Cognitive and linguistic factors associated with reading and spelling in 9-11-year-old multilingual Arabic-English-Turkish speaking poor readers and typically developing readers

Iman Salama Elshawaf
Student ID: ELS07054091

A thesis submitted in partial fulfilment of the requirements of the degree of

Doctor in Education (EdD)
International Programme

Institute of Education

University College London

2021

Declaration:
I, Iman Salama Elshawaf, confirm that the work presented is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Abstract

This study aimed to investigate whole-word *lexical* and phonological *sublexical* processes in reading in multilingual Arabic^{L1}-English^{L2}-Turkish^{L3} speaking children aged nine to eleven. Accuracy and times for reading words and nonwords were recorded in the three languages. Associations of the measures across languages were examined as well as associations with measures of phonological awareness, rapid automatized naming, and vocabulary knowledge that were assessed in the children's three languages. Participants were 38 typically developing readers (TD) and 12 poor readers (PR) who were recruited from the same school in Istanbul, Turkey. Examination of the findings from the typically developing readers indicated that word and nonword reading processes were associated for English and Turkish and for Arabic and Turkish, but it seemed that processes for reading words and nonwords in English and Arabic were, to an extent, independent. Correlations with the literacy-related variables (LRVs) revealed the scores for vocabulary were associated with all the reading measures except for Turkish nonword reading. Phonological awareness was associated with all measures, but the correlation was only marginally significant in the case of English exception words and Arabic words. RAN was associated with all measures except for Turkish nonwords.

The findings for the poor readers involved two sets of analyses. One was a group comparison, where the results in reading and LRVs for the poor reader were compared with those of the typically developing readers. Findings indicated that, as a group, the poor readers had weaknesses in both lexical and sublexical reading processes across the three languages. The second set of analyses involved creating individual profiles for the poor readers in terms of deficits in reading words and nonwords and LRVs. The majority of children had deficits of both word and nonword reading in English and Turkish. There was one child with a selective lexical deficit in English and one with a sublexical deficit in Turkish. For Arabic, three children had deficits of both word and nonword reading and a further four children had sublexical deficits only. There did not appear to be any consistent association of deficits across the children's three languages and nor were there any consistent associations of reading deficits with those in the literacy-related variables.

The findings are discussed in relation to previous research looking at reading development across different orthographies and in multilingual children with reading difficulties.

Impact Statement

Arabic and its different dialects are spoken by around 422 million speakers and Turkish is spoken by around 220 million speakers. Arabic is spoken by 1.2% of the Turkish population. Moreover, and following the Arab spring in 2011, thousands of families from different Arabic-speaking countries have settled in Turkey. The result is that Turkey has now become the home to the world's largest refugee population, most of which are Arabic speakers who continue to use their mother tongue language in their daily lives at home. While these learners are obliged to learn the Turkish language, which is the language of the hosting country, and that is needed for their daily life, a significant portion of this population is targeting English Education for their children because it is the first international language. As a result, these children are now trilingual learners with varying abilities and needs that should be met. Nevertheless, there is a surprising lack of research into reading acquisition and reading difficulties in multilinguals children speaking these languages.

I believe my study has concrete educational implications, for instruction, assessment and interventions, and can help to highlight how we can meet the needs of this growing population of trilinguals. The study also contributes to the field by showing how we might increase our understanding of subtypes of developmental dyslexia among learners reading in different orthographies representing varying degrees of transparency. Moreover, research on impairments in literacy-related skills underlying reading difficulties has mainly been conducted from a monolingual perspective, therefore, widening the scope of investigations by conducting research on impairments in literacy-related skills in multilinguals has theoretical implications, since cross-language findings inform theory.

It is hoped that the research reported will provide an impetus to investigate different underlying causes for reading difficulties in the largely neglected writing systems of Turkish and Arabic, as well as providing means for assessing the component skills involved in being a skilled reader in these languages.

Reflective statement

In 2010, I obtained my MA from the Institute of Education (IOE) in Inclusive Education. Being an educator for more than 25 years, I claim that learning disabilities are common among Arabic- speaking learners just as they are among non-Arabic speaking learners. I have witnessed a variety of challenges faced by different learners, including members in my own family, not to mention the difficulties I have experienced. However, I am convinced that the more we understand how to meet the unique needs of all learners, and the more we raise awareness about different types of learning difficulties, we will be better able to overcome any challenges that might be faced.

In addition to the academic knowledge I have gained, my personal and family experiences with learning difficulties made me conscious of the many issues learners face despite their intelligence and outstanding personalities. I decided to implement what I have learned, and I managed to become a pioneer Special Educational Needs Coordinator (SENCo) in a mainstream school in Cairo, a job that has only existed among international schools where only very rich families could afford to enroll their children in such schools. Furthermore, and immediately after I have obtained my MA, I embarked on the EdD (October 2010).

Given my background, where the personal and the professional have merged into a true passion for helping others, I wanted to further what I learned in the field of Inclusive Education. In order to contribute to and enrich the limited pool of resources about Arabic learners with reading difficulties, this is my motivation, encouraging me to go further within this field.

As an international student in the EdD program, I had to visit London at least four times during the first two years, and to spend two weeks on each visit to attend the four taught modules that were a prerequisite to complete the first two years of the program.

During the first year, I obtained an A grade in the first modules, Foundations of Professionalism (FOP), after I wrote an essay on the SENCo as an unrecognized profession. I was awarded a B grade on my second module Methods of Enquiry (MOE1).

During my second year, following the Egyptian Revolution, I have experienced some problems with my progress, and I then faced the challenge of having to submit two overlapping assignments. I was able to earn an A on one of them (Specialist Course in International Education) but I had to ask for a deferral on the second (MOE2).

Working on (MOE2), which aims to develop research skills through conducting a small scale study, continued to be a challenge for me, especially in handling the results section, due to my poor knowledge in the area of SPSS at that time. However, I dedicated a special visit to

London and with the help of my supervisor, I joined a related course on data analysis and a related workshop which was included in the research week programme offered by the IOE in 2012.

I also attended a workshop on Endnote and a workshop to enhance my academic writing skills. My supervisor and I agreed that enhancing my weakness in data analysis and other related fields should continue to be among my priorities.

During the third year, and after passing (MOE2), I started preparing for the Institution Focused Study (IFS) and had been in contact with other researchers in Egypt who are involved in investigating early literacy and language skills in Arabic speaking children. I began building up a network of contacts. Nevertheless, my experience in conducting the Institution Focused Study (IFS) was the most challenging part of this journey. Instead of spending one year on this piece of research, I became distracted by my family commitments and my job requirements, in addition to problems in carrying on the study. My study was conducted at the school where I used to work in Cairo. It took me some time and many meetings to meet the required criteria before I could submit the ethical approval form on time. The approval process at my school proceeded slowly since this was the first time the school administration has ever agreed to support this kind of research. I also struggled to find the related literature through the limited research conducted in Arabic and to present the results section in a presentable form.

One of the paradoxes that occurred to me during this period, and because of the pressures of work (outside the home) and my family obligations and the problems I face as a mother of two teenagers and one young daughter (inside the home), I found myself in a serious state of depression, doing nothing but watching a Turkish T.V. show. I couldn't send the IFS study on time because I had not completed the discussion section. Moreover, I was disappointed to find out that studies about dyslexia in Arabic, which was the topic I was planning to work on for my final thesis, had no longer become a unique contribution to the field and have been studied by several researchers during the first three years of my EdD.

My husband encouraged me to contact my supervisor to discuss the option of having a pause in my studies. I did and she was very understanding when I told her that I am going through hard time. We had a very friendly and productive conversation that I am sure that not all supervisors would offer to their students. Within this talk, she asked me if they have dyslexia in Turkish and I answered with a smile that I know nothing about this language except that they have several words that I could understand because they have an Arabic origin.

This was not only a turning point in my academic route, as I was encouraged to resume my journey and overcome my depression, but it was also a turning point in my life as I started reading about Turkish, learning the basics of the Turkish language and contacting professors in the field of special education in Turkey.

In 2015, I decided I would include Turkish in my final thesis and was able to join the Exchange program (Erasmus), and travel to Ankara University in Turkey. I spent one year studying the language and attending lectures in the field of special education in Turkish. By that time, and due to the consequences of the Arab Spring in 2011, thousands of families from different Arabic-speaking countries had settled in Turkey, most of which continue to use their mother tongue language in their daily lives at home. Surprisingly, a significant portion of this population sought English language education for their children. This, in turn, created the need to establish an international school in Turkey which provides education in English and offers an internationally recognized certificate, while retaining elements of Arabic and Islamic cultural identity, to meet the needs of this growing population.

In September 2015, an international school was established in Istanbul (Turkey) to fulfil this need and I was fortunate to be a SENCo in that school. I was inspired by the multilingualism of the students (who speak Arabic, English, and Turkish), and with the encouragement of my supervisor, I modified my proposal and started investigating learning difficulties in this population. I believe this is the first study of Arabic-speaking children who are poor readers, and who have been exposed to a third language in their school and surrounding social context (Turkish). Moreover, I intended to focus on universal predictors and language specifics in typically developing children and poor readers who are learning to read more than one language. This should lead to a better understanding of normal and abnormal reading development and to answer unanswered questions beyond the scope of English research.

I know it has taken me a very long time to reach the stage of submission and defending my thesis; however, I think my work reflects a unique contribution that will be of value in the academic world and beyond. I also did my best to balance between my family commitments and my work requirements, both of which are important to me as a mother, an educator, and a breadwinner.

Table of Contents

1 Chapter 1: Literacy in English and Other Languages

Introduction

Reading and writing are critical to an individual and community's well-being. They are prerequisites to success in any society where a great deal of information is communicated in a written form. By reading, we gain knowledge and expand our understanding through learning information about the world around us. By writing, we send messages to our audience and transfer the knowledge we want to share. Thus, reading and writing are tools to communicate effectively. Learning to spell is also part of this communication process and is an initial step in writing. This ability to read and write is what we call literacy, and the development of this ability can be referred to as "literacy acquisition". As a complex cognitive skill and a measure of literacy acquisition, reading could be viewed as decoding or reading for meaning (Gleason & Ratner, 2009). In this current study, the emphasis is on reading aloud single words (word recognition). In fact, efficient word recognition, which is characterized by accurate and fluent word reading, is an important precondition for good reading comprehension (Karageorgos, Müller & Richter, 2019). Reading accuracy, in some studies called decoding accuracy or word decoding, refers to the number of items read correctly. The definition of "reading fluency", on the other hand, varies depending on the theoretical perspective of what the term represents (Wolf & Katzir-Cohen, 2001), however, and in its general sense, reading fluency refers to reading efficiency and combines accuracy and speed of reading (e.g., the number of correctly read words per minute), a sense in which the term will be used throughout this thesis.

Research on literacy acquisition in typically developing (TD) readers and those with reading difficulties has gained momentum over the past few decades. Studies in both fields have led to significant theoretical advances, which has helped broaden the understanding of how literacy develops and what cognitive and linguistic factors may anticipate this development, or otherwise predict literacy difficulties).

A large body of research is now focusing on developmental dyslexia as a general term for reading impairment (Friedmann & Haddad-Hanna, 2014). According to the International Dyslexia Association, dyslexia is a 'specific language-based disorder' characterised by difficulties with single word decoding, which usually reflects insufficient phonological abilities. These difficulties are 'unexpected' in relation to the child's age and cognitive and academic abilities and do not result from a general development disability or a sensory impairment. Dyslexia, according to English language-based research, is generally manifested in difficulties with reading, writing, and spelling familiar words and unfamiliar word-like letter strings (nonwords).

This study aimed to investigate whole-word (or lexical) and sub word phonological (or sublexical) processes in reading in multilingual Arabic-English-Turkish children aged nine to eleven years. Reading processes were examined in relation to selected literacy-related variables, namely phonological awareness, rapid automatised naming and vocabulary knowledge, to see whether patterns of association would be the same across the different languages of the children. The study involved TD children as well as poor readers (PR). I examined individual profiles of reading difficulty in the PR group and whether manifestations of the subtypes previously described, mainly in English, be observed, and whether these would differ in the three languages of the children, considering the different degrees of transparency.

This literature review chapter is divided into three parts. In the first, the focus is on literacy acquisition in typically developing children, and in the second, on reading difficulties. In the final section, the context of the present study is outlined, the research questions are put forward, and the contribution of the research to the field of reading development and reading disorders is discussed.

1.1 Part One: Reading in Typically Developing Children

This section will introduce the most influential word recognition models and the most prominent approaches for understanding the processes involved in typical literacy acquisition. Then, the cognitive and linguistic factors (determinants) of literacy acquisition investigated in the current study will be visited. This will help understand how deficits in them are related to the manifestation of dyslexia in English. Next, literacy acquisition in other orthographies and the nature and features of Arabic and Turkish will be outlined. The final part will shed light on how the phenomenon of reading in more than one language (multilingualism) is perceived in the literature.

1.1.1 Theoretical models of word recognition and reading aloud

Readers of English can read aloud regular and exception (irregular) words without any problems. A word is considered regular if using the spelling-sound/ grapheme-phoneme correspondence (GPC) rules yields the correct pronunciation and is considered an exception if using the same rule produces a wrong pronunciation. These readers also use the learned GPC rules to pronounce novel words that they have not seen before.

In order to understand the processes underlying reading aloud English words, which in turn allows for determining the origin of the failure when children experience reading difficulties (the focus of the second part of the chapter), a number of developmental models have been proposed. Among the most popular theories concerning the course of development of reading and spelling is the stage model presented by Frith (1985) and Ehri (1995)'s phase theory. According to Frith, readers acquire reading in three stages. Frith's model shows reading development as a developmental shift from reading using salient graphic cues, to slow phonological decoding, to automatic recognition of whole-word forms. Ehri (1995)'s model, on the other hand, assumes that although developing fluent reading skill requires children to proceed through some phases, the shift from one phase to another is rather transitional. Nevertheless, despite some differences, both models seemed to adopt the same sequence in the development of literacy skills (i.e., the start is usually a pre-phonological stage, then the alphabetic stage where grapheme-phoneme correspondences (GPCs) are gained and practised, and finally comes the complete understanding of the orthographic system). Also, in both theories, the availability of strategies in the later stages is assumed to depend on the development of earlier strategies. Moreover, both emphasised the importance of letter-sound mapping in literacy acquisition and argued that it serves as the basis for more complex orthographic rules. More recent models of word recognition that have received greater empirical support as frameworks for considering the development of literacy and reading difficulties are covered next.

The popularity of developmental models had opened the door for psychologists and neuroscientists to develop computational models. Among the most influential of these for word recognition are the Dual Route Cascaded (DRC) model (e.g., Coltheart, Rastle, Perry, Langdon & Ziegler, 2001) and connectionist Triangle Models (e.g., Plaut, McClelland & Seidenberg, 1996; Seidenberg & McClelland, 1989). These two models were chosen here because, in addition to their position as influential models in the field of reading, they provide contrasting frameworks. While the DRC model is more consistent with the idea that word identification is guided by linguistic "rules" that are used to access a word's pronunciation from its orthography, the triangle model is more compatible with describing this process as one in which different types of lexical information provide mutual "soft" constraints on the pronunciations that are generated during word identification (Rayner & Reichle, 2010).

The DRC model

In English, as mentioned above, we are familiar with the notions of regular (e.g., <gave>) and exception words (e.g., <have>). The presence of regular and exception words was behind the development of the dual-route theory of reading (e.g., Coltheart, 1978; Morton & Patterson, 1980), which proposes two main mechanisms for single word recognition (Figure 1.1).

The first mechanism is the sublexical route in which reading proceeds via grapheme-to-phoneme conversion. This route operates in a serial and sequential way by transforming visual symbols (graphemes) into sounds (phonemes) through GPC rules. These rules support accessing lexical information or semantic information, as well as effective pronunciation of unfamiliar printed words or made-up letter strings (e.g., <mave>), which will be referred to as nonwords throughout this paper. This sublexical route may also be critical to the ability to learn entirely new words from print (de Jong & Share, 2007). However, it has the disadvantage, in the case of some orthographies, not all, that if a word's orthography-phonology mapping deviates from the regular mapping, then it may be impossible to derive a phonological form that meets the proper pronunciation of that word. Therefore, a different mechanism by which words with unfamiliar orthographic representation (exception words) could be directly mapped as words was needed.

The second mechanism is the lexical route, which includes the orthographic input lexicon and the phonological output lexicon. The orthographic input lexicon is where information about the written form of known words is stored. The lexical route processes all words known to the individual, including exception words. However, the phonological information of the known spoken words is stored in the phonological output lexicon (e.g., consonants, vowels, stress position, number of syllables). Through the lexical route, which is the link between these two lexicons, written words can be directly mapped onto mental representations, and hence converted to phonological forms without an intermediary step of GPC rules, which would otherwise result in reading errors due to the generalisation of these rules when applied to exception words (e.g., <have> misread as /heiv/).

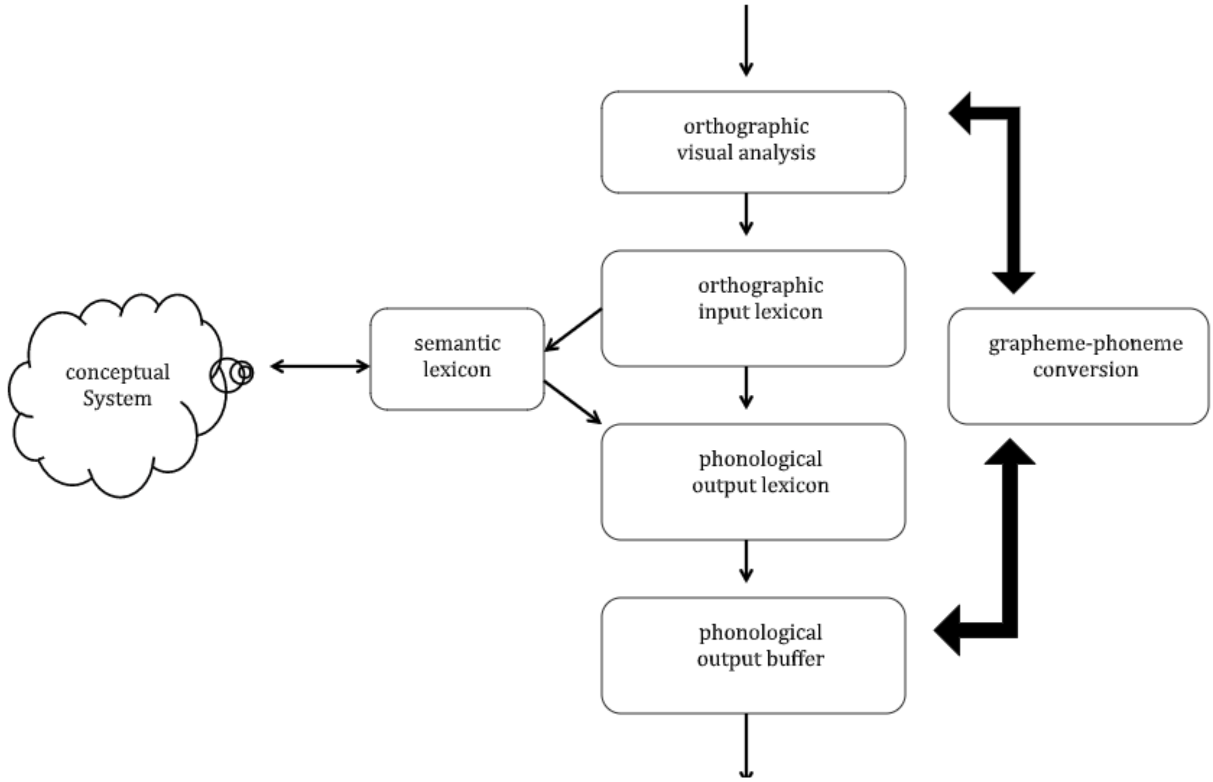

Figure 1.1 The Two Routes in the Dual Route Model of Word Recognition (from Friedmann & Haddad-Hanna, 2014)

The first stage that precedes the two routes is the orthographic-visual analysis stage. Through this process, letter identities are encoded, the position of letters within words are encoded, and letters are blended into words. In addition to the lexical and sublexical routes, the model includes a connection between the orthographic input lexicon and the semantic lexicon and conceptual system. This access to semantics allows for the comprehension of written words.

The development of both routes is believed to be necessary to become a proficient reader in English. Moreover, while different types of dyslexia indicate impairments in one of these two main routes (or both), a deficit in any of the three functions of orthographic-visual analysis or impairment to the connection between the orthographic input lexicon and the conceptual-semantic system may also lead to different types of dyslexia (Friedmann & Haddad-Hanna, 2014), as will be discussed in the second part of this chapter.

Wile and Borowsky (2004) argued that the way in which the different types of English letter strings are best read is clear; that is, exception words must be read using the lexical pathway, nonwords using the sublexical pathway, and regular words by way of either processing route. However, some studies had found that nonwords were not always processed via the sublexical route but rather (at least partially) via the lexical route, as this might happen when the nonword shows similarity to an orthographic neighbour that has irregular spelling.

Coltheart et al. (2001) developed a computational version of the DR theory. There are two fundamental assumptions of this dual route computational model (DRC) as outlined by Coltheart et al. (2001) – pronunciations of printed words can either be generated by using linguistic rules transforming graphemes to assemble the pronunciation, or by retrieving the whole word pronunciation from the lexicon (the two routes of DR theory), and that both the orthographic and phonological forms of words are represented holistically. Moreover, lexical and sublexical routes run in parallel in the DRC model, with the pronunciation of any given word in most cases being jointly generated by the outcome of both routes.

Several studies with English-speaking children and adults provided evidence for the relevance of the two routes. Albeit that most of the available evidence comes from studies with English-speaking children, successful utilisation of the model was seen in other orthographies, e.g., German (Ziegler, Perry & Coltheart, 2000), Arabic (Friedmann & Haddad-Hanna, 2014), and Turkish (Güven & Friedmann, 2019).

The distinction between visual and phonological processes in the dual route model has been opposed by some researchers, who argued that a single cognitive mechanism may be capable of processing all types of words. Furthermore, they claimed that their approach (visited in the next section) helps avoid the problematic distinction of regularity of printed words (Seidenberg & McClelland, 1989).

Connectionist models of printed word recognition (Triangle models).

The second influential theoretical framework in the field is the connectionist approach, comprising mechanistic models that were implemented as computer programs (Hulme & Snowling, 2009). This approach has gone through several levels of development, aiming to represent an alternative to the dual route approach. The original version had one non-semantic route between orthographic representations and phonological representations and one semantically mediated route (e.g., Harm & Seidenberg, 1999; Plaut, McClelland, Seidenberg & Patterson, 1996). The ability of this type of model to account for reading of all types of letter-strings was contested by many authors (e.g., Besner, Twilley, McCann, & Seergobin, 1990; Coltheart et al., 2001). Among the most popular attempts to develop the approach was the connectionist model proposed by Seidenberg and McClelland (1989). The model was a three-layer, feedforward neural network. The framework was referred to as the “triangle” model, referring to the three components of the model – orthographic input, phonological output, and a semantic component.

Generalisation (i.e., reading of nonwords) was a major limitation of the early versions of the connectionist models, and hence was the strongest evidence for the dual route model. Also, triangle models were criticised for the use of monosyllabic stimuli and for ignoring the development of phonological skills prior to formal reading instruction. Therefore, Seidenberg and McClelland (1989) attempted to improve their model by increasing training and adjusting ways to model human performance (see Figure 1.2). Moreover, the model was substituted by that of Plaut et al. (1996). This later model, although still a single-route model, was claimed to be able to read monosyllabic nonwords at a level of performance like that of humans.

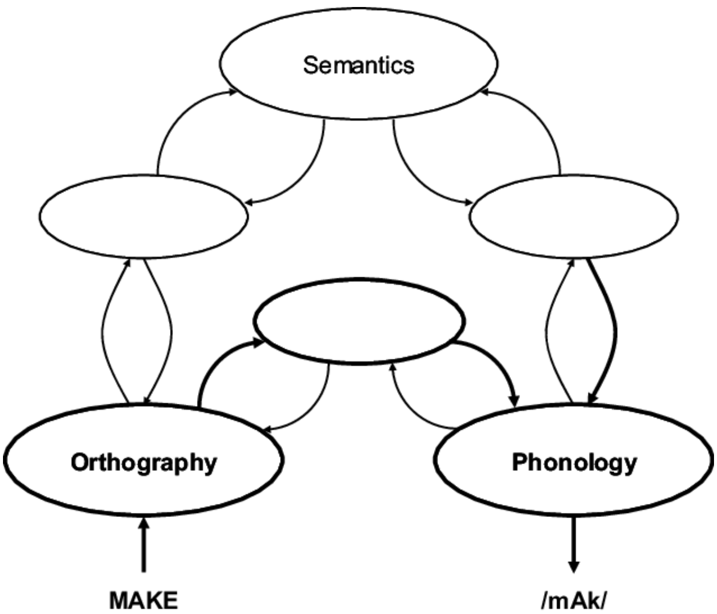

Figure 1-2 The Triangle Model of Reading Aloud (from Seidenberg & McClelland, 1989).

Unlike DRC theory, the triangle models avoid any notion of rules and lexicons, and instead propose an integrated, indivisible mechanism for handling both regular and exception items. Orthographic, phonological and semantic information are connected by means of “hidden units” mediating between inputs and outputs in word reading, and words and nonwords are read by a single mechanism (Plaut et al., 1996). Furthermore, the two fundamental hypotheses of the triangle models are the opposite of those of the DRC model. First, the input that allows identifying printed words is contained in a “single” set of input-to-output connections, in a manner that allows it to influence the pronunciation of each word. Second, the triangle models do not assume that lexical information is represented by separate processing units, but instead assume that such information is contained in the connections that mediate between the orthographic input and phonological output.

In spite of the differences, the two influential approaches explain a large number of important ‘benchmark’ phenomena such as the finding that frequent words are identified more rapidly than infrequent (nonwords), and that this frequency effect is typically larger for exception word pronunciation than for regular words (Rayner & Reichle, 2010). For beginning readers, almost every word in its written form is new and unknown, so they may not use the lexical route, however, when these words are frequently recognised the mental lexicon is then ready to be accessed either in a direct way (lexical route, according to the DRC model, Coltheart et al., 2001) or side by side with phonological and meaning representations (according to the triangle model, Plaut et al., 1996).

As for developmental reading difficulties, both the DRC and the more recent versions of the triangle models were found to be able to demonstrate the patterns of behavioural deficit that are observed with surface and phonological subtypes of dyslexia, yet in very different ways. However, the Dual Route model has been extensively used as a framework for investigating developmental reading difficulties in many writing systems (e.g., Rowse & Wilshire, 2007 ; Douklias, Masterson & Hanley, 2009; Sotiropoulos & Hanley, 2017). Moreover, Friedmann and Haddad-Hanna (2014) employed the dual route model to account for the various types of dyslexia that can be observed in Arabic. Given that Arabic is one of the languages studied in the present research, the Dual Route model was used in this study as the theoretical framework. In the next section, three factors that are considered to have an important role in literacy acquisition, phonological awareness (PA), rapid automatised naming (RAN) and vocabulary knowledge, will be discussed.

1.1.2 Literacy-related variables associated with effective reading acquisition

There is joint agreement that reading is a complex skill involving the orchestration of several components (Rayner & Reichle, 2010). Therefore, the task for typically developing readers consists of mastering these components and will be dependent on the integrity of a variety of linguistic and cognitive processes (Vellutino, Fletcher, Snowling, & Scanlon, 2004). For example, in terms of alphabetic writing systems, research has indicated that effective skills in predicting reading development are phonological awareness (PA), which denotes sensitivity to the sound structure of words, and Rapid Automatized Naming (RAN), which reflects the ability to name familiar stimuli quickly (e.g., Caravolas, Lervåg, Defior, Málková & Hulme, 2013; Moll Ramus, Bartling, Bruder, Kunze, Neuhoff, …., & Landerl, 2014; Ziegler, Bertrand, Tóth, Csépe, Reis, Faísca …. & Blomert, 2010). Other cognitive and linguistic skills that have been reported to contribute are letter knowledge, verbal memory, vocabulary, and morphological knowledge (e.g., Nation and Snowling, 2004; Carroll, Snowling, Hulme & Stevenson, 2003; Gillon, 2004; Landerl and Wimmer, 2008; Muter, Hulme, Snowling & Stevenson, 2004; Savage, Lavers & Pillay 2007; Snowling, Gallagher, & Frith, 2003).

