1016/j.jecp.2012.01.005

Bosman, A. M., & Van Orden, G. C. (1997). Why spelling is more difficult than reading. In C.A. Perfetti, L. Rieben, M. Fayol (Eds.) *Learning to Spell: Research, Theory, and Practice across Languages* (pp. 173-194). Mahwah, NJ: Lawrence Erlbaum.

Bowey, J. A., & Hansen, J. (1994). The development of orthographic rimes as units of word recognition. *Journal of Experimental Child Psychology, 58*(3), 465-488. doi:10.1006/jecp.1994.1045

Bowey, J. A., & Miller, R. (2007). Correlates of orthographic learning in third-grade children's silent reading. *Journal of Research in Reading*, *30*(2), 115-128. doi: 10.1111/j.1467-9817.2007.00335.x

Bowey, J. A., & Muller, D. (2005). Phonological recoding and rapid orthographic learning in third-graders' silent reading: A critical test of the self-teaching hypothesis. *Journal of Experimental Child Psychology, 92*(3), 203-219. doi:10.1016/j.jecp.2005.06.005

Bowey, J. A., & Underwood, N. (1996). Further evidence that orthographic rime usage in nonword reading increases with word-level reading proficiency. *Journal of Experimental Child Psychology, 63*(3), 526-562. doi:10.1006/jecp.1996.0061

Brett, A., Rothlein, L., & Hurley, M. (1996). Vocabulary acquisition from listening to stories and explanations of target words. *The Elementary School Journal*, *96*(4), 415-422. doi:10.1086/461836

Brinchmann, E. I., Hjetland, H. N., & Lyster, S. -A. H. (2015). Lexical Quality Matters: Effects of Word Knowledge Instruction on the Language and Literacy Skills of

Third-and Fourth-Grade Poor Readers. *Reading Research Quarterly*, 51(2) 165–180. doi:10.1002/rrq.128

Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language*, *20*(2), 136-163. http://www.nflrc.hawaii.edu/rfl/October2008/brown/brown.pdf

Brown, S. W. (1997). Attentional resources in timing: Interference effects in concurrent temporal and nontemporal working memory tasks. *Perception and Psychophysics*, *59*(7), 1118-1140. doi:10.3758/bf03205526

Brusnighan, S. M., & Folk, J. R. (2012). Combining contextual and morphemic cues is beneficial during incidental vocabulary acquisition: Semantic transparency in novel compound word processing. *Reading Research Quarterly*, *47*(2), 172-190. doi:10.1002/rrq.015

Brusnighan, S. M., Morris, R. K., Folk, J. R., & Lowell, R. (2014). The role of phonology in incidental vocabulary acquisition during silent reading. *Journal of Cognitive Psychology*, *26*(8), 871-892. doi:10.1080/20445911.2014.965713

Bürki, A., Spinelli, E., & Gaskell, M. G. (2012). A written word is worth a thousand spoken words: The influence of spelling on spoken-word production. *Journal of Memory and Language*, *67*(4), 449-467. doi:10.1016/j.jml.2012.08.001

Bus, A. G., Takacs, Z. K., & Kegel, C. A. (2015). Affordances and limitations of electronic storybooks for young children's emergent literacy. *Developmental Review*, *35*, 79-97. doi: 10.1016/j.dr.2014.12.004

Bialystok, E., Luk, G., & Kwan, E. (2005). Bilingualism, biliteracy and learning to read: Interactions among languages and writing systems. *Scientific Studies of Reading*, 9(1), 43-61. doi:10.1207/s1532799xssr0901_4

Cain, K. (2010). *Reading Development and Difficulties*. Chichester, UK: John Wiley & Sons.

Cain, K., & Oakhill, J. (2011). Matthew effects in young readers: Reading comprehension and reading experience aid vocabulary development. *Journal of Learning Disabilities*, *44*(5), 431-443. doi:10.1177/0022219411410042

Cain, K., Oakhill, J. V., & Elbro, C. (2003). The ability to learn new word meanings from context by school-age children with and without language comprehension difficulties. *Journal of Child Language*, *30*(03), 681-694. doi:10.1017/s0305000903005713

Cain, K., Oakhill, J., & Lemmon, K. (2004). Individual differences in the inference of word meanings from context: The influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of Educational Psychology, 96*(4), 671-681. doi:10.1037/0022-0663.96.4.671

Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan, & G. Miller (Eds.), *Linguistic Theory and Psychological Reality* (pp. 264-293). Cambridge, MA: MIT Press.

Catts, H. W., Fey, M. E., Zhang, X., & Tomblin, J. B. (1999). Language basis of reading and reading disabilities: Evidence from a longitudinal investigation. *Scientific Studies of Reading*, *3*(4), 331-361. doi:10.1207/s1532799xssr0304_2

Chaffin, R. (1997). Associations to unfamiliar words: Learning the meanings of new words. *Memory and Cognition*, *25*(2), 203-226. doi:10.3758/bf03201113

Chaffin, R., Morris, R. K., & Seely, R. E. (2001). Learning new word meanings from context: a study of eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(1), 225. doi:10.1037/0278-7393.27.1.225

Chang, A. C. -S. (2009). Gains to L2 listeners from reading while listening vs. listening only in comprehending short stories. *System, 37*(4), 652-663. doi:10.1016/j.system.2009.09.009

Chang, A. C. -S. (2011). The Effect of Reading While Listening to Audiobooks: Listening Fluency and Vocabulary Gain. *Asian Journal of English Language*

Teaching, *21*, 43–64. Retrieved from: http://www.cuhk.edu.hk/ajelt/vol21/contents.htm

Charles-Luce, J., & Luce, P. A. (1990). Similarity neighbourhoods of words in young children's lexicons. *Journal of Child Language*, *17*(01), 205-215. doi: 10.1017/s0305000900013180

Charles-Luce, J., & Luce, P. A. (1995). An examination of similarity neighbourhoods in young children's receptive vocabularies. *Journal of Child Language*, *22*(03), 727-735. doi: 10.1017/s0305000900010023

Chereau, C., Gaskell, M. G., & Dumay, N. (2007). Reading spoken words: Orthographic effects in auditory priming. *Cognition*, *102*(3), 341-360. doi:10.1016/j.cognition.2006.01.001

Chomsky, C. (1976). After decoding: what? *Language Arts*, *53*(3), 288-314. Retrieved from: http://www.jstor.org/stable/41404149

Clark, E.V. (1993). *The lexicon in acquisition*. Cambridge: Cambridge University Press.

Clarke, M. G., & Leonard, L. B. (1996). Lexical comprehension and grammatical deficits in children with specific language impairment. *Journal of Communication Disorders*, *29*(2), 95-105. doi:10.1016/0021-9924(94)00036-0

Clifton, C., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R.P.G. van Gompel, M.H. Fischer, W.S. Murray, & R.L. Hill (Eds.), *Eye Movements: A Window on Mind and Brain* (pp. 341–371). New York: Elsevier

Coady, J. (1997). L2 vocabulary acquisition through extensive reading. In J. Coady, & T. Huckin (Eds.), *Second Language Vocabulary Acquisition: A Rationale for Pedagogy* (pp. 225 – 299). Cambridge: Cambridge University Press.

Coady, J. A., & Aslin, R. N. (2003). Phonological neighbourhoods in the developing lexicon. *Journal of Child Language*, *30*(02), 441-469. doi:10.1017/s0305000903005579

Cocchini, G., Logie, R. H., Della Sala, S., MacPherson, S. E., & Baddeley, A. D. (2002). Concurrent performance of two memory tasks: Evidence for domain-specific working memory systems. *Memory and Cognition*, *30*(7), 1086-1095. doi:10.3758/bf03194326

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407. doi:10.1037/0033-295x.82.6.407

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, *8*(2), 240-247. doi:10.1016/s0022-5371(69)80069-1

Coltheart, M. (1981). The MRC Psycholinguistic Database. *The Quarterly Journal of Experimental Psychology*, 33A, 497-505. doi:10.1080/14640748108400805

Coltheart, M. (2005).Modeling reading: The dual-route approach. In M. J. Snowling & C. Hulme (Eds.), *The Science of Reading: A Handbook* (pp. 6 –23). Oxford, UK: Blackwell Publishing Ltd.

Comalli, P. E., Wapner, S., & Werner, H. (1962). Interference effects of Stroop color-word test in childhood, adulthood, and aging. *The Journal of Genetic Psychology, 100*(1), 47-53. doi:10.1080/00221325.1962.10533572

Coyne, M. D., Simmons, D. C., Kame'enui, E. J., & Stoolmiller, M. (2004). Teaching vocabulary during shared storybook readings: An examination of differential effects. *Exceptionality*, *12*(3), 145-162. doi:10.1207/s15327035ex1203_3

Cossu, G., Gugliotta, M., & Marshall, J. C. (1995). Acquisition of reading and written spelling in a transparent orthography: Two non-parallel processes? *Reading and Writing*, *7*(1), 9-22. doi:10.1007/bf01026945

Cunningham, A. E. (2006). Accounting for children’s orthographic learning while reading text: Do children self-teach? *Journal of Experimental Child Psychology*, *95*(1), 56-77. doi:10.1016/j.jecp.2006.03.008

Cunningham, A. E., Perry, K. E., Stanovich, K. E., & Share, D. L. (2002). Orthographic learning during reading: Examining the role of self-teaching. *Journal of*

Experimental Child Psychology, *82*(3), 185-199. doi:10.1016/s0022-0965(02)00008-5

Cutting, J. C., & Ferreira, V. S. (1999). Semantic and phonological information flow in the production lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(2), 318-344. doi:10.1037/0278-7393.25.2.318

Dale, E. (1965). Vocabulary development of the underprivileged child. *Elementary English*, *42*(7), 778-786. Retrieved from: http://www.jstor.org/stable/41385894

Danan, M. (2004). Captioning and subtitling: Undervalued language learning strategies. *Meta: Journal des traducteurs, 49*(1), 67-77. doi:10.7202/009021ar

de Jong, P. F., & Share, D. L. (2007). Orthographic learning during oral and silent reading. *Scientific Studies of Reading*, *11*(1), 55-71. doi:10.1207/s1532799xssr1101_4

De Temple, J., & Snow, C. E. (2003). Learning words from books. In A. van Kleeck, S. A. Stahl, & E. B. Bauer (Eds.), *On Reading Books to Children: Parents and Teachers* (pp. 16-36). Mahwah, NJ: Lawrence Erlbaum.

Dean, R. S., Yekovich, F. R., & Gray, J. W. (1988). The effect of modality on long-term recognition memory. *Contemporary Educational Psychology*, *13*(2), 102-115. doi:10.1016/0361-476x(88)90010-0

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283-321. doi:10.1037/0033-295x.93.3.283

Dell, G. S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language*, *27*(2), 124-142. doi:10.1016/0749-596x(88)90070-8

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1996). A connectionist model of naming errors in aphasia. In J. Reggia., R. Bendt, & E. Ruppin (Eds.), *Neural Modelling of Cognitive and Brain Disorders* (pp. 135-156). Singapore: World Scientific Publishing.

Diao, Y., & Sweller, J. (2007). Redundancy in foreign language reading comprehension instruction: Concurrent written and spoken presentations. *Learning and Instruction*, *17*(1), 78-88. doi:10.1016/j.learninstruc.2006.11.007

Dickinson, D. K. (1984). First impressions: Children's knowledge of words gained from a single exposure. *Applied Psycholinguistics*, *5*(4), 359-373. doi:10.1017/s0142716400005233

Dockrell, J., & McShane, J. (1990). Young children's use of phrase structure and inflectional information in form-class assignments of novel nouns and verbs. *First Language*, *10*(29), 127-140. doi:10.1177/014272379001002903

Doctor, E. A., & Coltheart, M. (1980). Children's use of phonological encoding when reading for meaning. *Memory and Cognition*, *8*(3), 195-209. doi:10.3758/bf03197607

Druks, J. (2002). Verbs and nouns - a review of the literature. *Journal of Neurolinguistics*, *15*(3), 289-315. doi:10.1016/s0911-6044(01)00029-x

Duckett, P. (2003). Evisioning story: The eye movements of beginning readers. *Literacy Teaching and Learning*, *7*(1/2), 77–89. Retrieved from: http://files.eric.ed.gov/fulltext/EJ966146.pdf

Duff, D., Tomblin, J. B., & Catts, H. (2015). The influence of reading on vocabulary growth: A case for a Matthew Effect. *Journal of Speech, Language, and Hearing Research*, *58*(3), 853-864. doi:10.1044/2015_jslhr-l-13-0310

Duff, F. J., Reen, G., Plunkett, K., & Nation, K. (2015). Do infant vocabulary skills predict school-age language and literacy outcomes? *Journal of Child Psychology and Psychiatry*, *56*(8), 848-856. doi:10.1111/jcpp.12378

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language*, *27*(4), 429-446. doi:10.1016/0749-596x(88)90066-6

Dumay, N., & Gaskell, M. G. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological Science*, *18*(1), 35-39. doi:10.1111/j.1467-9280.2007.01845.x

Dunn, L. M. (1959). *Peabody Picture Vocabulary Test (PPVT)*. Circle Pines, MN, USA: AGS Inc.

Dunn, L. M., Dunn, D. M., & NFER. (2009). *British Picture Vocabulary Scale - 3 (3rd ed.)*. London: GL Assessment Ltd.

Durand, V. N., Loe, I. M., Yeatman, J. D., & Feldman, H. M. (2013). Effects of early language, speech, and cognition on later reading: a mediation analysis. *Frontiers in Psychology, 4,* 586. doi:10.3389/fpsyg.2013.00586

Durso, F. T., & Shore, W. J. (1991). Partial knowledge of word meanings. *Journal of Experimental Psychology: General*, *120*(2), 190-202. doi:10.1037/0096-3445.120.2.190

d'Ydewalle, G., Praet, C., Verfaillie, K., & Van Rensbergen, J. (1991). Watching subtitled television. *Communication Research*, *18*(5), 650-666. doi:10.1177/009365091018005005

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 (Eds.), *Reading Acquisition* (pp. 107-144). Hillside, N.J.: Lawrence Erlbaum.

Ehri, L. C. (2005a). Development of sight word reading: Phases and findings. In M. J. Snowling and C. Hulme (Eds.), *The Science of Reading: A Handbook*. Oxford, UK: Blackwell Publishing Ltd.

Ehri, L. C. (2005b). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading*, *9*(2), 167-188. doi:10.1207/s1532799xssr0902_4

Ehri, L. C. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Scientific Studies of Reading*, *18*(1), 5-21. doi:10.1080/10888438.2013.819356

Ehri, L. C., & Robbins, C. (1992). Beginners need some decoding skill to read words by analogy. *Reading Research Quarterly, 27*(1), 13-26. doi:10.2307/747831

Ehrlich, S. F. (1981). Children's word recognition in prose context. *Visible Language*, *15*(3), 219-244. Retrieved from: https://s3-us-west-2.amazonaws.com/visiblelanguage/pdf/V15N2_1981_E.pdf

Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior*, *20*(6), 641-655. doi:10.1016/s0022-5371(81)90220-6

Ehrlich, M. F., & Remond, M. (1997). Skilled and less skilled comprehenders: French children's processing of anaphoric devices in written texts. *British Journal of Developmental Psychology*, *15*(3), 291-309. doi:10.1111/j.2044-835x.1997.tb00522.x

Elbro, C., Rasmussen, I., & Spelling, B. (1996). Teaching reading to disabled readers with language disorders: A controlled evaluation of synthetic speech feedback. *Scandinavian Journal of Psychology*, *37*(2), 140-155. doi:10.1111/j.1467-9450.1996.tb00647.x

Elleman, A. M., Lindo, E. J., Morphy, P., & Compton, D. L. (2009). The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. *Journal of Research on Educational Effectiveness, 2*(1), 1-44. doi:10.1080/19345740802539200

Elley, W. B. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly*, *24*(2),174-187. doi:10.2307/747863

Evans, M. A., & Saint-Aubin, J. (2005). What children are looking at during shared storybook reading: Evidence from eye movement monitoring. *Psychological Science*, *16*, 913–920. doi:10.1111/j.1467-9280.2005.01636.x

Evans, M. A., & Saint-Aubin, J. (2013). Vocabulary acquisition without adult explanations in repeated shared book reading: An eye movement study. *Journal of Educational Psychology*, *105*(3), 596-608. doi:10.1037/a0032465

Evans, M. A., Williamson, K., & Pursoo, T. (2008). Preschoolers' attention to print during shared book reading. *Scientific Studies of Reading*, *12*(1), 106-129. doi:10.1080/10888430701773884

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*(5), 1-185. doi:10.2307/1166093

Fernald, A., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology, 27*(2), 209-221. doi:10.1037/0012-1649.27.2.209

Ferrand, L., & Grainger, J. (1993). The time course of orthographic and phonological code activation in the early phases of visual word recognition. *Bulletin of the Psychonomic Society, 31*(2), 119-122. doi:10.3758/bf03334157

Field, A. (2009). *Discovering statistics using SPSS*. Sage publications.

Fien, H., Santoro, L., Baker, S. K., Park, Y., Chard, D. J., Williams, S., & Haria, P. (2011). Enhancing teacher read alouds with small-group vocabulary instruction for students with low vocabulary in first-grade classrooms. *School Psychology Review, 40*(2), 307-318. Retrieved from: http://www.nasponline.org/publications/periodicals/spr/volume-40/volume-40-issue-2

Forster, K. I. (1976). Accessing the mental lexicon. In R. J. Wales & E. Walker (Eds.), *New Approaches to Language Mechanisms* (pp. 231-256). Amsterdam: North Holland.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior, 12*(6), 627-635. doi:10.1016/s0022-5371(73)80042-8

Foster, H. (2007). *Single word reading test 6-16*. GL Assessment Limited.

Freebody, P., & Anderson, R. C. (1983). Effects of vocabulary difficulty, text cohesion, and schema availability on reading comprehension. *Reading Research Quarterly, 18*(3), 277-294. doi:10.2307/747389

Frisch, S. A., Large, N. R., & Pisoni, D. B. (2000). Perception of wordlikeness: Effects of segment probability and length on the processing of nonwords. *Journal of Memory and Language*, *42*(4), 481-496. doi:10.1006/jmla.1999.2692

Frishkoff, G. A., Perfetti, C. A., & Collins-Thompson, K. (2011). Predicting robust vocabulary growth from measures of incremental learning. *Scientific Studies of Reading*, *15*(1), 71-91. doi:10.1080/10888438.2011.539076

Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. Patterson, M. Coltheart & J. C. Marshall (Eds.), *Surface Dyslexia* (pp. 301-330). London: Routledge.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: true issues and false trails. *Psychological Bulletin*, *123*(1), 71-99. doi:10.1037/0033-2909.123.1.71

Gallo, D.A., McDermott, K.B., Percer, J. M., & Roediger III, H. L. (2001). Modality effects in false recall and false recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *27*(2), 339-353. doi:10.1037/0278-7393.27.2.339

Garza, T. J. (1991). Evaluating the use of captioned video materials in advanced foreign language learning. *Foreign Language Annals*, *24*(3), 239-258. doi:10.1111/j.1944-9720.1991.tb00469.x

Gaskell, M. G., & Marslen-Wilson, W. D. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes*, *12*(5-6), 613-656. doi:10.1080/016909697386646

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. A. Kuczaj II (Ed.), *Language development, Vol. 2, Language, Thought, and Culture* (pp. 301-334). Hillsdale, NJ: Lawrence Erlbaum.