PA and RAN were targeted in the present study since, of all the language and cognitive processes that have been investigated concerning reading development in children of the age range covered in this study, these are arguably the most widely studied. Vocabulary knowledge was also targeted due to its particular importance, as will be shown in the following sections, as the participants in this present study learn more than one language. These three factors will be reviewed from the perspective of research conducted in English in the following subsections. Then, in the next section (reading in other orthographies) the variables will be revisited to see what the effect of different orthographies on each of them is. Studies of PA, RAN, and vocabulary in the target languages for the present study are covered in the section after.

Phonological awareness (PA)

Phonological processes are an integral part of reading models, as discussed earlier. For example, in the Dual Route model, phonological information is accessed through an assigned pathway (the sublexical route). In the triangle model, phonological processing is recruited, together with orthographic and meaning representations whenever words are read (Knoepke, Richter, Isberner, Naumann & Neeb, 2014). Phonological processing skills have been widely reported to be related to the acquisition of literacy (e.g., Rack, Hulme, Snowling, & Wightman, 1994; Stanovich & Siegel, 1994), leading some to argue that they are the "sine qua non" for successful reading acquisition (Ziegler & Goswami, 2005). Phonological processing skills include PA and phonological memory.

PA is defined as sensitivity to, or explicit awareness of, the phonological structure of words in the language. Thus, it encompasses an awareness of individual words in sentences, syllables, onset and rime segments, and awareness of individual phonemes in words. For this reason, Castles and Coltheart (2004) emphasised the significance of "awareness" as an essential component of the definition. In short, PA refers to the ability to reflect, process, conceptualise and manipulate sub-lexical speech segments (Elbeheri & Everatt, 2006). Furthermore, the development of PA is thought to enhance the ability to map between graphemes and phonemes. Accordingly, many studies have reported that PA is one of the most potent concurrent and longitudinal predictors of subsequent reading ability in English (e.g., Wagner & Torgesen, 1987).

PA was more strongly linked to Nonword reading ability, which are reflected more in reading accuracy than in reading fluency. Because of this, much of the available research on reading disabilities has focused on PA deficits and how to "alleviate them" (Jacobs, 2010). However, some researchers have argued that the relationship might be the other way around and that the recognition of the internal phonemic structure of spoken words is a result of learning to read in an alphabetic orthography (e.g., (Oktay & Aktan, 2002; Castles & Coltheart, 2004; Peterson, Arnett, Pennington, Byrne, Samuelsson, & Olson 2018).

PA may be assessed in rhyme judgment tasks, syllable awareness (blending, segmentation, deletion of syllables), and phoneme awareness (blending, segmentation, deletion of phonemes). These tasks may vary across orthographies and studies. However, Castles and Coltheart (2004) reported that as reading and spelling skills develop, PA tasks may increasingly be carried out using orthographic information rather than purely phonological information. So, children may use an orthographic strategy to perform tasks with a moderate level of difficulty, such as syllabic and phonemic awareness tasks, using orthographic images. Nevertheless, phonemic awareness has been the primary focus of the studies investigating the relation between PA and progress in word recognition in beginning readers (Durgunoğlu et al., 1993; Gillon, 2004). Moreover, there is evidence that difficulties with phonemic awareness predict subsequent reading problems (e.g., Wallach & Wallach, 1976; Pratt & Brady, 1988; Liberman & Liberman, 1990), mainly based on training studies (Bradley & Bryant, 1983). Likewise, Melby-Lervåg, Lyster & Hulme (2012) showed that independent of orthography, children with dyslexia had a significant deficit in phonemic awareness, more than in other PA measures, compared to their typically developing peers.

Rapid Automatized Naming (RAN)

RAN has also been reported to be an essential associate of reading ability. It refers to the ability to quickly name sequentially and repeatedly presented sets of stimuli, such as letters, digits, colour patches, or pictures of familiar objects (Moll et al., 2014). These tasks are thought to tap a range of cognitive skills, including speed of processing, visual and verbal integration, executive function, and access to phonological representations (Alves, Siqueira, Ferreira, Alves, Lodi, Bicalho & Celeste 2016).

The exact association between RAN and reading is not fully agreed upon (e.g., Lafrance & Gottardo, 2005), and some competing theories aim to explain the underlying mechanism driving the relationship (da Silva, Engel de Abreu, Laurence, Nico, Simi, Tomás & Macedo, 2020). Initially, deficits in rapid naming were viewed as part of the 'phonological core deficit' in poor readers (Catts, 1996; Wagner, Torgesen, Laughon, Simmons, & Rashotte, 1993). These deficits have been widely explained in terms of problems in accessing phonological codes in memory. However, Swanson, Trainin, Necoechea, and Hammill (2003) reported that RAN constitutes a cognitive mechanism underpinning reading development that is mainly independent of PA and memory. Moreover, findings from a large-scale longitudinal study including different European orthographies indicated that RAN and PA were separate factors (Moll et al., 2014) and were reliable predictors of reading skill, with equal relative importance (Caravolas et al., 2013).

RAN tasks can be separated according to the stimuli used. A common distinction is between RAN tasks that use alphanumeric stimuli (letters and digits) and pictorial or non-alphanumeric stimuli (objects and colours). Non-alphanumeric RAN has been shown to predict reading for children who have not entered school yet (da Silva et al., 2020). On the other hand, as suggested by research, alphanumeric RAN tasks (letters and digits) are related to lexical processing. For example, Manis, Seidenberg, and Doi (1999) reported that alphanumeric RAN was most strongly related to three measures emphasising knowledge of orthographic patterns: Orthographic Choice, Word Likeness Judgment, and Exception Word Reading, while it was only weakly associated with later nonword reading skill. Further evidence was reported by Wile & Borowsky (2004), where their findings supported the hypothesis that performance in RAN letters reflected exception word processing.

Some studies have found that RAN is more closely associated with reading fluency than reading accuracy. Its role during fluent reading is more important than during spelling processes across all alphabetic orthographies (Moll et al., 2014). Per contra, Powell and Atkinson (2020) conducted a longitudinal study to assess pictorial RAN and PA in very young non-reading English children. Children were reassessed on RAN, PA, and word-level reading a year and a half later and then a further year later. They found that RAN was a significant predictor of word reading ability, both in terms of accuracy and fluency.

Semantic knowledge (Vocabulary)

Semantic knowledge is the feature of language knowledge that involves word meanings. Its role as one of the necessary stages in printed word recognition has been stressed in several models (Adams, 1990; Plaut et al., 1996). For example, in the triangle model, semantic knowledge as well as phonological and orthographic knowledge were believed to work together in cases where GPC processes were unsuitable for an accurate reading, as in reading exception words (Nation & Snowling, 2004; Ricketts, Nation, & Bishop, 2007). And in the Dual Route model, reading via the lexical route involves drawing from a repertoire of coexisting semantic representations, sometimes referred to as vocabulary knowledge.

Vocabulary may be defined as "all the words known and used by a particular person" (Cambridge Dictionary). Receptive vocabulary involves words that are recognised when they are seen or heard, although the person may not necessarily be able to produce them (Harmer, 1991), while expressive or productive vocabulary involves words that are used in communication (McDowell, Carroll & Ziolkowski, 2013).

Some studies have reported that vocabulary knowledge was not a predictor of single word reading. Muter et al. (2004), for example, conducted a two-year longitudinal study with 90 British children at their school entry to investigate the role of phonological, grammatical, and vocabulary skills as predictors of single word reading and reading comprehension. They found that word reading skills were predicted by earlier measures of letter knowledge and phoneme sensitivity but not by vocabulary knowledge, rhyme skills, or grammatical skills, while reading comprehension was predicted by prior word recognition skills, vocabulary knowledge, and grammatical skills. Other researchers reported that vocabulary seems to be more important for reading comprehension rather than for decoding (e.g., Nation & Snowling, 2004; Roth et al., 2002). In contrast, Vellutino, Tunmer, Jaccard, and Chen (2007) used the expressive vocabulary task from the WISC-R in the primary school years. They found a stronger effect of semantic knowledge on single word reading in older (sixth and seventh grade) rather than in younger (second and third grade) children. Ouellette and Beers (2010) looked at the influence of vocabulary (assessed by a receptive task and an expressive task) on nonword and exception word reading. Receptive vocabulary was shown to predict both nonword reading and exception word reading for grade six children. Vocabulary also predicted exception word reading but not nonword reading in grade one. For grade one children, nonword reading was predicted by PA measures only. The link between nonword reading and receptive vocabulary in older children in Ouellette and Beers (2010)'s findings was in accord with their proposal that lexical size is the characteristic that influences nonword reading, and that as children become more skilled at nonword reading, the contribution of vocabulary becomes more evident. Similarly, some researchers reported that the effect of word reading on vocabulary growth is not trivial (Duff, Tomblin & Catts, 2015), other studies indicated that through learning to read and increasing vocabulary, children acquire literacy-related phonological skills (e.g., Metsala, 1999).

The discrepancy in findings from studies investigating the association between vocabulary and reading accuracy and fluency might be a result of the difference in age of the participants, the measures used as an index of vocabulary, or the nature of the orthography being examined. In addition to being investigated as one of the predictors of literacy development, vocabulary has been commonly used as a background variable for identifying the verbal ability of participants in research studies. As this thesis investigated literacy in children speaking more than one language (multilinguals), it was of importance to assess vocabulary across all the languages considered. In the next section, literacy acquisition and the contribution of the literacy related variables (LRVs) to literacy acquisition in languages different than English, which has been the orthography in which most of the research has been conducted, will be explored.

1.1.3 Reading and Reading-Related Variables in Other Orthographies

As acknowledged earlier, literacy happens when a written form of any writing system is successfully converted into the oral form of the language and vice versa. Orthographies have been categorised as logographic, syllabic, or alphabetic, depending on how the structure of the written form is related to the language. According to the "central processing" (universalist) perspective, the development of word-based processes in different languages is shaped by the existence of general cognitive processes that affect reading in alphabetic orthographies (Seidenberg & McClelland, 1989). The "script dependent" perspective, rather postulated that various cognitive processes might contribute differently to reading, depending on the nature of the orthographic system and as a function of orthographic regularity (Share, 2008; Katz & Frost, 1992).

In what follows, I examine the effect of the nature of the orthography being read on reading development and the selected reading-related variables, to see whether processes involved in reading may vary across different orthographies.

Studies investigating literacy acquisition in different orthographies

The dominance of English in this field encouraged scientists to research different languages to test the applicability of the findings that were based on English-language studies. For example, Ziegler et al. (2010) conducted a cross-sectional cross-linguistic study that investigated the role of PA, phonological short-term memory (PSTM), vocabulary, RAN, and nonverbal intelligence. Their sample consisted of 1,265 children in grade two learning to read in Finnish, Hungarian, Dutch, Portuguese, and French. They found that PA, assessed with a classic phoneme deletion task, was important yet not the most important predictor in all the tested languages (PA was not the most important in the Finnish orthography), and was rather more important in less consistent orthographies (French and Portuguese) and less important in more consistent orthographies (mainly Hungarian), because PA scores reached ceiling quite early in more consistent orthographies. In addition, and unlike what they expected, RAN was not a strong predictor of accuracy, nor did it have a significant relation with the consistency of the orthography. They attributed this finding to the use of pictorial RAN and acknowledged that pictorial RAN tended to have lower correlations with reading performance than alphanumerical RAN. As for vocabulary, where this study was among the very few studies that included vocabulary within its assessments, using an expressive vocabulary test, they found that the relationship between vocabulary and reading was weak. Only Finnish, which has a very consistent orthography, showed strong correlations of vocabulary with both word and nonword accuracy. Ziegler et al. concluded that the significance of PA was modulated by the ceiling effect, and the impact of RAN was limited to fluency (both words and nonwords).

Likewise, Moll et al. (2014) conducted a cross-sectional cross-linguistic study to compare the associations of PA, memory (PSTM and working memory) and RAN, with word and nonword reading (fluency and accuracy) and word spelling in five alphabetic orthographies covering the full range of orthographic consistency. The participants were 1062 children from Grade two to Grade five for the three consistent orthographies Finnish, Hungarian and German, and from Grade three to Grade seven for the less consistent orthographies French and English. The reading measures were fluency and accuracy presented in language-specific material for each country. RAN (assessed with the digit and picture naming) was the best predictor of reading fluency, and PA (assessed with phoneme deletion tasks) and memory predicted reading accuracy and spelling. In this study, which was reported as the analysis of PA and RAN in reading and spelling beyond the initial stages of literacy development, Moll et al. concluded that the predictive patterns were broadly comparable across orthographies. However, the overall influence of PA and RAN on word and nonword reading accuracy and fluency, as well as on spelling, was stronger in English than in other orthographies.

A third cross-sectional study was conducted by Tobia and Marzocchi (2014) to investigate predictors of reading fluency in transparent Italian. They recruited 651 Italian children in Grades one through five and examined the contribution of PA and RAN to fluency. The children were asked to read a text aloud within a four-minute time limit. Only fluency score (in terms of syllables per second) was analysed as accuracy was very high in all grades. They also included expressive vocabulary, and verbal short-term memory, which they considered controversial variables. The researchers hypothesised a decreasing effect of PA and an increasing effect of RAN in older grades. However, their findings showed that RAN was a stronger predictor than PA in younger children, while the predominance of RAN faded for third to fifth graders, and PA and memory were the most important independent predictors of reading fluency in this age group. The study confirmed that PA and RAN were influential predictors of reading fluency in Italian orthography for typical readers throughout primary school. Furthermore, and beside RAN and PA, the findings indicated a significant effect of vocabulary. Precisely, vocabulary and memory played a significant role in the reading fluency of third, fourth, and fifth graders. Tobia and Marzocchi attributed the contribution of vocabulary to the nature of the reading task used: text reading is a complex task that involves not only word-level decoding and lexical strategies but semantic and syntactic processes that may support fluent reading. According to Tobia and Marzocchi, the findings of their study could be generalised to text reading in various alphabetic orthographies, however, the weight of each predictor would be modulated by orthographic transparency, i.e., the importance (and weight) of these predictors may vary depending on the nature of the writing system tested.

In addition to the cross-sectional studies, longitudinal studies were also conducted in scripts varying in the degree of consistency. For example, Landerl and Wimmer (2008) followed a group of 115 students in Grade one learning to read in (transparent) German for 8 years. They assessed letter knowledge, PSTM, PA (using a special task of imitation of phoneme segmentation), pictorial RAN, and nonverbal IQ to examine their influence on word reading fluency and spelling. To assess fluency, the number of reading errors and reading time for word lists (for younger students) or text (for older students) were recorded. They found no significant contribution of PA to reading level after Grade one when preliteracy letter knowledge was considered. The contribution of RAN to reading fluency was significant in all grades. The findings revealed that word reading fluency was mainly predicted by early RAN performance rather than PA, and that PA was the only predictor for spelling. Given that German is a transparent writing system, these findings were applicable to consistent or transparent writing systems.

Another longitudinal study with another transparent writing system was conducted by Lervåg, Bråten and Hulme (2009). The study followed 228 Norwegian children at the beginning of Grade one before they started learning to read (average age was 6 years 4 months) for two years. The researchers investigated the role of letter knowledge, PA (using phoneme isolation, phoneme segmentation, and phoneme deletion), visual–verbal paired-associate learning, RAN (Pictorial in Time one then both pictorial and alphanumeric afterwards), short-term memory, and verbal and nonverbal ability. They related these skills to later measures of single word reading fluency, orthographic choice, text reading, and nonword reading accuracy. They found that the best predictors of early reading skill were letter knowledge and PA. The findings were similar to those reported for English, which has an inconsistent letter-sound relation (Muter et al., 2004). They revealed that Grade one PA significantly affected Grade one reading, but not Grade two or three reading, after prior reading level was controlled due to ceiling effects.

Beside longitudinal studies that tested one writing system, longitudinal studies have been carried out across several orthographies. A recent study by Landerl, Freudenthaler, Heene, De Jong, Desrochers, Manolitsis, Parrila & Georgiou (2019) was the first to apply the same predictor and reading measures in five orthographies. They followed 1120 children across grades one and two, where the participants were learners of orthographies with varying degrees of consistency (English, French, German, Dutch, and Greek). The measures of PA were phoneme elision with real words and nonwords, and RAN assessed with colour and digit naming tasks. The English version of the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999) and similar forms in the other languages were used to assess reading fluency and accuracy. There were differences in predictive patterns between the languages. These variations were not only observed at the first assessment point at the beginning of grade one where most of the German-speaking children were still non-readers, whereas French- and English-speaking children had comparably advanced reading skills (with Dutch and Greek children somewhere in between), but also in the last assessment point at the end of grade two. At this point, PA did not predict reading in Greek and Dutch, the most consistent orthographies, nor did it predict reading in English, the orthography with the most inconsistent orthography. As for German, an interactive pattern (meaning that the development of PA and reading was mutual) was apparent in Grade two, and in French, this interactive pattern was evident right from the start. These findings did not support the "classic" view that PA predicts reading, but rather that the relationship between PA and reading relied on many factors like task characteristics, developmental status, and the consistency of the orthographies. Therefore, the results suggested that instead of being a prerequisite for reading, PA might function as a co-requisite skill for typical reading development. On the other hand, they conclude that RAN was a consistent robust predictor of reading fluency in all five orthographies (independent of consistency).

In short, Moll et al. (2014)'s study showed that the pattern of relations between the predictors in English were different from the pattern of relations among the same predictors in other orthographies, particularly, RAN which showed a negligible association with reading accuracy and spelling in more consistent orthographies, but was a significant predictor in English for reading accuracy and spelling. However, RAN was a consistent predictor of reading fluency in all orthographies, which is a similar finding to that of Zeigler et al. (2010).

Similarly, Landerl and Wimmer (2008) found that word reading fluency was mainly predicted by early RAN performance rather than PA for German speaking children in grades 1 to 8. Tobia and Marzocchi (2014) in an Italian study, found that RAN was a stronger predictor than PA in younger children, while the predominance of RAN faded for third to fifth graders, and PA and memory were the most important independent predictors of reading fluency for all grades. Tobia and Marzocchi showed that the robust effect found for PA can extend to higher primary school grades if the PA task is sufficiently sensitive and does not produce a ceiling effect. Additionally, Lervåg et al. (2009) in a study in Norwegian, found that PA was the main predictor of word reading accuracy.

The most studied reading measure in consistent orthographies is reading fluency because reading accuracy reaches ceiling at an early age in these orthographies. Moreover, the predictors of reading, at least in alphabetic languages, were found to be relatively universal and there was a robust association between RAN and PA and reading. Variation has been found regarding whether PA predicts reading beyond the first year or so of learning to read. Results of more recent studies indicate that the reason PA is not a significant associate of reading for older grades is because PA scores reach ceiling quite early on for transparent writing systems (e.g., Landerl & Wimmer, 2008, and Tobia & Marzocchi, 2014). On the other hand, results indicated that the relationship between RAN and reading can continue to be detected in higher grades because it is a measure of speed. Nevertheless, variations in the effects of RAN were also apparent, specifically whether the influence of RAN is limited to reading fluency, and whether alphanumerical RAN is better in predicting reading fluency than pictorial RAN.

As for vocabulary, and unlike Ziegler et al. (2010)'s findings that indicted a weak effect of vocabulary in reading in young participants, Tobia and Marzocchi (2014) found a significant effect of vocabulary. They indicated that the effect of vocabulary is affected by the type of reading task chosen; when reading text, the effect of vocabulary is stronger than in a single word reading task because the sentence structure and the context of the passage provide clues to the reader. Because older children know a greater number of words, they can use this lexical knowledge to anticipate the next word in a sentence they are reading.

From the studies discussed, it appears that similarities exist across languages in the patterns of association of reading with literacy-related variables at the early stages of learning to read and later in primary school. However, there still appear to be some uncertainties, and it is difficult to draw strong conclusions about the reasons for inconsistent findings when studies use different reading measures, different assessments of PA, RAN, and vocabulary and participants are drawn from different age groups.

Reading-related variables in different orthographies

Besides findings from English orthography, as noted above, PA was found to be a powerful concurrent predictor of subsequent reading ability in other orthographies (e.g., Moll et al., 2014; Caravolas et al., 2013). In addition, PA was found to be a "longitudinal predictor" of reading ability in some regular orthographies, at least over the first year of learning to read (e.g., de Jong & van der Leij, 1999 [Dutch]); Öney & Durgunoğlu, 1997 [Turkish]; Silven, Poskiparta, Niemi, & Voeten, 2007 [Finnish]; Wimmer, Landerl, Linortner, & Hummer, 1991 [German]).

Some research has suggested that the contribution of PA to literacy ability varies depending on the nature of the writing system (Ziegler et al., 2010). For example, in inconsistent orthographies, PA was slightly more critical than in consistent orthographies (Zeguers et al., 2018). In addition, reading accuracy measures, which are more linked to PA than those of reading fluency, as mentioned above, tend to quickly approach ceiling in readers of consistent orthographies (Wimmer, Mayringer & Landerl, 2000; Powell & Atkinson, 2020). Accordingly, the effect of PA is less strongly and enduringly evident in reading accuracy with more consistent orthographies. Findings from some consistent orthographies (e.g., Spanish) showed that features of the linguistic system such as differentiated syllable boundaries and salience of the initial phoneme could also help the development of strategies that further facilitate children's performance in terms of accuracy (Serrano & Defior, 2008). Findings from some consistent orthographies also suggested that early PA deficits might not affect literacy acquisition until later stages (Wimmer et al., 2000).

In transparent languages, the contribution of PA to reading fluency was found to be stronger at the beginning of reading development but then declined with reading expertise (Germano, Reilhac, Capellini, & Valdois, 2014). Differently, several studies have shown a greater role for PA in word reading fluency in English than orthographies with a more consistent nature (e.g., Landerl et al, 2019). However, Ziegler et al. (2010) claimed that PA would show identical effects on reading in consistent and inconsistent orthographies if the PA tasks are at a sufficiently difficult level in consistent orthographies.

Regarding RAN, concurrent and longitudinal studies have shown that it correlates with reading accuracy and early reading fluency (Norton & Wolf, 2012; Araújo, Reis, Petersson, & Faísca, 2015) and that it is predictive of developmental outcomes in reading across languages (Farran, 2010). Furthermore, it has been argued by some that RAN is the predominant predictor of individual differences in reading in consistent orthographies (Mayringer & Wimmer, 2000), but others have shown different results. For example, Patel, Snowling, and de Jong (2004) looked at reading skills (single word reading accuracy and latency) in 67 English and 40 Dutch-speaking children, aged six to eleven and found a similar pattern for the two orthographies: PA was a significant predictor of individual differences in reading skills in both languages, while RAN did not enter the regression model as a significant predictor. The unexpected finding that RAN did not contribute as a predictor of reading skill was attributed to the relatively large age range in association with the relatively small sample size, or the use of pictorial RAN rather than RAN digits and letters. Patel et al. also claimed that this finding might support the more parsimonious explanation, which is that the effects of RAN were subsumed by phoneme awareness.

Vaessen et al. (2010), on the other hand, investigated the effect of PA and RAN (letter, digits, and objects) on reading fluency cross-sectionally in a large sample of Hungarian, Dutch, and Portuguese children from grades one to four and found that the impact of RAN increased with grade level across the whole sample indicating a universal effect regardless of the nature of the orthography (at least for alphabetic scripts). Moreover, Papadopoulos, Spanoudis and Georgiou (2016) followed 286 Greek children from Grade one to Grade two to examine the relation between RAN and reading fluency. They included both type of RAN tasks (pictorial vs. alphanumeric). PA and orthographic processing were also assessed. The findings of both concurrent and longitudinal analyses indicated that RAN was a unique predictor of oral reading fluency. Papadopoulos et al. indicated that using alphanumeric or non-alphanumeric RAN did not particularly affect the RAN-reading relationship.

Regarding vocabulary, and as shown in the previous section, studies including vocabulary in other orthographies showed a similar discrepancy in findings, as those found in English-language based studies. For example, Caravolas, Volin, and Hulme (2005) compared the predictors of the development of reading fluency, spelling, and reading comprehension between Czech- and English-speaking children. The average age of the Czech children was 9 years 8 months, and the average age of the English children was 9 years 2 months. They administered the WISC-III Vocabulary subtest (Wechsler, 1992, 1996) to both language groups to provide an estimate of verbal ability. Scores from this expressive vocabulary assessment correlated significantly with PA scores, and vocabulary was found to be a significant predictor of single word spelling and reading comprehension in both languages. Vocabulary was also a significant predictor of reading fluency in Czech.

Moreover, and as noted earlier, the relationship between vocabulary and reading fluency was found to be weak in students in the second grade in a cross-language investigation that included five European languages varying in orthographic consistency (Ziegler et al., 2010). The influences of PA and vocabulary were modulated such that PA was a more critical component in inconsistent orthographies, whereas vocabulary was a more critical component in orthographies with consistent GPC. In contrast, the relationship between vocabulary knowledge and text reading fluency was described by Tobia and Marzocchi (2014) as particularly important, because knowing the meaning of the words could help readers understand the context and, consequently, predict subsequent words, enabling them to read faster.

The Orthographic Depth Hypothesis

The transparency of a writing system has been hypothesised to systematically influence the cognitive skills associated with reading development and to influence fundamental aspects of skilled reading. Moreover, as noted earlier, it has been shown that the level of transparency of an orthography modulates the ease with which beginning readers acquire it (Seymour et al., 2003).

Early evidence for a relationship between orthographic consistency (or depth) and lexical access was obtained in studies of word naming and lexical decision in adult skilled readers of English and Serbo-Croatian. However, to increase the range of orthographic depth examined in a single study, Frost, Katz and Bentin (1987, cited in Katz and Frost, 1992), conducted a three-way comparison of Hebrew, Serbo- Croatian, and English. The study was based on the rationale that in shallow orthographies the lexicon plays only a minor role in the naming process compared to its role in the lexical decision process. The findings revealed that in Serbo- Croatian, nonword inclusion had no effect on naming latencies, indicating the use of a phonological strategy, while much larger effects were found in English and Hebrew. This supported the hypothesis that the shallower the orthography, the greater the amount of phonological recoding that is carried out in naming. The 'orthographic depth hypothesis' thus explained different reading strategy use according to writing system transparency. According to this, the one-to-one mapping relationship between letters and sounds in transparent orthographies strongly supports the use of phonological decoding processes, while opaque orthographies encourage the visual recognition of words, due to the incorrect results that grapheme–phoneme mappings produce.

The orthographies in the study of Frost, Katz and Bentin are very similar to the languages in the current study. That is to say that Hebrew is like Arabic in that writing without diacritics is usually sufficient to indicate the exact intended (i.e., spoken) word if it is supported by a phrasal or sentential context, and Serbo-Croatian, on the other hand, is like Turkish in that each letter represents only one phoneme, and each phoneme is represented by only one letter.

To sum up, the real challenge for children learning to read in an opaque orthography is to derive the pronunciation of exception words, because the sublexical information is inconsistent (Katz & Frost, 1992). The psycholinguistic grain size hypothesis (Ziegler & Goswami, 2005) proposes one possible solution to this problem for the reader.
It was claimed that the hypothesis offered a framework for considering how different lexical, phonological, and structural factors contribute to explaining cross-language differences in reading acquisition (Goswami, 2010). According to this, readers of inconsistent orthographies such as English develop reading skills using larger grain sizes (e g., rimes) from the beginning of phonological decoding, whereas children learning to read in consistent languages develop reading skills using small grain sizes (Goswami, Ziegler, Dalton, & Schneider, 2001; Ziegler & Goswami, 2005). Thus, the hypothesis proposes that the less transparent the orthography is, the more its readers rely on developing a reliance on larger

units. Consequently, reading in English, which is highly inconsistent at the small grain level (single graphemes and clusters), is expected to lead to the dependence on larger units, such as onsets and rimes, which are less inconsistent (Treiman, Mullennix, Bijeljac-Babic & Richmond-Welty, 1995). In contrast, readers of transparent orthographies, such as Turkish, are expected to rely on smaller units, such as letters or graphemes, and still achieve high reading accuracy.

Ziegler, Perry, Ma-Wyatt, Ladner, and Schulte-Körne (2003) interpreted stronger word length effects in reading in German than English as support for the psycholinguistic grain size hypothesis. Given that the hypothesis predicts that readers of consistent orthographies rely on sublexical decoding using small units, then the number of letters should matter. Recalling that the orthographic depth hypothesis assumes that the nature of the sublexical correspondences in opaque orthographies, by definition, impedes the process of sublexical decoding, then this leads to relatively greater reliance on lexical processes in opaque compared to consistent orthographies. Consequently, readers of consistent orthographies should exhibit relatively stronger reliance on sublexical processing, which would also manifest as stronger length effects in consistent than in opaque orthographies.

Daniels and Share (2018) reported that the case of "transparent" Arabic demonstrates the inability to apply the orthographic depth approach and the psycholinguistic grain size hypothesis outside the context of European alphabets. They indicated that the entire discussion of grain size is centred on the varying spelling–sound consistency of European (Roman) alphabets, and that there was no consideration of grain size or grain sizes in non-European Roman alphabets. The inapplicability of these theories to Arabic was said to be due to sources of complexity other than orthography-phonology inconsistency, such as diglossia and orthographic factors that will be discussed in the next section.

1.1.4 Nature and Characteristics of Arabic and Turkish writing systems

As noted earlier, learning the GPC rules is insufficient to learn to read and spell accurately in English. It has an opaque orthography, with complex GPCs and complex syllable structures, sometimes composed of a string of consonant clusters, such as CCVCCC (C: consonant, V: vowel, e.g., <stands>, <trench>). Spelling in English is far more difficult than in more transparent orthographies such as French (Caravolas, Bruck & Genesse, 2003).

The orthographies involved in this study, English, Arabic and Turkish, differ in degree and nature of transparency. English can be considered the most inconsistent language, while Arabic is medium (when vowelised) and low (when unvowelized). On the other hand, Turkish is characterised by highly consistent relationships between letters and sounds in reading and spelling. Despite this consistency, and although Turkish is written with the Roman alphabet, like English, it is not among the orthographies that have been well researched. However, analogies between Turkish and other transparent orthographies enable the development of predictions. Therefore, since the research on reading development in Arabic and Turkish is so sparse, an outline of the main characteristics of the Turkish and Arabic writing systems is presented in the following sections.

Characteristics of Arabic orthography

Arabic is the most widely spoken Semitic language[1]. It is the official language of 27 countries and the fifth most common language in the world, with over 300 million speakers. It differs considerably from English, "on phonetic, phonologic, morpho-syntactic and semantic levels" (Ibrahim & Aharon-Peretz, 2005). Arabic is read and written from right to left and upper-case versions of the letters do not exist, rather all letters are written only in cursive script. The Arabic script consists of 17 characters, which, with the addition of dots that are placed above or below various letters, make up the 28 letters of its alphabet (Elbeheri & Everatt, 2007). The dots differ in their number (one, two or three) and in their position (below or above the letter). The Arabic letters, their realisation using the International Phonetic Alphabet, and English approximations are given in Appendix 1A. The graphemic form of 22 of the letters is modified according to their position within a word, that they ligate to both preceding and following letters, while the remaining six letters ligate to preceding ones only. This compulsory connecting of some letters depending on their positions (initial, medial, or final) in words, which is one of the features of Arabic, creates spaces within words. Thus, words may include unclear boundaries that cannot be distinguished by inexperienced readers (i.e., a word may contain a space that may be interpreted as the end of one word and the beginning of another).

The letters in Arabic represent consonants, but three of them serve dual purposes, acting as both long vowels and consonants as well. Beside the three long vowels, which are part of the alphabet, there are three short vowels, which are not a part of the alphabet, but instead they are marks written below or on top of consonants and sometimes over or below a

---

[1] Most scripts used to write Semitic languages are abjads. Abjad is a type of alphabetic script that omits some or all the vowels and only consonants are represented.

long vowel, known as the diacritical signs (Abu-Rabia, 2001). These diacritics vowelise words and hence enable the reader to infer the specific pronunciation of the written word, especially when the reader encounters new and non-contextual words. Accordingly, when Arabic text is fully vowelised, the grapheme-phoneme correspondence is highly regular, and it fits a transparent orthography condition. Authors such as Saiegh-Haddad and Henkin-Roitfarb (2014) argue that when Arabic text appears without the diacritical signs (unvowelized), it fits an opaque orthography condition. They suggest that while in English, orthographic opacity stems from the ambiguity or lack of systematicity in the mappings between graphemes and phonemes, in Arabic opacity stems from the absence of graphemes that represent phonological information. Such orthographic opacity "necessitates reliance in reading and spelling on large grain-size units", which are mainly lexical representations (Taha, 2016). Although Arabic unvowelized orthography is claimed to be opaque, the exception words produced from this opacity are different from exception words found in English. That is to say that, unlike English, which is inconsistent in terms of grapheme-phoneme mapping (e.g., <ea> is pronounced <ay> or <ee> or <e> in different words), Arabic, even in its unvowelized form, has consistent grapheme-phoneme mappings since the consonant letters map invariantly to specific phonemes. Accordingly, exception words in Arabic are the (unvowelized) words that may be read aloud differently (in two or more ways) due to the absence of diacritics or long vowel (letters), representing the absence of GPC information, rather than any deviation from rules.

One of the characteristics of Semitic languages is that Semitic words that are related by derivation have a common core (usually three consonants), called the root. Although the vowels in each of the different relatives may be quite different and although there may be additional consonants in a prefix or suffix, there remains an invariant series of this root. For example, the same consonantal letters of the root (k.t.b) are used to represent a number of words: <kataba> "he wrote", <kutiba> "was written", <kutub> "books", and <ka:tib> "writer". The different diacritics (representing vowels) in these words provide phonological, morpho-syntactic, and semantic information that is necessary for reading accuracy, meaning, and grammar (Saiegh-Haddad & Henkin-Roitfarb, 2014). The written texts and materials for beginning readers are usually presented with the fully vowelised form of the words, while by the end of elementary school, unvowelized words are used in the reading materials (Taha, 2016a)[2]. Table 1 shows the names, placement and sounds of the Arabic short vowels.

Table 1

Arabic Short Vowels

Short Vowel	◌ِ – تِ	◌َ- تَ	◌ُ – ثُ
Vowel name	Kasrah	Fat' hah	Dammah
Vowel placement	Bottom	Top	Top
Vowel sound	I	A	U
Equivalent English sound	/i/ in *sit*	/a/ in *bat*	/u/ in *put*

It is important to note that the definition of exception words in Arabic is thus different from the one for English words. Words are exceptional in English when the GPC deviates from the most common rules of the language. However, in Arabic, words are exceptional when the GPC information is not there. So, many (unvowelized) printed Arabic words can be read aloud differently (in two or more ways) due to the absence of diacritics, as mentioned earlier. In other words, the underrepresentation of short vowels, can result in reading certain words as other real words sharing the same consonant string. For example, the word (جَزَرْ) which means *carrot* can be read as the word (جُزُرْ) which means *islands* if the word is represented without diacritics (جزر). In short, for unvowelized words, phonological and semantic access depends on inferring from the context in which the word appears on each occasion; however, fully vowelised words do not need to rely on further resources to infer the pronunciation, as all the phonological information needed for accurate pronunciation is present (Abu Rabia, 2001).

[2] The underrepresentation of short vowels (diacritics) usually happens between the third grade of school until the end of elementary school in almost all Arabic speaking countries, and texts thereafter are written without diacritics. The participants in this study were from grades 4 and 5 where they have been used to unvowelized texts, although the measures used are vowelized.

Arabic has several versions, including Classical Arabic (CA), which is the language of the Quran, and local dialects (the colloquial language) that can differ considerably. However, there is a more common form of Arabic, Modern Standard Arabic (MSA), that is spoken by most individuals in the Arab world and that forms the basis of the writing system. MSA is now considered the official Arabic and is the direct descendent of Classical Arabic. It is used in writing and in formal speaking, for example, newspapers and in education across all Arab countries. Moreover, children are also exposed to it through many television programmes. On the other hand, colloquial Arabic is the language used for daily verbal communication. The two forms have a different structure across all language domains: phonology, morphology, semantics, and syntax (Assad & Eviatar, 2013). This makes "diglossia" an important feature of Arabic in all Arabic speaking countries. Diglossia is evident when the spoken and written languages are substantially different (Tahan, Cline & Messaoud-Galusi, 2011). This means that children read a language with which they are unfamiliar.

Although fully vowelised words refer to where every letter is given a diacritic sign, reading assessments sometimes use a partly / semi-vowelised form. This means that diacritics (short vowels) are given to the first letter or two to assure the correct pronunciation of the printed word. While vowelised Arabic appears to be orthographically shallow, in the sense of transparency of grapheme–phoneme relationships, it has other characteristics, including allography, ligaturing, and diglossia, that make it challenging to learn (Tibi & Kirby, 2018).

Characteristics of the Turkish orthography

The Turkish language is classified as an Altaic language, so it is different from Indo-European languages (like English) and Semitic languages (like Arabic). Turkish is spoken in many countries around the world; some immigrant communities also speak Turkish in non-Turkish speaking countries, such as Germany, the Netherlands, and Sweden. Essentially, the language is spoken mostly where the Ottoman Empire used to exist. In total, there are more than 100 million speakers of the Turkish language around the world, with most of those speakers, around 85 million, being native speakers. The original written form of Turkish used to be a modified version of the Arabic alphabet until 1928, when Ataturk's Turkish language reform replaced the Arabic script with the Latin script, and implemented a very systematic writing system, with an invariant one-to-one mapping between phonemes and their spellings. It is worth mentioning here that the Arabic influence on the Turkish vocabulary still exist, that more than 30% of the Turkish vocabulary is actually Arabic words written with Latin Alphabet.

In spite of the fact that Turkish and English share some features of the Latin script, such as being read from left to right, having an upper and lower case of each letter, and that it can be written in a cursive as well as print form, Turkish differs markedly from English in some aspects of phonological, morphological, and syntactic structure (Öney & Durgunoğlu, 1997). Compared with English, Turkish has a better defined syllable structure as well as a narrower band of possible syllable types (Oktay & Aktan, 2002). For example, 98 percent of all Turkish syllables belong to one of the four simple syllable forms (V, VC, CV, CVC) while the most common syllable form is CV structure covering almost 50 percent of all Turkish syllables. Also, the Turkish writing system consists of 29 letters, seven of which have been modified from their Latin originals for the phonetic requirements of the language. These are (Ç, Ğ, I, İ, Ö, Ş, and Ü).

Most importantly, Turkish differs from English in that it has a transparent writing system. In fact, and unlike many transparent alphabetic writing systems, such as German, Greek, and Dutch, where the PGC can be many-to-one, rendering spelling more complex than reading, both reading and spelling in Turkish orthography are relatively symmetrically transparent (Babayiğit & Stainthorp, 2010; Bingöl, 2003). Furthermore, young typically developing Turkish readers recognise nonwords and spell them just as accurately as words, indicating that GPC is used in these processes regardless of the lexicality of the item (Durgunoğlu & Öney, 1999; Durgunoğlu & Öney, 2000). Due to the high level of transparency, it is hard to find irregular words for reading in Turkish, except for some loan words.

Aside from loan words, the letter <ğ> otherwise referred to as <yumuşak ge> /jụmuʃak ge̥/, which is the ninth letter in the Turkish alphabet, can be a source of some irregularity in spelling, but not reading, as one of its functions is to lengthen the preceding vowel, which is normally short in Turkish. This letter cannot be the initial letter of a word but rather must follow a vowel. If the <ğ> is not recognized in dictating for spelling, <dağ> /da:/ (mountain), for example could be spelled as <da> / da/ (in). Therefore, unfamiliar words that include the silent letter <ğ> can be considered irregular in spelling.

Moreover, although a double vowel sequence is not allowed within words, vowels exist together in some loanwords. In this case, each vowel retains its individual sound (e.g., <saat> / sa:t / (hour, clock) which is originally Arabic: <sa'at>). This is particularly the case in words with syllable-initial <ğ, yumuşak ge> where VV superficially come together due to the zero realization of the <ğ> (e.g., <değil> / dejil/ (not), < mağaza> /mʌ'ɣʌzʌ/ (shop), < çoğu> /'tʃo. ɣ/ (most)). The Turkish printed letters, their realization using the International Phonetic Alphabet, and English approximations are given in Appendix 1B.

In addition to transparency, and beside the fact that there are no diphthongs in Turkish, one of the most distinctive characteristics of the Turkish language is agglutination. Agglutination is the formation of words from morphemes that retain their original forms and meanings with little change during the combination process. In other words, inflections are attached to the ends of words, and as new morphemes are attached, the syllable structure at the end of a word can be reformed. This means that adding multiple suffixes to the end of a noun or a verb may form words of infinite length. Such long inflected words may indicate grammatical functions such as number, possessiveness, preposition, or tense. Often, a single word with various suffixes can convey the meaning of an entire phrase or sentence (e.g., <you should go to sleep> in Turkish is <uyumalısın> / ṷjṷmałɯsɯn/). In addition, morphemes can change form as a function of vowel harmony (outlined below). Although agglutination can be a challenge, the rules of the Turkish language are very regular. For example, verbs usually come at the end of the sentence and the basic word order is subject-object-verb. Also, Turkish morphemes have a certain order of attachment, such as plural suffix (-ler) preceding the possessive suffix (-im), for example, ev (*home*) → evler (*homes*) → evlerim / evlɛrim/ (*my homes*) but NOT evimler (Durgunoğlu, 2006).

Turkish also differs from English in that there is no expressed grammatical gender: the pronoun (o) means 'he', 'she', or 'it'. Additionally, vowel harmony involves changing vowels in words so that the language flows more smoothly and is one of the basic concepts essential to the Turkish language, a feature unknown to English speakers, but rather a feature of all Ural-Altaic tongues. When adding suffixes in Turkish, the final vowel of the stem is looked at and the suffix that best matches this vowel is chosen. If multiple suffixes are added, each new suffix must harmonise with the suffix immediately before it. Hence, from a very young age, typically developing children hear morphemes that change one or two sounds to match the preceding vowels (Oktay & Aktan, 2002). Such a manipulation in spoken language is assumed to enable the speakers to hear individual phonemes in words more quickly.

In the following two sections, studies investigating the contribution of PA, RAN and vocabulary to early reading skill in Arabic and Turkish are discussed.

Studies investigating literacy acquisition in Arabic

Asadi, Khateb, Ibrahim and Taha (2017) conducted a cross-sectional study with Arabic speaking children (N=1305) from the first to the sixth grades, living in Israel. The researchers investigated the contribution of eight predictors to reading accuracy and reading fluency of vowelised words. The predictors were alphanumeric and pictorial RAN, PA (phoneme deletion and segmentation), and vocabulary measured by receptive and expressive tasks. The other predictors that Asadi et al. investigated were visual perception, morphological knowledge, syntactic knowledge, memory (using digit span and PSTM), and orthographic knowledge, assessed with parsing and orthographic choice tasks. The findings indicated a weak correlation (across grades) between PA and RAN. Moreover, RAN had no contribution to reading accuracy, a finding that contradicted the argument that in transparent orthographies, the contribution of RAN to reading accuracy is stronger than that of PA (Ziegler et al., 2010). However, the contribution of RAN to fluency was highly significant in almost all grades. This was consistent with results from both opaque and transparent orthographies as presented earlier.

Findings for PA showed that its contribution to accuracy was more significant than to fluency, which is also a result that fits with findings from English and other alphabetic orthographies. Although the contribution of PA in transparent orthographies was frequently reported to be more crucial in lower grades and to disappear in higher grades (e.g., Shatil & Share, 2003; Landerl & Wimmer, 2008), in Asadi et al.'s study the contribution of PA to reading accuracy increased in the higher grades (four to six). Asadi et al. discussed this finding, considering the characteristics of the Arabic orthography which forces the reader to rely on phonological information, regardless of whether the script is vowelised or unvowelized. They argued that although using the short vowels (diacritical marks) allows transparency by adding phonological information, however, it was argued to be a source of visual density and orthographic complexity. The reason is that the visual similarity between the short vowels makes their automatic perception and the processing of orthographic patterns a difficult process. This in turn necessitates the use of GPC rules to consider the punctuation marks even in higher grades in addition to diglossia.

The measures of visual processing, used in Asadi et al.'s study, were symbol search and Beery's test of visual perception. These failed to predict reading accuracy and fluency, while morphological knowledge was highly correlated with both measures in first and fourth grades (and in fifth grade to accuracy only). Also, orthographic knowledge correlated with both accuracy and fluency. The findings also revealed that phonological working memory, where the student was asked to repeat the stimulus heard after reversing the order of its sounds, showed higher correlation with accuracy than did digit span. However, memory measures did not predict fluency in the first and sixth grades. Finally, vocabulary did not contribute to reading accuracy or fluency in any of the grades. This finding fits with findings from the Hebrew orthography, where it was proposed that lexical abilities may contribute to higher processes such as reading comprehension. It also corresponds with findings of Muter et al. (2004), for the independence of vocabulary and rhyme skills. The study reinforced the hypothesis that different predictors vary in their importance depending on the characteristics of specific languages. Asadi et al. concluded that the findings indicate that Arabic readers do not rely greatly on lexical information, but rather go from the orthographic channel to pronunciation via the phonological channel. They attributed this to the presence of diglossia (see section 1.1.4.1 for details about diglossia and specific features of Arabic).

While the literacy measures in Asadi et al. did not include nonword reading, a recent cross-sectional study did involve this measure. Tibi and Kirby (2018) included nonword reading accuracy, word reading accuracy, word reading fluency and text reading fluency as measures of reading ability in Arabic. The researchers tested 201 typically developing Arabic-speaking children from Grade three in Dubai. They also assessed nonverbal ability, receptive vocabulary, RAN digits and objects, and PA (phoneme deletion, syllable deletion and word blending). Results revealed that both PA and RAN showed unique contributions to each measure. Particularly, PA had a larger unique effect than RAN for word and nonword reading accuracy, while effects of RAN were largest for word reading fluency and text fluency. These findings are in line with the findings from Asadi et al. (2017). They argued that the strength of the PA effects in their results was greater than those in some of the transparent orthography studies but consistent with previous studies in Arabic.

As for vocabulary, Tibi and Kirby found that it was a significant predictor of both word reading accuracy and fluency, a finding that contradicts what Asadi et al. (2017) reported. Tibi and Kirby stated that the correlation between vocabulary and reading comprehension was expected and attributed the discrepant findings regarding the correlation of vocabulary and word reading accuracy and fluency to the use of oral (expressive) vocabulary measures in the study of Asadi et al. They suggested that oral vocabulary may be affected by the lack of familiarity due to the difference between the children’s spoken dialect and standard written Arabic (i.e., due to diglossia). This, in turn, may hold back some children’s performance negatively.

In a third cross-sectional study, Saiegh-Haddad and Taha (2017) examined spelling as well as reading in native Arabic-speakers in Israel. Participants were 80 poor readers and 80 age matched TD children from first to fourth grade. The researchers assessed morphological awareness using tasks of derivational morphology, and PA using deletion, segmentation and blending tasks. The literacy measures involved word and nonword spelling and reading accuracy for vowelised words, unvowelized words and nonwords. In their findings, the gap between the good and poor readers revealed an early emerging morphological awareness deficit. However, PA was found to be the primary predictor of word and nonword reading in both groups. Both PA and morphological awareness significantly predicted spelling, although PA explained more unique variance in nonword spelling than in word spelling in both groups, whereas morphological awareness explained more unique variance in word spelling than in nonword spelling in both groups. Saiegh-Haddad and Taha found that while TD children showed high spelling ability from Grade one, the poor readers were only able to show this by Grade four. The researchers argued that the accurate early spelling of the TD readers was a reflection of the transparent mapping between phonemes and graphemes in Arabic. In contrast, the delayed development in spelling of the poor readers was interpreted in terms of the marked deficit in phonological skills, which affected orthographic learning.

Another Arabic cross-sectional study that included assessment of spelling in participants in an age range similar to the age of the current study was that of Taha (2016b). He used the case of the transparent Arabic to examine the assumption of whether word reading, and word spelling skills in transparent orthographies develop in an equal manner. He hypothesised that if reading and spelling of vowelised Arabic words depends only on sublexical strategies, then the development of reading skills would reach the development of spelling skills in equal and parallel manners for both beginner and advanced readers. 143 typically developing native Arab students, from second, fourth, and sixth grades were tested with reading, spelling, and orthographic decision tasks. His findings for students in Grades two and four supported the assumption. However, for students from the sixth grade he found that spelling performance exceeded reading.

In Asadi et al. (2017)'s study the researchers found that RAN had no contribution to reading accuracy while its contribution to fluency was highly significant in almost all grades. The contribution of RAN to reading fluency reached its maximum in the fourth grade. This is the age where children usually make the shift from reading vowelised scripts to reading unvowelized Arabic, hence reading vowelised material might be accompanied by high fluency due to reliance on sublexical strategies. On the other hand, the contribution of PA to accuracy was highly significant and it increased in the higher grades (four to six). Both findings fit with others from alphabetic languages and contradict observations from studies in transparent orthographies where the contribution of RAN to reading accuracy is stronger than that of PA, and the observation that the contribution of PA was more important in lower grades. One possible explanation for this deviation from findings from transparent orthographies might be that unlike in transparent orthographies, lexical processes play an essential role in later stages in Arabic readers as vowelized text is no more used. However, the sublexical strategies continue to be used in Arabic speaking students, to deal with the vowelisation signs (Taha, 2016a), which are no more used after grade four.

Findings from Tibi and Kirby (2018) were in line with the findings from Asadi et al. (2017) except in the role of vocabulary which was found to show no contribution in the first study while it contributed significantly to reading in the second study. The inconsistent result regarding vocabulary, despite that fact that the age of the participants was similar, were attributed to the effect of diglossia which might have manifested due to the use of expressive vocabulary assessments in the study of Asadi et al. In Saiegh-Haddad and Taha (2017) study, PA was found to be the primary predictor of word and nonword reading, just as it was in the two other studies, but they extended the scope to show that this applied to both typically developing and poor readers. Their findings supported the centrality of PA in reading development and disability, and in spelling.

The findings from Taha (2016b)'s study showed that reading and spelling of vowelised words in Arabic could depend only on sublexical strategies at the early stages, as they showed that they were developing equally (up to grade four). Their findings in grade six, where performance of spelling, that does not require the use of vowel signs, went better and faster than reading vowelized words, indicated that both lexical and sublexical strategies supported spelling as orthographic knowledge had become richer. This reflected that Arabic, even in its transparent form, contains other features of orthographic depth. Notably, these findings are in line with findings from Asadi et al. (2017) where PA continued to make a significant contribution to reading with older participants.

Studies investigating literacy acquisition in Turkish

Studies of reading development in Turkish are rarer than those in Arabic. One of the pioneer longitudinal studies was that of Öney and Durgunoğlu, (1997). They aimed to see if their findings would show similar patterns to those found with German readers where word decoding, and phonemic awareness were acquired rapidly due to the transparency of the writing system. They tested 30 Turkish children in Grade one in three sessions over a period of nine months (the average age for children at the start of the study was 6;04). They assessed correlations between PA, letter knowledge, word and nonword reading, syntactic awareness given as an oral task of correcting syntactic mistakes, spelling, and listening and reading comprehension. PA was assessed using phoneme deletion, phoneme blending, syllable blending, phoneme segmentation, syllable segmentation, and a rhyming task. They reported that word and nonword reading were highly correlated as were word and nonword spelling. Also, reading and spelling showed a dramatic increase in performance across the three testing points, a finding which they attributed to the transparent nature of Turkish. Additionally, they reported that the PA tasks were not correlated very strongly with each other and that the beginning readers in their study were highly successful in performing syllable segmentation while the syllable blending task was (unexpectedly) more difficult, also tasks with syllables seemed to be easier than with phonemes (as in studies with other alphabetic languages). Their findings indicated that PA contributed to the mastery of word and nonword reading and of spelling in Turkish, but that this effect was limited to the early stages of word reading and spelling as the Turkish children were at ceiling on both reading and spelling by the end of first grade.

Durgunoğlu and Öney (1999) conducted a study where they compared the development of PA across a group of 138 Turkish-speaking and English-speaking children in kindergarten and in Grade 1. The children were tested in letter knowledge, PA and a word reading task. Four measures of PA were given, syllable segmentation, phoneme segmentation, initial-phoneme deletion, and final-phoneme deletion. The children in Grade 1 in both languages did better than those in Kindergarten in PA tasks, a finding that replicated previous findings showing that as children become literate, their PA level increases. Moreover, due the consistently defined syllable structure in the Turkish language, the researchers found (as they predicted) that the Turkish-speaking group, even in kindergarten, were able to perform in syllable tasks more accurately than the English-speaking group. Also, the Turkish-speaking group performed more accurately on phoneme tasks, and they were significantly better at final phoneme deletion than at initial phoneme deletion, most probably because inflections are attached to the ends of words and follow vowel harmony (please see the section on Turkish language and orthography). The findings indicated that PA develops as a function of the characteristics of the spoken language.

A longitudinal study was conducted by Babayiğit and Stainthorp (2007) who followed 56 Turkish preschool children into Grade two. The children were assessed in nonverbal IQ, PSTM, letter knowledge, PA, word and nonword reading fluency, text reading ability, spelling, and an arithmetic task. PA was measured using syllable tapping, syllable deletion, rhyme awareness, onset awareness, initial phoneme deletion and final phoneme deletion. They reported that PA was the most powerful predictor of spelling, however, unexpectedly, it was not a significant predictor of reading fluency. Memory, on the other hand was (also unexpectedly) found to be a powerful predictor of reading fluency. In line with Öney and Durgunoğlu (1997), their findings indicated that reading accuracy was at ceiling level at the end of Grade one, suggesting that the extreme simplicity of the orthography left no room for PA to make any meaningful impact on reading accuracy by the end of Grade one.

In another longitudinal study, Babayiğit and Stainthorp (2010) investigated the role of PA (using syllable and phoneme deletion tasks), grammatical knowledge and RAN (alphanumeric and pictorial) in reading and spelling development for children followed from Grade 1 to Grade 2. The researchers controlled for the effect of IQ and expressive vocabulary and found that RAN was a powerful longitudinal predictor of reading fluency while PA as well as grammatical knowledge, predicted spelling accuracy. Due to ceiling effects in the reading accuracy measure, they dropped reading accuracy from their analyses.