Godfroid, A., Boers, F., & Housen, A. (2013). An eye for words: Gauging the role of attention in incidental l2 vocabulary acquisition by means of eye-tracking. *Studies* *in Second Language Acquisition*, *35*(3), 483-517. doi:10.1017/s0272263113000119

Golinkoff, R. M., Hirsh-Pasek, K., Bailey, L. M., & Wenger, N. R. (1992). Young children and adults use lexical principles to learn new nouns. *Developmental Psychology, 28*(1), 99-108. doi:10.1037/0012-1649.28.1.99

Golinkoff, R. M., Mervis, C. B., & Hirsh-Pasek, K. (1994). Early object labels: The case for a developmental lexical principles framework. *Journal of Child Language*, *21*(1), 125-155. doi:10.1017/s0305000900008692

Goswami, U. (1986). Children's use of analogy in learning to read: A developmental study. *Journal of Experimental Child Psychology*, *42*(1), 73-83. doi:10.1016/0022-0965(86)90016-0

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, *7*(1), 6-10. doi:10.1177/074193258600700104

Gough, P. B., & Wren, S. (1999). Constructing meaning: The role of decoding. In J. Oakhill, & R. Beard (Eds.), *Reading Development and the Teaching of Reading* (pp. 59-78). Mahwah, NJ: Lawrence Erlbaum.

Haigh, C. A., & Jared, D. (2007). The activation of phonological representations by bilinguals while reading silently: Evidence from interlingual homophones. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(4), 623-644. doi:10.1037/0278-7393.33.4.623

Hartman, F. R. (1961). Single and multiple channel communication: A review of research and a proposed model. *Educational Technology Research and Development*, *9*(6), 235-262. Retrieved from: http://www.jstor.org/stable/30218701

Hashimoto, N., McGregor, K. K., & Graham, A. (2007). Conceptual organization at 6 and 8 years of age: Evidence from the semantic priming of object decisions. *Journal of Speech, Language, and Hearing Research*, *50*(1), 161-176. doi:10.1044/1092-4388(2007/014)

Henderson, J. M., Dixon, P., Petersen, A., Twilley, L. C., & Ferreira, F. (1995). Evidence for the use of phonological representations during transsaccadic word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(1), 82-97. doi:10.1037/0096-1523.21.1.82

Henderson, L. M., Devine, K., Weighall, A., & Gaskell, G. (2015). When the daffodat flew to the intergalactic zoo: Off-line consolidation is critical for word learning from stories. *Developmental Psychology*, *51*(3), 406-417. doi:10.1037/a0038786

Henderson, L. M., Weighall, A. R., Brown, H., & Gaskell, M. G. (2012). Consolidation of vocabulary is associated with sleep in children. *Developmental Science*, *15*(5), 674-687. doi:10.1111/j.1467-7687.2012.01172.x

Henderson, L. M., Weighall, A.R., & Gaskell, G. (2013). Learning new vocabulary during childhood: Effects of semantic training on lexical consolidation and integration. *Journal of Experimental Child Psychology*, *116*(3), 572-592. doi:10.1016/j.jecp.2013.07.004

Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. (2009). Longitudinal analysis of early semantic networks: Preferential attachment or preferential acquisition? *Psychological Science*, *20*(6), 729-739. doi:10.1111/j.1467-9280.2009.02365.x

Hintzman, D. L. (1984). MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments, and Computers, 16*, 96–101. doi:10.3758/bf03202365

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing*, *2*(2), 127-160. doi:10.1007/bf00401799

Horst, J. S., Parsons, K. L., & Bryan, N. M. (2011). Get the story straight: Contextual repetition promotes word learning from storybooks. *Frontiers in Psychology*, *2*, 17. doi:10.3389/fpsyg.2011.00017

Houston-Price, C., Howe, J. A., & Lintern, N. J. (2014). Once upon a time, there was a fabulous funambulist…: what children learn about the “high-level” vocabulary they encounter while listening to stories. *Frontiers in Psychology*, *5*, 75. doi:10.3389/fpsyg.2014.00075

Houston-Price, C., & Law, B. (2013). How Experiences with Words Supply All the Tools in the Toddler's Word-Learning Toolbox. In L. Gogate & G. Hollich (Eds.) *Theoretical and Computational Models of Word Learning: Trends in Psychology and Artificial Intelligence: Trends in Psychology and Artificial Intelligence* (pp. 81-108). Hershey, PA: Information Science Reference.

Houston-Price, C., Plunkett, K., & Duffy, H. (2006). The use of social and salience cues in early word learning. *Journal of Experimental Child Psychology*, *95*(1), 27-55. doi:10.1016/j.jecp.2006.03.006

Hu, C. F. (2008). Use orthography in L2 auditory word learning: Who benefits? *Reading and Writing*, *21*(8), 823-841. doi:10.1007/s11145-007-9094-6

Inhoff, A. W., & Radach, R. (1998). Definition and computation of oculomotor measures in the study of cognitive processes. In G. Underwood (Eds.) *Eye Guidance in Reading and Scene Perception* (pp. 29-53). Oxford, England: Elsevier Science Ltd.

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception and Psychophysics*,*40*(6), 431-439. doi:10.3758/bf03208203

Inhoff, A. W., Solomon, M., Radach, R., & Seymour, B. A. (2011). Temporal dynamics of the eye–voice span and eye movement control during oral reading. *Journal of Cognitive Psychology*, *23*(5), 543-558. doi:10.1080/20445911.2011.546782

Ivey, G. (2003). The intermediate grades:" The teacher makes it more explainable" and other reasons to read aloud in the intermediate grades. *The Reading Teacher*, *56*(8), 812-814. Retrieved from: http://www.jstor.org/stable/20205300

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*(4), 434-446. doi:10.1016/j.jml.2007.11.007

James, E., Gaskell, M. G., Weighall, A., & Henderson, L. (2017). Consolidation of vocabulary during sleep: the rich get richer? *Neuroscience and Biobehavioral Reviews, 77*, 1-13. doi:10.1016/j.neubiorev.2017.01.054

Jenkins, J. R., Stein, M. L., & Wysocki, K. (1984). Learning vocabulary through reading. *American Educational Research Journal*, *21*(4), 767-787. doi:10.2307/1163000

Johnston, M., McKague, M., & Pratt, C. (2004). Evidence for an automatic orthographic code in the processing of visually novel word forms. *Language and Cognitive Processes*, *19*(2), 273-317. doi:10.1080/01690960344000189

Joseph, H. S., & Nation, K. (2017). *Examining incidental word learning during reading in children: The role of context.* Manuscript submitted for publication.

Joseph, H. S., Wonnacott, E., Forbes, P., & Nation, K. (2014). Becoming a written word: Eye movements reveal order of acquisition effects following incidental exposure to new words during silent reading. *Cognition*, *133*(1), 238-248. doi:10.1016/j.cognition.2014.06.015

Jusczyk, P. W., & Aslin, R. N. (1995). Infants′ detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, *29*(1), 1-23. doi:10.1006/cogp.1995.1010

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychological review*, *87*(4), 329-354. doi:10.1037/0033-295x.87.4.329

Justice, L. M., Meier, J., & Walpole, S. (2005). Learning new words from storybooks: An efficacy study with at-risk kindergartners. *Language, Speech, and Hearing Services in Schools*, *36*(1), 17-32. doi:10.1044/0161-1461(2005/003)

Justice, L. M., Pullen, P. C., & Pence, K. (2008). Influence of verbal and nonverbal references to print on preschoolers' visual attention to print during storybook reading. *Developmental Psychology*, *44*(3), 855-866. doi:10.1037/0012-1649.44.3.855

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, *38*(1), 23-31. doi:10.1207/s15326985ep3801_4

Kalyuga, S., & Sweller, J. (2014). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.) *The Cambridge Handbook of Multimedia Learning* (pp. 247-262). Cambridge, UK: Cambridge university press.

Kennison, S. M., & Clifton, C. (1995). Determinants of parafoveal preview benefit in high and low working memory capacity readers: implications for eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(1), 68-81. doi:10.1037/0278-7393.21.1.68

Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*(5), 363. doi:10.1037/0033-295x.85.5.363

Koda, K. (1990). The use of L1 reading strategies in L2 reading. *Studies in Second Language Acquisition*, *12*(4), 393-410. doi:10.1017/s0272263100009499

Korat, O. (2010). Reading electronic books as a support for vocabulary, story comprehension and word reading in kindergarten and first grade. *Computers and Education*, *55*(1), 24-31. doi:10.1016/j.compedu.2009.11.014

Kruger, J. L., & Steyn, F. (2014). Subtitles and eye tracking: Reading and performance. *Reading Research Quarterly*, *49*(1), 105-120. doi:10.1002/rrq.59

Kyte, C. S., & Johnson, C. J. (2006). The role of phonological recoding in orthographic learning. *Journal of Experimental Child Psychology*, *93*(2), 166-185. doi:10.1016/j.jecp.2005.09.003

Landi, N. (2010). An examination of the relationship between reading comprehension, higher-level and lower-level reading sub-skills in adults. *Reading and Writing*, *23*(6), 701-717. doi:10.1007/s11145-009-9180-z

Landi, N., Perfetti, C. A., Bolger, D. J., Dunlap, S., & Foorman, B. R. (2006). The role of discourse context in developing word form representations: A paradoxical relation between reading and learning. *Journal of Experimental Child Psychology*, *94*(2), 114-133. doi:10.1016/j.jecp.2005.12.004

Laubrock, J., & Kliegl, R. (2015). The eye-voice span during reading aloud. *Frontiers in Psychology*, *6*, 1432. doi:10.3389/fpsyg.2015.01432

Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, 21(1), 59-80. doi:10.1348/026151003321164627

Leung, C. B. (2008). Preschoolers' acquisition of scientific vocabulary through repeated read-aloud events, retellings, and hands-on science activities. *Reading Psychology, 29*(2), 165-193. doi:10.1080/02702710801964090

Levelt, W.J.M. (1989). *Speaking: From intention to articulation.* Cambridge, MA: MIT Press.

Levy-Schoen, A. (1981). Flexible and/or rigid control of oculomotor scanning behavior. In D. F. Fisher, R. A. Monty, & J. W. Senders (Eds.), *Eye movements: Cognition and Visual Perception* (pp. 299-314). New York: Erlbaum.

Lewis, P. A., & Durrant, S. J. (2011). Overlapping memory replay during sleep builds cognitive schemata. *Trends in Cognitive Sciences*, *15*(8), 343-351. doi:10.1016/j.tics.2011.06.004

Lezak, M. (1995). Neuropsychological Assessment (3rd edition). Oxford: Oxford University Press.

Li, H., Zhang, J., Ehri, L., Chen, Y., Ruan, X., & Dong, Q. (2016). The role of orthography in oral vocabulary learning in Chinese children. *Reading and Writing*, *29*(7), 1363-1381. doi:10.1007/s11145-016-9641-0

Li, P., Zhao, X., & Mac Whinney, B. (2007). Dynamic self-organization and early lexical development in children. *Cognitive Science, 31*(4), 581-612. doi:10.1080/15326900701399905

Liberman, I. Y., Shankweiler, D., Liberman, A. M., Fowler, C., & Fischer, F. W. (1977). Phonetic segmentation and recoding in the beginning reader. In A.S. Reber & D. Scarborough (Eds.) *Towards a Psychology of Reading: The Proceeding of the CUNY conferences* (pp. 207-226). Hillsdale, NJ: Lawrence Erlbaum.

Lin, E. L., & Murphy, G. L. (2001). Thematic relations in adults' concepts. *Journal of Experimental Psychology: General*, *130*(1), 3-28. doi:10.1037//0096-3445.130.1.3

Lin, L. C. (2014). Learning Word Meanings from Teachers’ Repeated Story Read-Aloud in EFL Primary Classrooms. *English Language Teaching*, *7*(7), 68-81. doi:10.5539/elt.v7n7p68

Lucas, R., & Norbury, C. F. (2014). Orthography facilitates vocabulary learning for children with autism spectrum disorders (ASD). *The Quarterly Journal of Experimental Psychology*, *67*(7), 1317-1334. doi:10.1080/17470218.2013.859714

Luce, P. A., & Large, N. R. (2001). Phonotactics, density, and entropy in spoken word recognition. *Language and Cognitive Processes*, *16*(5-6), 565-581. doi:10.1080/01690960143000137

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, *19*(1), 1-36. doi:10.1097/00003446-199802000-00001

Lund, R. J. (1991). A comparison of second language listening and reading comprehension. *Modern Language Journal 75*, 196-204. doi:10.2307/328827

Marchman, V. A., & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language*, *21*, 339-339. doi:10.1017/s0305000900009302

Mark, L. S., Shankweiler, D., Liberman, I. Y., & Fowler, C. A. (1977). Phonetic recoding and reading difficulty in beginning readers. *Memory and Cognition*, *5*(6), 623-629. doi:10.3758/bf03197408

Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science*, *14*(1), 57-77. doi:10.1207/s15516709cog1401_4

Marslen-Wilson, W., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, *8*(1), 1-71. doi:10.1016/0010-0277(80)90015-3

Martin, A. W., & Wiechers, J. E. (1954). Raven's Coloured Progressive Matrices and the Wechsler Scale for children. *Journal of Consulting Psychology*, *18*(2), 143-144. doi:10.1037/h0055913

Masterson, J., Stuart, M., Dixon, M., Lovejoy, D., & Lovejoy, S. (2003). *The children's printed word database.* Retrieved from: https://www1.essex.ac.uk/psychology/cpwd/.

Mayer, R. E. (2014). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.) *The Cambridge Handbook of Multimedia Learning* (pp. 43-71). Cambridge: Cambridge university press.

Mayer, R. E., Moreno, R., Boire, M., & Vagge, S. (1999). Maximizing constructivist learning from multimedia communications by minimizing cognitive load. *Journal of Educational Psychology*, *91*(4), 638-643. doi:10.1037/0022-0663.91.4.638

McBride–Chang, C., Wagner, R. K., Muse, A., Chow, B. W. Y., & Shu, H. (2005). The role of morphological awareness in children's vocabulary acquisition in English. *Applied Psycholinguistics*, *26*(3), 415-435. doi:10.1017/s014271640505023x

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1-86. doi:10.1016/0010-0285(86)90015-0

McClelland, J. L., McNaughton, B. L., & O'reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*(3), 419-457. doi:10.1037/0033-295x.102.3.419

McGregor, K. K. & Duff, D. (2013). Promoting diverse and deep vocabulary development. In T. Ukrainetz (Ed.), *School-age Language Intervention: Principles and Procedures for Oral and Written Language Success.* Austin, TX: Pro-Ed.

McGregor, K. K., Newman, R. M., Reilly, R. M., & Capone, N. C. (2002). Semantic representation and naming in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 45*(5), 998-1014. doi:10.1044/1092-4388(2002/081)

McGregor, K. K., Sheng, L., & Ball, T. (2007). Complexities of expressive word learning over time. *Language, Speech, and Hearing Services in Schools*, *38*(4), 353-364. doi:10.1044/0161-1461(2007/037)

McKague, M., Davis, C., Pratt, C., & Johnston, M. B. (2008). The role of feedback from phonology to orthography in orthographic learning: an extension of item-based accounts. *Journal of Research in Reading, 31*(1), 55-76. doi:10.1111/j.1467-9817.2007.00361.x

McKague, M., Pratt, C., & Johnston, M. B. (2001). The effect of oral vocabulary on reading visually novel words: A comparison of the dual-route-cascaded and triangle frameworks. *Cognition*, *80*(3), 231-262. doi:10.1016/s0010-0277(00)00150-5

McLeod, A. N., & McDade, H. L. (2011). Preschoolers' incidental learning of novel words during storybook reading. *Communication Disorders Quarterly, 32*(4), 256-266. doi:10.1177/1525740109354777

Medler, D. A., & Binder, J. R. (2005). *MCWord: An on-line orthographic database of the English language.* Retrieved from www.neuro.mcw.edu/mcword/.

Mengoni, S. E., Nash, H., & Hulme, C. (2013). The benefit of orthographic support for oral vocabulary learning in children with Down syndrome. *Journal of Child Language*, *40*(01), 221-243. doi:10.1017/s0305000912000396

Menne, J. M., & Menne, J. W. (1972). The relative efficiency of bimodal presentation as an aid to learning. *AV Communication Review*, *20*(2), 170-180. Retrieved from: http://www.jstor.org/stable/30219753

Messer, D. J. (1981). The identification of names in maternal speech to infants. *Journal of Psycholinguistic Research, 10*(1), 69-77. doi:10.1007/bf01067362

Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology, 91*(1), 3-19. doi:10.1037//0022-0663.91.1.3

Metsala, J.L. & Walley, A.C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In J.L. Metsala & L.C. Ehri (Eds.), *Word Recognition in Beginning Literacy* (pp. 89–120). Mahwah, NJ: Lawrence Erlbaum.

Miles, K. P., Ehri, L. C., & Lauterbach, M. D. (2016). Mnemonic value of orthography for vocabulary learning in monolinguals and language minority English-speaking college students. *Journal of College Reading and Learning*, *46*(2), 99-112. doi:10.1080/10790195.2015.1125818

Mol, S. E., Bus, A. G., & de Jong, M. T. (2009). Interactive book reading in early education: A tool to stimulate print knowledge as well as oral language. *Review of Educational Research*, *79*(2), 979-1007. doi:10.3102/0034654309332561

Monaghan, P. J., Chang, Y. N., & Welbourne, S. (2017). *Different processes for reading words learned before and after onset of literacy*. In Proceedings of the 39th Annual Conference of the Cognitive Science Society (p. 811). Cognitive Science Society: Austin, Texas.

Morton, J. (1979). Facilitation in word recognition: experiments causing change in the logogen model. In P. A. Kolers, M. Wrolstamd, & H. Boumah (Eds.), *Processing of Visible Language* (pp. 259-268). New York: Plenum Press.

Muneaux, M., & Ziegler, J. (2004). Locus of orthographic effects in spoken word recognition: Novel insights from the neighbour generation task. *Language and Cognitive Processes*, *19*(5), 641-660. doi:10.1080/01690960444000052

Murphy, M. L. (2003). *Semantic relations and the lexicon: antonymy, synonymy and other paradigms*. Cambridge: Cambridge University Press.

Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development:

evidence from a longitudinal study. *Developmental Psychology*, *40*(5), 665-681. doi:10.1037/0012-1649.40.5.665

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, *17*(2), 357-374. doi:10.1017/s0305000900013817

Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal*, *24*(2), 237-270. doi:10.2307/1162893

Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly 20* (2), 233-253. doi:10.2307/747758

Nagy, W. E., & Scott, J. A. (2000). Vocabulary processes. In M.L. Kamil, P.B. Mosenthal, P.D. Pearson, & R. Barr (Eds.), *Handbook of Reading Research, Vol. 3* (pp. 269-284). Mahwah, NJ: Lawrence Erlbaum.

Närhi, V., & Ahonen, T. (1995). Reading disability with or without, attention deficit hyperactivity, disorder: Do attentional problems, make a difference? *Developmental Neuropsychology*, 11(3), 337-349. doi:10.1080/87565649509540624

Närhi, V., Räsänen, P., Metsäpelto, R. L., & Ahonen, T. (1997). Trail Making Test in assessing children with reading disabilities: A test of executive functions or content information. *Perceptual and Motor Skills*, *84*(3_suppl), 1355-1362. doi:10.2466/pms.1997.84.3c.1355

Nation, I. S. P. (2013). *Teaching and learning vocabulary*. Boston: Heinle Cengage Learning.

Nation, K. (2017). Nurturing a lexical legacy: reading experience is critical for the development of word reading skill. *Npj Science of Learning*, *2*, 3. doi:10.1038/s41539-017-0004-7

Nation, K., Angell, P., & Castles, A. (2007). Orthographic learning via self-teaching in children learning to read English: Effects of exposure, durability, and context. *Journal of Experimental Child Psychology*, *96*(1), 71-84. doi:10.1016/j.jecp.2006.06.004

Nation, K., & Snowling, M. J. (1998). Individual differences in contextual facilitation: Evidence from dyslexia and poor reading comprehension. *Child Development*, *69*(4), 996-1011. doi:10.2307/1132359

Nation, K., & Snowling, M. J. (1999). Developmental differences in sensitivity to semantic relations among good and poor comprehenders: Evidence from semantic priming. *Cognition*, *70*(1), B1-B13. doi:10.1016/s0010-0277(99)00004-9

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading*, *27*(4), 342-356. doi:10.1111/j.1467-9817.2004.00238.x

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G.W. Humphreys (Eds.), *Basic Processes in Reading: Visual Word Recognition* (pp. 264-336). Hillsdale, NJ: Lawrence Erlbaum.

Nelson, J. R. (2004). *Vocabulary Learning in Modality-Specific Context: Building Complete Representations* (Doctoral dissertation, University of Pittsburgh, USA). Retrieved from: http://d-scholarship.pitt.edu/10276

Nelson, J. R., Balass, M., & Perfetti, C. A. (2005). Differences between written and spoken input in learning new words. *Written Language and Literacy, 8*(2), 25-44. doi:10.1075/wll.8.2.04nel

Neuman, S. B., & Koskinen, P. (1992). Captioned television as comprehensible input: Effects of incidental word learning from context for language minority students. *Reading Research Quarterly 27*(1), 94-106. doi:10.2307/747835

Oakhill, J. (1983). Instantiation in skilled and less skilled comprehenders. *Quarterly Journal of Experimental Psychology, 35*(3), 441-450. doi:10.1080/14640748308402481

Oetting, J. B., Rice, M. L., & Swank, L. K. (1995). Quick incidental learning (QUIL) of words by school-age children with and without SLI. *Journal of Speech, Language, and Hearing Research*, *38*(2), 434-445. doi:10.1044/jshr.3802.434

Olson, R. K., Forsberg, H., & Wise, B. (1994). Genes, environment, and the development of orthographic skills. In V. Wise Berninger (Eds.), *The Varieties of Orthographic Knowledge* (pp. 27-71). Springer Netherlands.

Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology, 98*(3), 554-566. doi:10.1037/0022-0663.98.3.554

Ouellette, G., & Fraser, J. R. (2009). What exactly is a yait anyway: The role of semantics in orthographic learning. *Journal of Experimental Child Psychology*, *104*(2), 239-251. doi:10.1016/j.jecp.2009.05.001

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist, 38*(1), 1-4. doi:10.1207/s15326985ep3801_1

Paivio, A. (1986). *Mental representations: A dual coding approach.* Oxford, UK: Oxford University Press.

Paribakht, T. S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In J. Coay & T. Huckin (Eds.), *Second Language Vocabulary Acquisition: A Rationale for Pedagogy* (pp. 174-200). Cambridge, UK: Cambridge University Press.

Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly*, *40*(2), 184-202. doi:10.1598/rrq.40.2.3

Patterson, K., Graham, N., & Hodges, J. R. (1994). The impact of semantic memory loss on phonological representations. *Journal of Cognitive Neuroscience, 6*(1), 57-69. doi:10.1162/jocn.1994.6.1.57

Penno, J. F., Wilkinson, I. A., & Moore, D. W. (2002). Vocabulary acquisition from teacher explanation and repeated listening to stories: Do they overcome the

Matthew effect? *Journal of Educational Psychology*, *94*(1), 23-33. doi:10.1037/0022-0663.94.1.23

Perea, M., & Pollatsek, A. (1998). The effects of neighborhood frequency in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 767-779. doi:10.1037/0096-1523.24.3.767

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). Hillside, N.J.: Lawrence Erlbaum.

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*(4), 357-383. doi:10.1080/10888430701530730

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: John Benjamins.

Pfost, M., Hattie, J., Dörfler, T., & Artelt, C. (2014). Individual differences in reading development: A review of 25 years of empirical research on Matthew effects in reading. *Review of Educational Research*, *84*(2), 203-244. doi:10.3102/0034654313509492

Pick, A. D., Unze, M. G., Brownell, C. A., Drozdal Jr, J. G., & Hopmann, M. R. (1978). Young children's knowledge of word structure. *Child Development 49*(3), 669-680. doi:10.2307/1128234

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*, 56 –115. doi:10.1037/0033-295x.103.1.56

Plunkett, K., Sinha, C., Møller, M. F., & Strandsby, O. (1992). Symbol grounding or the emergence of symbols? Vocabulary growth in children and a connectionist net. *Connection Science*, *4*(3-4), 293-312. doi:10.1080/09540099208946620

Pollatsek, A., Lesch, M., Morris, R. K., & Rayner, K. (1992). Phonological codes are used in integrating information across saccades in word identification and reading. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(1), 148-162. doi:10.1037/0096-1523.18.1.148

Pollatsek, A., Perea, M., & Binder, K. S. (1999). The effects of "neighborhood size" in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance, 25*(4), 1142-1158. doi:10.1037/0096-1523.25.4.1142

Pollatsek, A., Slattery, T. J., & Juhasz, B. (2008). The processing of novel and lexicalised prefixed words in reading. *Language and Cognitive Processes*, *23*(7-8), 1133-1158. doi:10.1080/01690960801945484

Powell, D., Plaut, D. C., & Funnell, E. (2006). Does the PMSP connectionist model of single word reading learn to read in the same way as a child? *Journal of Research in Reading, 29*(2), 229-250. doi:10.1111/j.1467-9817.2006.00300.x

Priya, K., & Wagner, R. K. (2009). The roles of fluent decoding and vocabulary in the development of reading comprehension. In R.K. Wagner, C. Schatschneider, & C. Phythian-Sence (Eds.), *Beyond Decoding: The Behavioral and Biological Foundations of Reading Comprehension* (pp. 124-139). New York: The Guilford Press.

R Core Team (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL: http://www.R-project.org/

Radach, R., Huestegge, L., & Reilly, R. (2008). The role of global top-down factors in local eye-movement control in reading. *Psychological Research, 72*(6), 675-688. doi:10.1007/s00426-008-0173-3

Raney, G. E., & Rayner, K. (1995). Word frequency effects and eye movements during two readings of a text. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, 49*(2), 151-172. doi:10.1037/1196-1961.49.2.151

Rasinski, T. V. (1989). Fluency for everyone: Incorporating fluency instruction in the classroom. *The Reading Teacher, 42*(9), 690-693. Retrieved from: http://www.jstor.org/stable/20200275

Rasinski, T. V. (1990). Effects of repeated reading and listening-while-reading on reading fluency. *The Journal of Educational Research*, 147-150. doi:10.1080/00220671.1990.10885946

Rastle, K., & Brysbaert, M. (2006). Masked phonological priming effects in English: Are they real? Do they matter? *Cognitive Psychology, 53*(2), 97-145. doi:10.1016/j.cogpsych.2006.01.002

Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology, 41*(2), 211-236. doi:10.1016/0022-0965(86)90037-8

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372-422. doi:10.1037/0033-2909.124.3.372

Rayner, K., Pollatsek, A., Ashby, J., & Clifton Jr, C. (2012). *Psychology of Reading*. New York: Psychology Press.

Rayner, K., Rotello, C. M., Stewart, A. J., Keir, J., & Duffy, S. A. (2001). Integrating text and pictorial information: eye movements when looking at print advertisements. *Journal of Experimental Psychology: Applied, 7*(3), 219-226. doi:10.1037/1076-898x.7.3.219

Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin and Review, 3*(4), 504-509. doi:10.3758/bf03214555

Reichle, E. D., Liversedge, S. P., Drieghe, D., Blythe, H. I., Joseph, H. S., White, S. J., & Rayner, K. (2013). Using EZ Reader to examine the concurrent development of eye-movement control and reading skill. *Developmental Review, 33*(2), 110-149. doi:10.1016/j.dr.2013.03.001

Reichle, E. D., & Perfetti, C. A. (2003). Morphology in word identification: A word-experience model that accounts for morpheme frequency effects. *Scientific Studies of Reading*, *7*(3), 219-237. doi:10.1207/s1532799xssr0703_2

Reitan, R., & Wolfson, D. (1993). *The Halstead-Reitan Neuropsychological Test Battery: Theory and clinical interpretation.* Tucson, AZ: Neuropsychology Press.

Rice, M. L., (1990). Preschoolers' QUIL: Quick incidental learning of words. In G. Conti-Ramsden & C. Snow (Eds.), *Children's Language, Volume 7* (pp. 171-195). Hillsdale, NJ: Lawrence Erlbaum.

Rice, M. L., Buhr, J. C., & Nemeth, M. (1990). Fast mapping word-learning abilities of language-delayed preschoolers. *Journal of Speech and Hearing Disorders*, *55*(1), 33-42. doi:10.1044/jshd.5501.33

Riches, N. G., Tomasello, M., & Conti-Ramsden, G. (2005). Verb learning in children with SLI: Frequency and spacing effects. *Journal of Speech, Language, and Hearing Research*, *48*(6), 1397-1411. doi:10.1044/1092-4388(2005/097)

Ricketts, J., Bishop, D. V., & Nation, K. (2009). Orthographic facilitation in oral vocabulary acquisition. *The Quarterly Journal of Experimental Psychology*, *62*(10), 1948-1966. doi:10.1080/17470210802696104

Ricketts, J., Bishop, D. V., Pimperton, H., & Nation, K. (2011). The role of self-teaching in learning orthographic and semantic aspects of new words. *Scientific Studies of Reading*, *15*(1), 47-70. doi:10.1080/10888438.2011.536129

Ricketts, J., Dockrell, J. E., Patel, N., Charman, T., & Lindsay, G. (2015). Do children with specific language impairment and autism spectrum disorders benefit from the presence of orthography when learning new spoken words? *Journal of Experimental Child Psychology*, *134*, 43-61. doi:10.1016/j.jecp.2015.01.015

Ricketts, J., Nation, K., & Bishop, D. V. (2007). Vocabulary is important for some, but not all reading skills. *Scientific Studies of Reading*, *11*(3), 235-257. doi:10.1080/10888430701344306

Robbins, C., & Ehri, L. C. (1994). Reading storybooks to kindergartners helps them learn new vocabulary words. *Journal of Educational Psychology*, *86*(1), 54-64. doi:10.1037/0022-0663.86.1.54

Rosenthal, J., & Ehri, L. C. (2008). The mnemonic value of orthography for vocabulary learning. *Journal of Educational Psychology*, *100*(1), 175-191. doi:10.1037/0022-0663.100.1.175

Rosenthal, J., & Ehri, L. C. (2011). Pronouncing new words aloud during the silent reading of text enhances fifth graders' memory for vocabulary words and their spellings. *Reading and Writing*, *24*(8), 921-950. doi:10.1007/s11145-010-9239-x

Ross, N. M., & Kowler, E. (2013). Eye movements while viewing narrated, captioned, and silent videos. *Journal of Vision*, *13*(4), 1-19. doi:10.1167/13.4.1

Roth, F. P., Speece, D. L., & Cooper, D. H. (2002). A longitudinal analysis of the connection between oral language and early reading. *The Journal of Educational Research*, *95*(5), 259-272. doi:10.1080/00220670209596600

Roy-Charland, A., Saint-Aubin, J., & Evans, M. A. (2007). Eye movements in shared book reading with children from kindergarten to grade 4. *Reading and Writing*, *20*(9), 909-931. doi:10.1007/s11145-007-9059-9

Rust, J. (2008). *Coloured Progressive Matrices and Chrichton Vocabulary Scale Manual*. London: Pearson.

Samuels, S. J., & Flor, R. F. (1997). The importance of automaticity for developing expertise in reading. *Reading and Writing Quarterly*, *13*(2), 107-121. doi:10.1080/1057356970130202

Scarborough, D. L., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance*, *3*(1), 1-17. doi:10.1037/0096-1523.3.1.1

Schatz, E. K., & Baldwin, R. S. (1986). Context clues are unreliable predictors of word meanings. *Reading Research Quarterly*, *21*(4), 439-453. doi:10.2307/747615

Schmitt, N. (1998). Tracking the incremental acquisition of second language vocabulary: A longitudinal study. *Language Learning*, *48*(2), 281-317. doi:10.1111/1467-9922.00042

Schneeberg, H. (1977). Listening while reading: A four year study. *The Reading Teacher*, *30*(6), 629-635. Retrieved from: http://www.jstor.org/stable/20194347

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime (Version 2.0). Computer software and manual.* Pittsburgh, PA: Psychology Software Tools Inc.

Schwanenflugel, P. J. (1991). Contextual constraint and lexical processing. *Advances in Psychology*, *77*, 23-45. doi:10.1016/s0166-4115(08)61528-9

Schwanenflugel, P. J., Stahl, S. A., & Mcfalls, E. L. (1997). Partial word knowledge and vocabulary growth during reading comprehension. *Journal of Literacy Research*, *29*(4), 531-553. doi:10.1080/10862969709547973

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*(4), 523 - 568. doi:10.1037/0033-295x.96.4.523

Seidenberg, M. S., & Tanenhaus, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory*, *5*(6), 546-554. doi:10.1037/0278-7393.5.6.546

Semel, E., Wiig, E., & Secord, W. (2006). *Clinical Evaluation of Language Fundamentals–4th Ed (CELF - 4 UK)*. London:Pearson.

Sénéchal, M. (1997). The differential effect of storybook reading on preschoolers' acquisition of expressive and receptive vocabulary. *Journal of Child Language*, *24*(01), 123-138. doi:10.1017/s0305000996003005

Sénéchal, M., & Cornell, E. H. (1993). Vocabulary acquisition through shared reading experiences. *Reading Research Quarterly*, *28*(4), 360-374. doi:10.2307/747933

Sénéchal, M., Thomas, E., & Monker, J. A. (1995). Individual differences in 4-year-old children's acquisition of vocabulary during storybook reading. *Journal of Educational Psychology, 87*(2), 218-229. doi:10.1037//0022-0663.87.2.218

Sereno, S. C., O'Donnell, P. J., & Rayner, K. (2006). Eye movements and lexical ambiguity resolution: Investigating the subordinate-bias effect. *Journal of Experimental Psychology: Human Perception and Performance*,*32*(2), 335-350. doi:10.1037/e537052012-045

Shamir, A., Korat, O., & Fellah, R. (2012). Promoting vocabulary, phonological awareness and concept about print among children at risk for learning disability: can e-books help? *Reading and Writing*, *25*(1), 45-69. doi:10.1007/s11145-010-9247-x

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

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

Shefelbine, J. L. (1990). Student factors related to variability in learning word meanings from context. *Journal of Reading Behavior*, *22*(1), 71-97. doi:10.1080/10862969009547695

Silverman, R., & Hines, S. (2009). The effects of multimedia-enhanced instruction on the vocabulary of English-language learners and non-English-language learners in pre-kindergarten through second grade. *Journal of Educational Psychology*, *101*(2), 305-314. doi:10.1037/a0014217

Siskind, J. M. (1996). A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition*, *61*(1), 39-91. doi:10.1016/s0010-0277(96)00728-7

Slowiaczek, L. M., Soltano, E. G., Wieting, S. J., & Bishop, K. L. (2003). An investigation of phonology and orthography in spoken-word recognition. *The*

Quarterly Journal of Experimental Psychology: Section A, *56*(2), 233-262. doi:10.1080/02724980244000323

Smeets, D. J. H., & Bus, A. G. (2015). The interactive animated e-book as a word learning device for kindergartners. *Applied Psycholinguistics*, *36*(4), 899-920. doi:10.1017/s0142716413000556

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, *81*(3), 214-241. doi:10.1037/h0036351

Smith, L. B., Jones, S. S., & Landau, B. (1996). Naming in young children: A dumb attentional mechanism? *Cognition*, *60*(2), 143-171. doi:10.1016/0010-0277(96)00709-3

Smith, L.B., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*(3), 1558-1568. doi:10.1016/j.cognition.2007.06.010

Snowling, M. J., Stothard, S. E., Clarke, P., Bowyer-Crane, C., Harrington, A., Truelove, E., & Hulme, C. (2009). *York Assessment of Reading for Comprehension*. London: GL Assessment.

Solomon, R. L., & Howes, D. H. (1951). Word frequency, personal values, and visual duration thresholds. *Psychological Review*, *58*(4), 256-270. doi:10.1037/h0058228

Stahl, S. (1983). Differential word knowledge and reading comprehension. *Journal of Literacy Research*, *15*(4), 33-50. doi:10.1080/10862968309547495

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, and Computers*, *31*(1), 137-149. doi:10.3758/bf03207704

Stanley, P. D., & Ginther, D. W. (1991). The effects of purpose and frequency on vocabulary learning from written context of high and low ability reading comprehenders. *Literacy Research and Instruction*, *30*(4), 31-41. doi:10.1080/19388079109558059

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*(4), 360-407. doi:10.1598/rrq.21.4.1

Stanovich, K. E., Nathan, R. G., West, R. F., & Vala-Rossi, M. (1985). Children's word recognition in context: Spreading activation, expectancy, and modularity. *Child Development, 56*(6), 1418-1428. doi:10.2307/1130461

Steffler, D. J., Varnhagen, C. K., Friesen, C. K., & Treiman, R. (1998). There's more to children's spelling than the errors they make: Strategic and automatic processes for one-syllable words. *Journal of Educational Psychology, 90*(3), 492-505. doi:10.1037//0022-0663.90.3.492

Stemberger, J. P. (1985). An interactive activation model of language production. In A. Ellis (Ed.), *Progress in the Psychology of Language*, Vol. 1 (pp. 143-186). London: Lawrence Erlbaum.