In a third longitudinal study, Babayiğit and Stainthorp (2011) followed Turkish Cypriot children from Grades 2 and 4 to Grades 3 and 5. Reading fluency, spelling accuracy, reading comprehension, and narrative text writing skills were examined in 109 children. The test battery included measures of IQ, PA, alphanumeric RAN, expressive vocabulary, listening comprehension, and STM. PA measures were sound oddity, phoneme deletion, and spoonerism tasks. The results replicated the previous findings of Babayiğit and Stainthorp (2010) showing that RAN was the most powerful predictor of word reading fluency and PA was the strongest predictor of spelling skills. Vocabulary in Babayiğit and Stainthorp (2010) was not found to be a significant predictor of reading fluency or spelling, but rather predicted compositional writing. However, vocabulary in the study of Babayiğit and Stainthorp (2011), using the same expressive measure as in Babayiğit and Stainthorp (2010), was only found to predict reading comprehension. Furthermore, the results suggested that even when the spelling system is highly consistent, it is still possible to find reliable relationships between RAN and spelling, independent of PA skills, a finding that was in line with with previous findings in English (e.g., Savage, Pillay, & Melidona, 2008) and German (Landerl & Wimmer, 2008).

The two studies of Öney and Durgunoğlu (1997), and Durgunoğlu and Öney (1999), emphasized the role of PA in Turkish literacy acquisition but indicated it was limited to early years and attributed to the nature of the orthography. Turkish children used syllables easier than phonemes in the PA tasks.

In the study of Babayiğit and Stainthorp (2007), where RAN was not included in the measures, PA was not a significant predictor of reading fluency. When RAN was assessed, in Babayiğit and Stainthorp (2010, 2011)'s studies, the findings showed that RAN was the most powerful predictor of word reading fluency and PA was the strongest predictor of spelling skills, and that this relation can be observed even among older children in a highly consistent writing system. Therefore, the findings are in line with the literature on reading development in English that emphasise the role of RAN in reading fluency and that PA contributed to the mastery of word and nonword reading at early stages (e.g., Bowers & Newby-Clark, 2002; Ehri, 1997).

1.1.5 Multilingualism and reading development

This section will address considerations related to the effects of multilingualism on reading development; Particularly, linguistic transfer which is considered an important element in investigating predictors of word reading.

Bilingualism (knowing two languages) (Grosjean, 2010; Valdez & Figueora, 1994), and multilingualism (the development of an additional non-native language by individuals who speak two or more languages) (Schwartz, Geva, Share & Leikin, 2007) have become common linguistic phenomena (Mirza, Gottardo, & Chen, 2017). Research on multilingualism has been used to gain knowledge about cross-linguistic relations among languages and among literacy skills, which in turn helps to understand more about reading development. The participants in the current study are multilinguals, but the literature often refers to bilingual children with the same term. As learners who had any amount of experience with languages other than their mother tongue (no matter how little) were classified as 'multilinguals', more and more researchers started to feel encouraged to consider even small amounts of language experience in the participant profiles, which in turn contributed to the growing body of literature on multilingualism.

Two hypotheses have been the focus of research studies that examined the transfer of linguistic skills between languages: the Interdependence Hypothesis and the Script Dependent Hypothesis. The linguistic interdependence hypothesis, put forward by Cummins

(1979, 1981), assumes that the basic cognitive knowledge common to all languages allows the transfer of linguistic skills from one language to another and that learning two languages will improve the learning process of both, regardless of the orthography type. By contrast, the script-dependent hypothesis, proposed by Geva (2008), suggests that reading development in a certain language will be constrained by the orthographic features and the degree of transparency of that language. This means that children's strengths and difficulties when learning to read in a particular language do not necessarily influence their ability to read a second language (i.e., each language develops independently from another). Consequently, when children learn to read languages that differ in orthographic transparency, they must resolve two mapping problems, one for each language.

Taking both the linguistic interdependence hypothesis and the script-dependent hypothesis in consideration, the challenges faced by bilingual children likely stem from either the transfer across languages, including deficiencies in various language components or language-specific constraints which impose limitations on learning to read. Together, they provide a comprehensive understanding of learning to read two or more languages. Research on bilingualism involved studying subskills essential for reading across languages with similar alphabetic scripts (e.g., Spanish and English) (Proctor, August, Carlo, & Snow, 2006) and across different alphabetic scripts (e.g., English and Hebrew) (Abu-Rabia, 2001). Moreover, studies have been extended to examine relations across alphabetic and non-alphabetic languages (e.g., Japanese and English) (Wydell & Butterworth, 1999).

As mentioned earlier, general PA skills are thought to be learned only once (Durgunoğlu, 2002) and transfer positively across languages. Consequently, several research studies have shown associations between first language (L1) PA and second language (L2) reading abilities (Chow et al., 2005; Durgunoğlu, 2002). For example, Durgunoğlu, Nagy, and Hancin-Bhatt (1993) found that PA in Spanish, the L1 of the children tested, predicted how children learned to read new L2 English words. Similarly, Comeau, Cormie, Éric and Diane (1999) found evidence for cross-language transfer of PA skills and reading performance in children with English as L1 and French as L2. They showed that L1 PA skills transferred to L2 reading. Moreover, the researchers found that the relationship between PA in both languages and L1 and L2 reading skills was equivalent. Similar findings regarding the transfer of PA exist for diverse languages, such as Portuguese and English (DaFontoura & Siegel, 1995) and Chinese and English (Gottardo et al., 2001). However, and to not deviate from the scope of the present study, a closer look is given to studies with students speaking Arabic as it is the mother tongue of the participants of this study.

Saiegh-Haddad and Geva (2008) tested the relationship between PA and morphological awareness and how relevant they were to word and nonword reading accuracy, and complex word reading fluency in forty-three English-Arabic bilingual children in Canada. The results showed a significant correlation between PA in English and in Arabic. PA predicted reading cross-linguistically. Moreover, while both PA and morphological awareness in English predicted independent unique variance in English word reading, only PA in Arabic predicted Arabic word reading. Complex word reading fluency was predicted by morphological awareness within both languages. Similarly, in both languages, PA was the single factor predicting nonword reading accuracy. The findings coincided with the findings of others supporting the transfer of PA ability from an L1 to an L2 (Abu-Rabia & Siegel, 2003; Verhoeven, 2007; Wade-Woolley & Geva, 2000).

In addition to the direct transfer from L1 to other language(s), some studies had found that skills may also be transferable the other way round. Transferring linguistic skills from L3 to both L1 and L2 is termed "Cognitive Retroactive Transfer" (CRT). Abu-Rabia and Shakkour, (2014) conducted a study to test the CRT of language skills from L3 to L1 and L2. The participants were sixty trilingual Arabic-Hebrew-English learners from the sixth grade who were poor achievers in Arabic, Hebrew, and English, with their average grade in all three languages below 60 percent. They were divided into an experimental and a control group. The researchers examined orthographic knowledge, PA, morphological awareness, syntactic awareness, reading accuracy and reading comprehension in all three languages. The experimental group took part in an intervention in English but not in Arabic or Hebrew, while the control group received no intervention. The intervention program focused on the linguistic skills for which the students were examined in the study, which included orthographic, phonological, morphological, syntactical, reading, and reading comprehension skills. After completing the intervention program, both groups took the same series of tests in Arabic, Hebrew and English alike. The aim was to make sure that skills practiced in the program was reflected on the scores in English. Redoing the tests in Arabic and Hebrew also aimed to examine any changes in students' achievements in Arabic and/or Hebrew when compared with achievements before the English intervention program. The findings revealed a significant improvement in all language skills for English, Arabic and Hebrew in the experimental group, except there was no cross-language transfer of orthographic knowledge. The improvement in reading accuracy in Abu-Rabia and Shakkour's study was in line with findings from other studies that supported the cross-language transfer of word-identifying skills (e.g., Geva & Siegel, 2000; Morfidi, et al., 2007), and the improvement in PA. This finding also indicated that PA was not a language-specific skill, and that increased exposure and practice of the sounds of any language may help improve phonological abilities generally. The researchers concluded that their findings validated the CRT hypothesis.

The two studies discussed here indicated that both the Interdependence Hypothesis and the Script Dependent Hypothesis could be supported, depending on the skill being considered. That is, morphological awareness in the study of Saiegh-Haddad and Geva (2008) and orthographic knowledge in the study of Abu-Rabia and Shakkour (2014) seemed to be language-specific skills, supporting the Script Dependent Hypothesis. On the other hand, PA seems not to be a language-specific skill. This suggests that learning linguistic skills of a language will improve reading skills, as they are transferable between alphabetic languages (Durgunoğlu et al., 1993). This also suggests that multilingual readers showing difficulties in one writing system due to poor PA, will most likely show difficulty in the other languages acquired. In the next section I will focus on research looking at difficulties in reading and studies related to multilingualism in poor readers.

1.2 Part Two: Specific Learning Difficulties (Dyslexia)

Unlike general difficulties, specific difficulties involve disorders where there is a deficit in just one or a small number of skills, with typical functioning in other areas. Specific learning difficulty means that the difficulty occurs in a restricted domain which must be learned. Dyslexia is one of the best-known examples of a specific learning difficulty. This section will cover research looking at underlying causes of dyslexia, dyslexia in multilingual readers, subtypes of dyslexia in English, and studies investigating dyslexia in other languages.

1.2.1 The Search for Underlying Causes of Specific Learning Difficulty (Dyslexia)

According to Hulme and Snowling (2009), around 5% of children have a significant difficulty learning to read even though they have normal intelligence, have no known neurological or psychological problems, and have had normal reading instruction. This "specific" learning difficulty is often called developmental dyslexia, where dyslexia is "a disorder that is characterized by difficulties with accurate and/ or fluent word recognition and by poor spelling and decoding abilities" (IDA, 2002).

Research on developmental dyslexia attempts to identify the specific deficits that characterize the disturbance. Different theories regarding the causes of developmental dyslexia have been proposed, including cognitive, biological, and language-based deficits. Most of the research conducted on developmental dyslexia has focused mainly on the cognitive characteristics of dyslexia and its manifestation in English-speaking dyslexic children. One of the most influential theories has been the phonological deficit hypothesis. The hypothesis of the phonological deficit theory is that children with developmental dyslexia have extreme difficulties in acquiring basic reading skills such as word identification and phonological (letter-sound) decoding and that these poor readers will perform worse than normal readers, in PA, RAN, verbal learning, and verbal memory (e.g., Snowling, 2001; Joanisse et al., 2000; Boets *et al.*, 2007). More recent research has indicated alternative non-phonological underlying causes, such as deficits of visual short-term memory and visual attention (e.g., Facoetti, Lorusso, Paganoni, Cattaneo, Galli , Umilta`& Mascetti, 2003; Ram-Tsur, Faust & Zivotofsky, 2008). Before this linguistic-based hypothesis of dyslexia began to develop, the two most prevalent cognitive deficit theories were theories considering low-level auditory or visual deficits.

Nowadays, and according to the International Dyslexia Association (2002), dyslexia is defined as a specific learning disability that is neurobiological in origin. It is often unexpected in relation to other cognitive abilities and the provision of effective classroom instruction; it is not the result of generalized developmental disability or sensory impairment. Dyslexia is a disorder that generally persists throughout life; however, the presenting symptoms differ depending on the age of the person, the severity of the disorder and the availability of interventions and compensatory strategies. This lifelong persisting literacy disorder is called developmental to distinguish it from various types of acquired dyslexia, in which the reading difficulties are related to causes such as strokes, tumours, or accidents that have led to brain damage (Rahbari & LeFevre, 2015).

Given that the poor readers in the present research presented with the features of dyslexia, namely at least average intelligence, and proper opportunity to learn, this raises the possibility that their reading deficit indicates dyslexia. However, the poor readers are referred to as such in this thesis, rather than dyslexics, due to the lack of awareness about dyslexia among most Arabic speaking families, which hinders official screening. More details about the context of the current study are given in the *Present Study* section at the end of this chapter.

1.2.2 Dyslexia in multilingual children

Miller Guron and Lundberg (2003) stated that despite the widespread use of reading assessments in schools, evidence suggests an under-representation of multilingual students among those identified as having specific learning difficulties/dyslexia. One possible reason put forward for this was that practitioners assume that poor reading performance is a natural consequence of poor oral language skills. This assumption, the researchers noted, is not supported by studies, but could lead to an imbalance in the numbers of multilingual compared to monolingual individuals who are selected for assessment, and consequently low numbers of multilinguals among those diagnosed as dyslexic and referred for remedial intervention.

This paucity of studies investigating dyslexia among multilinguals is likely due to methodological difficulties such as ensuring assessment materials are of equal ease/difficulty across languages and ensuring that participants have comparable levels of proficiency in their languages. It is often difficult for researchers to show that tasks used are comparable in level of difficulty and cognitive demands across different languages (Lundberg, 2002). In addition, Cline (2000) suggested that among the major challenges facing those working in research on multilingualism and dyslexia concerns the many different dimensions of bilingualism.

Miller-Guron & Lundberg (2000) conducted a special study to closely look at reading proficiency of L2 in dyslexics and whether it could act as an exceptional alternative route to improved literacy. For this purpose, the researchers recruited three groups with ten participants in each group: a group of dyslexic young adults who preferred to read English; a second group of dyslexics with no special preference for L2 reading, and a group of typical readers matched on age group and educational level.The researchers systematically looked at the differential reading scores and compared components of L1 and L2 literacy by examining the level of PA, orthographic development, single word reading, text oral reading and reading comprehension. Findings showed a considerable degree of variance in L2 word recognition skills and reading aptitude between the two dyslexic groups. This, in turn, suggested a broad scale of L2 reading efficiency despite L1 deficits. One possible interpretation was that the Swedish writing system, may place higher demands on phonological skills than English (Lundberg 1998). Weak early phonological skills may encourage the individual, paradoxically, to develop a preference for a deeper orthography, and to employ word recognition skills based on larger orthographic structures, a technique that is less affected by phonological processing difficulties (Miller-Guron & Lundberg, 2000).

That being the case, areas of difficulty can be identified that differentiate bilingual children with reading impairment, based on knowledge about the reading process across languages, even when children have not fully acquired both of their languages (Geva, 2000). The researchers have argued that employing assessments of PA and RAN, as well as checking if there is a gap between reading and listening comprehension can lead to high levels of accuracy in identifying bilingual children with reading impairment (Miller Guron & Lundberg, 2003; Frederickson and Frith, 1998; Everatt, Smythe, Adams & Ocampo, 2000).

1.2.3 Subtypes of Developmental Dyslexia

Studying profiles of dyslexia started with reports of individuals who were previously competent readers but lost this skill due to acquired brain damage. As noted previously, this is usually referred to as acquired dyslexia. The background to subtypes of dyslexia was studies on acquired dyslexia, which indicated that reading procedures can be selectively impaired (Holmes, 1973; Marshall & Newcombe, 1973; Shallice, 1981). According to the Dual Route framework (outlined in section 1.1.1.1) a dysfunction in the nonlexical route would lead to impairment in nonword reading, but performance on familiar words could be unimpaired, as these words can still be processed by the lexical route. This pattern of performance was known as acquired phonological dyslexia. On the other hand, a selective impairment in the lexical route, would lead to poor performance on irregular words relative to regular words and nonwords. This profile was referred to as acquired surface dyslexia. The subtypes were subsequently reported among children with developmental dyslexia (e.g., Temple & Marshall, 1983; Coltheart et al., 1983). When both lexical and sub-lexical reading processes were affected, this was referred to as a mixed profile of dyslexia.

Castles and Coltheart (1993) reported a subtyping study with English-speaking children aged seven to fourteen years where they used the regression outlier technique to identify dyslexic children as having one profile or another. They tested reading aloud of regular and irregular words and nonwords in 56 typically developing (TD) readers and 53 dyslexic children. The dyslexic group and the TD group were matched in age. The relationship of age and accuracy in reading the irregular words and nonwords in the TD children was used as the basis for the classification of the dyslexic children. Ten dyslexic children exhibited the pattern of surface dyslexia (poor irregular word reading) and eight showed the pattern of phonological dyslexia (poor nonword reading). The rest of the dyslexic children showed impairment in both irregular word and nonword reading; however, they were more impaired in one than the other. Castles and Coltheart's findings supported those from earlier single case studies (e.g., Temple & Marshall, 1980) showing distinct varieties of developmental dyslexia. Subtyping studies have been repeated by different researchers, and with other orthographies than English, as outlined in the next section.

1.2.4 Specific Reading Difficulty and Subtypes of Dyslexia in Different Languages

When studying dyslexia in an orthography other than English, one major question is whether it will be possible to find the same underlying deficits or whether the level of transparency of the writing system will cause deficits to vary. Wimmer (1993) and other researchers reported that developmental dyslexia in transparent orthographies is more associated with slow than with inaccurate reading (e.g., Italian: Zoccolotti, De Luca, Di Pace, Judica, Orlandi, & Spinelli, 1999; Dutch: Van den Bos, 1998; Norwegian: Lundberg & Hoien, 1990; German: Wimmer, 1993; Spanish: Gonzalez & Valle, 2000; Hungarian: Smythe, Salter & Everatt, 2003). The view came to be prevalent that reading in a transparent orthography facilitates the development of PA, which in turn facilitates reading accuracy (Durgunoğlu & Öney, 1999). It was also argued that exposure to systematic literacy instruction in a transparent orthography helps children with low PA skills to be able to compensate for a deficit (Landerl & Wimmer, 2000; Wimmer, Mayringer & Landerl 2000). More recent research, as discussed next, has brought this unidimensional view of dyslexia in transparent orthographies into question.

Douklias, Masterson and Hanley (2009) conducted a study with Greek poor readers aged 9–12 years. Greek is transparent for reading, so irregular words do not exist, however, due to historical reasons, phoneme-to-grapheme inconsistencies exist for spelling. The researchers employed the regression outlier technique, as used by Castles and Coltheart (1993), but using word reading latency as a measure of lexical reading skill (since irregular word reading could not be used), and nonword reading accuracy as a measure of sublexical skill. Their hypothesis was that reading latency should be slow in surface dyslexic children since they have impaired lexical processes and need to rely on time-consuming sublexical processes. Douklias et al. found two poor readers with this pattern - accurate nonword reading plus slow word reading relative to controls, and two poor readers with inaccurate nonword reading plus relatively fast word reading relative to the age matched controls (equivalent of phonological dyslexia). Additional testing revealed significant difficulties with spelling irregular words but not nonwords in the children with the profile of surface dyslexia, while the profile of phonological dyslexia was associated with poor nonword spelling. In addition, the two children with the profile of phonological dyslexia exhibited worse performance in PA tasks (blending and deleting syllables and phonemes, spoonerisms) than the control children. On the other hand, the two children with the profile of surface dyslexia showed worse performance in rapid naming tasks than the control children, while the phonological dyslexics were unimpaired in rapid automatized naming tasks.

Douklias et al. speculated, in line with previous suggestions of Manis et al. (1999), that rapid automatized naming deficits and surface dyslexia may reflect the same underlying difficulty. Their findings indicated that both types of developmental dyslexia (surface and phonological) can be observed in Greek and that surface dyslexia can exist in writing systems with consistent grapheme-phoneme correspondence. Douklias et al.'s study was one of the first attempts to look for manifestations of lexical and sublexical deficits in a transparent language.

Niolaki, Terzopoulos, and Masterson (2014) conducted a study with Greek dyslexic children to replicate and extend the findings of Douklias et al. (2009). They investigated whether the phonological dyslexia profile would be associated with a selective deficit of phonological abilities, and whether surface dyslexia would be associated instead with a deficit of rapid naming. They used nonword reading accuracy as a measure of sublexical processing, and word reading latency as a measure of lexical processes, as in Douklias et al., and applied the same regression outlier technique using data from TD children. However, they included a wider range of non-literacy assessments than Douklias et al., to look at co-occurring deficits. Participants were a group of nine dyslexic children with a mean age of 9.09 years. Thirty-three TD readers from the same region in Greece were used as controls and were matched in terms of chronological age, non-verbal ability, and receptive vocabulary to the dyslexic students. The results for the dyslexic group revealed several characteristics previously reported for children with dyslexia in studies with transparent and opaque orthographies. The dyslexic children were slower than the TD children at reading both text and single words, although their word reading was relatively accurate, which is a finding that has been reported in several studies with transparent writing systems, as discussed earlier. The subtyping findings revealed that three of the dyslexic children exhibited the profile of phonological dyslexia and two children the profile of surface dyslexia. The remaining four children had a mixed profile. The findings, therefore, supported the existence of distinct types of developmental dyslexia in Greek. The results also revealed an association between developmental phonological dyslexia and a phonological deficit, and developmental surface dyslexia and a deficit of visual attention span, as assessed in a letter report task. Niolaki et al. did not find that the children with the profile of surface dyslexia had a rapid naming deficit, as reported by Douklias et al.

A third subtyping study with Greek dyslexic participants was conducted by Sotiropoulos and Hanley (2017). They assessed 34 Greek-English bilingual students attending UK universities to see whether corresponding subtypes would be evident in both the students' languages. They recruited 25 students to serve in a control group and 30 dyslexic students. They used modified t-tests to compare the individual scores of the dyslexics with the scores of the controls. Seven dyslexic students when assessed in Greek showed slow word reading and a deficit in irregular word spelling, that is, the profile of surface dyslexia in Greek as specified in the study of Douklias et al. (2009). All these students were found to have inaccurate reading and spelling of irregular words in English. Two further students when assessed in Greek showed the profile of phonological dyslexia, and they both showed impaired nonword reading and spelling in English, although they had relatively normal word reading times and relatively accurate spelling of irregular words in both languages. The findings ran counter to the view that reading difficulties are manifested solely by slow reading in transparent writing systems. Instead, they indicated that lexical and sublexical processes can be selectively impaired among individuals with literacy difficulties when the writing system is transparent, just as when it is opaque.

Dyslexia in Arabic

Research on developmental dyslexia in Arabic is sparse, however, a notable exception is the publication of a paper outlining different profiles of dyslexia in Arabic in 2014 by Friedmann and Haddad-Hanna. These authors identified seven subtypes of dyslexia in Arabic resulting from deficits in different reading processes, six of which were said to be reported for the first time. The researchers used the dual route model as their framework and distinguished between subtypes that are caused by impairment at the orthographic-visual analysis stage (see Figure 1.1, section 1.1.1.1) which they called *peripheral dyslexias*, and included letter position dyslexia, attentional dyslexia, visual dyslexia, and neglect dyslexia. Subtypes resulting from impairments at later stages, which the authors called *central dyslexias,* included surface dyslexia, phonological dyslexia, vowel dyslexia, deep dyslexia, and access to semantics dyslexia. Of these latter five subtypes, only three were evident among the participants in the Friedmann and Haddad-Hanna study, namely, surface dyslexia, vowel dyslexia which is a selective impairment in vowel processing in the sublexical route, and deep dyslexia, which is caused by a deficit in both routes, leading to semantic and morphological errors.

Friedmann and Haddad-Hanna's study included 74 Arabic-speaking dyslexic children from Israel. A group of 26 TD readers in third to fifth grade was included to provide control data. The diagnosing of developmental dyslexia was done based on the Arabic TILTAN screening test (Friedmann & Haddad-Hanna, 2009). This included 207 words, 27 non-words and 23-word pairs. The materials were selected so that the various types of dyslexia could be detected.

Phonological dyslexia was not detected in Friedmann and Haddad-Hanna's study, however, vowel dyslexia was reported. Vowel dyslexia, as noted above, involves a selective deficit in the sublexical route that affects the processing of vowels. Individuals with vowel dyslexia may omit, add, or migrate vowel letters when they read unfamiliar words or nonwords. Many of the vowel addition and omission errors were in fact shortening of a long vowel, or elongation of a short vowel. They also make short vowels long, and long vowels short in spelling. The researchers noted that the presence of this subtype was supported by some findings from Spanish, French, English, and Thai studies (Winskel, 2011), which showed that vowels and consonants are treated separately within the sublexical route. Moreover, vowel dyslexia, which was never reported in Arabic or in any other language except Hebrew before, was found to be the most frequent type of dyslexia among their participants.

Surface dyslexia, on the other hand, was assessed by using unvowelized words that, when read out of context, could be read as other real words. In this case, and due to the absence of diacritics that represent short vowels, readers must guess the appropriate vowel missing from the words. Readers with surface dyslexia either produce nonwords or they might produce a possible real word for the target provided, yet this newly produced word cannot stand as a single word on its own. For example, (فتحه) which is a name of one of the diacritics (Fat' hah), could be read as a verb attached to the pronoun (*he opened),* which cannot legally be realised without an object. The researchers reasoned that it is this sort of information retrieval that requires the involvement of the lexical route, which is a weakness in the case of surface dyslexia. The researchers made their assumptions based on the characteristics of surface dyslexia in English-speaking cases, and the fact that in Arabic irregular words result from script ambiguity, which happens only when the printed words are not vowelized. Surface dyslexia was found in nine of the participants in Friedmann and Haddad-Hanna's study.

As mentioned earlier, the third type of observed *central dyslexia* in the study was deep dyslexia. Deep dyslexia refers to a deficit in both the sublexical and the lexical route (Coltheart, Patterson, & Marshall, 1987; Stuart & Howard, 1995). This type is characterized by the production of semantic and morphological errors, better reading of nouns than verbs and adjectives, and better reading of imageable and concrete words compared to abstract words. Nonword reading is also impaired. Five of the participants in the study showed the reading pattern of developmental deep dyslexia.

Dyslexia in Turkish

Despite the expansion of research investigating types of dyslexia in different languages, almost no studies had been carried out for Turkish until Güven and Friedmann (2019) conducted a study reporting letter position dyslexia in Turkish. They aimed to examine the view, discussed in section 1.2.3, that dyslexia in a transparent writing system such as Turkish mainly affects reading fluency. Participants in the study were 24 Turkish-speaking children in 4th grade, who showed letter position dyslexia (LPD). The control group comprised 71 TD 4th graders. For initial assessment and identification of LPD they used the screening test from the FRİGÜ test battery (Güven & Friedmann, 2014). LPD is a deficit in the function that encodes the relative positions of letters within words, and results in letter migration within words, while the first and the last letters are relatively immune to migrations. LPD is manifested in reading aloud with words in which letter migration creates other existing words (such as (for English) reading *flies* as "files", *slept* as "spelt").

1.3 Part Three: The Present Study

1.3.1 Description of The Present Study

The aim of the present study was to investigate the use of lexical and sublexical processes in reading, in Arabic-Turkish-English trilingual children aged nine to eleven years. I looked at possible interrelations of lexical and sublexical processing across the languages of the children, and the patterns of relation each reading process showed regarding PA, RAN and vocabulary knowledge. The analyses were conducted first for a group of 38 typically developing (TD) children and subsequently for a group of 12 poor readers (PR). I also aimed to see whether subtypes of developmental dyslexia could be observed in the PR group and whether the manifestations of the subtypes might differ across the three languages of the children, considering the different degrees of transparency. For this purpose, results from a subgroup from the TD sample with equivalent non-verbal ability and vocabulary knowledge to the PR group were used for comparison purposes.

In the current study I was able to examine the role of the reading-related processes in literacy in three languages within a single sample of children, which has several advantages over previous research, discussed below. This section distils the main relevant findings from the previous literature and outlines the aims of the current research and how I went about addressing the aims.

As outlined in the literature review, the main predictors of early literacy skill have been shown to be letter knowledge, PA, RAN, verbal memory, vocabulary, and morphological knowledge (e.g., Carroll et al., 2003; Gillon, 2004; Landerl & Wimmer, 2008; Muter et al., 2004; Savage et al., 2007; Snowling et al., 2003). Although, as noted, the early research was carried out mostly in English, later studies looked at other languages. Comparable investigations were conducted for Arabic. Recent studies in Arabic are those of Asadi et al. (2017), Tibi and Kirby (2018), Saiegh-Haddad & Taha (2017) and Taha (2016b).

For Turkish where there is a scarcity of translated and published research, five studies have been visited, a pioneer study by Öney and Durgunoğlu (1997), another study by the same researchers Durgunoğlu and Öney (1999), and three studies conducted by Babayiğit and Stainthorp (in 2007, 2010 and 2011).