Sternberg, R. J., & Powell, J. S. (1983). Comprehending verbal comprehension. *American Psychologist, 38*(8), 878-893. doi:10.1037/0003-066x.38.8.878

Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science, 29*(1), 41-78. doi:10.1207/s15516709cog2901_3

Storkel, H. L. (2001). Learning new words: Phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research, 44*(6), 1321-1337. doi:10.1044/1092-4388(2001/103)

Storkel, H. L., Armbrüster, J., & Hogan, T. P. (2006). Differentiating phonotactic probability and neighborhood density in adult word learning. *Journal of Speech, Language, and Hearing Research, 49*(6), 1175-1192. doi:10.1044/1092-4388(2006/085)

Storkel, H. L., & Rogers, M. A. (2000). The effect of probabilistic phonotactics on lexical acquisition. *Clinical Linguistics and Phonetics, 14*(6), 407-425. doi:10.1080/026992000415859

Stuart, M., Masterson, J., & Dixon, M. (2000). Spongelike acquisition of sight vocabulary in beginning readers? *Journal of Research in Reading*, *23*(1), 12-27. doi:10.1111/1467-9817.00099

Stuss, D., Stethem, L., Hugenholtz, H., & Richard, M. (1989). Traumatic brain injury: A comparison of three clinical tests, and analysis of recovery. *Clinical Neuropsychologist*, *3*(2), 145-156. doi:10.1080/13854048908403287

Suggate, S. P., Lenhard, W., Neudecker, E., & Schneider, W. (2013). Incidental vocabulary acquisition from stories: Second and fourth graders learn more from listening than reading. *First Language*, *33*(6), 551-571. doi:10.1177/0142723713503144

Swanborn, M. S., & de Glopper, K. (1999). Incidental word learning while reading: A meta-analysis. *Review of Educational Research*, *69*(3), 261-285. doi:10.2307/1170540

Swanborn, M. S., & de Glopper, K. (2002). Impact of reading purpose on incidental word learning from context. *Language Learning*, *52*(1), 95-117. doi:10.1111/1467-9922.00178

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, *22*(2), 123-138. doi:10.1007/s10648-010-9128-5

Sydorenko, T. (2010). Modality of input and vocabulary acquisition. *Language, Learning and Technology*, *14*(2), 50-73. Retreved from: http://llt.msu.edu/vol14num2/sydorenko.pdf

Takacs, Z. K., Swart, E. K., & Bus, A. G. (2014). Can the computer replace the adult for storybook reading? A meta-analysis on the effects of multimedia stories as compared to sharing print stories with an adult. *Frontiers in Psychology*, *5*, 1366. doi:10.3389/fpsyg.2014.01366

Tamura, N., Castles, A., & Nation, K. (2017). Orthographic learning, fast and slow: Lexical competition effects reveal the time course of word learning in developing readers. *Cognition*, *163*, 93-102. doi:10.1016/j.cognition.2017.03.002

Tannenbaum, K. R., Torgesen, J. K., & Wagner, R. K. (2006). Relationships between word knowledge and reading comprehension in third-grade children. *Scientific Studies of Reading*, *10*(4), 381-398. doi:10.1207/s1532799xssr1004_3

Taylor, J. S. H., Duff, F. J., Woollams, A. M., Monaghan, P., & Ricketts, J. (2015). How word meaning influences word reading. *Current Directions in Psychological Science*, *24*(4), 322-328. doi:10.1177/0963721415574980

Tenpenny, P. L. (1995). Abstractionist versus episodic theories of repetition priming and word identification. *Psychonomic Bulletin and Review*, *2*(3), 339-363. doi:10.3758/bf03210972

Terrell, S. L., & Daniloff, R. (1996). Children's word learning using three modes of instruction. *Perceptual and Motor Skills*, *83*(3), 779-787. doi:10.2466/pms.1996.83.3.779

Thorn, A. S., & Frankish, C. R. (2005). Long-term knowledge effects on serial recall of nonwords are not exclusively lexical. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(4), 729-735. doi:10.1037/0278-7393.31.4.729

Tiffin-Richards, S. P., & Schroeder, S. (2015). Word length and frequency effects on children’s eye movements during silent reading. *Vision Research*, *113*, 33-43. doi:10.1016/j.visres.2015.05.008

Tincoff, R., & Jusczyk, P. W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, *10*(2), 172-175. doi:10.1111/1467-9280.00127

Tincoff, R., & Jusczyk, P. W. (2012). Six-Month-Olds comprehend words that refer to parts of the body. *Infancy, 17*(4), 432-444. doi:10.1111/j.1532-7078.2011.00084.x

Tomasello, M. (2000). The social-pragmatic theory of word learning. *Pragmatics, 10*(4), 401–413. doi:10.1075/prag.10.4.01tom

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Test of Word Reading Efficiency*. Austin: Texas: Pro-ed.

Vadasy, P. F., & Sanders, E. A. (2015). Incremental learning of difficult words in story contexts: the role of spelling and pronouncing new vocabulary. *Reading and Writing*, *28*(3), 371-394. doi:10.1007/s11145-014-9529-9

van der Lely, H. K. (2005). Domain-specific cognitive systems: insight from Grammatical-SLI. *Trends in Cognitive Sciences*, *9*(2), 53-59. doi:10.1016/j.tics.2004.12.002

Van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology*, *67*(6), 1176-1190. doi:10.1080/17470218.2013.850521

Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory and Cognition*, *15*(3), 181-198. doi:10.3758/bf03197716

Vellutino, F. R., Tunmer, W. E., Jaccard, J. J., & Chen, R. (2007). Components of reading ability: Multivariate evidence for a convergent skills model of reading development. *Scientific Studies of Reading*, *11*(1), 3-32. doi:10.1207/s1532799xssr1101_2

Ventura, P., Morais, J., Pattamadilok, C., & Kolinsky, R. (2004). The locus of the orthographic consistency effect in auditory word recognition. *Language and Cognitive Processes*, *19*(1), 57-95. doi:10.1080/016909603440001

Verhallen, M. J., & Bus, A. G. (2010). Low-income immigrant pupils learning vocabulary through digital picture storybooks. *Journal of Educational Psychology*, *102*(1), 54-61. doi:10.1037/a0017133

Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning*, *61*(1), 219-258. doi:10.1111/j.1467-9922.2010.00593.x

Vitevitch, M. S., & Luce, P. A. (2004). A web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, Instruments, & Computers, 36*(3), 481-487. doi:10.3758/s13428-017-0872-z

Vlach, H., & Sandhofer, C. (2010). Desirable difficulties in cross-situational word learning. *Proceedings of the Cognitive Science Society, 32,* 2470-2475. Retrieved from: http://escholarship.org/uc/item/2151507t

Walley, A.C. (1993). The role of vocabulary development in children's spoken word recognition and segmentation ability. *Developmental Review, 13,* 286–350. doi:10.1006/drev.1993.1015

Walley, A. C., Metsala, J. L., & Garlock, V. M. (2003). Spoken vocabulary growth: Its role in the development of phoneme awareness and early reading ability. *Reading and Writing*, *16*(1), 5-20. doi:10.1023/a:1021789804977

Walsh, B. A., & Blewitt, P. (2006). The effect of questioning style during storybook reading on novel vocabulary acquisition of preschoolers. *Early Childhood Education Journal, 33*(4), 273-278. doi:10.1007/s10643-005-0052-0

Wang, H. C. [Hsueh-Cheng], & Pomplun, M. (2012). The attraction of visual attention to texts in real-world scenes. *Journal of Vision, 12*(6), 26. doi:10.1167/12.6.26

Wang, H. C. [Hua-Chen], Castles, A., Nickels, L., & Nation, K. (2011). Context effects on orthographic learning of regular and irregular words. *Journal of Experimental Child Psychology, 109*(1), 39-57. doi:10.1016/j.jecp.2010.11.005

Waters, G. S., & Seidenberg, M. S. (1985). Spelling-sound effects in reading: Time-course and decision criteria. *Memory and Cognition*, *13*(6), 557-572. doi:10.3758/bf03198326

Wechsler, D. (2005). *Wechsler Individual Achievement. (WIAT II).* The Psychological Corp, London.

Weekes, B. S. (1997). Differential effects of number of letters on word and nonword naming latency. *The Quarterly Journal of Experimental Psychology: Section A, 50*(2), 439-456. doi:10.1080/027249897392170

Wegener, S., Wang, H. C., de Lissa, P., Robidoux, S., Nation, K., & Castles, A. (2017). Children reading spoken words: Interactions between vocabulary and orthographic expectancy. *Developmental Science,* e12577. doi:10.1111/desc.12577

Weighall, A. R., Henderson, L. M., Barr, D. J., Cairney, S. A., & Gaskell, M. G. (2017). Eye-tracking the time-course of novel word learning and lexical competition in adults and children. *Brain and Language*, *167*, 13-27. doi:10.1016/j.bandl.2016.07.010

Weizman, Z. O., & Snow, C. E. (2001). Lexical output as related to children's vocabulary acquisition: Effects of sophisticated exposure and support for meaning. *Developmental Psychology, 37*(2), 265. doi:10.1037//0012-1649.37.2.265

Whaley, C. P. (1978). Word—nonword classification time. *Journal of Verbal Learning and Verbal Behavior*, *17*(2), 143-154. doi:10.1016/s0022-5371(78)90110-x

Wilhelm, I., Prehn-Kristensen, A., & Born, J. (2012). Sleep-dependent memory consolidation–what can be learnt from children? *Neuroscience and Biobehavioral Reviews, 36*(7), 1718-1728. doi:10.1016/j.neubiorev.2012.03.002

Williams, R., & Morris, R. (2004). Eye movements, word familiarity, and vocabulary acquisition. *European Journal of Cognitive Psychology*, *16*(1-2), 312-339. doi:10.1080/09541440340000196

Williams, S. E., & Horst, J. S. (2014). Goodnight book: sleep consolidation improves word learning via storybooks. *Frontiers in Psychology*, *5*, 184. doi:10.3389/fpsyg.2014.00184

Wilkinson, K. S., & Houston-Price, C. (2013). Once upon a time, there was a pulchritudinous princess...: The role of word definitions and multiple story contexts in children's learning of difficult vocabulary. *Applied Psycholinguistics*, *34*(3), 591-613. doi: 10.1017/s0142716411000889

Wise, J. C., Sevcik, R. A., Morris, R. D., Lovett, M. W., & Wolf, M. (2007). The relationship among receptive and expressive vocabulary, listening comprehension, pre-reading skills, word identification skills, and reading comprehension by children with reading disabilities. *Journal of Speech, Language, and Hearing Research, 50*(4), 1093-1109. doi:10.1044/1092-4388(2007/076)

Wochna, K. L., & Juhasz, B. J. (2013). Context length and reading novel words: An eye-movement investigation. *British Journal of Psychology*,*104*(3), 347-363. doi:10.1111/j.2044-8295.2012.02127.x

Wodka, L.E., Loftis, C., Mostofsky, H.S., Prahme, C., Larson, C.G.J., Denckla, B.M., Mahone, M. (2008). Prediction of ADHD in boys and girls using the D-KEFS. *Archives of Clinical Neuropsychology, 23*(3), 283–293. doi:10.1016/j.acn.2007.12.004

Wolter, J. A., & Apel, K. (2010). Initial acquisition of mental graphemic representations in children with language impairment. *Journal of Speech, Language, and Hearing Research, 53*(1), 179-195. doi:10.1044/1092-4388(2009/07-0130)

Woodward, A.L., & Markman, E.M. (1998). Early word learning. In W. Damon (Series Ed.), D. Kuhn & R. Siegler (Eds.), *Handbook of Child Psychology, Vol. 2, Cognition, Perception, and Language* (pp. 371-420). New York: Wiley.

Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, *18*(5), 414-420. doi:10.1111/j.1467-9280.2007.01915.x

Yuill, N., Oakhill, J., & Parkin, A. (1989). Working memory, comprehension ability and the resolution of text anomaly. *British Journal of Psychology*, *80*(3), 351-361. doi:10.1111/j.2044-8295.1989.tb02325.x

Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin and Review*, *5*(4), 683-689. doi:10.3758/bf03208845

Ziegler, J. C., Ferrand, L., & Montant, M. (2004). Visual phonology: The effects of orthographic consistency on different auditory word recognition tasks. *Memory and Cognition*, *32*(5), 732-741. doi:10.3758/bf03195863

Ziegler, J. C., Van Orden, G. C., & Jacobs, A. M. (1997). Phonology can help or hurt the perception of print. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(3), 845-860. doi:10.1037/0096-1523.23.3.845

Appendices:

Appendix A: word lists for Study 1 and their features

Set	**Word**	**Length**	**frequency MRC**	**CELEX COBUILD frequency**	**frequency SUBTLEX-UK**	**zipft SUBTLEX-UK**
Normans	motte	5	--	--	58	2.47
	palisade	8	4	6	30	2.19
	furrier	7	0	5	8	1.65
	trencher	8	0	2	5	1.47
	hauberk	7	2	0	--	--
	destrier	8	--	--	--	--
	wain	4	3	8	124	2.79
	pottage	7	5	4	42	2.33
India	mausoleum	9	3	23	122	2.79
	verandah	8	0	163	13	1.84
	viceroy	7	9	28	105	2.72
	teapoy	6	0	--	16	1.93
	dhoti	5	--	6	3	1.3
	gavial	6	--	--	27	2.14
	palanquin	9	0	9	9	1.7
	toddy	5	0	3	138	2.84
control	ashram	6	--	--	18	1.97
	atrium	6	0	--	105	2.72
	augur	5	1	9	26	2.13
	triclinium	10	--	--	2	1.17
	wimple	6	0	11	18	1.97
	palfrey	7	3	1	55	2.54
	gharry	6	--	0	--	--
	hippocras	9	--	--	--	--

Appendix B: definitions of the target words presented in the stories in Study 1

Story	**Word**	**Definition**
Normans	motte	a hillock built by men, on top of which a castle is built
	furrier	a hunter and dealer of animals for their fur
	hauberk	a piece of armour that covers the upper part of the body, made of metal rings and chains
	destrier	horse used for fighting
	palisade	a barrier made of wooden stakes
	wain	an open cart, pulled by an animal
	pottage	a thick liquid made with vegetables and meat, used as food by farmers
	trencher	a flat wooden dish
India	mausoleum	an impressive construction, housing the body of a famous person after his death
	viceroy	the governor of the country of India chosen by the king
	dhoti	a piece of cloth, wrapped around the waist and the legs, worn by men
	gavial	a type of crocodile with an elongated muzzle, that eats fish
	verandah	a covered gallery along the front of a house
	palanquin	a large box moved by four people, holding two poles, used by rich people to be carried around
	toddy	an alcoholic drink made with the sap of palm trees
	teapoy	a small three legged table

Appendix C: stories written for Study 1

<u>Normans Story:</u>

"**How Fred became a knight**

Jen lived in a cottage with her father, a furrier, who is a hunter and dealer of animals for their fur. They were neither very rich nor very poor, but they had an important friend: Fred, the knight who lived in the castle at the top of the motte, which is a hillock built by men, on top of which a castle is built.

None of Jen's friend had such important friends, so one day, before going to bed, she asked her father: "*Dad, how did you and Fred became friends? It's unusual for a knight to be friends with people like us!*"

"*It's very simple!*" he replied: "*Fred has not always been a knight!*"

And he told her the story of how his friend Fred became a knight:

"*When he was young, Fred lived in a rickety house. He was a farmer. All he owned was an old cottage and a garden, enclosed by a palisade, a barrier made of wooden stakes. In the garden he had an old donkey, and a small wain, an open cart, pulled by an animal.*

Fred had always dreamt of becoming a knight. He wanted to fight dragons and ride a fine horse, but he couldn't imagine that one day he would live his dream.

But one day something changed. The king had lost many soldiers during the war, and he was looking for young men to fight with him, and become his knights.

Fred was extremely excited by the news, but I enjoyed my job as a furrier, so I wasn't interested.

"I will set off today, and I will reach the king's castle in less than a month!" he said.

"You are mad" I replied: "It is a dangerous journey! You will have to pass through the mountains, and then through the valley of the dragons!"

But Fred had always been brave and stubborn, and I couldn't change his mind. I decided to go with him, because I was worried that something bad would happen to him if he went on his own.

The next morning we tied the donkey to the wain, and we set off in the direction of the castle.

In every village we passed, we heard people discussing the big news, that the king was looking for new knights. We also saw many men on their way to the castle.

We saw a man riding a destrier, a horse used for fighting. "I wish I had a strong horse like that!" Said Fred: "How can I become a knight without a proper horse?"
A few hours later we saw a knight who was wearing a hauberk, a piece of armour that covers the upper part of the body, made of metal rings and chains. "That is a very expensive thing to wear" said Fred. "I don't have enough money to buy one. How can I fight like a knight without armour?"
At that point Fred became very sad. Compared to the other men's possessions, his own seemed unsuitable for a knight! I felt sorry for him, because I knew he would be the best of knights, as he was brave and kind!
We continued on our journey, but we soon finished the food we had brought with us, so we stopped in a village to work in exchange for food. While we were asking around for work we met an old man. He seemed very upset.
"Can you help me?" he asked.
"Of course, what can we do for you?" replied Fred.
"I live at the top of the hill, where it's very windy! I am a farmer and have a bull and some cows in my garden. Every night the wind blows so hard that my palisade flies away, and every day I have to rebuild it. I am too old and sick to keep doing this. I'm looking for someone to build it for me" the man replied back.
"Don't worry, you have found just the men you need. We will rebuild the fence for you!" Fred said, eager to help.
"You are very kind, but I don't have any money to pay you!" the old man replied.
"Don't worry. We only need a place to stay for the night, and something to eat!" Fred said.
"I will cook you my famous pottage! It is a thick liquid made with vegetables and meat, used as food by farmers!" replied the old man, happy that Fred and I could help him out.
We worked all day, and we built a sturdy palisade around the old man's garden. We did such a good job that our fence was still there the next day. But the farmer's pottage was not a very filling dish, and we woke up the next morning extremely hungry! Fortunately for us, in the next village a kind woman brought us some roasted pork on a trencher, which is a flat wooden dish.
A few days later we started climbing up a mountain that we had to pass to reach the king's palace. After a few turns we had to abandon the old wain, because the donkey was unable to heave it up the steep path. At long last we reached a castle where a very

old knight lived. His servants said that the man was very sick, and that the doctor had declared that he was dying. The servants saw how very tired and hungry we were, and offered us a place to stay, and some of the pottage they were eating.

When the knight found out we were staying in his castle, he invited us to visit him in his chamber.

"My last wish is to see my son before I die!" he said: "But he lives at the top of another steep mountain. My servants can't reach his castle. Will you go there and ask him to come and visit me?"