The current research, as mentioned earlier, involved investigating the patterns of relations for reading in relation to selected literacy-related skills. The literature reviewed showed that similarities exist across languages in the patterns of association of reading with LRVs at the early stages of learning to read and later in primary school. However, there still appear to be some uncertainties, and it is difficult to draw strong conclusions about the reasons for inconsistent findings when studies use different reading measures, different assessments of PA, RAN, and vocabulary and participants are drawn from different age groups. Moreover, the transparency of the language was found to affect these relations. In opaque writing systems, such as English, poor PA, for example, was found to be directly reflected in poor reading accuracy (Vellutino et al, 2004) but in transparent writing systems, such as German and Turkish, poor PA was accompanied with poor reading accuracy of nonwords (Wimmer et al, 2000). Poor PA had no association with reading time (fluency) in opaque languages, however it manifested in slow yet accurate reading in transparent languages. RAN, on the other hand, seemed to be not related to accuracy but rather to reading fluency in transparent orthographies such as Greek (Papadopoulos, Spanoudis & Georgiou, 2016). Vocabulary, which is mainly related to reading exception words due to its association with lexical abilities, was also found to be related to reading fluency in opaque languages much more than in transparent languages (Tainturier, Robers & Leek, 2011). However, the discrepancy in the role of vocabulary found across studies visited supports the need to include vocabulary in the present study.

This study also involved looking at deficits in lexical and sublexical processes and reading-related processes in Arabic-English-Turkish trilingual children who were from the same environment as the TD children. Given that different patterns of developmental dyslexia can result either from impairments in the multiple skills involved in reading, or due to the specific properties of the different writing systems in which these individuals read (Babayiğit & Stainthorp, 2011; Friedmann & Haddad-Hanna, 2014), and that multilingual readers showing difficulties in one writing system due to poor PA, will most likely show difficulty in the other languages acquired (Saiegh-Haddad and Geva, 2008), the present study aimed to shed light on the relationship between these two factors. I investigated whether sub-types of dyslexia that have previously been reported for English (e.g., by Castles & Coltheart, 1993) could be observed in the poor readers in their different languages, that represent distinctly varying degrees of transparency. On the other side, findings from Güven & Friedmann (2019)'s study were of interest to be considered while looking at the reading profiles of the PR group in Arabic and Turkish in the current study. However, there are very few studies of dyslexia in Arabic and Turkish and very rare (if any) where the children speak three languages. Likewise, studies that supported the cross-language transfer of word-identification skills (e.g., Geva & Siegel, 2000; Morfidi, et al., 2007), were also of particular interest to the scope of the present study.

1.3.2 Research Rationale, Questions and Predictions

In this study, I aimed to answer four research questions. For the first two research questions, I looked at typically developing children (TD). Following this, and to answer the last two questions, the reading profiles of the poor readers were examined in detail and in relation to the LRVs.

RQ1: What are the interrelationships in reading processes across the three languages?

For RQ1, I investigated reading accuracy and reading time for words and nonwords in the children's three languages and analysed the relationships among the lexical and sublexical processes. I predicted that there would be a relationship among all three languages for accuracy and reading time for reading nonwords since the predominant process for reading nonwords is sublexical.

Reading English regular words should show significant correlations with reading Arabic and Turkish words because Arabic and Turkish word reading should recruit sublexical processes due to writing system transparency, and similarly for regular English words. Reading regular English words, Turkish words and Arabic words is likely to recruit lexical processes as well (cf. evidence from studies of transparent writing systems such as Italian, showing lexical involvement), but perhaps reading words in Turkish and Arabic might be more highly correlated than with English regular words, due to shared vocabulary and morphology for Turkish and Arabic. Reading regular words might also significantly correlate with nonwords across all three orthographies due to the recruitment of sublexical processes for these letter string types.

As for reading English exception words, I predicted that this would rely predominantly on lexical processes and so correlations should be weak with Turkish and Arabic word reading although maybe not entirely insignificant since real word reading is thought to involve an element of sublexical processing as well as lexical; and the relation with English regular word reading should be significant due to the shared lexical component of regular and exception word reading.

RQ2 Given that the degree of transparency is different between English, Turkish and Arabic, what are the patterns of association of LRVs with reading accuracy and reading time for words and nonwords? And would the patterns of association be different across the languages?

PA2, I expected different patterns of associations of LRVs with reading accuracy and reading time for words and nonwords. Phonological awareness (as a sublexical marker) should be associated with regular word and nonword reading accuracy in English, and with word and nonword reading accuracy in Turkish. PA might also be associated with word and nonword reading accuracy in Arabic due to its transparency. I did not expect PA to be related to reading accuracy for exception words in English. On the other hand, I expected that RAN would be associated with reading times for all letter strings across the three languages. It might be associated with reading accuracy for English exception words, and Arabic real words (cf. indication in the literature that it may be a lexically related variable) but less so with English regular words and Turkish words and I did not expect it to have a relation with accuracy for nonwords in any of the three languages. Regarding the third literacy-related variable, vocabulary knowledge, I expected that it would be strongly associated with reading accuracy and reading times for both English exception word and with Arabic word reading. Vocabulary knowledge might be least strongly associated with Turkish word reading accuracy since the script is the most transparent.

RQ3 Are the deficits in literacy and reading-related variables in the PR group across their three languages predictable from RQ2?

Predominant deficits when the poor readers are considered as a group, might be as follows: For English, which is opaque with a mixture of decodable and non-decodable printed words, we might expect both RAN and vocabulary deficits to be associated with poor reading of exception words, while a PA deficit would be associated with poor nonword reading. In Arabic and Turkish on the other hand, we might expect a lexical reading deficit to be manifest in slow reading speed for words and associated with a vocabulary or RAN deficit, while poor reading of nonwords might be associated with a deficit of PA.

RQ4 Given that the most prevalent subtyping approach for reading disability was derived with English poor readers, are comparable reading profiles observed across the different languages of Arabic-English-Turkish poor readers?

Previous dyslexia research (e.g., Douklias et al., 2009) indicated underlying deficits are consistent across languages, so a sublexical deficit in English, as indexed by poor nonword reading, should be found for Turkish and Arabic, and a lexical deficit in English, as indexed by poor word reading, should also be found for Turkish and Arabic. However, the associations may well be modulated by orthography-specific features. That was shown in the TD data which indicated that Arabic and Turkish reading are strongly associated (due to morphology/vocabulary, etc.) while Arabic and English are not, and English nonword reading is associated with Turkish word and nonword reading (due to the same writing script, etc.). So, these features may affect the associations (for example a sublexical deficit in Arabic may also be apparent in Turkish but not in English). Moreover, and as studies on the transfer of skills in multilingual readers with specific learning difficulties showed unexpected discrepancies between abilities across languages of the readers, the transfer of a lexical deficit might be less consistent across languages.

Research indicates that, for English, difficulty in reading and spelling irregular words, in association with poor vocabulary and sometimes poor RAN, are features of surface dyslexia. Given that Arabic words in the present study were presented in their transparent form (almost fully vowelized), surface dyslexia in Arabic would not be detected through inaccurate word reading and spelling. It was expected that, if surface dyslexia were to be identified in English, slow reading of Arabic words, and relatively accurate nonword reading could be observed. No impairment in PA would be expected. On the other hand, if phonological dyslexia is a selective impairment of the sublexical phonological recoding route, then this should lead to inaccurate nonword reading responses, or substantially slower nonword reading responses if a lexical analogy is used to compensate, as was shown by the phonological dyslexic undergraduate student reported by Campbell and Butterworth (1985), who read *bant* correctly but took ages and said that she thought of the word Bantu and then took off the final "oo" sound. If phonological dyslexia was to be identified in English, then, phonological dyslexia in Arabic would be expected to be characterized by unimpaired word reading time, but inaccurate nonword reading and spelling. Moreover, an associated deficit of PA would be expected.

Unlike the situation in Greek (as discussed in section 1.2.4) where irregular words do not exist for reading while they do for spelling, irregular words for spelling in Turkish do not exist, except for some loan words which are challenging for spelling (section 1.1.4.2 on language features). Loan words were included in the Turkish assessments in the present study. Therefore, it was hypothesised that if surface dyslexia was identified in English, slow reading of words, relatively accurate nonword reading and inaccurate irregular-word spelling would be observed in Turkish. Also, no impairment in PA would be expected. On the other hand, if phonological dyslexia was identified in English, then developmental phonological dyslexia in Turkish was expected to manifest in unimpaired word reading time and inaccurate nonword reading. Moreover, an associated deficit of PA would be expected.

1.3.3 Details of the context of the study

Following the Arab spring in 2011, thousands of families from different Arabic speaking countries (e.g., Egypt, Libya, and Iraq), not to mention millions of Syrians and Yemenis, have settled in Turkey. Reasons may vary, yet the result is that Turkey has now become the home to the world's largest refugee population, most of which are Arabic speakers who continue to use their mother tongue language in their daily lives at home. Surprisingly, a significant portion of this population sought English Education for their children because it is the first international language. This, in turn, created the need for establishing an international school in Turkey which provides education in English, offers an internationally recognized certificate, and teaches what can keep the Arabic and Islamic identity, to meet the needs of this growing population. For that reason, the school where the current study took place was established.

The school is in the centre of the European side of Istanbul, the second capital of Turkey, in an urban upper-middle-class socioeconomic area. Seven hundred eighty students of mainly upper-middle-class SES and around thirty nationalities are enrolled in the school, accredited by an American organization, and providing the American diploma as its end of school certificate. The school, as mentioned above, teaches core subjects (i.e., Language Arts or Literature, Math, Science and Social Studies) in English and teaches two secondary subjects in Arabic. Moreover, the school teaches Turkish as its compulsory third language to all students.

1.3.4 Original contribution

While very few studies had studied trilingual students who speak Arabic, (e.g., Abu-Rabia and Siegel, 2003; Mirza et al., 2017), this study seems to be the first study of Arabic-speaking children who are poor readers and who are exposed to a third language in their school and surrounding context (Turkish). Moreover, given that theories of normal reading development and reading difficulties have been derived mainly from research on English and a few other European alphabets, this study focuses on universal predictors and language specifics in typically developing children and poor readers who are learning to read more than one language. This should lead to better understanding of normal and abnormal reading development and to answer unanswered questions beyond the scope of English research. Research into the effect of orthographic depth on reading in children has largely involved conducting monolingual studies and comparing findings across languages. This has frequently raised the problem that characteristics of the participants such as age and grade in school often differ from one study to another, or that socio-cultural characteristics of the participants may differ. Other potential differences include school system, teaching methods or the age at which children are taught to read. Thus, there was a need for a cross- language research that overcomes these difficulties. In the current study, orthographic depth is 'manipulated' within- participants, in contrast to many studies that have compared different groups of readers of different orthographies. The current study extended the previous research whereby the role of linguistic and cognitive variables was examined in three languages at the same time. In addition, the three languages were of interest due to the variation in transparency, from highly transparent Turkish, through opaque Arabic to highly opaque English.

2 Chapter Two: Method

Introduction

The methods chapter begins with details of the participants and the selection criteria by which they were chosen for the poor reader and comparison groups. Next comes an outline of the measures and assessments used. Following this is a description of the procedures used for data collection and finally, a discussion of the ethical issues that were addressed in carrying out the study.

2.1 Participants

Participants in this study were 50 native Arabic speakers aged 9:00 to 11:00 who were originally from different Arab countries. Participants were divided into two groups. Twelve children were poor readers (PR group), and 38 were typically developing readers (TD group) from the same age and grade level as the PR group. All participants were recruited from grades 4 and 5 in the same private "International"[3] school located in an urban area of Istanbul in Turkey. The school teaches core subjects (i.e., Language Arts or Literature, Math, Science and Social Studies) in English; therefore, English is considered the medium of instruction. Two secondary subjects are taught in Arabic (Arabic Language and Islamic studies). Moreover, the school teaches Turkish as its compulsory third language to all students.

As noted above, the participants were all native Arabic speakers, which might guarantee their proficiency in Arabic (as stated by the parents/carers). Moreover, they had all been learning in English medium schools since the beginning of their academic lives (according to the school records). This number of years spent in English Language schools is considered a suitable measure of their proficiency in English. As for the Turkish language, the participants' proficiency in Turkish varied depending on the amount of external exposure to the daily life requirements, which varies from one case to another. However, the students in the school are given a Turkish placement test at the beginning of each academic year. If the student scored below 40% in this test, he must be pulled out from Turkish classes and

[3] Most schools in Turkey provide instruction in the Turkish language (whether they are public schools- no fees are required- or private schools where parents / carers pay tuition fees for a prestigious or a better service). Since 2012, the 12 years of compulsory education for children in Turkey have been divided equally between the three stages of education, elementary, middle, and high. A few foreign institutions are given licenses to open a special type of private school called international schools, and they teach other languages to non-Turkish citizens or Turkish citizens who are holders of a second nationality and are enrolled as non-Turkish students.

certain activity classes to be enrolled in an intensive Turkish Language course for 20-weeks (2 trimesters). This intensive course was designed to help students who started school with limited Turkish to bridge the gap they have in the Turkish Language foundational knowledge. Hence, it allows them to merge back with their class peers, who scored above the 40% in the placement test, in one unified Turkish class and take a suitable end of year Turkish test during the third trimester of the academic year. Therefore, administering Turkish assessments for the present study was not scheduled before the third academic trimester to ensure that all students have been exposed to an acceptable level of proficiency regarding the Turkish language. Accordingly, the children participating in the study could all be considered multilinguals.

The age range of the participants (9-11) was decided concerning the age range covered by the assessments available for the skills of interest. A further reason for selecting this age range was that the fourth grade was when unvowelized scripts were used to replace vowelized ones for Arabic reading and spelling. Although all the experimental measures for Arabic were partly vowelized (i.e., almost entirely vowelized), it has been of interest to make sure that the participants were exposed to both the transparent form (the vowelized) and the opaque one (the unvowelized form) to test the effect of the orthographic depth hypothesis and to assess both the lexical and non-lexical processes which are essential theories to this present study, as mentioned in the previous chapter. A summary of the demographic information for the participants in the PR and TD groups and their mean scores in background assessments are presented in Table 2 and in Table 3 respectively.

Table 2

Summary of Gender and Chronological Age for The PR Group and The TD Group (Standard Deviations are in Parentheses)

	PR Group N=12		TD Group N=38	
	Grade 4	Grade 5	Grade 4	Grade 5
Number of students	4	8	9	29
Gender	2M, 2F	7M,1F	6M,3F	17M,12F
Age in months	114 (4.96)	126.5 (4.69)	115.4 (4.69)	127 (4.83)

Teachers teaching English in these two grade levels were asked to nominate students with poor English reading skills for the PR group and typically developing readers for the TD comparison group. The researcher was employed as the Special Educational Needs Coordinator (SENCo) in the school (as explained in the ethical consideration section). Table 3 gives a summary of background measures for both groups. Details of the assessments are given in *Materials*.

Table 3

Mean and Standard Deviations for Age and Scores in Background Assessments for The PR Group and The TD Group (Standard Deviations are in Parentheses)

	PR group	TD group
	Mean	Mean
Age in months	122.33 (7.65)	124.50 (6.85)
MAT	103.83 (11.53)	112.60 (14.83)
Digit Span	12.33 (2.39)	13.33 (2.70)
BPVS	59.75 (14.50)	80.88 (17.27)
APVT	70.42 (20.59)	51.30 (26.57)
TPVT	23.83 (15.79)	28.08 (19.47)

Note: MAT – Matrix Analogies Test of nonverbal reasoning, BPVS – British Picture Vocabulary Scale, APVT – Arabic Picture Vocabulary Test, TPVT – Turkish Picture Vocabulary Test

2.1.1 The PR Group

As noted above, teachers teaching English nominated students with reading problems (from Grades 4 and 5). The teachers referred 28 children with poor reading scores. Out of the 28 referrals, parents/carers of 14 poor readers consented to their participation in the study. Students with neurological, behavioural, sensory impairments, socio-emotional difficulties, or other difficulties such as irregular reading instruction (i.e., low school attendance) were excluded. Moreover, children were also excluded if they did not score within the average or above range in non-verbal ability as assessed by the Matrix Analogy Test (Naglieri, 1985). Correspondingly, out of the fourteen students, two students were not recruited as they did not meet the selection criteria. The remaining 12 children were included in the PR group. They included two children (both in grade 5) who had been diagnosed with developmental dyslexia through professional centres in two different countries outside Turkey. The other ten students referred by their English teachers as poor readers were screened and diagnosed by the SENCo as being at risk of dyslexia based on obtaining 70 (TOWRE, Rashotte, Torgesen and Wagner, 1999) and (CTOPP, Wagner, Torgesen & Rashotte, 1999). All 12 children were right-handed and none of them had a history of neurological disorder, or sensory impairments. Background information about these students, their original countries, and the number of years they had been in Turkey is provided in Table 4.

It is essential to consider that the participants' proficiency in Turkish was not straightforwardly indicated by the number of years they had been living in Turkey. This is because most families used the Arabic language for communicating at home, which resulted in varying degrees of proficiency in Turkish depending on the amount of external exposure. Therefore, the receptive vocabulary assessments scores in all three languages, nonverbal ability, and verbal short-term memory (STM) for each student, were also included in Table 4.

Table 4

Participant Information for The Poor Reader Group

Number	Name	Gender	Original country	Grade	# Of years in Turkey	Nonverbal ability	STM
1	YB	M	Libya	4	3	123	12
2	LT	F	Syria	4	2	95	12
3	AM	M	Yemen	4	4	99	9
4	ZD	F	Palestine	4	4	100	11
5	ASH	M	Syria	5	4	105	9
6	AA	M	Iraq	5	3	105	16
7	SA	M	Libya	5	3	103	13
8	ZI	M	Libya	5	5	99	11
9	ST	M	Syria	5	4	105	13
10	AG	M	Tunisia	5	2	129	12
11	AD	M	Syria	5	4	94	17
12	JM	F	Egypt	5	4	89	13

2.1.2 The TD and the smaller CG Groups

In addition to the PR participants, teachers teaching English at the school were asked to nominate students from Grades 4 and 5 with average reading skills in English. The total enrolment for the two grades was 145 pupils. The same exclusionary criteria were used as for the PR group. Consent to participate was received from 40 typically developing readers, two of whom were above the assigned age (9-11) and were therefore excluded. Accordingly, 38 typically developing multilingual readers from the same age and the same grade level and with the same language background as the poor readers were recruited for the TD group.

The findings from the TD group (N=38) were used to investigate the first and second research questions and those from the PR group were used to address the third and fourth research questions. Examination of the reading and LRV scores for the PR group and TD groups revealed that the scores for the former group were substantially lower than those of the TD children for the English BPVS assessment. It is the case that poor readers have been found to have lower levels of vocabulary knowledge than TD children due to less exposure to new vocabulary through reading, however, we were concerned to examine the performance of the poor readers in relation to TD children of the same level of vocabulary when it came to group level analyses. It was therefore considered important to select a subsample of the TD group so that the English vocabulary scores were not significantly different from those of the PR group. Accordingly, a smaller comparison group (N=20) was chosen from the TD group. The participants in this comparison group (CG) were selected to match the PR group based on age and BPVS scores. Measures are presented in the Results chapter for the two groups.

2.2 Materials

Two sets of materials were used in this study, background assessments and experimental assessments. In this section, background assessments and experimental (empirical) assessments are outlined.

2.2.1 Background assessments

For background data a nonverbal ability test and phonological short-term memory measure were used. In addition, receptive vocabulary tests given in all three languages) were used as an indicator of the level of the background language. Receptive vocabulary was also one of the measures used in the experimental assessments.

Non-verbal ability

The Matrix Analogy Test (MAT, Naglieri, 1985) was used as a nonverbal standardized measure of general abilities. The test is composed of figural matrix items, which is administered in a one-to-one testing setting. It uses a language-free format and contains two parallel forms consisting of 72 test items and two trial items. Participants were presented with an incomplete matrix for each item and were asked to select a tile that completes the matrix. Testing stopped when the examinee provided four consecutive incorrect responses and was usually administered in around 25-30 min. As this test aimed to measure nonverbal ability and hence very little oral output was required (only the number of the correct tile), there was no need to administer this test in all three languages. As the English language is the primary language of instruction in school, the test was administered in English only. The test was also suitable for students from different cultural and linguistic groups, which eliminated the need for an equivalent tool in other languages investigated.

Phonological short-term memory

The memory for digits subtest from the Comprehensive Test of Phonological Processing (CTOPP, Wagner, Torgesen & Rashotte, 1999) was used to assess short-term memory. The whole test comprises thirteen subtests that are appropriate for participants ranging in age from 5 years to 24 years, 11 months. The CTOPP was normed on a sample of 1656 people from all over the USA. The Memory for Digits subtest of the CTOPP contains 21 items that measure the extent to which an individual can correctly repeat a series of numbers, ranging in length from two to eight digits, presented at a rate of two per second, in the same order in which they were heard. Testing was discontinued when the ceiling criterion was reached, and for this subtest, the ceiling was to miss three test items in a row. Just as with the nonverbal ability test, the test was administered in English only as there was no need for an equivalent tool in other languages investigated and the English language was chosen because it is the primary language of instruction in the school.

Receptive Vocabulary

Unlike nonverbal ability and phonological short-term memory, which were administered in English only, receptive vocabulary was assessed in participants' three languages as it was used as a proxy measure of language level, but also as one of the literacy-related variables associated with typically developing reading and as an indicator of underlying difficulty associated with a reading deficit in the poor reader analyses.

For English, the British Picture Vocabulary Scale-II edition (BPVS, Dunn, Dunn, Whetton & Burley, 1997) was used. This test is an updated version of the Peabody Picture Vocabulary Test, PPVT, used as a screening test to measure receptive vocabulary for Standard English. Each trial consisted of four pictures, one of which was the match for the orally presented word and three distractors, and the participant was asked to point to the picture that they thought would be best corresponded to the oral word pronounced by the examiner. There were five practice items followed by 168 items covering basal to ceiling levels. Testing was terminated when the child provided eight incorrect responses in a set of twelve items. The BPVS test was administered to all participants in both groups, TD and PR.

The standardized Arabic Picture Vocabulary Test (APVT, Abu Allam & Hady, 1998) assessed receptive vocabulary in Arabic. This test is the Arabic version of the Peabody Picture Vocabulary Test and is appropriate to assess receptive vocabulary skills of 3 to 15 years for Arabic speaking learners. Administration for the Arabic version was very similar to that for the BPVS, despite some minor differences between the two versions. For example, the total number of practice items was four, not five as in the English version and testing items in the Arabic version were 184. Unlike the English test that contained 14 sets where each set had an equal number of items, there were only seven sets in the Arabic version. The items were not distributed equally between the sets, but each set had a different number of items. Testing was terminated when the child provided six incorrect responses within any eight consecutive items.

For the Turkish language, the standardized Turkish Picture Vocabulary Test Peabody Resim Kelime Testi (TPVT; Katz, Önen, Demir, & Perihan. 1974), was used. This test is a translated and adapted version of the Peabody test and is appropriate to assess receptive vocabulary skills of 2 to 12-year-old Turkish children. The test had three practice items and 100 testing items organized in 9 sets, two of which have 15 items each. The other seven sets had ten items each. However, and just as was the case in the English and the Arabic test, a word said to the child, and he was asked to show the appropriate picture out of the four pictures on each card the children were only tested over the critical range of items that matched their age. Testing was terminated when the child provided eight incorrect responses within any ten consecutive items.

For the standardization of the English test, 2571 students from 152 schools took part, while 2833 students participated in the standardization process of the Arabic test. Moreover, the norm data for the Turkish test were collected from 3755 (nationally representative), according to the test manual. For all the vocabulary assessments tests given (English, Arabic and Turkish), each correct answer was given 1 point. Then the sequence number of the child's last known word was noted (after considering the number of allowed mistakes given in the Instruction manual of each language as shown above

2.2.2 Experimental assessments

For the experimental assessments, participants were assessed in word and nonword reading (accuracy and reading time), real word spelling, phonological awareness (PA), and rapid automatized naming (RAN) for letters and digits using the following tasks.

Word and nonword reading (accuracy and time) and word spelling

The Diagnostic Test of Word Reading Processes (DTWRP; FRLL, 2012) was employed for the assessment of word and nonword reading in English. The DTWRP consists of three reading subtasks: nonword reading, exception word reading and regular word reading. Each subtest was presented in a separate presentation using PowerPoint software. There were 30 items in each subtest, and items become increasingly longer and orthographically complex throughout each subtest. In constructing the test, the authors matched regular and exception words for printed word frequency using the Children's Printed Word Database (Masterson, Stuart, Dixon, Lovejoy, & Lovejoy, 2003). The nonwords were constructed by combining fragments of the regular and irregular words to ensure that orthographic familiarity was similar for words and nonwords (e.g., the nonword *wem* is a combination of the words *well* and *them*). The test provides an overall reading standard score and stanine scores (M = 5, SD = 2) for the three subtests.

For Arabic, words and nonwords were taken from The Test of Reading for elementary school children in Arabic (Haitham Taha, 2009). It consists of real words and nonwords subtests. The real word test consists of 15 simple words and 15 complex words (which were longer and less familiar than the simple words). The words used were partially vowelized - this serves to narrow the possible words that can be produced from the string of letters, as explained in the section on characteristics of the Arabic orthography (Chapter 1). Words were graded in difficulty and were presented in using the PowerPoint software (as for the English assessment). Participants were asked to read each word aloud. The list of nonwords also contained 30 items partially vowelized, varying in length and phonological complexity (15 simple and 15 complex) and was also presented through a PowerPoint presentation. The nonwords were orthographically comparable to words in aspects such as length and familiarity of at least one part or chunk in each word.

The Kelime Okuma Bilgisi Testi'nin (KOBT, Babür, Haznedar, Erçetin, Özerman & Çekerek, 2012) was used for reading words and nonwords in Turkish. The Turkish adaptation of the Test of Word Reading Efficiency (TOWRE-2, Torgesen, Wagner, & Rashotte, 2012) is considered a suitable tool to assess reading among children aged 6-11. The KOBIT was administered to more than 300 elementary school students (grades 1-5) in Istanbul. The results based on 283 students' data was used for the standardization. The subtests were presented in two separate PowerPoint presentations, just as the case in English and Arabic. To make the assessment comparable with the assessments used in the other two languages, the first 30 words from the list of the real words and the first 30 items from the nonword list, from the available 40 words in each list, were used.

Both accuracy and reading times were recorded for word and nonword reading in the three languages. Reading time rather than reading fluency was used as some students read lists of words provided in less than a minute, and reading fluency is usually measured by finding the number of words read correctly per minute.

To obtain reading time for words or nonwords in each list that had been correctly read, the Audacity software program (available at http://audaci- tyteam.org/) was used from the onset of the word on the computer screen to the end of the audio-recording to detect the sound wave recording for each word and thereby record the reading time in milliseconds (see *Procedure* section for calculation details). The reading time values used in the study were the average reading time per word in the list.

The assessments used in the three languages for reading were also employed for spelling, but with at least an eight-week gap before the spelling assessment was administered (please see *Procedure* section). It is worth noticing that sound-to-spelling correspondences are not necessarily the opposite of spelling-to-sound correspondences in English. For example, the regular GPC of EA is /i:/ (as in *tea*), the regular GPC of EE is /i:/ (as in *tee*), the regular GPC of IE is /i:/ (as in *thief*), and the regular GPC of E_E is /i:/ (as in *theme*). These words are all regular for reading, but they can't all be regular for spelling! This means that some regular words for reading may not necessarily be regular for spelling.

Phonological awareness

Phonological awareness (PA) was assessed with elision (deletion) tasks and blending for each of the children's three languages. The elision and blending subtest from the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 1999) assessed PA in English. The participant's score consisted of correctly blended or omitted words (maximum 20 per subtest). The 20-item Elision subtest requires participants to omit specified segments from words. The items were presented in increasing order of difficulty. They demanded the omission of a component word from a compound word (e.g., popcorn, omit /pop/), the omission of a syllable (e.g., spider, omit /de/), and the omission of a phoneme (e.g., farm, omit /f/). Testing was terminated when the participant made a specified number of errors (namely, three mistakes in a row). The blending subtest also consists of 20 trials. On each trial, the child was presented with a string of phonemes via audiotape and was asked to blend these to form a real word. As the test proceeded, the items became longer and contained more consonant clusters. The ceiling of this subtest was also achieved when the participant made three successive errors.