Fred agreed to fetch the knight's son, and immediately set off on his own. We waited for his return for many days. I made myself useful to the old knight by taking him trenchers of meat, but I soon started to worry whether I would ever see Fred again. Luckily, after a week, he returned, bringing the knight's son with him. Fred was given a beautiful destrier as a reward for his help, and we continued our journey.

We soon reached a very dangerous place called the dragon's valley.

"There's a motte! We must have reached the king's castle!" I said, but I was wrong. Another knight lived in that castle.

"There's a vicious dragon in these parts that is terrifying my farmers!" said the knight: "I will pay you a handsome reward and give you anything you want, if you get rid of him!"

I was too scared to fight the dragon, but Fred accepted the task. The knight gave him an old hauberk and a sword, and told him: "You can have your reward when you return with the dragon's head!".

Fred set off into the fields, and soon found the dragon, and engaged him in battle. All we saw of the battle, from our safe distance, was a lot of flames, but after a few days my friend did indeed return with the dragon's head! The knight kept his promise and rewarded Fred with a bag of gold.

At last we reached the motte on which the king's castle was built. By now, Fred really looked like a knight, with his hauberk, sword and destrier.

The king admitted us to his presence, and accepted Fred's request to become his knight. The ceremony took place that night, and during it the king said to Fred:

"I have received many letters about your bravery and kindness. You are one of the bravest knights I have ever met!" and he invited us to dine at his table, on which there was a huge trencher piled with meat.

"And that" Jen's father finished *"is how your father, a mere furrier, has a knight for a friend!*" "

<u>India Story:</u>

"**How Uncle Jack got his house refurbished**

One cloudy Sunday afternoon Jim and Tom were very bored.

"*What shall we do?*" Jim asked Tom.

Tom replied: "*Let's go to the park.*"

"*It's too cloudy*" Jim replied. "*It looks like it's going to rain, and if it does, we will get wet. We will have no fun at all!*"

"*Well then, let's go and see Uncle Jack at his house. He's a pirate, and he always has exciting stories to tell us!*" suggested Tom.

"*I do like Uncle Jack's stories, but he always wants to sit outside, even in the rain, and I don't feel like getting wet today.*"

"*Oh, but he has had some work done on his house. He has a new verandah, which is a covered gallery along the front of a house. So we can sit outside, without getting wet!*"

"*I didn't know that. How did he get such a thing?*" asked Jim, curiously.

"*Didn't you hear? He tricked someone into building it for him!*" replied Tom, and while they were walking to Uncle Jack's house, Tom told Jim the story:

"*A few months ago, when Uncle Jack was in India, he was swimming in a pond, and a gavial, a type of crocodile with an elongated muzzle, that eats fish, tried to bite his feet off. He must have mistaken Uncle Jack's feet for fish! Uncle Jack was lucky that day. A kind shopkeeper, who happened to be passing by, helped him to escape from the animal. To help Uncle Jack recover from his frightening experience with the gavial, the shopkeeper invited him to his house.*

When they were there the shopkeeper offered him some toddy, an alcoholic drink made with the sap of palm trees.

Uncle Jack agreed to join him for a drink and the two became good friends. Then the shopkeeper told him about his troubles: he was having problems selling his goods. Nobody wanted to buy them anymore. The viceroy, who is the governor of the country of India chosen by the king, had said that his fruit and vegetables were rotten, and now everybody was refusing to buy them from him.

"But why would he say such a thing, my dear friend?" asked the pirate.

"It's just because one day he overheard me laughing about the way he dresses! He was very upset by my remark, and from that moment he decided to get his own back!" answered the shopkeeper in despair.

Uncle Jack knew Sir Cattermole, the viceroy, and agreed that he was a rather strange man. Sir Cattermole could be found in the strangest places and wore very peculiar clothes. Moreover he was so lazy, he went around the city in a palanquin, which is a large box moved by four people, each one holding one end of two poles, used by rich people to be carried around. Uncle Jack was cross that Sir Cattermole had told everyone that the shopkeeper's goods were rotten, and decided to play a trick on him, to teach him a lesson.

The next day Uncle Jack went to Sir Cattermole's palace, and Sir Cattermole invited him to have some tea. While they were sitting around the teapoy, a small three legged table, Uncle Jack offered his host a challenge. He told Sir Cattermole that he could have his entire cellar of toddy, a very fine stock indeed, if he could prove that the people of the city loved him.

"That's a very kind offer!" said Sir Cattermole. "But what will you get in return, if I lose the bet? What will happen if I can't prove that the people love me?"

The pirate thought to himself. "I know, I have always wanted a verandah on my house. I would like you to build it!" he replied.

"Well, I don't understand what you are up to, pirate, but I'm quite sure my people love me, so I shall take your bet!" agreed Sir Cattermole.

The next day, Sir Cattermole dressed up to look just like his Indian servant. He put on an old shirt, and a white dhoti, which is a piece of cloth, wrapped around the waist and the legs, worn by men. Then he left the palace by the back door, and met his friend, the pirate, at the marketplace.

"Nice dhoti!" laughed Uncle Jack.

"So what shall I do now?" asked the other man.

"Just go around and ask people what they think of Sir Cattermole. Let's see what they say!" said the pirate, cheerfully. "Let's see if they really love you!"

So Sir Cattermole started asking around. The things people said about him were not at all what he expected.

"Sir Cattermole is a really creepy person!" exclaimed a boy: "He spends a lot of time in the weirdest places. He can often be seen visiting the mausoleum, which is an

impressive construction, housing the body of a famous person after his death. Who would want to spend his time there? I wouldn't!"

"Maybe he wants one for himself." answered Sir Cattermole, still pretending to be his servant.

"A mausoleum! Why spend so much money on something so useless?" asked the boy, baffled.

Another man, a very old one, said to Sir Cattermole: "I think the governor is a really lazy man! He does not walk around the town, like everybody else. He does not even use a horse, like other rich men. He uses a palanquin! I am not as lazy as him, and I'm 70 years old!"

Sir Cattermole thought he might have better luck asking the opinion of a woman. When he found a woman to ask, she looked around, as if she were frightened, and whispered: "If you want to know my opinion, we'll have to go somewhere quieter!" She took Sir Cattermole to a small teashop, and, when they were sitting at a teapoy with two hot drinks, she said: "You see, we have to be careful what we say. That man gets upset very quickly. The last time he heard someone saying something nasty about him, he took revenge on him. He told everyone not to buy his goods! I don't want that to happen to me! Sir Cattermole is a very bad-tempered man!"

That night Sir Cattermole invited Uncle Jack to dinner, and admitted his defeat.

"They don't love me, and they think I'm weird!" said the poor man sadly.

"Well, of course they do! You are weird, my friend!" said the pirate laughing: "Who else would spend so much time in a mausoleum? But it doesn't matter that you are a bit odd. You should accept who you are!"

"I suppose you are right!" answered Sir Cattermole, feeling a little bit better: "Well, you've won the bet. When shall the construction work on your house begin?"

"As soon as we arrive in England!" exclaimed the pirate.

And that's the story of how Uncle Jack got the viceroy of India to build him a verandah!"

Tom and Jim had by this time arrived at Uncle Jack's house.

"*You skipped the most interesting part of the story, Tom!*" said his uncle, having overheard the last part of the tale. "*I made the poor man wear his dhoti while he did the job, and invited everyone I know to see the show. It was fun to see Sir Cattermole doing some work, for once!*"

The three of them laughed, and sat outside, around the teapoy. The two boys drank hot chocolate, and the pirate drank a glass of his famous toddy.

"*What's that?*" exclaimed Jim, suddenly noticing a strange object in the garden.

"*Oh, it's a palanquin! Sir Cattermole decided he didn't want people to make fun of him anymore, so he gave it to me! But don't think of getting inside. I bought myself a little present…. Do you remember the gavial that tried to bite off my leg? It turns out that it makes a very nice pet!*" "

Appendix D: word pronunciations and phonological distractors for the phonological task (following IPA transcription) for Study 1

Set	Word	Word pronunciation	Phonological distractors
Normans	motte	mɒt	mɒtɪ
	furrier	fʌrɪə	fʊrɪə
	hauberk	hɔːbək	hɑːbək
	destrier	dɛstrɪə	dɛstraɪə
	palisade	palɪseɪd	palɪsɑːd
	wain	weɪn	waɪən
	pottage	pɒtɪdʒ	pɒtɑːdʒ
	trencher	trɛntʃə	trɪntʃə
India	mausoleum	mɔːsəliːəm	mɔːsəlʌm
	viceroy	vaɪsrɔɪ	vɪsrɔɪ
	dhoti	dəʊtɪ	dɔɪtɪ
	gavial	geɪvɪəl	gævɪəl
	verandah	vərændə	vɛərændə
	palanquin	paləŋkwɪn	paləŋkwiːn
	toddy	tɒdɪ	tuːdɪ
	teapoy	tiːpɔɪ	tepɔɪ
control	ashram	ɑːʃrəm	ɑːʃrɑːm
	augur	ɔːgə	eɪgə
	wimple	wɪmpl	waɪmpl
	palfrey	pɔːlfri	pʌlfri
	atrium	eɪtrɪəm	ætrɪəm
	gharry	gærɪ	gɛərɪ
	hippocras	hɪpəʊkræs	haɪpəʊkræs
	triclinium	traɪklɪnɪəm	trɪklɪnɪəm

Appendix E: words and orthographic distractors for the orthographic task for Study 1

Set	**Word**	**Orthographic distractors**
	motte	mot
	furrier	fourrier
	hauberk	horberk
Normans	destrier	destriar
	palisade	palisaid
	wain	waine
	pottage	potage
	trencher	trencha
	mausoleum	moseleum
	viceroy	visroy
	dhoti	dottie
India	gavial	gaviel
	verandah	varandah
	palanquin	pelanquin
	toddy	toddie
	teapoy	teapoi
	ashram	ashran
	augur	orga
	wimple	wimpol
control	palfrey	polfrey
	atrium	atriam
	gharry	gurry
	hippocras	hippocrase
	triclinium	tryclinium

Appendix F: Headteacher information sheet and consent form for Study 1

Researcher:
Alessandra Valentini
Email:
a.valentini@pgr.reading.ac.uk

Supervisors:
Carmel-Houston Price
Tel: 0118 378 5378
Email:
c.houston-price@reading.ac.uk

Jessie Ricketts
Tel: 01183782796
Email:
j.ricketts@reading.ac.uk

Head Teacher information sheet

Research Project: Learning new words through listening to and reading stories

Project Team Members: Miss Alessandra Valentini
Dr. Carmel Houston-Price
Dr. Jessie Ricketts

Dear Head Teacher

I am writing to invite your school to take part in a research study on vocabulary development.

What is the study?
The study is being conducted by the University of Reading. The project is organised by Alessandra Valentini, as part of her PhD research, under the supervision of Dr Carmel Houston-Price and Dr Jessie Ricketts, and is funded by a University of Reading Social Science Research Studentship. The study aims to investigate how children learn new words in the classroom. In particular, we are interested in how words are learned through stories, both when they are read to the child, and read by the child.

Why has this school been chosen to take part?
This school was chosen because it is close to the University of Reading.

Does the school have to take part?
It is entirely up to you whether you give permission for the school to participate. You may also withdraw your consent to participation at any time during the project, without any repercussions to you, by contacting Alessandra Valentini by email on a.valentini@pgr.reading.ac.uk.

What will happen if the school takes part?
With your agreement, your participation would involve us asking Year 4 teachers to present three specially designed stories to their class. Each story would be read once per week for three consecutive weeks so the whole process will take a total of nine weeks. The procedure will be discussed fully with teachers at the beginning of the study. Each week, this will take up approximately 30 minutes of class time. The stories will provide children with an opportunity to learn new words. These will be unfamiliar words that are linked to the secondary curriculum.

In addition to the stories, we hope to work individually with children in Year 4 in seven sessions, each lasting approximately 30 minutes to 1 hour. In a preliminary session we will assess language, reading and cognitive abilities using standardized tests that are often used in school settings. In the remaining six sessions, we will investigate the extent to which children have learned the unfamiliar words that were embedded in the stories. For each story, learning will be assessed immediately after the third presentation of the story and also after a two-week delay.

We will be sending a letter home to give parents and pupils the opportunity to opt out of the research, if they do not wish to take part in it. We have chosen to use opt out consent process as we hope to gather information about a wide range of children. Teachers will also be given an information sheet and will be asked to consent to take part.

What are the risks and benefits of taking part?
The information and the data collected in the course of the study will remain confidential and will only be seen by the research team listed at the beginning of this letter. Neither you, the children, nor the school will be identifiable in any published report resulting from the study.

In similar studies, schools have found it interesting to take part. We anticipate that the findings of the study will be useful for planning how to increase vocabulary learning in the classroom. We will use a range of vocabulary that children are likely to meet later in the course of their education. An electronic summary of the findings of the study can be made available to you by contacting Alessandra Valentini.

Researcher:
Alessandra Valentini
Email:
a.valentini@pgr.reading.ac.uk

Supervisors:
Carmel-Houston Price
Tel: 0118 378 5378
Email:
c.houston-price@reading.ac.uk

Jessie Ricketts
Tel: 01183782796
Email:
j.ricketts@reading.ac.uk

What will happen to the data?
Any data collected will be held in strict confidence and no real names will be used in this study or in any subsequent publications. The records of this study will be kept private. No identifiers linking you, the children or the school to the study will be included in any sort of report that might be published. Participants will be assigned a number and will be referred to by that number in all records. Research records will be stored securely in a locked filing cabinet and on a password-protected computer and only the research team will have access to the records. The data will be destroyed securely once the findings of the study are written up, after five years. The results of the study may be presented at national and international conferences, and in written reports and articles.

Who has reviewed the study?
This project has been reviewed following the procedures of the University Research Ethics Committee and has been given a favourable ethical opinion for conduct. I possess an enhanced disclosure certificate from the Disclosure and Barring Service. The University has the appropriate insurances in place. Full details are available on request.

What happens if I change my mind?
You can change your mind at any time without any repercussions. If you change your mind after data collection has ended, we will discard the school's data.

What happens if something goes wrong?
In the unlikely case of concern or complaint, you can contact Dr. Carmel Houston-Price at the University of Reading by email on c.houston-price@reading.ac.uk.

Where can I get more information?
If you would like more information, please contact Alessandra Valentini by email on a.valentini@pgr.reading.ac.uk.

What do I do next?
We do hope that you will agree to participate in the study. If you do, please complete the attached consent form and return it to us as soon as possible.

Thank you for your time.

Yours sincerely,

Alessandra Valentini

Researcher:
Alessandra Valentini
Email:
a.valentini@pgr.reading.ac.uk

Supervisors:
Carmel-Houston Price
Tel: 0118 378 5378
Email:
c.houston-price@reading.ac.uk

Jessie Ricketts
Tel: 01183782796
Email:
j.ricketts@reading.ac.uk

Head Teacher Consent Form

I have read the Information Sheet about the project and received a copy of it.

I understand what the purpose of the project is and what is required of me. All my questions have been answered.

Name of Head Teacher: ______________________________________

Name of school: ____

Please tick as appropriate:

I consent to the involvement of my school in the project as outlined in the Information Sheet ☐

Signed:__________________________

Date: ___________________________

Appendix G: Parent information sheet and consent form for Study 1

University of Reading

Researcher:
Alessandra Valentini
Email:
a.valentini@pgr.reading.ac.uk

Supervisors:
Carmel-Houston Price
Tel: 0118 378 5378
Email:
c.houston-price@reading.ac.uk

Jessie Ricketts
Tel: 01183782796
Email:
j.ricketts@reading.ac.uk

Parent/carer information sheet

Research Project: Learning new words through listening to and reading stories

Project Team Members: Miss Alessandra Valentini
Dr. Carmel Houston-Price
Dr. Jessie Ricketts

Dear Parent/Carer

We would like to invite your child to take part in a research study on vocabulary development that will take place in his/her school.

What is the study?
The study is being conducted by the University of Reading. The project is organised by Alessandra Valentini, as part of her PhD research, under the supervision of Dr Carmel Houston-Price and Dr Jessie Ricketts, and is funded by a University of Reading Social Science Research Studentship. The study aims to investigate how children learn new words in the classroom. In particular, we are interested in how words are learned through stories, both when they are read to the child, and read by the child.

Why has my child been chosen to take part?
Your child has been invited to take part in the project because he or she is currently in Year 4. All students who are in Year 4 have been asked to participate in this project.

Does my child have to take part?
It is entirely up to you whether your child participates. You may also withdraw your consent to participation at any time during the project, without any repercussions to you or your child, by contacting Alessandra Valentini by email on a.valentini@pgr.reading.ac.uk.

What will happen if my child takes part?
If you give your consent, your child will listen to and read some stories in his or her classroom. Your child will also be seen individually by Alessandra Valentini on seven occasions, with each session lasting approximately 30 minutes to 1 hour. In a preliminary session we will assess language, reading and cognitive abilities using standardized tests that are often used in school settings. In the remaining six sessions, we will investigate the extent to which children have learned the unfamiliar words that were embedded in the stories. For each story, learning will be assessed immediately after the third presentation of the story and also after a two-week delay. The Principal Researcher will work with the teachers to ensure that these sessions will not interfere with other important school activities that your child might be involved in.

What are the risks and benefits of taking part?
The information you give will remain confidential and will only be seen by the research team listed at the start of this letter. Neither you, the children or the school will be identifiable in any published report resulting from the study. Taking part will in no way influence the grades your child receives at school. Information about individual children will not be shared with the school.

Participants in similar studies have found it interesting to take part. We hope that the findings of this study will be useful for teachers in planning how to increase vocabulary learning in the classroom. Your child will be exposed to words that he or she is likely to encounter later on during the course of his or her

University of Reading

Researcher:
Alessandra Valentini
Email:
a.valentini@pgr.reading.ac.uk

Supervisors:
Carmel-Houston Price
Tel: 0118 378 5378
Email:
c.houston-price@reading.ac.uk

Jessie Ricketts
Tel: 01183782796
Email:
j.ricketts@reading.ac.uk

education. An electronic summary of the findings of the study can be made available to you by contacting Alessandra Valentini.

What will happen to the data?
Any data collected will be held in strict confidence and no real names will be used in this study or in any subsequent publications. The records of this study will be kept private. No identifiers linking you, your child or the school to the study will be included in any sort of report that might be published. Children will be assigned a number and will be referred to by that number in all records. Research records will be stored securely in a locked filing cabinet and on a password-protected computer and only the research team will have access to the records. The data will be destroyed securely once the findings of the study are written up, after five years. The results of the study may be presented at national and international conferences, and in written reports and articles.

Who has reviewed the study?
This project has been reviewed following the procedures of the University Research Ethics Committee and has been given a favourable ethical opinion for conduct. I possess an enhanced disclosure certificate from the Disclosure and Barring Service. The University has the appropriate insurances in place. Full details are available on request.

What happens if I change my mind?
You can change your mind at any time without any repercussions. During the research, your child can stop completing the activities at any time. If you change your mind after data collection has ended, we will discard your child's data.

What happens if something goes wrong?
In the unlikely case of concern or complaint, you can contact Dr. Carmel Houston-Price at University of Reading by email on c.houston-price@reading.ac.uk

Where can I get more information?
If you would like more information, please contact Alessandra Valentini by email on a.valentini@pgr.reading.ac.uk.

What do I do next?
We do hope that you will agree to your child's participation to the study. <u>If you are happy for your child to take part you do not need to do anything</u>. If, however, you do not wish for your child to take part you need to complete and return the consent form on the next page to the school office (reception) as soon as possible.

Thank you for your time.

Yours sincerely,

Alessandra Valentini

Researcher:
Alessandra Valentini
Email:
a.valentini@pgr.reading.ac.uk

Supervisors:
Carmel-Houston Price
Tel: 0118 378 5378
Email:
c.houston-price@reading.ac.uk

Jessie Ricketts
Tel: 01183782796
Email:
j.ricketts@reading.ac.uk

Research Project: Learning new words through listening to and reading stories

IF YOU ARE HAPPY FOR YOUR CHILD TO TAKE PART THEN YOU DO NOT NEED TO TAKE ANY FURTHER ACTION

IF YOU DO NOT WANT YOUR CHILD TO BE INCLUDED THEN PLEASE FILL IN THE FORM BELOW AND RETURN IT TO THE SCHOOL OFFICE (RECEPTION)

Parent/Carer Consent Form

I **DO NOT** give consent for my child to take part in the research.

Name of child: ____

Name of school: ______________________________________

Signed:____________________________

Date: _____________________________

Appendix H: Children information leaflet for Study 1

What happens next?

Your parents have been sent a letter asking for their permission for you to take part in this project.

Before we do the tasks, we will check that you are happy to help us with our project.

If you have any questions please speak to your class teacher, or you can contact Miss Valentini by email at a.valentini@pgr.reading.ac.uk

This application has been reviewed following the procedures of the University of Reading Research Ethics Committee and has been given a favourable ethical opinion for conduct.

University of Reading

School of Psychology
Earley Gate, Whiteknights Road
RG6 6AL

Institute of Education
London Road Campus
RG1 5EX

Research Project: An invitation for children.

What do I learn from storybooks?

Information about the project

Miss Valentini is doing a project that will help her to understand what children learn when they listen to or read a new story.

Miss Valentini would like you to help her with the project.

Your teacher will present some stories to you, and Miss Valentini will ask you some questions about those stories.

We have already asked your parents and Head Teacher if they are happy for you to help us.

Why have I been invited to take part?

You have been invited to take part because you are in Year 4 at your school.

What will I have to do if I agree to take part?

You will be asked to read or listen to the stories presented in class.

Miss Valentini will ask you to do some enjoyable exercises before and after the stories.

Will anyone know about my answers?

Only the people working on the project will know about your answers.

Will it help me if I take part?

We think you will find it interesting and fun to do the activities and listen to the stories. Also, there will be an opportunity to learn useful things from the stories. We hope that your answers will help your teachers and us to know the best ways to boost language in the classroom.

Do I have to take part?

No, not at all. You can stop at any time if you want to.

Just ask your teacher, Miss Valentini or your parents if you would like to stop.

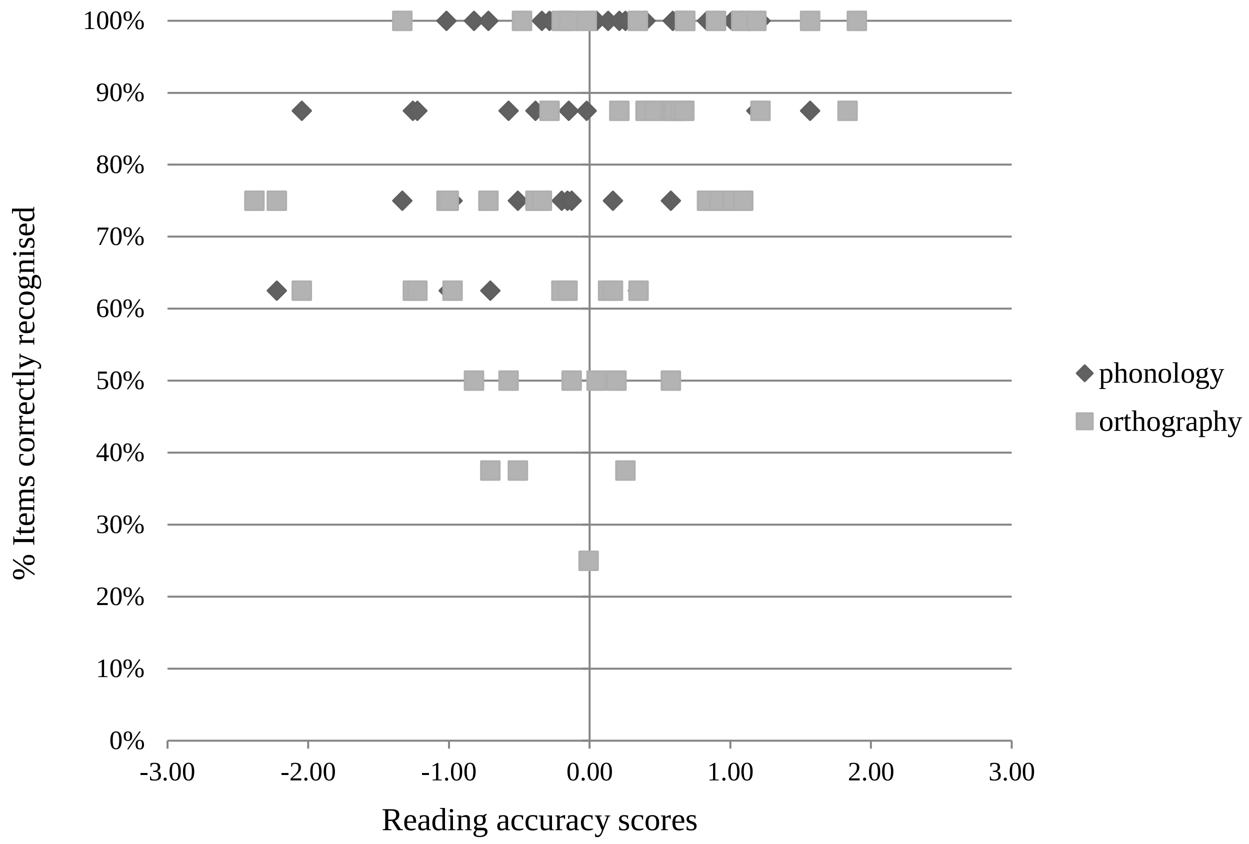
Appendix I: details of Principal components analyses carried out to compute the reading accuracy factor in Study 1

For the first PCA a rotation was necessary to maximise the loadings of the variables onto one factor, and minimise their loading on the remaining factor. A principal component analysis with oblique rotation (direct oblimin) which led to similar results was computed. Given the results and the conclusions obtainable by the two analyses were similar, only one analysis has been reported. The Kaiser-Meyer-Olkin measure was used to verify the sampling adequacy for first PCA, KMO= .82 ("great" according to Field, 2009), and all KMO values for individual items are higher than .72, which is well above the limit of .50 (Field, 2009). Bartlett's test of sphericity ($\chi^2(28)$ = 312.33 , p < .001) indicated that correlations between measures were sufficiently large for PCA.

Given that a value higher than 1 can result in problematic interpretations of the scores, and the correlation between factors computed with the two PCAs were very highly correlated, it was decided to use the factor loadings of the second PCA to compute factor scores, as reported in the text. The correlation between the factor computed in the second PCA and the first component created by the first PCA was r = .97, p < .001. This high correlation confirmed that using the second PCA to compute scores would not result in differences in associations with other measures.

The Kaiser-Meyer-Olkin measure for the second PCA verified the sampling adequacy (KMO= .84, "great" according to Field, 2009), and all KMO values for individual items were higher than .80, which is well above the limit of .50 (Field, 2009). Bartlett's test of sphericity ($\chi^2(6)$ = 195.98 , p < .001) indicated that correlations between measures were sufficiently large for PCA.

Appendix J: distribution of scores in the phonological and orthographic task in the combined condition in Study 1 at post-test 1 depending on reading accuracy.



Appendix K: target and control word pairs properties and comparisons for Study 2

Pair	Type	Word	L	zipf	PRE-TEST		Adults ORTH[1]		Adults PHON[1]	
					C	I	C	I	C	I
animal	target	destrier	8	-	10	80	2	6	10	0
	control	gavial	6	2.14	18	72	0	8	6	4
	χ^2				2.71		-.38		-.50	
	p				.099		.467		.087	
part of a house	target	palisade	8	2.19	34	56	4	4	10	0
	control	atrium	6	2.72	34	56	6	2	9	1
	χ^2				-		.26		-.23	
	p				-		.608		.999	
building	target	motte	5	2.47	19	71	1	7	4	6
	control	catacomb	8	1.90	18	72	4	4	7	3
	χ^2				.03		.40		.30	
	p				.854		.282		.370	
Food or drink	target	pottage	7	2.33	29	61	6	2	9	1
	control	toddy	5	2.84	7	83	7	1	10	0
	χ^2				16.81		.16		-.23	
	p				$< .001^*$		>.999		.999	
Target vs. control	target		7	2.33	23.00		3.25		8.25	
	control		6.5	1.97	17.75		3.00		8.50	
	U		25.00	21.00	32.00		28.50		30.00	
	p		.505	.755	>.999		.721		.878	

Pair	Type	Word	L	zipf	PRE-TEST		Adults ORTH[1]		Adults PHON[1]	
					C	I	C	I	C	I
job	target	furrier	7	1.65	20	70	4	4	6	4
	control	augur	5	2.13	26	64	0	8	6	4
	χ^2				1.05		.58		-	
	p				.305		.067		-	
object	target	trencher	8	1.47	29	61	5	3	10	0
	control	teapoy	6	1.93	41	49	4	4	10	0
	χ^2				3.37		.13		-	
	p				.067		.999		-	
vehicle	target	wain	4	2.79	7	83	3	5	10	0
	control	palanquin	9	1.70	14	76	6	2	10	0
	χ^2				2.64		.38		-	
	p				.104		.314		-	
clothing	target	hauberk	7		15	75	0	8	8	2
	control	dhoti	5		16	74	1	7	9	1
	χ^2				.04		.26		.14	
	p				.844		>.999		>.999	

Note. Type= type of word (control word or target word); L = Length of the word (number of characters forming the word); zipf = frequency derived using the Zipf scale from the SUBTLEX-UK; PRE-TEST = number of children correctly (C) or incorrectly (I) identifying the word at Pre-test in Study 1; Adults ORTH = number of adults who correctly (C) or incorrectly (I) spelled the word to dictation in Study 1; Adults PHON = number of adults who correctly (C) or incorrectly (I) pronounced the word in Study 1; χ^2 = χ^2 statistic; U = Mann-Whitney U; p= p-value associated with the χ^2 statistic or the Mann-Whitney U.
[1] Fisher Exact Probability Test computed instead of χ^2, due to small cell frequencies. * χ^2 or Fisher Exact Probability Test is significant.

Appendix L: definitions and clues in the story used in Study 2

Words	definition	Part 1 definition in text	Part 2 definition in text
Furrier	A hunter who sells animal's furs	experienced in hunting	stinky furrier father, working with furs all day
Motte	hill built by men, on top of which a castle is built.	Live on the … built by my ancestors	… on which the king's castle was built
Palisade	barrier made of wooden stakes	a garden, surrounded by a … … surrounds my garden	Wooden
Pottage	thick liquid food that farmers usually eat	watery	farmer's …
Wain	cart, pulled by an animal	we can use to transport the materials we need	the donkey was unable to heave it up the steep path
Destrier	horse used for fighting	horse we had been given	used to knightly fights
Hauberk	piece of armour that covers the top of the body, made of metal chains	to protect my chest	made of rusty chains
Trencher	a flat dish made of wood	roasted pork on a … …full of meat and pies	wooden trenchers

Appendix M: target word pairs properties and comparisons for words in Study 2

Pair	List	Word	Length	zipf SUBTLEX -UK	Def Lenght	Dist.	PRE-TEST		Adults ORTH[1]		Adults PHON[1]		Post-test Step 1 Comp		Post-test Step 2 Comp		Post-test Step 3 Comp		Post-test phon Combined		Post-test orth Combined	
							C	I	C	I	C	I	C	I	C	I	C	I	C	I	C	I
1	A	destrier	8	-	4	683	10	80	2	6	10	0	37	87	32	92	29	95	58	35	54	39
	B	furrier	7	1.65	9	672	20	70	4	4	6	4	82	42	47	77	42	82	51	42	59	34
	χ^2						4.00		.26		-.50		32.71		4.18		3.34		1.09		.56	
	p						.045*		.610		.087		<.001*		.040*		.067		.296		.454	
2	A	palisade	8	2.19	6	461	34	56	4	4	10	0	67	57	48	76	37	87	34	59	46	47
	B	trencher	8	1.47	4	415	29	61	5	3	10	0	76	48	57	67	44	80	59	34	52	41
	χ^2						.61		.13		-		1.34		1.34		1.85		13.44		.02	
	p						.434		.999		-		.247		.247		.173		<.001*		.887	
3	A	motte	5	2.47	13	549	19	71	1	7	4	6	55	69	32	92	29	95	57	36	46	47
	B	wain	4	2.79	7	528	7	83	3	5	10	0	69	55	62	62	56	68	57	36	43	50
	χ^2						6.47		.29		.65		3.16		15.42		13.05		0		.19	
	p						.011*		.560		.011*		.075		.001*		<0.001*		1		.663	
4	B	hauberk	7	-	18	45	15	75	0	8	8	2	85	39	73	51	64	60	55	38	42	51
	A	pottage	7	2.33	13	61	29	61	6	2	9	1	93	31	71	53	60	64	59	34	49	44
	χ^2						5.9		.77		.14		1.27		.07		.26		.36		1.05	
	p						.015*		.006*		.999		.260		.791		.610		.548		.306	
List A vs. List B	A		7	2.33	9.00	548.00	23.00		3.25		8.25		63.00		45.75		38.75		52.00		48.75	
	B		6.5	1.97	9.50	331.50	17.75		3.00		8.50		78.00		59.75		51.50		55.50		49.00	
	U		6.00	3.00	7.50	6.00	5.50		7.50		8.00		4.00		4.00		3.00		7.00		7.00	
	p		.686	.700	.886	.686	.486		.886		>.999		.343		.343		.200		.886		.886	

Note. Zipf SUBTLEX-UK = frequency derived using the Zipf scale from the SUBTLEX-UK; Def Length =number of words in the definition; Dist. =distance in words between the first mention of the word in the story and the first repetition; PRE-TEST = number of children correctly (C) or incorrectly (I) identifying the word at Pre-test in Study 1; Adults ORTH = number of adults who correctly (C) or incorrectly (I) spelled the word to dictation in Study 1; Adults PHON = number of adults who correctly (C) or incorrectly (I) pronounced the word in Study 1; Post-test Step 1 Comp =number of children who correctly (C) or incorrectly (I) identified the category for the word in Step 1 of the comprehension task in Study 1; Post-test Step 2 Comp =number of children who correctly (C) or incorrectly (I) identified the sub-category for the word in Step 2 of the comprehension task in Study 1; Post-test Step 3 Comp =number of children who correctly (C) or incorrectly (I) identified the definition for the word in Step 3 of the comprehension task in Study 1; Post-test phon Combined =number of children who correctly (C) or incorrectly (I) identified the phonological form of the word at post-test for the Combined Condition in Study 1; Post-test orth Combined =number of children who correctly (C) or incorrectly (I) identified the orthographic form of the word at post-test for the Combined Condition in Study 1; χ^2 = χ^2 statistic; U = Mann-Whitney U; p= p-value associated with the χ^2 statistic or the Mann-Whitney U.

[1] Fisher Exact Probability Test computed instead of χ^2, due to small cell frequencies. * χ^2 or Fisher Exact Probability Test is significant.

Appendix N: template story for Study 2

How Fred became a knight

Jen lived in a cottage with her father whose job was to be a furrier, a hunter who sells animal's furs. They were neither very rich nor very poor, but they had an important friend called Fred. Fred was the knight who lived on the motte, which is a hill built by men, on top of which a castle is built.

Nobody else in the village had such important friends, so one day, before going to bed, she asked her father: "*Dad, how did you and Fred become friends? It's unusual for a knight to be friends with people like us*"

"*It's very simple*" he replied: "*Fred has not always been a knight!*"

And he told her the story of how his friend Fred became a knight:

"*When he was young, Fred lived in an old house. He was a farmer. All he owned was an old cottage and a garden, surrounded by a palisade, a barrier made of wooden stakes. In the garden he had an old donkey, and a small wain, a cart, pulled by an animal.*

Fred had always dreamt of becoming a knight. He wanted to fight dragons and ride a fine horse, but he couldn't imagine that one day he would live his dream.

However, one day something changed. The king had lost many soldiers during the war, and he was looking for young men to fight with him, and to become his knights.

Fred was extremely excited by the news: "I will set off today, and I will reach the king's castle in less than a month!" he said.

"You are mad" I replied: "It is a dangerous journey! You will have to pass through the mountains, and then through the valley of the dragons!"

But Fred had always been brave and stubborn, and I couldn't change his mind. I decided to go with him, because I was worried that something bad would happen to him if he went on his own.

The next morning we hitched the donkey to the vehicle, and we set off in the direction of the castle.

In every village we passed, we heard people discussing the big news that the king was looking for new knights. We also saw many men on their way to the castle.

We saw a man riding a destrier, a horse used for fighting. "I wish I had a strong animal like that!" said Fred: "Every knight needs one, but they are very expensive and rare."
A few hours later we saw a knight who was wearing a hauberk, a piece of armour that covers the top of the body, made of metal chains. "That is a very expensive thing to wear" said Fred. "I don't have enough money to buy one. How can I fight like a knight without a hauberk to protect my chest?"
At that point Fred became very sad. Compared to the other men's possessions, his own seemed unsuitable for a knight. I felt sorry for him, because I knew he would be the best of knights, as he was brave and kind.
We continued on our journey, but we soon finished the food we had brought with us, so we stopped in a village to look for a temporary job to earn some money. While we were asking around for work we met an old man. He seemed very upset.
"Can you help me?" he asked.
"Of course, what can we do for you?" replied Fred.
"I live on a motte built by my ancestors, where it's very windy! I am a farmer and have a bull and some cows in my garden. Every night the wind blows so hard that the palisade that surrounds my garden flies away, and every day I have to rebuild it. I am too old and sick to keep doing this. I'm looking for someone to fix it for me" the man replied back.
"Don't worry, you have found just the men you need. We will help you!" Fred said. "We are only a farmer and a furrier, and we are more experienced in hunting than building works, but we have a wain we can use to transport the materials we need up the hill!"
"You are very kind, but I don't have any money to pay you" the old man replied.
"Don't worry. We only need a place to stay for the night, and something to eat" Fred said.
"I can offer you my famous pottage. It is a thick liquid food that farmers usually eat!" replied the old man, happy that Fred and I could help him out.
We worked all day, and we fixed the wooden palisade for the old man. We did a very good job, but the farmer's pottage was a poor reward for our work, and we woke up the next morning extremely hungry! Fortunately for us, in the next village a kind woman also brought us some roasted pork on a trencher, which is a flat dish made of wood.