For the Arabic language, the deletion and blending subtests from the unstandardized Test of Reading for elementary school children in Arabic (Haitham Taha, 2009) were used. The test originally contained 30 trials; however, only the first 20 items were used in both subtests (elision and blending) to balance with the number of words given in the English test. In other words, children were asked to omit a component word, a syllable, and a phoneme from the first 20 words and blend 20 words to get a raw score of 20, as it was in English and Turkish. The number of items pronounced correctly by the child was used as the measure for this task.

Given the lack of standardized tests of phonological awareness in Turkish (Babayiğit & Stainthorp, 2011), the study used unstandardized available assessments of phonological processing skills for Turkish. The unstandardized Turkish version of CTOPP, Kapsamlı Fonolojik Farkındalık Testleri (KFFT; Babür et al., 2013) was used to assess PA in Turkish through the Elision and the Blending subtests which followed the English tasks in administration, difficulty, and the number of items (as stated in the manual).

As both the Arabic and the Turkish versions were unstandardized tests of phonological awareness, the raw scores were used throughout all tests in all three languages.

Rapid Automatized Naming (RAN)

The digit and letters RAN subtests presented in all three languages were used to assess RAN for this study. For English, the rapid automatized naming subtests for digits and letters from the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 1999) were used. The stimuli for each task were presented on two pages, each page has an array of four rows, and nine columns of twelve randomly ordered digits or letters.

The Arabic version of the English Phonological Awareness Battery (PhAB, Taibah & Elbeheri, 2011) was employed to assess RAN in letters in Arabic; however, a RAN task, for Arabic digits, was devised for this study by translating the English version of the digit used into Arabic numbers.

The Turkish letters and digits subtests from the Turkish version of Rapid Naming from CTOPP, *Türkçe Hızlı Otomatik İsimlendirme (HOTI;* Bakır and Babür, 2009) were employed to assess RAN in Turkish. The stimuli for each task, in all three languages, were presented on two pages, each page had an array of four rows, and nine columns of twelve randomly ordered digits or letters. At the start of each task, participants (individually) were shown the individual items (digits or letters) on a separate sheet and were asked to name them to ensure that they were familiar with the stimuli. The digits in all three tests (English, Arabic and Turkish) were the same but written in the relevant script for the three versions of the digit subtest.

The participants were asked to name the test items as rapidly as possible, starting from the top row at the left for English, Turkish and Arabic digits, and top row on the right for Arabic letters (to emulate the right-to-left scanning employed for reading Arabic texts, in Arabic letters). The time to name the arrays on both pages (in seconds) was recorded for the digit, and letters RAN tasks separately. The total time taken to name items in each subtest was recorded. Errors were infrequent in the task and were not reported.

2.3 Procedures

Testing began after ethical approval was obtained and after consent forms were received from parents/carers. It lasted for a period of around 20 weeks. Participants were tested individually in a quiet room in their school, except for the spelling to dictation assessments where they were tested in a class setting. The assessments were administered in 10 separate sessions for each participant individually; each lasted between 10- to 30- minutes.

The background measures that were administered only in English (namely nonverbal ability and digit span) were administered in the first session, word reading, and vocabulary measures were administered in the following three sessions (one for each of the tested languages -English, Arabic and Turkish), the phonological measures (deletion, blending and RAN) were administered in sessions 5-7 with each language being assigned a separate session. Then, spelling to dictation was administered in the last three (group-administered) sessions. Task order was counterbalanced within sessions. The order of presentation of the assessments is provided in Table 5.

Table 5

Order of Assessments Presented for Each Participant

Session	Assessment
1st	Matrix Analogies Test and Digit Span (assessed in English)
2nd	English word and Nonword Reading and BPVT (English)
3rd	Arabic Word and Nonword Reading, APVT (Arabic)
4th	Turkish Word and Nonword Reading, TPVT (Turkish)
5th	English PA and RAN
6th	Arabic PA and RAN
7th	Turkish PA and RAN
8th -10th	Spelling (first English, then Arabic, then Turkish)

Testing began in the English language as it is the language of instruction. The DTWRP items and subtests were administered in a fixed order, with nonwords followed by exception words and then regular words, as in the instruction manual. The researcher administered testing reading in Arabic, as an Arabic native speaker and nonwords were presented first, followed by real words, as in the instruction manual and the case with the English items. However, the real words in Turkish were given before the nonwords, as in the instruction manual, and the testing was carried out in the last month of school to assure that all students were exposed to an adequate amount of the language even if this had been their first year to learn the language (see *Participants* section for more details). Moreover, the testing in Turkish tasks was administered by a Turkish native speaker.

Word and nonword reading times and reading accuracy were assessed for each participant and for each language separately. For this purpose, the items from the different word reading tests were presented on an Apple Mac PowerBook G3 computer, employing Microsoft Office PowerPoint software. The words/nonwords were presented in lower case in font size 44. The experimenter controlled the presentation of items by pressing a key on the keyboard. Participants were asked to read aloud the items as quickly and accurately as possible, starting with nonwords, then exception words and finally regular words, to follow the instruction manual. The reading responses were transcribed during the testing sessions and were also recorded using an iPhone 7 voice memos application, and the voice recordings were transferred to the Apple Mac PowerBook G3 computer and saved as MP3 files to be used for measuring reading times later.

To obtain reading times, procedures like those followed by Sotiropoulos and Hanley (2017) were used with version 2.2.2 of Audacity recording and editing software (Audacity team, 2018). Each slide in each PowerPoint presentation presenting a printed word/nonword on the computer screen was accompanied by an auditory tone (a short beep) that was used to indicate the onset of a new stimulus in the presentation. Reading time was the time (in milliseconds) recorded from the beep until the soundwave was completed, appearing as a soundwave signal. The time for all correct responses was calculated (in milliseconds) then an average time per word was used.

For the phoneme and syllable deletion tasks, children were presented with a spoken word and were asked to delete either the initial or the last sound or syllable and to pronounce the remaining part. Each of these tests consisted of 20 two-, three-, and four-phoneme/syllable words.

For the RAN tasks, the participants were first shown the individual items (digits or letters) on a separate sheet and were asked to name them to ensure that they were familiar with the stimuli. Then the participants were asked to name the test items that were presented on two pages as rapidly as possible. The time to name the arrays on both pages (in seconds) was recorded for the digit, and letters RAN tasks separately. There was no ceiling on this subtest.

As for spelling, all participants were gathered in one class in three different sessions, one for testing spelling for each language. A native Turkish speaker dictated the Turkish words and nonwords. For spelling assessment in each of the three languages, the experimenter read the items aloud. In the case of the real words, the tester repeated them in the context of a sentence (e.g., "picnic, every weekend we go on a picnic"). In the case of spelling nonwords (pseudowords), the children were instructed to spell each word just as they hear it. A pseudoword was considered "correct" if the response was phonologically plausible.

2.4 Design and data analysis

The study involved group analyses of reading and reading-related measures in typically developing readers and poor readers, as well as an investigation of the profiles of individual poor readers (multiple case studies, McCloskey & Caramazza, 1988). Studies in subtyping of literacy profiles that have looked for associations of cognitive/linguistic deficits with the profiles have employed different designs. These have involved using single case studies (e.g. Goulandris & Snowling, 1991; Hanley, Hastie, & Kay, 1992; Kohnen, Nickels, Brunsdon, &

Coltheart, 2008; Romani, Ward & Olson, 1999; Valdois, Bidet-Ildei, Lassus-Sangosse, Reilhac, N'guyen- Morel, Guinet, & Orliaguet, 2011), multiple case studies with a few participants, (e.g. Hanley & Gard, 1995, N=2; Rowse, & Wilshire, 2007, N=2; Valdois et al., 2003, N=2; Wang, Nickels, & Castles, 2015, N=2; Douklias et al., 2009, N=4; Niolaki et al., 2014, N=9), and group studies (e.g. Castles & Coltheart, 1993; Manis et al., 1996; McArthur et al., 2013; Sprenger-Charolles et 2000).

Different statistical techniques have been used in the case of each approach. For example, in the case of single case studies, McNemar's test has been used when comparison of performance in categories of stimuli are of interest; assessing for impaired performance relative to the scores of a control group has been conducted using z scores or regression (Sprenger-Charolles & Serniclaes, 2003), or modified t-tests (as recommended by Crawford & Howell, 1998). For multiple case studies, the same techniques have been used. Still, a series of single cases are reported in the same study, while in group studies, dyslexics were subtyped and assigned to groups, with analyses of potential cognitive deficits conducted with statistical techniques suitable for group comparisons, such as ANOVA or between-groups t-tests.

Z scores were employed in the present study to indicate the extent of deficit in the reading and reading-related assessments for the individual poor readers. This technique was used with word and nonword reading measures for each child in the PR group and for each language separately to obtain profiles for each of the child's languages. Following that, z scores for PA, RAN, and vocabulary were examined to see whether deficits in these variables were associated with the profiles of reading.

2.5 Ethical Considerations

Ethical approval for the study was obtained from the Ethics Committee of the UCL Institute of Education. The UCL REC ethical guidelines were adhered to throughout the research, particularly emphasizing respondents' voluntary nature of participation and their right to withdraw at any time.

After the approval from the university was obtained, permission to conduct the study in the school was obtained from school staff and from the school administration. Then, teachers were asked to refer students with poor performance in reading in English to the researcher. During this time, and as the study involves participants aged 9 to 11 years old who are thus potentially vulnerable, all the parents/carers of students enrolled in grades 4 and

5 were handed an approach letter and a consent including a description of the research to make sure that the participants and their parents/carers are given ample opportunity to understand the nature, purpose, and anticipated outcomes of participation in the research. The guarantee of anonymity and confidentiality were confirmed and their right to withdraw their child from the study at any stage. Meeting the researcher for clarifications was also offered and was taken by 13 of the parents/ carers, who met the researcher in person and got information about the study and were assured that the records would be confidential and that if the study is published, it will not be identifiable as theirs. In addition to this, participants were promised to be given information about the findings in general upon their request, which is different from the information debriefed and given to all anticipant participants. Accordingly, the recruitment of participants was restricted to only those who had signed the consent form and who approved participation. Testing both the typically developing children and children with reading difficulties began after parent/carer consent was obtained.

Furthermore, and even though children were assessed in familiar surroundings at their school with assessments that are of a format that is used in their daily routines to assure that the procedures should not cause stress or anxiety, there was still a possibility that some poor readers may have trouble in progression with some assessments. However, the researcher, an experienced SEN specialist, was ready to help, and she constantly reassured every participant in the study. The researcher was trained to be alert to any signs of potential distress and was ready to handle any situation if children experienced difficulties and was ready, for instance, to terminate assessments if the respondents showed any distress. Still, none of this took place during assessments.

The relation of the researcher with the institution was also considered. Because the researcher was employed as SENCo at the school, the parents/carers were provided with a clear explanation, in the information letter, that the researcher was neither a teacher at their child's grade level nor would unwillingness to participate have any consequence for their child's situation in school. Data storage and security of information were also thought about carefully and considered the researcher's responsibility, both during and after the research (including transfer, sharing, encryption, and protection).

3 Chapter 3: Results for Typically Developing Readers (TD)

Introduction

This chapter reports the findings for a group of 38 Arabic-English-Turkish trilingual typically developing (TD) children in reading and reading-related assessments, and in spelling. The first section of the chapter focuses on reading. First, the scores for word and nonword reading accuracy, word and nonword reading time, and the scores in the literacy-related assessments of PA, RAN and vocabulary knowledge for English, Arabic and Turkish are presented separately. Following this, a summary of the scores for the reading and the literacy-related variables (LRVs) for the three languages is presented for comparison purposes. Standardised scores, as well as raw scores, are provided for each of the assessments where these were available. In the final part of the reading section, a set of correlation analyses is reported. These were carried out with the aim of examining two issues, whether reading processes may be inter-related across the children’s languages, and whether patterns of association of reading with the LRVs might differ across the languages.

The second section of the chapter reports the results for spelling. As for the reading results, the spelling scores for all three languages are presented separately, and then a summary of the scores across the languages is presented for comparison purposes.

Before conducting inferential statistics, exploratory data analyses (EDA) were carried out, and data distributions were examined for normality. Nonparametric analyses were conducted when the assumption of normality was violated.

3.1 Reading and literacy-related results

As noted above, the scores obtained from the reading assessments and from assessments of LRVs are presented in the first section of this chapter.

3.1.1 Scores for the three languages separately

Results for word and nonword reading tasks (in terms of accuracy and reading times) and for assessments of LRVs (PA, RAN and vocabulary knowledge) are presented for each of the three languages, English, Arabic, and Turkish separately.

English assessments

Reading words and nonwords

Tests of normality revealed that most of the scores deviated from normality, except for: Eng nonword reading accuracy (skewness-.327, Shapiro-Wilk: *p=.194*), English RAN digits (skewness .612, Shapiro-Wilk: *p=.293*), English vocabulary (skewness .094, Shapiro-Wilk: *p=.718*), Arabic RAN letters (skewness -.059, Shapiro-Wilk: *p=.077*), Arabic RAN digits (skewness .243, Shapiro-Wilk: *p=.531*), Arabic vocabulary (skewness .334, Shapiro-Wilk: *p=.130*), Turkish word reading time (skewness .443, Shapiro-Wilk: *p=.294*), and Turkish nonword reading time (skewness .070 , Shapiro-Wilk: *p=.395*).

The results for reading accuracy with regular word, exception word and nonword subtests of the *Diagnostic Test of Word Reading Processes* (*DTWRP*) are summarized in Table 6. Stanine equivalents for the raw scores in the three subtests were obtained for each child from the *DTWRP* manual, and they are summarised in the second column of Table 6. Mean reading times in milliseconds for the three types of printed letter string are also presented in the table.

Table 6

Mean Accuracy and Stanine Scores and Reading Times for English Words and Nonwords for the TD Children (standard deviations are in parentheses)

Measures	Accuracy (Max.= 30)	Stanines	Reading times (msec.)
Regular words	28.13 (1.96)	5.92 (1.53)	1760 (544)
Exception words	24.97 (3.13)	5.34 (1.65)	1554 (495)
Nonwords	25.03 (2.91)	6.39 (1.39)	2132 (459)

To test for differences in accuracy and reading times between the different types of English letter string, Wilcoxon Signed Ranks tests were conducted. The difference between regular word and exception word accuracy was significant z=-4.870, p<.001, and the difference between regular word and nonword accuracy was significant, z =5.114, p<.001, in both cases in favour of regular words. The difference between exception word and nonword accuracy was not significant, z=-.352, p=.725. For reading times, the difference between regular words and exception words and the difference between exception words and nonwords were both significant, z=-3.459, p=.001 and z=-4.749, p<.001, respectively. In both cases exception word reading times were shorter. The difference between regular words and nonwords was also significant, z=-4.677, p<.001, with reading times for regular words shorter.

PA, RAN and vocabulary knowledge

The results for the PA, RAN and vocabulary tasks are summarised in Table 7. The scores for deletion and blending are in terms of accuracy, and those for RAN of letters and digits are in terms of seconds to complete the task. Subtest scaled scores (mean of 10 and a standard deviation of 3) were obtained for the phonological ability assessments from the *CTOPP* manual (Wagner et al., 1999) and are summarised in the table. For the vocabulary assessment, mean standard scores (with mean of 100 and standard deviation of 15) were obtained from the *BPVS* manual (Dunn et al., 1997) and are summarised in the table. The range for the raw scores obtained within the sample is also reported.

Table 7

Mean Raw Scores and Standard Scores for English Language PA, RAN and Vocabulary Assessments for the TD Children (Standard Deviations are in Parentheses)

Measures	Accuracy	Scaled / standard scores
Deletion (Max = 20)	18.00 (1.80)	12.21[a] (1.77)
Blending (Max = 20)	18.61 (1.24)	13.63[a] (1.60)
RAN Letters (Time in sec)	37.58 (7.67)	9.53[a] (2.10)
RAN Digits (Time in sec)	32.95 (7.18)	10.16[a] (2.46)
BPVS (Range 46-109)	80.18 (16.75)	83.03[b] (14.54)

Note: [a] subtest scaled scores from the *CTOPP* (Wagner et al., 1999), [b] standard scores from the *BPVS* (Dunn et al., 1997).

Difference tests were not conducted for the assessments of phonological abilities since the tasks are not comparable, however it can be seen from Table 7 that scores in the blending task were comparable to those in the deletion task, and time for RAN letters was slightly longer than for RAN digits. Scores in the BPVS produced an overall standard score indicating low average performance regarding the standardisation norms.

The BPVS provides a technical supplement with normative data for pupils with English as an Additional Language (EAL), however, these scores are only provided for children up to age 8:05, which is lower than the age of the participants in the current study. The technical supplement shows that the performance of the BPVS standardization sample was superior to that of the EAL sample - the difference was 10 months for the youngest group, almost 1.5 years for the Year 1 group (5-6 years old) and nearly two years for the oldest (Year 3, 7-8 years old) children in the sample. Thus, the performance of the children in the current study, considered in the light of the standardisation data, might be within the average range of performance for EAL children (the mean raw score in the present sample was 80 which ranged between 8:05 to 9:05 as the equivalent age under the 68% confidence band).

Arabic Assessments

Reading words and nonwords

The results for reading accuracy and reading times for words and nonwords are summarized in Table 8. Standard scores are not available for the Arabic reading tasks, which were adapted from the *Test of Reading for Elementary School Children in Arabic* (Taha, 2009) to meet the needs of the present study (see *Materials* in Chapter 2 for details).

Table 8

Mean Scores for Reading Accuracy and Reading Times for Arabic Words and Nonwords for the TD Children (Standard Deviations are in Parentheses)

Measures	Accuracy (Max. = 30)	Reading times (Msec)

Words	22.92 (5.66)	2409 (993)
Nonwords	19.26 (6.85)	2812 (1139)

Results of Wilcoxon Signed Ranks tests revealed that words were read significantly more accurately than nonwords, $z=-3.905$, $p<.001$, and reading times for words were significantly shorter than those for nonwords, $z=-4.285$, $p<.001$.

PA, RAN and vocabulary knowledge

The results for the PA, RAN and vocabulary tasks in Arabic are summarised in Table 9. The scores for deletion and blending are in terms of accuracy, and those for rapid naming of letters and digits are in terms of seconds to complete the task. Standard scores were not available for the PA and RAN assessments. For the vocabulary assessment, standard scores were obtained from the *APVT* manual (Abu et al., 1998) and are summarised in the table. The range for the raw scores obtained within the sample is also reported.

Table 9

Mean Raw Scores and Standard Scores for Arabic Language PA, RAN and Vocabulary Assessments for the TD Children (Standard Deviations are in Parentheses)

Measures	Accuracy	Standard score
Deletion (max =20)	18.32 [a] (1.74)	-
Blending (Max =20)	19.18 [a] (1.25)	-
RAN Letters (Time in sec)	62.03 [b] (15.85)	-
RAN Digits (Time in sec)	88.21 [b] (30.70)	-
APVT (Range 12-101)	53.66 (25.10)	95.92 [c] (24.82)

Note: [a] subtest raw scores from the *Test of Reading for Elementary School Children in Arabic* (Taha, 2009), [b] *PhAB*, Taibah & Elbeheri, 2011), [c] standard scores from the *APVT* (Abu et al., 1998).

The results in Table 9 reveal that scores in the blending task were marginally more accurate than in the deletion task, and time for RAN letters was shorter than for RAN digits. Scores in the vocabulary task produced an overall standard score in the average range. Normative data were not available for the *APVT* for children with additional languages.

Turkish Assessments

Reading words and nonwords

The results for reading accuracy and reading times for words and nonwords are summarized in Table 10. Standard scores were not available for the Turkish reading tasks, which were adapted from the KOBT (Babür, Haznedar, Erçetin, Özerman & Çekerek, 2012) to meet the needs of the present study (see *Materials* in Chapter 2 for details).

Table 10

Mean Scores for Reading Accuracy and Reading Times for Turkish Words and Nonwords for The TD Children (Standard Deviations are in Parentheses)

Measures	Accuracy (Max. = 30)	Reading times (Msec)
Words	25.71 (4.29)	2107 (534)
Nonwords	22.66 (5.62)	2478 (501)

Results of Wilcoxon Signed Ranks tests revealed that words were read significantly more accurately than nonwords, z=-4.226, p<.001, and reading times for words were significantly shorter than those for nonwords, z=-3.901, p<.001.

PA, RAN and vocabulary knowledge

The results for the PA, RAN and vocabulary tasks are summarised in Table 11. The scores for deletion and blending are in terms of accuracy, and those for rapid naming of letters and digits are in terms of time in seconds to complete the task. Standard scores were not available for these measures. The standard scores for the vocabulary assessment were obtained from the *TPVT* manual (Katz et al., 1974). The range for the raw scores obtained within the sample is also reported.

Table 11

Mean Raw Scores and Standard Scores for Turkish Language PA, RAN and Vocabulary Assessments for The TD Children (Standard Deviations are in Parentheses)

Measures	Accuracy	Standard score
Deletion (Max= 20)	18.50 [a] (1.59)	-
Blending (Max= 20)	18.05 [a] (2.01)	-
RAN letters (Time in sec)	32.45 [b] (8.59)	-
RAN Digits (Time in sec)	52.45 [b] (15.00)	-
TPVT (Range 9-81)	28.55 (19.81)	68.92 [c] (31.09)

Note: [a] subtest raw scores from (*KFFT*; Babür et al., 2013), [b] *(HOTI;* Bakır & Babür, 2009), [c] standard scores from the *TPVT* (Katz et al., 1974).

The results revealed that scores in the blending and deletion tasks were comparable, and time for RAN letters was shorter than for RAN digits. Vocabulary standard scores were below the low average range. Normative data were not available for children with additional languages for the *TPVT*.

Comparison of scores in reading and literacy-related assessments across all three languages

In this section, a summary of the reading and the literacy-related measures is presented for the TD group across the three languages.

Reading accuracy and reading times measures

In Table 12 a summary of accuracy scores for reading for the three languages is presented as percentages to facilitate comparison. Mean reading times in msecs are also given.

Table 12

Mean Percentage Reading Accuracy Scores and Mean Reading Times for The TD Children for All Three Languages (Standard Deviations are in Parentheses)

Measures	English	Arabic	Turkish
Regular Accuracy	93.77 (6.54)	76.40 (18.86)	85.70 (14.31)
Exception Accuracy	83.25 (10.44)	-	-
Nonword Accuracy	83.42 (9.70)	64.21 (22.84)	75.53 (18.72)
Regular reading times (msec)	1760 (544)	2409 (993)	2107 (534)
Exception reading times (msec)	1554 (495)	-	-
Nonword reading times (msec)	2132 (459)	2812 (1139)	2478 (501)

Regarding the accuracy of reading words across the three languages, the results revealed that English regular words were read most accurately, and Arabic words were read least accurately. For nonwords, English nonwords were read most accurately, with Arabic nonwords read least accurately. Regarding reading times, results revealed that English exception words had the shortest times and Arabic words had the longest. Reading times for English nonwords were shortest and Arabic the longest.

PA, RAN measures and vocabulary knowledge

In Table 13 percentage accuracy scores for the PA measures and average times for the RAN tasks across all three languages are given. Vocabulary assessment scores are also summarised in the table. Due to differences in length of the vocabulary assessments across the languages, standard scores are reported for this measure.

Table 13

Mean Percentage Accuracy Scores for PA Tasks, Mean Time for RAN Tasks, and Standard Scores for Vocabulary for The TD Children for All Three Languages (Standard Deviations are in Parentheses)

Measures	English	Arabic	Turkish
Deletion	90.00 (9.01)	91.58 (8.71)	92.50 (7.95)
Blending	93.03 (6.21)	95.92 (6.24)	90.26 (10.06)
RAN Letters (time in secs)	37.58 (7.67)	62.03 (15.85)	32.45 (8.59)
RAN Digits (time in secs)	32.95 (7.18)	88.21 (30.70)	52.45 (15.00)
Vocabulary (Standard scores)	83.03[a] (14.54)	95.92[b] (24.82)	68.92[c] (31.09)

[a] *BPVS* (Dunn et al.1997), [b] *APVT* (Abu et al., 1998) [c] *TPVT (*Katz et al., 1974)

The results for PA revealed a high level of accuracy. Across languages, scores were comparable in deletion and close to ceiling for Arabic in blending. For RAN, the times were faster for digits than letters for English, but they were faster for letters than digits for Arabic and Turkish. Vocabulary assessments indicated low average scores for English, average for Arabic and below average for Turkish.

Interim summary for reading scores

The results for reading accuracy, for the different types of letter string for each language, revealed that for English, regular words were read more accurately than exception words and nonwords. The difference between exception word and nonword accuracy was not significant. For Arabic and Turkish, words were read more accurately than nonwords in both languages. When the performance in reading accuracy for real words was compared across the three languages, English regular words were most accurate, while Arabic words were least accurate. For nonwords, English accuracy was highest, while Arabic nonwords were read least accurately.

For reading times, scores showed that for English, the difference between regular words and exception words and the difference between exception words and nonwords were both significant. In both cases exception word reading times were shorter. The difference between regular words and nonwords was also significant, with reading times shorter for regular words than nonwords. For Arabic and Turkish, reading times for words were significantly shorter than those for nonwords. When reading times for words was compared across the three languages, English exception words were the shortest and Arabic words the longest. For nonwords, English reading times were the shortest and Arabic the longest. Overall, performance in English reading was more accurate and faster than in the other two languages. Reading in Arabic was slowest and least accurate.

In tasks assessing PA, RAN and vocabulary, the results indicated a high level of accuracy in deletion and blending, with near ceiling level performance in Arabic for blending. RAN times were shorter for letters than digits for Arabic and Turkish while, for English, times for digits were slightly faster than for letters. Time for RAN letters and digits was slower in Arabic than in both English and Turkish. Standard scores from the vocabulary assessments indicated a higher level of vocabulary knowledge in Arabic than in English and Turkish.

3.1.2 Analysis of relationships between reading for the three languages, and of reading and literacy-related processes within languages

In this section, two sets of correlation analyses are reported. The first set involved the reading measures (accuracy and reading times) for English, Arabic and Turkish. The aim was to look for evidence of relationships in the processes used for reading in the children's languages. The second set of correlations looked at relationships between reading scores and the LRVs within each language. The aim here was to look at whether reading is associated with PA scores, RAN scores and vocabulary knowledge in the three languages studied, and whether the pattern of association might differ across the languages. Both sets of correlations were carried out controlling for chronological age (CA) and nonverbal ability (Matrix Analogies Test, MAT) scores, since these variables were found to be significantly associated with several of the reading and literacy-related measures, and therefore could be confounding variables unless controlled. Full correlation matrices including CA and MAT scores, revealing the significant associations of these two variables with the other measures, can be found at Appendix 2.

Correlations among reading measures for the three languages

To examine relationships in the processes used for reading in the children's languages, partial correlations across reading measures for English, Arabic and Turkish were conducted, controlling for CA and nonverbal ability, as mentioned above. Although the word and nonword within-language associations were all significant, these will not be mentioned below as they were expected, and the associations of interest in this set of analyses were between the languages. The results are shown in Table 14 for accuracy and in Table 15 for reading times.

Table 14

Correlations Between Word Reading Accuracy in The Three Languages, Controlling for CA, and Nonverbal Ability, for The TD Children

Measure	Regular word	Exception word	Nonword	Arabic word	Arabic NW	Turkish word	Turkish NW
Regular word	.						
Exception word	.312^{MS}	.					
Nonword	.608***	.371*	.				
Arabic word	.348*	-.256	.245	.			
Arabic NW	.230	-.028	.310^{MS}	.709***	.		
Turkish word	.570***	-.139	.532**	.509**	.532**	.	
Turkish NW	.317^{MS}	-.160	.344*	.360*	.399*	.740***	.