A few days later we started climbing up a mountain that we had to pass to reach the king's palace. After a few turns we had to abandon the old wain, because the donkey was unable to heave it up the steep path. At long last we reached a castle where a very old knight lived. His servants said that the man was very sick, and that the doctor had declared that he was dying. The servants saw how very tired we were, and offered us a place to stay, and some watery pottage.
When the knight found out we were staying in his castle, he invited us to visit him in his chamber.
"My last wish is to see my son before I die!" he said: "But he lives at the top of another steep mountain. My servants can't reach his castle. Will you go there and ask him to come and visit me?"
Fred agreed to fetch the knight's son, and immediately set off on his own. We waited for his return for many days. At that point I started to worry about whether I would ever see Fred again. Luckily, after a week, he returned, bringing the knight's son with him. Fred was given a beautiful destrier, used to knightly fights, as a reward for his help, and we continued our journey.
With the help of the horse we had been given, we moved much faster, and we soon reached a very dangerous place called the dragon's valley.
"I can see a big building. We must have reached the king's castle!" I said, but I was wrong. Another knight lived in that castle.
"There's a vicious dragon in these parts that is terrifying my farmers!" said the knight: "I will pay you a handsome reward and give you anything you want, if you get rid of him!"
I was too scared to fight the dragon, but Fred accepted the task. The knight gave him an old hauberk made of rusty chains, and a sword, and told him: "You can have your reward when you return with the dragon's head".
Fred set off into the fields, and soon found the dragon, and engaged him in battle. All we saw of the battle, from our safe distance, was a lot of flames, but after a few days my friend did indeed return with the dragon's head! The knight kept his promise and rewarded Fred with a bag of gold. He also offered us a rich meal, with trenchers full of meat and pies, before we set off again.
At last we reached the motte on which the king's castle was built. By now, Fred really looked like a knight, with his armour, sword and destrier.

The king welcomed us, and accepted Fred's request to become his knight. The ceremony took place that night, and during it the king said to Fred:
"I have received many letters about your bravery and kindness. You are one of the bravest knights I have ever met!" and he invited us to dine at his table, where there was a wondrous choice from the wooden trenchers.

"And that" Jen's father finished *"is how your stinky furrier father, working with furs all day, has a knight for a friend!*" "

Appendix O: story comprehension questions for Study 2

Questions	Right answer	Other alternatives
1 – What did Fred become at the end of the story?	A knight	A farmer A hunter A king
2 – Where di Fred need to go to become a knight?	The king's castle	The king's tent The king's ship The king's village
3 – Fred wanted to reach the king's castle. How long did he think the journey would last?	A month	A day A week One hour
4 – The second man that Fred helped was a dying knigh. Where did the knight live?	A mountain	A cave by the sea The countryside A city
5 – After the fight with the dragon, Fred came back with something. What was it?	The dragon's head	The dragon's tail The dragon's tooth The dragon's claw

Appendix P: control sample performance in the comprehension questions in Study 2

Question Number	Children's performance		Fisher Exact p-value
	C	I	
1	5	7	.668
2	11	1	.003*
3	3	9	> .999
4	1	11	.590
5	4	8	> .999

Note. Children's performance = number of children who choose the correct (C) or incorrect (I) answer for each specific question.
* Fisher Exact Probability Test is significant.

Appendix Q: information sheets and consent forms for Study 2

Faculty of Life Sciences
School of Psychology and Clinical Language Sciences
Earley Gate
Whiteknights Road
Reading, RG6 6AL

INFORMATION SHEET

Project title: Learning new words through listening and reading stories

Project Supervisors: Dr. Rachel Pye r.e.pye@reading.ac.uk Tel.: +44 (0)118 378 8526
Dr. Carmel Houston-Price c.houston-price@reading.ac.uk
Dr. Jessie Ricketts Jessie.Ricketts@rhul.ac.uk

Investigator: Alessandra Valentini a.valentini@pgr.reading.ac.uk

Dear Head Teacher

I am writing to invite your school to take part in a research study on vocabulary development.

What is the study?
The study is being conducted by the University of Reading. The project is organised by Alessandra Valentini, as part of her PhD research, under the supervision of Dr Rachel Pye, Dr. Carmel Hoston-Price and Dr. Jessie Ricketts and is funded by a University of Reading Social Science Research Studentship. The study aims to investigate how children learn new words. In particular, we are interested in how words are learned through stories, both when they are read to the child, and read by the child.

Does the school have to take part?
It is entirely up to you whether you give permission for the school to participate. You may also withdraw your consent to participation at any time during the project, without any repercussions to you, by contacting Alessandra Valentini by email on a.valentini@pgr.reading.ac.uk.

What will happen if the school takes part?
With your agreement, and the agreement of children's parents, your participation would involve us asking Year 4 children to participate in 3 sessions of 1 hour, for a total participation time of 3 hours per child. In a preliminary one to one session we will assess language, reading and cognitive abilities using standardized tests that are often used in school settings. In the remaining two sessions, children will be presented with a story that includes some rare words, and their knowledge of these words will be investigated at the end of the second session. The stories will provide children with an opportunity to learn the new words. These will be unfamiliar words that are linked to the secondary curriculum.
For the second and third session children will be assigned to a group, formed of up to 6 children. Each group will be seen on the same day for two consecutive weeks.
Once all the children participating in the study have completed the first preliminary session, the process of story presentation will take a total of two weeks.

We will be sending a letter home to give parents and pupils the opportunity to opt in to the research, if they wish to take part in it. Teachers will also be given an information sheet and the procedure will be discussed fully with them at the beginning of the study, to minimise disruption of their planned class activities.

Faculty of Life Sciences
School of Psychology and Clinical Language Sciences
Earley Gate
Whiteknights Road
Reading, RG6 6AL

What are the risks and benefits of taking part?

The information and the data collected in the course of the study will remain confidential and will only be seen by the research team listed at the beginning of this letter. Neither you, the children, nor the school will be identifiable in any published report resulting from the study.

In similar studies, schools have found it interesting to take part. We anticipate that the findings of the study will be useful for planning how to increase vocabulary learning in the classroom. We will use a range of vocabulary that children are likely to meet later in the course of their education. An electronic summary of the findings of the study can be made available to you by contacting Alessandra Valentini.

What will happen to the data?

Any data collected will be held in strict confidence and no real names will be used in this study or in any subsequent publications. The records of this study will be kept private. No identifiers linking you, the children or the school to the study will be included in any sort of report that might be published. Participants will be assigned a number and will be referred to by that number in all records. Research records will be stored securely in a locked filing cabinet and on a password-protected computer and only the research team will have access to the records. The data will be destroyed securely once the findings of the study are written up, after five years. The results of the study may be presented at national and international conferences, and in written reports and articles.

Who has reviewed the study?

This project has been reviewed following the procedures of the University Research Ethics Committee and has been given a favourable ethical opinion for conduct. I possess an enhanced disclosure certificate from the Disclosure and Barring Service. The University has the appropriate insurances in place. Full details are available on request.

What happens if I change my mind?

You can change your mind at any time without any repercussions. If you change your mind after data collection has ended, we will discard the school's data.

What happens if something goes wrong?

In the unlikely case of concern or complaint, you can contact Dr. Rachel Pye at the University of Reading by email on r.e.pye@reading.ac.uk.

Where can I get more information?

If you would like more information, please contact Alessandra Valentini by email on a.valentini@pgr.reading.ac.uk.

What do I do next?

We do hope that you will agree to participate in the study. If you do, please complete the attached consent form and return it to us as soon as possible.

Thank you for your time.

Yours sincerely,
Alessandra Valentini

University of Reading

Faculty of Life Sciences
School of Psychology and Clinical Language Sciences
Earley Gate
Whiteknights Road
Reading, RG6 6AL

CONSENT FORM

Project title: Learning new words through listening and reading stories

Project Supervisors: Dr. Rachel Pye r.e.pye@reading.ac.uk Tel.: +44 (0)118 378 8526
Dr. Carmel Houston-Price c.houston-price@reading.ac.uk
Dr. Jessie Ricketts Jessie.Ricketts@rhul.ac.uk

Investigator: Alessandra Valentini a.valentini@pgr.reading.ac.uk

I hereby declare that I have been clearly informed about the nature of the research and the methods used.

I have received and read a copy of the Information Sheet and have been given opportunities to ask questions about the study and these have been answered to my satisfaction.

I understand that all personal information will be anonymized and remain confidential to the investigators and arrangements for the storage and eventual disposal of any identifiable material have been made clear to me.

I hereby grant voluntary consent for participation in this research by my school. I retain the right to revoke this consent without having to provide any reasons for my decision.

Name of Head Teacher: ____

Name of school: ____

Signed:___________________________

Date: ____________________________

Faculty of Life Sciences
School of Psychology and Clinical Language Sciences
Earley Gate
Whiteknights Road
Reading, RG6 6AL

INFORMATION SHEET

Project title: Learning new words through listening and reading stories

Project Supervisors: Dr. Rachel Pye r.e.pye@reading.ac.uk Tel.: +44 (0)118 378 8526
Dr. Carmel Houston-Price c.houston-price@reading.ac.uk
Dr. Jessie Ricketts Jessie.Ricketts@rhul.ac.uk

Investigator: Alessandra Valentini a.valentini@pgr.reading.ac.uk

Dear Parent/Carer

We are a research group based at the University of Reading, and we are going to conduct a research study on vocabulary development in your child's school. We would like to ask your permission to include your child in this study. The study aims to investigate how children learn new words. In particular, we are interested in how words are learned through stories, both when they are read to the child, and read by the child.

Background to the study

School aged children can learn new words in many ways: they can learn new words while listening, but they can also learn new words while reading books. We would like to compare several ways of presenting stories to children, with new words embedded in them, to explore how many words the children learn from these stories. We will also explore if the children learn different features of the words, depending on how these words were presented. We anticipate that the findings of the study will be useful for planning how to increase vocabulary learning in the classroom. We will use a range of vocabulary that children are likely to meet later in the course of their education.

Who is running the study?

The study is being conducted by the University of Reading. The project is organised by Alessandra Valentini, as part of her PhD research, under the supervision of Dr Rachel Pye, Dr. Carmel Hoston-Price and Dr. Jessie Ricketts. All researchers involved have been checked by the Criminal Records Bureau (CRB) and approved to work with children. This application has been reviewed by the University Research Ethics Committee and has been given a favourable ethical opinion for conduct, which means that an independent group did not raise any objections to the study on ethical grounds and have permitted the study to proceed. The study is funded entirely by a University of Reading Social Science Research Studentship and no commercial or industrial parties are involved.

What does the study involve?

If you give your consent, your child will participate in three sessions led by Alessandra Valentini, each session lasting 1 hour. In the first session your child's language and reading abilities will be assessed, using standardized tests that are often used in school settings. In two remaining sessions your child will be asked to listen to or read a story. The story will contain rare words. After the story presentation, your child will participate in a few tasks, designed to ascertain how much he or she has learnt from the story, and how many new words he or she can remember. The Principal Researcher will work with the teachers to ensure that these sessions will not interfere with other important school activities that your child might be involved in. Your child will therefore participate in 3 sessions of 1 hour, for a total time of 3 hours.

Faculty of Life Sciences
School of Psychology and Clinical Language Sciences
Earley Gate
Whiteknights Road
Reading, RG6 6AL

What happens to the data?
All information collected will remain fully confidential and assigned an anonymous number. All data are kept safely locked at the University of Reading and only specified trained researchers will have access. The data will be used only for research purposes, and in accordance with the Data Protection Act of 1998, they will be destroyed 5 years after completion of the study.

Taking part will in no way influence the grades your child receives at school.
No information regarding your child, collected during this research, will be shared with the school, unless you provide explicit permission.

Why has my child been chosen to take part?
Your child has been invited to take part in the project because he or she is currently in Year 4. All students who are in Year 4 have been asked to participate in this project.

Does my child have to take part?
It is entirely up to you whether your child participates. You may also withdraw your consent to participation at any time during the project, without any repercussions to you or your child, by contacting Alessandra Valentini by email on a.valentini@pgr.reading.ac.uk.

Where can I get more information?
If you would like more information, please contact Alessandra Valentini by email on a.valentini@pgr.reading.ac.uk.

If you are happy for your child to take part in this study please complete the consent form enclosed and return it to the school office. Your permission would be very much appreciated.

Thank you for your time.

Yours sincerely,

Alessandra Valentini

University of Reading
Faculty of Life Sciences
School of Psychology and Clinical Language Sciences
Earley Gate
Whiteknights Road
Reading, RG6 6AL

CONSENT FORM

Project title: Learning new words through listening and reading stories

Project Supervisors: Dr. Rachel Pye r.e.pye@reading.ac.uk Tel.: +44 (0)118 378 8526
Dr. Carmel Houston-Price c.houston-price@reading.ac.uk
Dr. Jessie Ricketts Jessie.Ricketts@rhul.ac.uk

Investigator: Alessandra Valentini a.valentini@pgr.reading.ac.uk

I understand that my child's participation in this study is voluntary and that I may withdraw him/her at any time without giving any reason.

I have read the accompanying information sheet about this study and understand what will be required of my child.

I have received a copy of the information sheet and the consent form.

I understand that this application has been reviewed by the University Research Ethics Committee and has been given a favourable ethical opinion for conduct.

I have been given the opportunity to ask any questions that I may have about the study and these have been answered to my satisfaction.

I am / am not* **happy for you to share my child's results with my child's school.**

**Please delete as appropriate*

Parent/Guardian's signature	--
Name (in capitals)	--
Date	--
Child's Name	--
Child's Date of Birth	--

What happens next?

Your parents have been sent a letter asking for their permission for you to take part in this project.

Before we do the tasks, we will check that you are happy to help us with our project.

If you have any questions please speak to your class teacher, or you can contact Miss Valentini by email, at a.valentini@pgr.reading.ac.uk

This application has been reviewed following the procedures of the University of Reading Research Ethics Committee and has been given a favourable ethical opinion for conduct.

University of Reading

School of Psychology
Earley Gate, Whiteknights Road
RG6 6AL

Research Project: An invitation for children.

What do I learn from storybooks?

Information about the project

Miss Valentini is doing a project that will help her to understand what children learn when they listen to or read a new story.

Miss Valentini would like you to help her with the project.

She will present to you some stories, and she will ask you some questions about those stories.

We have already asked your parents and Head Teacher if they are happy for you to help us.

Why have I been invited to take part?

You have been invited to take part because you are in Year 4 at your school.

What will I have to do if I agree to take part?

You will be asked to read or listen to the stories presented.

Miss Valentini will ask you to do some activities before and after the stories (for example answering questions about a story, reading words, choosing pictures, and short computer games).

Will anyone know about my answers?

Only the people working on the project will know about your answers.

Will it help me if I take part?

We think you will find it interesting and fun to do the activities and listen to the stories. Also, there will be an opportunity to learn useful things from the stories. We hope that your answers will help your teachers and us to know the best ways to boost your English language learning.

Do I have to take part?

No, not at all. You can stop at any time if you want to.

Just ask your teacher, Miss Valentini or your parents if you would like to stop.

Faculty of Life Sciences
School of Psychology and Clinical Language Sciences
Earley Gate
Whiteknights Road
Reading, RG6 6AL

CONSENT FORM FOR CHILDREN

Project title: Learning new words through listening and reading stories

Project Supervisors: Dr. Rachel Pye r.e.pye@reading.ac.uk Tel.: +44 (0)118 378 8526
Dr. Carmel Houston-Price c.houston-price@reading.ac.uk
Dr. Jessie Ricketts Jessie.Ricketts@rhul.ac.uk

Investigator: Alessandra Valentini a.valentini@pgr.reading.ac.uk

Please tick as appropriate:

1. I have read the Information Leaflet. ☐ YES ☐ NO

2. I know what I will have to do. ☐ YES ☐ NO

3. All my questions have been answered. ☐ YES ☐ NO

4. Remember, you can stop at any time if you want to.

NAME (CHILD) ________________________

Researcher ____________________

Signature ____________________ DATE _____________________

Appendix R: interaction between definition and background measures in the semantic task in Study 2

Factor	Estimate	St. Error	z values	
			z value	*p*
Category Recognition				
definition*reading accuracy	.00	.20	-.01	.999
definition*YARC comp	-.16	.20	-.81	.419
definition*BPVS	-.20	.20	-1.00	.318
definition*USP CELF	-.19	.19	-1.03	.305
definition*CPM	.09	.19	.51	.613
definition*TMT	.09	.20	.45	.650
Sub-category Recognition				
definition*reading accuracy	-.05	.19	-.24	.808
definition*YARC comp	-.07	.20	-.37	.708
definition*BPVS	.21	.20	1.08	.279
definition*USP CELF	.29	.20	1.51	.131
definition*CPM	.12	.19	.63	.531
definition*TMT	.16	.20	.76	.446
Definition Recognition				
definition*reading accuracy	-.14	.19	-.76	.450
definition*YARC comp	-.01	.19	-.07	.942
definition*BPVS	.11	.20	.56	.573
definition*USP CELF	-.07	.19	-.40	.686
definition*CPM	.30	.18	1.65	.099
definition*TMT	.04	.20	.22	.825

Appendix S: information sheets and consent forms for Study 3

Project title: Learning new words through listening and reading stories

Project Supervisors: Dr. Julie Kirkby jkirkby@bournemouth.ac.uk
Dr. Rachel Pye r.e.pye@reading.ac.uk Tel.: +44 (0)118 378 8526

Investigators: Alessandra Valentini a.valentini@pgr.reading.ac.uk
Kirsten Tracy i7733434@bournemouth.ac.uk
Sophie Clent i7245552@bournemouth.ac.uk

Dear parent/guardian,

We are a research group based at Bournemouth University and the University of Reading, and we are conducting a research study on vocabulary development in your child's school. We would like to ask your permission to include your child in this study. The study aims to investigate how children learn new words while listening to and reading stories.

During the course of the study your child's eye-movement will be recorded, while he or she reads the story, in order to investigate where children allocate their attention during reading.

Who is running the study?
The study is being conducted by Bournemouth University and the University of Reading. The project will be conducted by Alessandra Valentini, Kirsten Tracy and Sophie Clent, under the supervision of Dr. Rachel Pye (University of Reading) and Dr. Julie Kirkby (Bournemouth University).
All investigators on this project have had criminal records checks and have been approved to work with children. This application has been reviewed by the University of Reading Research Ethics Committee and has been given a favourable ethical opinion for conduct.