Note: * p<.05, ** p<.01, *** p<.001, ^{MS}p<.08

The results revealed that English regular word accuracy was significantly associated with that for Turkish words (highly significant), Arabic words, and marginally significantly with Turkish nonwords. English nonword accuracy was significantly associated with Turkish word and nonword accuracy, and marginally significantly with Arabic nonword accuracy. Arabic word accuracy was significantly associated with Turkish word and nonword accuracy. Arabic nonword accuracy also correlated significantly with Turkish word and nonword accuracy.

Table 15

Correlations Between Word Reading Times in The Three Languages, Controlling for CA, and Nonverbal Ability, for The TD Children

Measure	Regular Word	Exception word	Nonword	Arabic word	Arabic NW	Turkish word	Turkish NW
Regular word	.						
Exception word	.732***	.					
Nonword	.794***	.540**	.				
Arabic word	.041	-.042	-.008	.			
Arabic NW	.148	.145	.083	.798***	.		
Turkish word	.355*	.272	.261	.350*	.444**	.	
Turkish NW	.479**	$.299^{MS}$	.406*	$.325^{MS}$	.442**	.643***	.

Note: * p<.05, ** p<.01, *** p<.001, ^{MS}p<.08

The results revealed a significant correlation between English regular word reading times and those for Turkish words and nonwords. English exception words reading times were marginally significantly associated with those for Turkish nonwords. English nonword reading times correlated significantly with those for Turkish nonwords. Arabic word reading times correlated significantly with those for Turkish words and marginally significantly with those for Turkish nonwords. Arabic nonword reading times correlated significantly with reading times for Turkish words and nonwords.

Correspondence of significant associations for reading accuracy and reading times

The significant correlations for reading accuracy and reading times are presented together in Table 16, separated by '/', with reading accuracy presented first. Correlations in square brackets are within-language associations which are not of interest in this analysis.

Table 16

Correspondence of Significant Associations for Reading Accuracy and Reading Times in The Three Languages, Controlling for CA and Nonverbal Ability, for The TD Children

Measure	Regular word	Exception word	Nonword	Arabic Word	Arabic NW	Turkish Word	Turkish NW
Regular	-						
Exception word	[MS/***]	.					
Nonword	[***/***]	[*/**]	.				
Arabic Word	*/NS	-	-	.			
Arabic NW	-	-	MS/ NS	[***/***]	.		
Turkish Word	***/*	-	**/NS	**/*	**/**	.	
Turkish NW	MS/**	NS/ MS	*/*	*/ MS	*/**	[***/***]	.

Note: * p<.05, ** p<.01, *** p<.001, MS p<.08

English regular words showed a significant association with Arabic words for accuracy but not reading times, with Turkish words for accuracy and reading times. English regular words also showed a significant association with Turkish nonwords for reading times and marginally significantly for accuracy. English exception words correlated with Turkish nonwords marginally significantly for reading times. English nonwords correlated marginally significantly with Arabic nonwords for accuracy, with Turkish words for accuracy, and with Turkish nonwords for accuracy and reading times. Arabic words were significantly associated with Turkish words for accuracy and reading times and with Turkish nonwords for accuracy. Also, Arabic words were marginally significantly associated with Turkish nonwords reading times. Finally, Arabic nonwords were significantly associated with Turkish words and nonwords for both accuracy and reading times.

Interim discussion of findings and predictions for association of reading and LRVs

The descriptive statistics revealed slow reading times for Arabic and least accurate reading responses. This is notable since Arabic is the first language of the children in the present study and is likely due to the difficulties associated with spoken and written Arabic outlined in section 1.1.4.1.

Correlation analyses of the interrelationships in reading processes across the three languages revealed a robust significant association of Turkish and Arabic for accuracy and reading times for Arabic and Turkish word and nonword reading. This is in line with the prediction that Arabic and Turkish reading may be associated since both languages have strong reliance on sublexical processes in addition to the transparency of the writing system of both languages and some similar vocabulary (see section 1.1.4.2).

English regular words, which were the most accurately read type of letter strings, were associated with Turkish word and nonword reading accuracy and reading times, as was predicted. This association can be explained by dependence on sublexical processes, as well as similarity of script. English nonwords were associated with both types of Turkish letter strings for accuracy, while they associated with Turkish nonwords only for reading times. The predicted association between English nonwords and Turkish word reading time was not met in the results.

English exception words were notable for having the fastest reading times of all letter string types across languages. This can be interpreted in terms of dual route theory as reliance on fast lexical processes for this letter string type. English exception words were also notable for showing lack of association with the reading measures in the other languages, except for a marginally significant association with Turkish nonwords for reading times which could be related to the shared script with Turkish.

The associations of English and Arabic word and nonword reading were weaker, with a significant association for English regular word and Arabic word accuracy and a marginally significant association for English and Arabic nonword accuracy. The predicted association for English regular word and Arabic word accuracy, and marginal association of English nonword Arabic nonword accuracy, could also be due to reliance on sublexical processes for these letter string types.

It is notable that Turkish reading accuracy and reading times were both better than those for Arabic, although Turkish was the children’s third language, reinforced by Turkish vocabulary scores being the lowest across the three languages. This underlines the difficulties associated with reading in Arabic, as noted above, and the advantage for unskilled readers in reading a transparent script, as discussed in section 1.1.3.1. It may also be the case that practice with the Latin alphabet, conferred from English being the children’s main language of instruction, had a positive transfer effect on reading in Turkish.

Based on the results observed so far, we might expect the pattern of association of reading with the LRVs to be different across the three languages. PA can be expected to be strongly associated with operation of the sublexical route, and vocabulary knowledge with the operation of the lexical route (see section 1.1.2.3). Moreover, I tentatively expect RAN to be more related to the lexical processes. More precisely, we might predict that PA (as a sublexical marker) will be significantly associated with regular word and nonword reading in English, and with word and nonword reading in Turkish. Given the strong association between Arabic and Turkish, as well as the transparency of the Arabic script used in this study, we may also expect PA to be associated, to some extent, with word and nonword reading in Arabic.

It was predicted that when associations with LRVs were examined, English exception word reading might show a high degree of association with vocabulary, due to reliance on lexical processes because of the inability of sublexical processes to provide accurate responses in the case of this type of letter string. It was also predicted that, since reading times for exception words were the fastest of all the letter string types then RAN might also be a strong associate, due to the association of this LRV with reading speed. I also predicted that Arabic word reading might show association with vocabulary due to diglossia (section 1.1.4.3). Vocabulary knowledge might be least strongly associated with Turkish reading since the script is so transparent and is likely to encourage reliance on sublexical processes in young readers, and in addition the participants in the current study had relatively low levels of vocabulary knowledge for Turkish.

Finally, (as RAN is considered more as a lexical marker as mentioned earlier), we might predict that RAN will be related to exception word reading in English and word reading in Arabic. RAN might be least strongly associated with Turkish. I also expect an association between RAN and mostly all letter strings in reading times due to its relationship with reading fluency in previous studies (see section 1.1.2.2).

Correlations of reading processes and literacy-related abilities within languages

In this section, analyses that looked at the association of the word and nonword reading measures with the LRVs within each language are reported. As noted above, the aim of carrying out these analyses was to explore potential differences in patterns of association across the three languages. CA and nonverbal ability were controlled in the analyses, as in the analyses reported in the previous section. Due to very high levels of accuracy in the phonological awareness tasks in the three languages and non-normality of scores, the scores for blending and deletion were transformed prior to analysis and a combined PA variable was used in the correlation analyses. For clarity of exposition, the correlation matrices are reported in Appendix 3, and only the significant correlations of the reading scores and PA, RAN and vocabulary ability scores are presented in the tables below.

English correlations

In Table 17, the significant correlations of the reading measures with the LRVs for English are presented.

Table 17

Significant Correlations of Scores for Reading Measures and Literacy-Related Variables for English for The TD Children

Measure	Correlated with literacy-related measures
Regular word accuracy	PA .016, RAN letters .006, BPVS ^{MS} .072
Exception word accuracy	RAN letters .007, RAN digits .001, BPVS <.001

Nonword accuracy	PA .037, RAN letters 048, BPVS .017
Regular word reading time	RAN letters .011, RAN digits .018, BPVS <.001
Exception word reading time	PA^{MS} .064, RAN letters .035, RAN digits .002, BPVS .004
Nonword reading time	RAN letters .005, RAN digits .022 BPVS .001

Regular word reading accuracy was significantly associated with PA and RAN letters and was marginally significantly associated with vocabulary, exception word reading accuracy with RAN letters, RAN digits and vocabulary, nonword reading accuracy with PA, RAN letters and vocabulary. For reading times, regular word, exception word and nonword times were significantly associated with RAN letters, RAN digits and vocabulary. Reading time for exception words was marginally significantly associated with PA.

Arabic correlations

Table 18 shows the significant correlations of the reading measures with the LRVs for Arabic.

Table 18

Significant Correlations of Scores for Reading Measures and Literacy-Related Variables for Arabic for The TD Children

Measure	Correlated literacy-related measures
Word accuracy	PA ^{MS}.070, RAN letters .020, APVT .045
Nonword accuracy	RAN letters .002
Word reading time	RAN letters ^{MS}.060, APVT .029
Nonword reading time	PA .013, RAN digits ^{MS}.063, APVT .021

Word reading accuracy was significantly correlated with PA, RAN letters and vocabulary, and nonword accuracy with RAN letters. Word reading time was significantly correlated with vocabulary and marginally significantly with RAN letters. Nonword reading time was significantly correlated with PA, vocabulary and marginally with RAN digits.

Turkish correlations

Table 19 shows the significant correlations of the reading measures with the LRVs for Turkish.

Table 19

Significant Correlations of Scores for Reading Measures and Literacy-Related Variables for Turkish for The TD Children

Measure	Correlated literacy-related measures
Word accuracy	PA .004, RAN dig.019, TPVT.037
Nonword accuracy	PA .012
Word reading time	PA .002, RAN digits .002, TPVT <001
Nonword reading time	-

Turkish word reading accuracy was significantly associated with blending, deletion, RAN digits and vocabulary and nonword accuracy with blending. For reading times, word reading was significantly associated with blending, deletion, RAN digits and vocabulary. Turkish nonword reading times did not show significant associations with any of the literacy related variables.

Correspondence of significant associations for reading measures and for LRVs in all three languages

The significant correlations for the reading measures (reading accuracy and reading times) and for the LRVs (vocabulary, PA, and RAN) are presented in Table 20 separated by '/', with reading accuracy presented first. Correlation matrices including CA and MAT scores, revealing the significant association of these variables with the other measures, can be found in Appendix 4.

Table 20

Correspondence of Significant Associations for Reading Accuracy/Times and Literacy-Related Variables in The Three Languages, Controlling for CA, and Nonverbal Ability, for The TD Children

Measures	Vocabulary	PA	RAN
English regular words	MS/***	*/NS	**/*
English exception words	***/**	NS/MS	**/**
English nonwords	*/**	*/NS	*/**
Arabic words	*/*	MS /NS	*/MS
Arabic nonwords	NS/*	NS/*	**/MS
Turkish words	*/***	**/**	*/**
Turkish nonwords	NS/NS	*/NS	NS/NS

Note: * p<.05, ** p<.01, *** p<.001, MS p<.08, NS Not Significant

Evaluating associations of reading processes and literacy-related abilities within languages in the light of predictions

In this section, the findings from the correlation analyses for reading processes and literacy-related skills will be looked at considering predictions set out in the Interim Discussion of the reading results presented above (and in *The Present Study,* section 1.3.2).

Looking at the intercorrelations of reading words and nonwords across the three languages in section 3.1.2.1, it appeared that Arabic and Turkish showed the strongest levels of association of word and nonword reading. English regular word and nonword reading were also associated with Turkish word and nonword reading (but not Arabic word and nonword reading). English exception word reading did not show association with either Turkish or Arabic reading (apart from a marginally significant association of exception word RT and Turkish nonword RT).

First, and as I expected, different patterns of associations of LRVs were observed with reading accuracy and reading time for words and nonwords across the different languages, due to the difference in the degrees of transparency.

The significant associations of reading measures and LRVs showed that vocabulary was most closely associated with English exception word reading accuracy, which is in line with the prediction from the *Interim Discussion* section, where it was suggested that exception word reading should be strongly associated with the lexical marker vocabulary knowledge. Vocabulary was also closely associated with reading times for English regular words and Turkish words, confirming the prediction that these letter strings involve lexical processes in addition to sublexical processes.

The prediction that RAN is a lexical marker was supported by the finding that exception words were the most closely associated with RAN. Moreover, I expected RAN to be associated with Arabic word reading, which was borne out. However, contrary to what was expected, RAN showed an association with English regular words, Turkish words, English nonwords, and Arabic nonwords. This association with nonwords in English and Arabic may be due to recruitment of lexical processes for reading nonwords (there is further reflection on this issue in the *Discussion* section of this chapter).

As for PA, it was suggested that it can be considered a marker of sublexical processes and so should be associated with reading accuracy for regular word and nonword reading in English, word and nonword reading in Turkish and word and nonword reading in Arabic. English exception word reading (which is thought to rely predominantly on the lexical route) was not expected to be strongly associated with PA, and this prediction was met. On the other hand, the prediction regarding the association with PA for English and for Turkish word and nonword reading was borne out. The prediction of an association of PA with Arabic word reading was barely met with a marginally significant association. Moreover, the prediction of a significant association of PA with Arabic nonword reading, since nonwords recruit sublexical processes, was not met. The association of PA with reading was strongest for Turkish, and this was in line with the predictions set out earlier.

In the next section the findings for spelling are presented, and the results for reading and spelling are then discussed in the final section of the chapter.

3.2 Spelling

This section of the chapter reports the results for spelling words and nonwords. The spelling scores for all three languages are presented separately, and then a summary of the scores across the languages is presented for comparison purposes.

3.2.1 Scores for three languages separately

Results are presented for word and nonword spelling for each of the three languages separately.

English words and nonwords

A summary of accuracy in spelling the regular words, exception words and nonwords from subtests of the *Diagnostic Test of Word Reading Processes* (*DTWRP*) is given in Table 21.

Table 21

Mean Scores for Accuracy in Spelling English Regular Words, Exception Words and Nonwords for The TD Children (Standard Deviations are in Parentheses)

Measures	Accuracy (Max.= 30)
Regular words	19.95 (4.13)
Exception words	19.37 (3.87)
Nonwords	21.18 (4.76)

To test for differences in spelling accuracy for the different types of English letter string, Wilcoxon Signed Ranks tests were conducted. The difference between regular word spelling and exception word spelling was not significant, z=-1.293 p=.196. The difference between regular word spelling and nonword spelling, was also not significant z=-1.825 p=.068. However, the difference between exception word spelling and nonword spelling was significant z=-2.095 p=.036, in favour of nonwords.

Arabic words and nonwords

The results for accuracy in spelling Arabic words and nonwords (tasks adapted from the *Test of Reading for Elementary School Children in Arabic,* Taha, 2009) are summarized in Table 22.

Table 22

Mean Scores for Accuracy in Spelling Arabic Words and Nonwords for The TD Children (Standard Deviations are in Parentheses)

Measures	Accuracy (Max.= 30)
Words	23.39 (7.33)
Nonwords	19.21 (7.21)

Results of the Wilcoxon Signed Ranks tests revealed that words were spelled significantly more accurately than nonwords, z=-4.761, p<.001.

Turkish words and nonwords

The results for accuracy in spelling Turkish words and nonwords (tasks adapted from the *Kelime Okuma Bilgisi Testi'nin,* KOBT, Babür, Haznedar, Erçetin, Özerman & Çekerek, 2012) are summarized in Table 23.

Table 23

Mean Scores for Accuracy in Spelling Turkish Words and Nonwords for The TD Children (Standard Deviations are in Parentheses)

Measures	Accuracy (Max.= 30)
Words	22.39 (4.29)
Nonwords	20.74 (5.13)

The results of the Wilcoxon Signed Ranks revealed that words were spelled significantly more accurately than nonwords, z=-2.682 p=.007.

Comparison of scores in spelling across languages

In Table 24 a summary of accuracy in spelling across the three languages is presented as percentage scores to facilitate comparison.

Table 24

Mean Percentage Scores for Accuracy in Spelling for The TD Group for All Three Languages (Standard Deviations are in Parentheses)

Measures	English	Arabic	Turkish
Regular words	66.49 (13.75)	77.98 (24.42)	74.65 (14.30)
Exception words	64.56 (12.89)	-	-
Nonwords	70.61 (15.88)	64.04 (24.03)	69.12 (17.10)

Regarding the accuracy of spelling real words across the three orthographies, the results revealed that Arabic words were spelled most accurately, while both types of English words (regular and exception) were spelled at a comparable level and were the least accurately spelled type of real words. For nonwords, spelling accuracy was at a comparable level for English and Turkish. Arabic nonwords were the least accurately spelled nonwords across all three writing systems. Within languages, for English, nonwords were spelled more accurately than words while the reverse was the case for Arabic and Turkish.

Interim summary for spelling scores

The results for spelling, for the different types of letter strings for each language revealed that for English, nonwords were spelled more accurately that real words, and that regular words were slightly more accurately spelled than exception words. The only significant difference between these three types of letter strings was between exception words and nonwords in favour of nonwords. These findings were opposite to the findings from the reading scores in the sense that nonwords were the least accurately read type of the three types of letter strings. For Arabic and Turkish, words were spelled more accurate than nonwords and the difference showed a significant relation in favour of words in both orthographies. This finding was similar to the findings from reading accuracy in both Arabic and Turkish. When the performance in spelling was compared across all three languages, and unlike in reading, spelling of Arabic words was the most accurate, followed by spelling Turkish words. English words were the least accurately spelled words, although regular words were spelled slightly more accurately than exception words, which was the least accurately spelled type of real words across all languages. However, findings from spelling nonwords across the three languages were like those from reading, in that English nonwords were the most accurately spelled and Arabic nonwords were the least accurately spelled.

3.3 Discussion

This chapter reported the performance of TD trilingual children aged nine to eleven years in reading and spelling, and in literacy-related tasks in English, Arabic and Turkish. Correlation analyses were used to examine whether reading processes may be related across the children's three languages and whether patterns of associations with the LRVs might differ across the languages. The findings for LRVs are covered first, then those for reading, and finally for spelling.

Findings for literacy-related processes

The vocabulary assessments showed that scores were highest for Arabic, followed by English and then Turkish. These results are in line with the language background of the children for whom Arabic is the language of birth, English is the second language (main language of instruction), and Turkish is an additional language of instruction. Nevertheless, the reading scores showed that Arabic was read least accurately and slowest, English was read the most accurately and fastest, and Turkish was between the other two.

For RAN, times were fastest in English and slowest in Arabic. For Arabic and Turkish, times were faster for RAN letters than digits, while the opposite was the case for English, likely due to the correspondence of names and sounds for letters in Arabic and Turkish. Reading English exception words was significantly associated with RAN. This is in line with the predictions put forward for the association with LRVs, and this association was the strongest of all the reading-RAN pairings (please see Table 20).

The results for PA showed a high level of accuracy, especially in the deletion task: 96% for Arabic, 93% for English, and 90% for Turkish, although levels were comparably high in the PA blending task across English Arabic and Turkish: 90%, 92%, 93%, respectively, which is to be expected in children of this age, and those learning to read and spell in alphabetic and transparent writing systems. The predictions were largely supported, however, since accuracy levels in the PA tasks were so high, the results for this LRV need to be treated with some caution. It might have been the case that stronger associations (or a different pattern of associations) would have been observed if there had not been such high levels of accuracy in the PA tasks.

Findings for reading

The findings for reading revealed that English word and nonword reading was faster and more accurate than word and nonword reading in the other two languages. This may be because English has been the main language of instruction for the children. Reading in Arabic was slowest and least accurate. This finding highlights the difficulty in reading in Arabic, as discussed in section 1.1.4.3 of the literature review, and is especially notable since Arabic is the first language of the children in this study.

The results for English for reading the three types of letter string, regular words, exception words and nonwords, revealed that while regular words were read more accurately than exception words and nonwords, exception words had the shortest reading times. The difference between regular words and nonwords was also significant, with reading times for regular words shorter. These results can be interpreted in terms of the Dual Route theory (section 1.1.1.1 of the literature review), if we consider that regular words can be read using lexical and sublexical processes, while exception words are read predominantly using lexical processes, and nonwords are read predominantly using sublexical processes. This means that when lexical and sublexical processes are optimal, as in the case of TD readers, regular words have an advantage over the other two types of letter string for accuracy since they benefit from input from both lexical and sublexical processes. In terms of reading times, regular words are read with both lexical and (slower) sublexical processes while exception words are the fastest due to predominant use of the speedy lexical process. Nonwords are slowest of the three types of letter string, due to having to rely predominantly on sublexical processes.

The findings also revealed that for both Arabic and Turkish, words were read significantly more accurately and faster than nonwords, indicating use of lexical processes for words, even in the case of Turkish which could in principle be read solely with sublexical processes. The literature (reviewed in section 1.1.3.1) suggests that even for transparent writing systems lexical processes are used (as well as sublexical processes) for reading in children of the age sampled in the present research.

Correlational analyses for reading

The correlation analyses were carried out with the aim of looking for evidence of relationships in the processes used for reading in the children's languages. In terms of accuracy, English regular word scores were significantly associated with those for Arabic word reading and Turkish word and nonword reading. English nonword scores were significantly associated with Turkish word and nonword scores and marginally with Arabic nonwords. Arabic word accuracy correlated significantly with Turkish word and nonword accuracy, and Arabic nonword accuracy also correlated significantly with Turkish word and nonword accuracy.

The correlation analyses for reading times showed a significant correlation between English regular word reading times and Turkish word and nonword reading times. English exception word reading times correlated marginally significantly with Turkish nonword reading times. English nonword reading times correlated significantly with those for Turkish nonwords. Arabic word reading times correlated significantly with those for Turkish words and marginally significantly with Turkish nonwords. Arabic nonword times correlated significantly with times for Turkish words and nonwords. The results suggest that word and nonword reading processes were associated for English and Turkish and for Arabic and Turkish, but that processes for reading words and nonwords in English and Arabic were, to an extent, independent.

The second set of associations of reading with the LRVs reported in this section show that, for Turkish, word reading was associated with vocabulary, as well as PA and RAN, while for reading nonwords, PA was the only significantly associated variable. The results for English, to a certain extent, underscore those from the descriptive statistics noted above in *Findings for reading*. It was predicted that English nonword reading would be strongly associated with the sublexical marker PA, a prediction that was borne out, in addition to significant associations for nonword accuracy with vocabulary and RAN. The significant association with vocabulary could be related to the need to use lexical processes, as well as sublexical, when it comes to English reading (due to the presence of exception words). The involvement of lexical processes even when reading nonwords (as well as sublexical processes when reading exception words) has been noted by several researchers as discussed in section 1.1.3.1. Moreover, the fastest reading times were observed for exception words, and the correlation analyses revealed that the strongest association ($p<.001$) for this letter string type was with vocabulary, a lexically related variable. Reading English exception words is likely to draw on semantics because of the poor reliability of sublexical processes for providing correct responses, as noted earlier. Reading English exception words was also significantly associated with RAN, in line with the predictions put forward for the association with LRVs, and this association was the strongest of all the reading-RAN pairings, as noted above.

Regular word reading, which draws on both lexical and sublexical processes, showed an association with all three LRVs. Furthermore, for reading nonwords in English and Arabic, vocabulary was significantly associated (as well as PA and RAN). The association of nonword reading and vocabulary in English and Arabic could be due to habitual higher activation of lexical processes in these two languages, due to language-specific features (such as diglossia in Arabic, lack of consistency of GPCs in English), but it might also be due to easier activation of semantics, due to higher levels of receptive vocabulary than in Turkish, or else to higher degrees of word-likeness of the nonwords in Arabic and English compared to Turkish. The findings for Arabic suggest the strong involvement of top-down language processes, likely due to both the complexity involved in processing the script, and the diglossia, discussed in section 1.1.4.1.

Findings for spelling scores

The findings for spelling revealed that, for English, regular words and exception words were spelled with comparable levels of accuracy, while nonwords were spelled more accurately than regular words and exception words, although the difference was only significant in the case of exception words. In contrast, for Arabic and Turkish, words were spelled significantly more accurately than nonwords. A possible theoretical interpretation of the result that, in English, spelling accuracy was greater for nonwords than for words might be due to how "accuracy" is defined for nonwords and words, For example, the nonword /vi:f/ may be spelled in a phonologically plausible way as VEEF, VEAF, VIEF or VEFE, whereas there is only one correct spelling of "leaf", "beef", and "thief".

Unlike findings from reading, Arabic words were the most accurately spelled type of letter string across all languages. This might reflect the relation between vocabulary knowledge and spelling, (i.e., good vocabulary may give the boost to Arabic spelling), where Arabic is the language of birth of all the participants, and level of vocabulary knowledge in Arabic was higher than in the other two languages. Vocabulary is an important predictor of spelling according to the research covered in section 1.1.4.1. English regular words were spelled less accurately than Arabic and Turkish words, even though they were read more accurately than Arabic and Turkish words, in spite the fact that English is the main language of instruction and the fact that regular words should benefit from being able to draw on both lexical and sublexical processes. English nonwords were spelled more accurately than Turkish nonwords, and Arabic nonwords were spelled least accurately of all.

4 Chapter 4: Results for poor readers

Introduction

This chapter reports investigation of deficits in lexical and sublexical processes and in literacy-related variables (LRVs) in the 12 Arabic-English-Turkish trilingual poor readers in the PR group. In the first part of the chapter, the performance of the poor readers as a group is compared to that of 20 TD children, a subsample of the larger TD group reported in the last chapter. As reported in the *Participants* section, inspection of the results for the PR and TD children revealed that the scores for the former group were substantially lower than those of the TD group in English vocabulary knowledge (please see Table 3 in *Participants*). To overcome the problem of interpretation of the group comparisons of TD and PR results, where the poor readers were also poorer in terms of vocabulary knowledge, it was considered important to form a comparison group (CG) who were matched in English vocabulary knowledge to the PR children. Findings are reported for the word and nonword reading tasks, for the assessments of PA, RAN and vocabulary knowledge, and finally for spelling. As in the previous chapter, a summary of the scores of the reading measures and the LRVs across the three languages, and a similar summary of the spelling scores, are reported for comparison purposes.

In the second part of the chapter, the individual profiles of each of the poor readers in terms of reading and association with LRVs are presented. Z scores were used to indicate the extent of the deficit in each of the measures. These were derived from the scores of the whole sample of TD children and the PR children divided into Grade 4 and Grade 5, since associations of some of the measures with age were found in the TD results. The z scores of the PR children were used to provide reading profiles which were then examined for a) association across languages, and b) association with deficits in the LRVs.

4.1 Reading and literacy-related variables

In this section, the performance of the PR group and comparison group in word and nonword reading tasks and in the literacy-related assessments of PA, RAN and vocabulary knowledge are presented for English, Arabic and Turkish. Comparisons of the scores for the two groups were conducted using t-tests. Where data violated the assumptions for parametric tests, the non-parametric Mann-Whitney test was employed. Following this, a summary of all scores is presented to facilitate the comparison of performance across the three languages.

4.1.1 Scores for the PR and the CG for the three languages separately

Scores of the PR and CG in the word and nonword reading accuracy and reading times, as well as results from assessments of LRVs are reported for each language separately. Comparisons of the scores for the two groups were conducted using t-tests. Where data violated the assumptions for parametric tests, the non-parametric Mann-Whitney test was employed.