What will my child be asked to do?
During the study your child will participate in two sessions of approximately 45 minutes, for a total time of approximately 1 hour and 30 minutes.
In one session your child's language and reading abilities will be assessed, using standardized tests. In the other session your child will be asked to read and listen to two stories presented through a computer, while his/her eye-movements are being recorded from a camera placed near the computer screen. After each story presentation, your child will participate in a few tasks, designed to ascertain how much he or she has learnt from the story. The Principal Researcher will work with the teachers to ensure that these sessions will not interfere with any important school activities that your child might be involved in.

What are the possible risks or discomforts to my child?
The study does not have any physical or emotional risk to your child. You and your child can withdraw your consent to participate at any point during the study or withdraw consent to the use of the data after the study has ended. Throughout the tasks your child will be able to take breaks whenever they feel tired or want to stop.

What happens to the data?
All information collected will remain fully confidential and assigned an anonymous number. All data are kept safely locked and only specified trained researchers will have access. The data will be used only for research purposes, and in accordance with the Data Protection Act of 1998, they will be destroyed 5 years after completion of the study.

Taking part will in no way influence the grades your child receives at school, and no information regarding your child, collected during this research, will be shared with the school.
If you would like a copy of the materials being used in the study you can contact jkirkby@bournemouth.ac.uk.

Benefits of the study
This research is designed to discover optimal ways to help children learn vocabulary. Research has not yet clarified what is the best method to present stories to children, to enhance their learning, and what strategies children use to learn new words from stories. We therefore hope to clarify what is the most effective method to present stories to children, to enhance their vocabulary learning, and we anticipate that the findings of the study will be useful for planning how to increase vocabulary learning in the classroom.

If you are happy for your child to take part in this study please complete the consent form enclosed and return it to the school office. Your permission would be very much appreciated.

Thank you for your time.

Yours sincerely,

Alessandra Valentini
Kirsten Tracy
Sophie Clent

CONSENT FORM

Project title: Learning new words through listening and reading stories

Project Supervisor: Dr. Julie Kirkby: jkirkby@bournemouth.ac.uk Tel.:
Dr. Rachel Pye r.e.pye@reading.ac.uk Tel.: +44 (0)118 378 8526

Investigator: Alessandra Valentini a.valentini@pgr.reading.ac.uk
Kirsten Tracy: i7733434@bournemouth.ac.uk
Sophie Clent: i7245552@bournemouth.ac.uk

I understand that my child's participation in this study is voluntary and that I may withdraw him/her at any time without giving any reason.

I have read the accompanying information sheet about this study and understand what will be required of my child.

I have received a copy of the information sheet and the consent form.

I understand that this application has been reviewed by the University Research Ethics Committee and has been given a favourable ethical opinion for conduct.

I have been given the opportunity to ask any questions that I may have about the study and these have been answered to my satisfaction.

Parent/Guardian's signature --

Name (in capitals) --

Date --

Child's Name --

Child's Date of Birth --

What happens next?

Your parents have been sent a letter asking for their permission for you to take part in this project.

Before we do the tasks, we will check that you are happy to help us with our project.

If you have any questions please speak to your class teacher, or you can contact Miss Valentini by email, at a.valentini@pgr.reading.ac.uk

This application has been reviewed following the procedures of the University of Reading Research Ethics Committee and has given a favourable ethical opinion for conduct.

Bournemouth University

University of Reading

Research Project: An invitation for children.

What do I learn from storybooks?

Information about the project

Miss Valentini, Miss Tracy and Miss Clent are doing a project that will help them to understand what children learn when they listen to or read new stories.

They would like you to help them with the project.

They will show you some stories, and they will ask you some questions about those stories.

We have already asked your parents and Head Teacher if they are happy for you to help us.

Why have I been invited to take part?

You have been invited to take part because you are in Year 4 at your school.

What will I have to do if I agree to take part?

You will be asked to read or listen to the stories presented on a computer screen.
We will record where you look on the screen with a camera, while you read.

You will also do some activities with either Miss Valentini, Miss Tracy or Miss Clent.

Some examples of activities are: answering questions about a story, reading words, choosing pictures, and short computer games.

Will anyone know about my answers?

Only the people working on the project will know about your answers.

Will it help me if I take part?

We think you will find it interesting and fun to do the activities and read the stories. Also, there will be an opportunity to learn useful things from the stories. We hope that your answers will help your teachers and us to know the best ways to boost your English language learning.

Do I have to take part?

No, not at all. You can stop at any time if you want to.

Just tell your teacher, Miss Valentini or your parents if you would like to stop.

APPENDIX D

Bournemouth University

University of Reading

Project title: Learning new words through listening and reading stories

Project Supervisors: Dr. Julie Kirkby jkirkby@bournemouth.ac.uk Tel.:
Dr. Rachel Pye r.e.pye@reading.ac.uk Tel.: +44 (0)118 378 8526

Investigators: Alessandra Valentini a.valentini@pgr.reading.ac.uk
Kirsten Tracy i7733434@bournemouth.ac.uk
Sophie Clent i7245552@bournemouth.ac.uk

Please tick as appropriate:

1. I have read the Information Leaflet. ☐ YES ☐ NO

2. I know what I will have to do. ☐ YES ☐ NO

3. All my questions have been answered. ☐ YES ☐ NO

4. Remember, you can stop at any time if you want to.

NAME (CHILD) ________________________

Researcher ____________________

Signature ____________________ DATE _____________________

Appendix T: items used in Study 3

Item	Correct pronunciation	Incorrect pronunciation	List	Category	Type of Item
thundelp	θʌndɛlp	θu:ndɛlp	a	animal	target
cynthor	sɪnθɚ	saɪnθɚ	b	animal	target
tellitry	telɪtrɪ	telɪtraɪ	a	building	target
concipan	kɑnsɪpan	kɑnkɪpan	b	building	target
kaptim	kɑptɪm	kæptɪm	a	clothing	target
progus	proʊgʌs	proʊdʒʌs	b	clothing	target
piclin	pɪklɪn	pɪslɪn	a	food	target
tifpote	tɪfpoʊt	tɪfpot	b	food	target
chimpister	t͡ʃɪmpɪstə	kɪmpɪstə	a	job	target
winniver	wi:nɪvə	wi:naɪvə	b	job	target
lutacrift	lʌtʌcrɪft	lʌtʌgrɪft	a	object	target
felnadit	fɛlnʌdɪt	fɛlnaɪdɪt	b	object	target
slunbort	slʌnboɚt	slu:nboɚt	a	animal	control non-word
chundelt	t͡ʃʌndɛlt	ʃʌndɛlt	b	animal	control non-word
anecoil	anɛcɔ:ɪl	eɪncɔ:ɪl	a	building	control non-word
heleager	həligə	həlidʒə	b	building	control non-word
tegwop	tɛgwɑp	tɛgwu:p	a	clothing	control non-word
bemwip	bɛmwɪp	bɪ:mwɪp	b	clothing	control non-word
filpin	fɪlpɪn	fɪlpaɪn	a	food	control non-word
netrich	nɛtrɪt͡ʃ	nɛtrɪʃ	b	food	control non-word
wilderdote	wɪldɜdəʊt	wɪldɜpəʊt	a	job	control non-word
marzentrate	mɑzɛntreɪt	mɔzɛntreɪt	b	job	control non-word
banifice	banɪfaɪs	banɪfɪs	a	object	control non-word
broganoft	brogænɑft	brogænu:ft	b	object	control non-word
peacock	pi:kɒk	pi:kəʊk	a	animal	real word
leopard	lepəd	li:opəd	b	animal	real word
cathedral	kəθi:drəl	kəθidrɑl	a	building	real word
parliament	pɑ:ləmənt	pɑ:lɪamənt	b	building	real word
apron	eɪprən	ɑprən	a	clothing	real word
sandal	sændl	sændɑl	b	clothing	real word
cabbage	kæbɪʤ	kæbɑʤ	a	food	real word
mustard	mʌstəd	mu:stəd	b	food	real word
architect	ɑ:kɪtɛkt	ɑ:kaɪtɛkt	a	job	real word
scientist	saɪəntɪst	skaɪəntɪst	b	job	real word
saxophone	sæksəfəʊn	sæksəfəʊnɪ	a	object	real word
umbrella	ʌmbrɛlə	ʌmbrələ	b	object	real word

Appendix U: Stories and questions used in Study 3 (Pirate story – list a; Knight story – list b).

Pirate Story: The pirate's bet

One cloudy Sunday afternoon Jill and Tom were very bored. They talked about going to the park, but they changed their minds. It was going to rain. So they decided to go and see Uncle Jack. He was a pirate, and always had interesting stories to tell.

Did Jill and Tom decide to go to the park? N

Uncle Jack was on his porch, smoking a cigar. Jill and Tom joined him on the porch and asked: "Uncle Jack, do you have a new adventure to tell us?"
He thought for a while: "Well…" he said. "I guess I could tell you about the bet with the governor of India. The poor man is going to remember that for a while."

Was Uncle Jack smoking a cigar? Y

This is the story of Uncle Jack's bet with the governor of India. On his last trip to India Uncle Jack lost his ship. A storm threw the ship onto some rocks, and the ship sank.
A kind chimpister, someone who sells furs, found Uncle Jack on the shore the morning after the storm. He brought him home and nursed him.
Since he had a chimpister's stand at the market, he had to leave early each morning, and each night he came back to look after Uncle Jack.
The man was an expert nurse. Under his watch Uncle Jack soon regained his strength.
He soon started to spend most of his days outside, to get away from
the hairy mess inside the chimpister's house.

Did Uncle Jack's ship sink? Y

After his recovery, Uncle Jack was feeling restless at home. He wanted to explore the town. His friend lent him his black kaptim, a dress worn by men. The pirate's clothes were too easy to spot. He did not want to be recognised as a pirate.

“Wow, a kaptim is worn by everyone!” Uncle Jack realised at the market. He was pleased about his disguise.
While he was shopping he saw Sir Cattermole. He was the governor of India. Uncle Jack noticed that people avoided the man. They seemed afraid of him. People whispered when he was near them.
Uncle Jack asked an Indian with a manly kaptim why they were whispering. The man did not answer.

Is Sir Cattermole the governor of India? Y

Uncle Jack asked his friend: “Why is everyone afraid of Sir Cattermole?”
The two friends were at dinner. Uncle Jack was eating a huge piclin, a potato wrapped in ham.
The friend said: “Sir Cattermole has put a wine merchant in jail. The man had sold him bad wine by accident. Now everybody is afraid to suffer the same fate.”
The piclin was a tasty treat. Uncle Jack finished eating it before replying: “I’m going to do something about it. I don’t like bullies!”
He knew Sir Cattermole. He was sure he could find a way to punish him. The next day, he managed to get an invitation to dinner at his house.
While eating a meat-wrapped piclin at the
governor’s house, he talked Sir Cattermole into a bet.

Did Sir Cattermole put Uncle Jack’s friend in jail? N

The next day, the two met in front of the town’s enormous tellitry, a grave for many people. They decided to ask the people of the town if they liked the governor. Uncle Jack bet that the people did not love Sir Cattermole. If the governor lost, he would do the pirate’s bidding for a day. If he won, he would get Uncle Jack’s treasure.
They left the tellitry to the resting bodies, and went to the market. They were disguised as farmers. Sir Cattermole started asking around. The things people said about him were not what he expected.
“Sir Cattermole is really creepy!” a boy said.“He spends a lot of time at the funerary tellitry outside the town. They say he keeps his enemies there!”

Did Sir Cattermole offer all his treasures if he lost the bet? N

A very old man said to Sir Cattermole: “I think the governor is a very lazy man! He doesn’t sit to eat his meals. When the governor eats, he lies on a very long lutacrift, a sofa used during meals.”
“He may have a medical reason to do so” replied the governor.
“I don’t think so! The man is just so lazy, he needs to use a lutacrift to eat!”
Many people refused to say anything. They were too scared that the governor would find out what they said. Sir Cattermole was unhappy. He was not a bad person at heart.
Uncle Jack laughed at the memory of Sir
Cattermole lying on the lutacrift the previous night.

Did the old man think that Sir Cattermole was lazy? Y

The next day Uncle Jack and Sir Cattermole met at the stables. There were a few black horses, and a grey thundelp, an elephant that pulls carriages, inside.
“I lost the bet!” said Sir Cattermole. He agreed to release the wine merchant. He admitted he hadn’t been fair to him.
Uncle Jack was trying to decide how to punish the governor. He looked at the thundelp’s huge body.
“Now, I’ll ask you to do one more thing” said the pirate, grinning.
Five minutes later all the people of the city saw the governor dressed as a jester. He was going around the town on a cart tied to a thundelp.
The people laughed at the governor. They were not afraid of him any longer.

Did the people see Sir Cattermole dressed as a woman? N

Knight Story: How Fred became a knight

Jen lives in a cottage with her father, a simple farmer. They are neither very rich nor very poor, but they have an important friend called Fred. Fred is the knight who lives in the castle next to Jen’s village.

Does Fred live in a castle? Y

Nobody else in the village has such an important friend, so one day, Jen asks her father, "Dad, how did you and Fred become friends? It's unusual for a knight to be friends with people like us."

"It's very simple" he replies. "Fred has not always been a knight!"

And he tells her the story of how his friend Fred became a knight.

Has Fred always been a knight? N

This is the story of how Fred became a knight. Fred had always dreamt of becoming a knight. He wanted to fight dragons and do good deeds. But he was only a simple winniver, someone who colours leather.

Fred had many friends. "I like what I do. I like being a winniver and mixing different dyes," he used to say to them. "But I'd rather be a knight."

His friends were happy with their simple, normal lives, but they thought that Fred was very brave, and deserved to become a knight, rather than being a leather-smelling winniver in a small town.

Did Fred's friends think he could be a good knight? Y

Fred was going home from work one day, when he heard something. Somebody was passing by. A soldier in a shiny progus, a shirt made of chains, was riding a horse and talking to another rider.

"The king lost many men during the war," the one with a progus covering his chest said. "He is looking for new knights to fight with him!"

Fred was extremely excited by the news.

He went home in a hurry and gathered his clothes.

"I will set off tomorrow, and reach the king's castle in a month!" he said to himself.

He opened an old trunk, and pulled out his father's old rusty progus and some clothes for the journey.

Was the soldier looking for new knights? N

The next morning Fred set off very early. After riding day and night he was very hungry. He had only eaten some tifpote, a soup eaten by farmers, in two days. Fortunately he met a kind old man on the road, who offered him food in exchange for work.

Fred worked all day to repair the old man’s house, but the tifpote he was offered was a farmer’s meal, and didn’t fill his stomach.

Fred continued his journey. He stopped to work and help people in several places. He ran errands for sick people and helped old men and women.

There was little food during the journey. Most nights Fred only had some watery tifpote to eat.

Did Fred help many people during the journey? Y

One day Fred arrived in a small village on the edge of a very big forest. The forest was cursed and had a very ancient concipan, a tower with no windows, built in the middle of it. People were sure that they could hear voices coming from it, and see light inside it during the night, even though they knew that the concipan had only solid walls.

Fred ventured into the forest.

“Am I going to find a wizard?” he wondered, “Or maybe it is all just a story.”

In any case the forest was the quickest way to the king’s castle. He was not going to delay his journey. He reached the tall concipan, and entered it. Nobody was there.

Did Fred dare to enter the cursed forest? Y

Fred continued his journey through the forest, and he soon met an old ragged man.

He had a shiny felnadit, a spear made of gold, in his hand: “I know your dreams, young man, but do you have the courage to become a knight?” he asked.

“I do!” Fred said.

“Can you retrieve this felnadit, this sharp tool, from a bear’s cavern, then?” the wizard asked, disappearing.

Fred thought he could do it, and he wandered around to find the cavern. Once he found it, he hid outside till the bear came out to hunt. At that point he sneaked into the cavern.
Inside the cave it was very dark, but the shiny felnadit was easy to spot.
He came out of the cavern and waited for the wizard, but he never reappeared.

Did Fred fight with the bear? N

Fred continued his journey. He came out of the forest, and reached a city. No one was in sight.
An enormous cynthor, a dragon that eats sheep, was roaming in the sky. Fred discovered that the people of the city were too frightened to come out of their houses.
"I'll try to free you from it," offered Fred.
"The cynthor is a gigantic creature, but I am not afraid."
Fred set off into the fields and engaged it in battle. People looked at the battle from a safe distance. They couldn't see much, just a lot of flames, but after a few days Fred returned. He was alive.
Nobody saw the meat-eating cynthor ever again.
Once the king heard of Fred's brave adventures, he allowed him to join his knights.

Were people happy and safe when Fred reached the city? N

Appendix V: definitions used in the stories in Study 3 and their features

Category	**Story**	**Definition**	**Definition type**	**Plausibility**	**Predictability proportion**	**Length**
Animal	a	dragon that eats sheep	that	2.73	0.08	22
Animal	b	elephant that pulls carriages	that	2.87	0.13	29
Building	a	tower with no windows	prep	1.67	0.25	21
Building	b	grave for many people	prep	1.67	0.08	21
Clothing	a	shirt made of chains	participle	2.87	0.00	20
Clothing	b	dress worn by men	participle	2.27	0.00	17
Food	a	soup eaten by farmers	participle	1.40	0.00	21
Food	b	potato wrapped in ham	participle	1.53	0.04	21
Job	a	someone who colours leather	who	2.47	0.00	27
Job	b	someone who sells furs	who	1.13	0.00	22
Object	a	spear made of gold	participle	1.67	0.00	18
Object	b	sofa used during meals	participle	2.20	0.00	22

Note: Plausibility = Mean plausibility computed from the adult sample (each judging plausibility on a scale from 1 – very implausible to 5 – very plausible); Predictability proportion = proportion of adults correcting predicting the final word of the definition from the previous ones; Length = length of the definition in characters.

Appendix W: generalized linear mixed models for accuracy in the phono-orthographic task in Study 3, considering gaze duration and re-reading time on the three repetitions of the target non-words when controlling for control non-word and real word performance

Factors	**Model 1: gaze duration**		**Model 2: re-reading time**		**Model 3: total reading time**	
	Estimate	**p-value**	**Estimate**	**p-value**	**Estimate**	**p-value**
(intercept)	-.28	.168	-.28	.167	-.28	.153
gaze duration	.01	.848				
re-reading time			.04	.655		
total reading time					.05	.581
condition	.54	< .001	.54	< .001	.55	< .001
reading accuracy	.33	.043	.32	.043	.33	.042
vocabulary	-.10	.453	-.10	.451	-.10	.434
non-verbal abilities	.24	.073	.24	.068	.26	.056
control non-words	.20	.148	.20	.147	.20	.146
real words	.17	.215	.17	.216	.16	.251
gaze duration* condition	-.05	.726				
re-reading time* condition			-.03	.826		
total reading time* condition					-.05	.722
Final Model *vs.* Empty Model	$\chi^2(8) = 85.63$, $p = .001$		$\chi^2(8) = 85.72$, $p < .001$		$\chi^2(8) = 39.92$, $p < .001$	