English assessments

Reading words and nonwords

The results for reading accuracy in the regular word, exception word and nonword subtests of the *Diagnostic Test of Word Reading Processes* (*DTWRP*) for the PR and CG groups are summarised in Table 25. Mean reading times in milliseconds for the three types of printed letter string are also summarised in the table. For the PR group, tests of normality revealed that scores were normally distributed, except for regular word reading times (skewness 1.90, Shapiro-Wilk: $p=.015$). For the CG scores deviated from normality for regular word reading accuracy (skewness -1.269, Shapiro-Wilk: p <.021), regular word reading times (skewness 2.138, Shapiro-Wilk: p =.001), exception word reading times (skewness 2.654, Shapiro-Wilk: p < .001), and nonword reading times (skewness .740, Shapiro-Wilk: p= .048). The results of tests comparing the scores for the two groups are presented in the final column of the table.

Table 25

Mean Scores for Reading Accuracy and Reading Times for Regular Words, Exception Words and Nonwords for The PR and CG (Standard Deviations are in Parentheses)

Measures	PR	CG	Difference
Reg. accuracy (Max. = 30)	20.67 (4.92)	27.55 (1.99)	U=13.5, p<.001, z= -4.175
Exc. accuracy (Max. = 30)	15.58 (3.94)	23.00 (2.97)	t (30) = 6.043, p<.001, d =2.13
Non. accuracy	13.58	24.50	

(Max. = 30)	(4.34)	(2.14)	t (30) = 9.551, p<.001, d =3.19
Reg. reading times (Msec)	3207 (1458)	2003 (517)	U=36.0, p=.001, z= -3.270
Exc. reading times (msec)	2453 (693)	1708 (572)	U= 50.0, p=.005, z= -2.725
Non. reading times (msec)	3203 (772)	2288 (484)	U= 38.0, p=.001, z = -3.192

The comparisons revealed worse performance for the PR group for all letter string types and for both accuracy and reading times. Tests were also conducted within the groups for the difference between letter string types. For the PR group, in reading accuracy, the difference between regular words and exception words, and between regular words and nonwords were both significant, t(11)=3.436, p=.006, and t(11)=4.838, p=.001 respectively, both in favour of regular words. The difference between exception word and nonword reading accuracy was not significant, t(11)=1.453, p=.174. For reading times, the difference between regular words and exception words was significant, z=-2.118, p=.034, where exception words reading time was shorter. The difference between regular words and nonwords was not significant, z=-.392, p=.695, and finally the difference between exception word and nonword reading times was significant, t(11)=-2.850, p=.016, with exception word reading times shorter.

For the CG, in reading accuracy, the difference between regular word and exception words, and between regular words and nonwords, z=-3.732, p<.001 and z=-3.747, p<.001, respectively, were both significant and in favour of regular words in both cases. The difference between exception word and nonword reading accuracy was also significant, t(19)=-2.276, p=.035, with a higher score for nonwords. For reading times, the difference between regular and exception words (z=-2.987, p=.003), and exception words and nonwords (z=-3.920, p<.001) was significant, with exception word reading times shorter in both cases. The difference for regular words and nonwords, z=-3.883, p<.001 was also significant, regular words reading times were shorter.

The results for reading accuracy for the PR group were like the within-group differences in the TD group (reported in Chapter 3), except that the difference in reading times for regular words and nonwords in the PR group was not significant, unlike the difference between these two types of letter strings in the TD group where regular word reading times were significantly shorter.

The differences for the TD group and the smaller CG were the same in all measures of accuracy and reading times, except for the difference between English exception words and English nonword reading accuracy which was found to be significant in the smaller CG, but not in the wider TD group, in favour of nonword reading accuracy (see section 3.1.1 in Chapter 3).

PA, RAN, and vocabulary knowledge

The results for the PR and CG groups are summarised in Table 26. The scores for vocabulary, deletion and blending are in terms of accuracy, and those for RAN letters and digits are in terms of seconds to complete the task. Also presented in Table 26, in the final column, is the result of tests comparing the scores of the two groups.

Table 26

Mean Accuracy in Deletion and Blending, Standard Scores in Vocabulary and Times in RAN for The English Language for The PR and CG (Standard Deviations are in Parentheses)

Measures	PR	CG	Difference
Deletion (Max = 20)	13.42[a] (2.84)	17.50[a] (2.06)	t(30) = 4.699, p<.001, d =1.644,
Blending (Max =20)	16.42[a] (1.44)	18.30[a] (1.26)	U=39.50, p=.001, z =-3.254
RAN Letters (Time in sec)	46.42[a] (6.95)	37.95[a] (7.90)	t(30) = 3.066, p=.005, d =1.138,
RAN Digits (Time in sec)	44.00[a] (11.89)	34.10[a] (6.47)	U= 51.00, p =.007, z = -2.689
BPVS	66.75[b] (10.40)	71.95[b] (8.57)	t(30) = 1.520, p=.139, d= 0.546

Note: [a] subtest scaled scores from the *CTOPP* (Wagner et al., 1999), [b] standard scores from the *BPVS* (Dunn et al., 1997).

Scores for the PR group deviated from normality for RAN digits (skewness .976, Shapiro-Wilk: *p* =.046). Scores for the CG deviated from normality for the blending task (skewness .242, Shapiro-Wilk: *p* =.003), but not for the other measures. Tests for the effect of group revealed a significant difference, in favour of the CG, for all literacy-related measures except for vocabulary (as expected since the two groups were matched on this measure).

Arabic assessments

Reading words and nonwords

The results for the PR group and the CG for reading words and nonwords in Arabic, which were adapted from the *Test of Reading for Elementary School Children in Arabic* (Taha, 2009) are summarised in Table 27. Reading times for reading the words and nonwords are also summarised in the table. For the PR group, all tests showed normally distributed values. Scores for the CG deviated from normality for word reading accuracy (skewness -1.616, Shapiro-Wilk: *p* =.008), and nonword reading accuracy (skewness-1.033, Shapiro-Wilk: *p*=.021), but not for the word reading times or nonword reading times. Tests for the effect of group revealed a significant difference in word and nonword accuracy scores in favour of the CG. The result of tests comparing the scores of the two groups is presented in the final column of the table.

Table 27

Mean Reading Accuracy and Reading Times for Arabic Words and Nonwords for The PR and CG (Standard Deviations are in Parentheses)

Measures	PR	CG	Difference
Word accuracy (Max. = 30)	20.50 (3.37)	23.75 (5.67)	U=62.50, p=.024, z =-2.25
Nonword accuracy (Max. = 30)	12.33 (8.84)	19.15 (7.92)	U=55.50, p=.012 z= -2.516
Word reading times (msec)	2718 (918)	2268 (834)	t (30) = -1.426, p = .164, d=-0.513
Nonword reading times (msec)	2868 (1138)	2669 (770)	t (30) = -.591, p =.559, d = -0.205

Tests were also conducted within groups for the differences between the different types of letter string for each group. Difference tests revealed that, for reading accuracy for the PR group the difference between words and nonwords was significant, t(11)=3.153, p=.009, in favour of words. For reading times, the difference between words and nonwords was not significant, t(11)=-.849, p=.414. For the CG, for reading accuracy, the difference between words and nonwords was significant, z=-3.049, p=.002, in favour of words, and for reading times, the difference between words and nonwords was also significant, t(19)=-3.083, p=.006, reading times for words were faster.

The difference in reading times between Arabic words and nonwords in the PR group was not significant, unlike the finding for the TD group where the difference was significant in favour of words, which were faster. The results for the TD group and the smaller CG were the same for both accuracy and reading times for Arabic.

PA, RAN, and vocabulary knowledge

Table 28 gives a summary of the results for the PR the for the Arabic literacy-related tasks. The results for deletion, blending and vocabulary are in terms of accuracy, and those for RAN of letters and digits are in terms of seconds to complete the task. The results of tests comparing the scores for the two groups are presented in the final column of the table.

Table 28

Mean Accuracy in Deletion and Blending, Standard Scores in Vocabulary and Times in RAN for The Arabic Language for The PR and CG (Standard Deviations are in Parentheses)

Measures	PR	CG	Difference
Deletion (Max =20)	17.417[a] (2.78)	18.30[a] (1.84)	U = 101.50, p =.452, z= -.725
Blending (Max =20)	18.58[a] (1.56)	19.30[a] (1.13)	U = 87.00, p = .155, z = -1.422
RAN Letters (Time in sec)	62.50[b] (25.69)	63.85[b] (16.71)	U = 101.50, p = .471, z = -.721
RAN Digits (Time in sec)	101.58[b] (25.87)	90.40[b] (32.03)	t(30) = -1.024, p = .314,d= -0.384

APVT	112.50[c] (23.66)	97.80[c] (21.76)	t(30) = -1.791, p= .083, d=-0.646

Note: [a] subtest raw scores from the Test of Reading for Elementary School Children in Arabic (Taha, 2009), [b] PhAB, Taibah & Elbeheri, 2011), [c] standard scores from the APVT (Abu et al., 1998).

The scores of the poor reader group deviated from normality for deletion (skewness -.946, Shapiro-Wilk: *p* = .048), blending (skewness -.698, Shapiro-Wilk: *p* = .021) and RAN letters (skewness 1.311, Shapiro-Wilk: *p* = .041). The scores of the CG also deviated from normality for the deletion task (skewness -.775, Shapiro-Wilk: *p* =.002), the blending task (skewness -1.396, Shapiro-Wilk: *p* <.001) and RAN letters (skewness -.555, Shapiro-Wilk: p=.030). Tests for the effect of group revealed no significant difference in the PA and RAN measures in Arabic.

Turkish assessments

Reading words and nonwords

The results for the PR group and the CG for reading words and nonwords in Turkish are summarised in Table 29. Standard scores were not available for the Turkish reading tasks, which were adapted from the KOBT (Babür, Haznedar, Erçetin, Özerman & Çekerek, 2012). Scores for the CG deviated from normality for word reading accuracy (skewness -.971, Shapiro-Wilk: *p*.= 018). For the PR group, scores showed normally distributed values. Tests for the effect of group revealed a significant difference in both accuracy and reading times for words and nonwords in favour of the CG. Reading times for reading the words and the nonwords are also summarised. The result of tests comparing the scores for the two groups is presented in the final column of the table.

Table 29

Mean Reading Accuracy and Reading Times for Turkish Words and Nonwords for The PR and Comparison Groups (Standard Deviations are in Parentheses)

Measures	PR	CG	Difference

Word accuracy (Max=30)	18.92 (2.87)	25.30 (4.60)	U =34.50, p =.001, z = -3.337
Nonword accuracy (Max=30)	15.917 (4.94)	22.80 (4.69)	t (30) = 3.942 p<.001, d = 1.429
Word reading times (msec)	3769 (976)	2152 (552)	t (15.30) = -5.256, p<.001, d = -2.039
Nonword reading times (msec)	3995 (949)	2591 (420)	t(13.626)= -4.851, p<.001, d= -1.913

Tests for differences in accuracy and reading times between the different types of letter string for each group, revealed that for the PR group, the difference between word and nonword accuracy was significant, t(11)=2.514, p=.029, in favour of words. For reading times, the difference between words and nonwords was not significant, t(11)=-1.108, p=.292. For the CG, the difference between word and nonword accuracy was significant, z=-2.685, p=.007, with words more accurate than nonwords. For reading times, the difference between words and nonwords was also significant, t(19)=-4.566, p<.001, reading times for words were shorter. The differences for the wider TD group and the smaller CG, were the same for both accuracy and reading times in Turkish.

PA, RAN, and vocabulary knowledge

Table 30 gives a summary of the results for the PR group and the CG for the literacy-related tasks for Turkish. Results for vocabulary, deletion and blending are in terms of accuracy, and those for RAN of letters and digits are in terms of seconds to complete the task. The result of tests comparing the scores for the two groups is presented in the final column of the table.

Table 30

Mean Accuracy in Deletion and Blending, Standard Scores in Vocabulary and Times in RAN for The Turkish Language for The PR and CG (Standard Deviations are in Parentheses)

Measures	PR	CG	Difference
Deletion (Max= 20)	17.00[a] (2.83)	18.25[a] (1.62)	U = 85.50, p = .171, z = -1.370
Blending (Max= 20)	14.75[a] (2.38)	18.30[a] (1.72)	U = 30.00, p <.001, z = -3.547
RAN letters (Time in sec)	40.67[b] (9.48)	32.25[b] (7.55)	U =46.50, p = .004, z =-2.867
RAN Digits (Time in sec)	66.58[b] (16.14)	51.40[b] (15.78)	U = 48.00, p = .005, z = -2.805
TPVT	61.50[c] (24.96)	69.05[c] (30.49)	U =102.50, p =.495, z = -.683

Note: [a] subtest raw scores from (*KFFT*; Babür et al., 2013), [b] *(HOTI;* Bakır & Babür, 2009), [c] standard scores from the *TPVT* (Katz et al., 1974).

Scores for the CG deviated from normality for all the phonological ability measures, namely: deletion (skewness -.864, Shapiro-Wilk: p =.023), blending (skewness -.450, Shapiro-Wilk: p =.006), RAN letters (skewness 1.991, Shapiro-Wilk: p = .002), and RAN digits (skewness 1.468, Shapiro-Wilk: p =.009). The TPVT scores for the CG also deviated from normality (skewness 1.446, Shapiro-Wilk: p =.001). For the PR group, scores in most measures showed normally distributed values, however, vocabulary scores deviated from normality (skewness 1.730, Shapiro-Wilk: p =.021). Tests for the effect of group revealed a significant difference in blending, RAN for letters, and RAN for digits, all in favour of the CG.

Comparing reading and literacy-related scores across all three languages

In this section, percentage scores of the reading accuracy measures and the average scores of reading times (in msec) across all three languages as well as the average scores for the LRVs for the PR group and the CG are presented.

Reading words and nonword

In Table 31, percentage scores for the reading accuracy measures and the average scores of reading times (in msec) across all three languages are given. Asterisks are used to denote a significant difference in scores for the poor readers and the CG.

Table 31

Mean Percentage Scores for Reading Accuracy and Mean Scores for Reading Times in Words and Nonwords for The PR Group and CG Across the Three Languages

Measures	English		Arabic		Turkish	
	PR	CG	PR	CG	PR	CG
Reg. Accuracy	68.90	**91.83***	68.33	**79.17***	63.03	**84.33***
Exc. Accuracy	51.93	**76.67***	-	**-**	-	**-**
Non. Accuracy	45.27	**81.67***	41.10	**63.83***	53.03	**76.00***
Reg reading times (msec)	3207	**2003***	2718	**2268**	3769	**2152***
Exc reading times (msec)	2453	**1708***	-	**-**	-	-
Nonword reading times (msec)	3203	**2288***	2868	**2669**	3994	**2591***

Regarding the accuracy of reading words across the three languages, the results for the CG revealed that English regular words were read most accurately, and English exception words were read least accurately. For the PR group English regular words were also the most accurate and English exception words were the least accurate. As for nonword accuracy, for the CG English nonwords were the most accurate, and Arabic nonwords were the least accurate, while for the PR group Turkish nonwords were most accurate and Arabic nonwords least accurate.

For reading times, results showed that English exception words were the fastest for the CG and for the PR group. However, Arabic words the slowest for the CG, while Turkish were the slowest for the PR group. As for nonwords, for the CG English nonwords were fastest and Arabic nonwords slowest, while for the PR group Arabic nonwords were fastest and Turkish slowest.

PA, RAN, and vocabulary knowledge

Table 32 gives a summary of percentage accuracy scores for the PA measures and the average times for the RAN tasks for the PR and the CG, across all three languages. Percentages of raw scores of vocabulary assessments are also presented in the table to make comparison easier. Asterisks are inserted to denote significant group differences, as in the previous table.

Table 32

Mean Accuracy in Deletion, Blending, Standard Score in Vocabulary and Time in RAN Letter and Digit Tasks for The Poor Readers and Comparison Group Across the Three Languages

Measures	English		Arabic		Turkish	
	PR	Comparison	PR	Comparison	PR	Comparison
Deletion	67	**87.5***	87.05	**90.15**	85	**91.25**
Blending	82.1	**91.5***	92.9	**96.5**	73.75	**91.5***
RAN letters (Time in secs)	46.42	**37.95***	62.50	**63.85**	40.67	**32.25***
RAN Digits (Time in secs)	44.00	**34.10***	101.58	**90.40**	66.58	**51.40***
Vocabulary	66.75	**71.95**	112.50	**97.80**	61.50	**69.05**

The results for deletion and blending varied between the different languages for the two groups. For deletion, the PR performed better in Arabic, followed by Turkish and finally in English while the CG performed better in Turkish than in Arabic and had the English deletion task as the least accurate among all three languages. Blending Arabic words was the most accurate in both groups, then came English followed by Turkish in the PR while blending English and Turkish were at the same level for the CG. RAN for letters was faster in Turkish than for English and Arabic were the slowest in both groups. RAN for English digits was the fastest, then RAN digits for Turkish and finally Arabic, a common order for both groups again. Standard scores from the vocabulary assessments show a higher level of vocabulary knowledge in Arabic than in English and the least knowledge in Turkish for both groups, the PR, and the CG.

Summary for reading scores

The results for reading accuracy, for the different types of letter string for each language, revealed that, for English, regular words were read more accurately than exception words and nonwords by the CG and PR groups. A significant difference was found between the two groups in all types of letter strings in favour of the CG. For Arabic, words were read more accurately than nonwords in the CG and PR groups, and tests for the effect of group revealed a significant difference in word and nonword accuracy scores in favour of the CG. In Turkish, words were also read more accurately than nonwords in both groups, and tests for the effect of group also revealed a significant difference in word and nonword accuracy scores in favour of the CG.

When the performance in reading accuracy for real words was compared between the two groups across the three languages, English regular words were found to be the most accurately read by the CG and the PR group. English exception words were the least accurately read by the CG and PR groups. For nonwords, the two groups showed different performance. English nonwords were the most accurate for the CG while Turkish nonwords were the most accurate for the PR group. Nevertheless, Arabic nonwords were the least accurately read by both groups.

For reading times, scores showed that for English, exception words were read faster than the other two types of letter strings by both groups, and all three types were read significantly faster by the CG children. For Arabic, words were read faster than nonwords in both the CG and the PR group. The difference between the groups in reading time for both words and nonwords was not found to be significant. For Turkish, reading times for words were significantly faster than for nonwords in the CG but not the PR group, and tests for the effect of group revealed shorter reading times for words and nonwords in the CG.

When reading times for words was compared across the three languages English exception words were the fastest, for both groups. For the PR group the Turkish words were the slowest type of letter strings while for the CG Arabic words were the slowest. Reading times for nonwords also differed across the three languages. For the PR group, Arabic nonwords were the fastest and Turkish were the slowest and for the CG English nonwords were the fastest and Arabic were the slowest. Overall, English exception words were the most difficult to read (least accurate) yet they were the fastest across all the languages for both groups.

In the tasks measuring PA, RAN and vocabulary, scores showed that English deletion was the least accurate among all three languages for both groups. Blending Arabic words was the most accurate in both groups. RAN letters showed the shortest time in Turkish in both groups and RAN digits showed the shortest time in English. RAN letters and digits in Arabic were the slowest across all languages in both groups. Standard scores from the vocabulary assessments for both the PR and the CG indicated a higher level of vocabulary knowledge in Arabic than in English, while the lowest level of vocabulary knowledge was in Turkish. These results are in line with the language background of the children for whom Arabic is the language of birth. The findings that Turkish is the weakest language can be explained by the fact that it is the third and additional language of instruction.

Given that the CG was a subgroup from the TD group investigated in the previous chapter, it is not surprising that the analyses revealed the same significant differences for the two groups for all measures in all languages, except for the difference between English exception word and English nonword reading accuracy which was found to be significant only in the smaller CG, in favour of nonword reading accuracy.

In terms of differences in results for the PR and CG (and wider TD group), the only disparities that were observed were in terms of reading times: for the TD children the difference between English regular word and nonword reading times was significant, and the difference between Arabic words and nonwords, and that between Turkish words and nonwords, was significant, with words being faster than nonwords in each case. For the PR children, in contrast none of these differences was significant. This result seems to be important since we saw in the literature review that faster speed of reading real words has been interpreted as a marker of the use of lexical processes in languages with transparent writing systems. For English, the PR children had significantly faster reading times for exception words than nonwords, and I interpreted this difference in terms of the strong association of exception words with lexical processes. However, the regular word-nonword difference is more akin to the word-nonword difference in Arabic and Turkish, where both sets of items are decodable (i.e., can potentially be read using sublexical processes). Significantly shorter times for real words indicates the involvement of speedy lexical processes, as opposed to time-consuming sublexical processes. The lack of such a difference in the PR group (even though the differences in terms of accuracy were in favour of real words), suggest at least a partial deficit in lexical processes in this group.

The consistently poorer accuracy in reading nonwords in the PR group than the TD children indicates a deficit in sublexical processes, as well as the potential deficit in lexical processes noted above. These observations are for the poor readers as a whole though, in the section below I report analyses of individual profiles in the PR children, which aimed to see whether there were selective deficits in lexical or sublexical processes and whether these might be consistent across languages, as well as being consistently associated with diverse deficits of LRVs.

4.1.2 Spelling

In this section the results for the PR children and the CG in word and nonword spelling assessments are reported, followed by a summary of the scores across the three languages to make the comparison easier.

Results in each language separately

The spelling scores for the PR children and the CG in word and nonword assessments are presented for each of the three languages, English, Arabic, and Turkish separately.

English assessments

The results for spelling regular words, exception words and nonwords for the PR group and the CG are summarised in Table 33. The results comparing the scores for the two groups are presented in the final column of the table.

Table 33

Mean Scores for Spelling Regular Words, Exception Words and Nonwords for the Poor Readers and The Comparison Group (Standard Deviations are in Parentheses)

Measures	PR	CG	Differences
Regular words (Max.= 30)	10.08 (2.88)	18.45 (3.09)	t(30)= -7.61, p<.001, d=2.802
Exception words (Max.= 30)	7.92 (2.94)	18.45 (3.97)	U=8.00, p<.001, z=-4.372
Nonwords (Max.= 30)	12.25 (5.14)	19.75 (4.77)	U=32.50, p=.001, z=-3.415

Scores for the CG deviated from normality for exception word spelling (skewness-1.232, Shapiro-Wilk: p=.029), and nonword spelling (skewness -1.266, Shapiro-Wilk: p=.046). For the poor readers, all three types of letter string showed normally distributed values. Tests for the effect of group revealed a significant difference, in favour of the CG, for all the spelling measures.

Arabic assessments

The results for word and nonword spelling for the poor readers and the CG in Arabic are summarized in Table 34. The result of tests comparing the scores for the two groups is presented in the final column of the table.

Table 34

Mean Scores for Spelling Arabic Words and Nonwords for The Poor Readers and The Comparison Group (Standard Deviations are in Parentheses)

Measures	PR	CG	Differences
Words (Max.= 30)	21.25 (6.51)	23.60 (7.07)	U=89.00, p=.225, z=-1.215
Nonwords (Max.= 30)	16.67 (7.64)	18.30 (6.71)	U=109.50, p=.681, z=-.410

Scores for the CG deviated from normality for word spelling (skewness-1.904, Shapiro-Wilk: p<.001), and nonword spelling (skewness -1.196, Shapiro-Wilk: p=.004). For the poor readers, all spelling tests showed normally distributed values. Tests for the effect of group revealed that the difference was not significant for the spelling of both types of letter string.

Turkish assessments

The results for spelling of words and nonwords for the PR group and the CG in Turkish are summarized in Table 35. The result of tests comparing the scores for the two groups is presented in the final column of the table.

Table 35

Mean Scores for Spelling for Turkish Words and Nonwords for The Poor Readers and The Comparison Group (Standard Deviations are in Parentheses)

Measures	PR	CG	Differences
Words (Max.= 30)	10.67 (3.34)	22.00 (4.12)	t(30)= 8.062, p<.001, d=3.021
Nonwords (Max.= 30)	13.33 (3.89)	20.65 (5.37)	t(30)= 4.104, p<.001, d=1.561

Scores in spelling for the CG and for the poor readers all showed normally distributed values. Tests for the effect of group revealed a significant difference, in favour of the CG, for both the spelling measures.

Comparison of spelling results across the three languages

A summary of the spelling scores across all three languages for the PR group and the CG is presented in Table 36.

Table 36

Mean Percentage Scores for Spelling for The PR and CG for All Three Languages (Standard Deviations are in Parentheses)

Measures	English		Arabic		Turkish	
	PR	CG	PR	CG	PR	CG
Reg.	33.61 (9.58)	**61.50*** (10.29)	71.11 (22.03)	**79.00** **(23.40)**	36.67 (11.46)	**73.50*** (13.83)
Exc.	26.67 (9.85)	**61.50*** (13.22)	-	-	-	**-**
NW	40.83 (17.12)	**65.83*** **(15.89)**	53.03 (25.48)	**61.00** (22.35)	44.44 (12.97)	**68.83*** (17.91)

The results revealed a generally worse performance of the PR compared to the CG in all letter strings in all languages, particularly in the case of English and Turkish words and nonwords - the group difference for Arabic words and nonwords was not significant. Moreover, the patterns of performance in spelling words for the two groups were similar. That is, for the poor readers, Arabic words were spelled most accurately, and English exception words were least accurately spelled. For the CG, Arabic words were also spelled most accurately, and English regular and exception words, which were comparable, were spelled least accurately.

As for nonwords, Arabic nonwords were most accurate and English nonwords least accurate for the PR group, while Turkish nonwords were the most accurate and Arabic nonwords were the least accurate for the CG.

Summary for spelling

The results for spelling accuracy showed that for English, nonwords were spelled more accurately than real words. This was the same for both TD groups as well as for the PR group. Tests for the effect of group revealed a significant difference in spelling measures in English in favour of the CG. For Arabic, and as was the case with the TD group, words were spelled more accurately than nonwords by both the CG and the PR group, but no significant difference was found between the groups. For Turkish, the PR group spelled nonwords more accurately than words, while the CG, in a similar pattern to the main TD group, spelled the words more accurately than nonwords. Tests for the effect of group revealed a significant difference for Turkish words and nonwords, in favour of the CG.

When the performance in spelling accuracy for real words was compared across the three languages, the results revealed that for the PR group, Arabic words were the most accurately spelled, which is a finding that is in line with the relatively high level of Arabic vocabulary knowledge in this group. The association between spelling and vocabulary knowledge was discussed in the literature review. English exception words, on the other hand, were the least accurately spelled type of words. For the CG, again Arabic words were spelled most accurately, and English regular and exception words, which were spelled at a comparable level, were the least accurately spelled.

As for nonwords, the PR group spelled Arabic nonwords most accurately while English nonwords were the least, while the CG spelled Turkish nonwords most accurately and Arabic nonwords were the least accurately spelled.

4.2 Individual Poor Reader Profiles

In this section of the chapter individual profiles for the PR children are presented in terms of strengths and weaknesses in lexical and sublexical reading processes and in LRVs. Z scores were used to provide an extent of the deficit or strength in each measure. The reading deficits were examined for a) association across languages[4], and b) association with deficits in the LRVs.

Z scores were used for the analyses since they provide the deviation from the mean for the group (of TD and PR children) in terms of standard deviations. The results for the TD children reported in Chapter 3 revealed that there were associations for some of the reading and LRVs in terms of grade level (English nonword reading accuracy, Arabic word reading time and nonword reading accuracy, Arabic word spelling accuracy), and so for the present analyses the z scores were derived from the sample of TD and PR children divided into Grade 4 and Grade 5 groups.

On the basis of research discussed in the literature review chapter, it was expected that a selective deficit in sublexical reading processes would be associated with a deficit of PA, and that a selective deficit of lexical processes may be associated with a deficit of vocabulary knowledge or RAN.

Word and nonword reading profiles

The results for word and nonword reading accuracy for individual poor readers in terms of z-scores are given in Appendix 5. For clarity of exposition in the current section, poor performance is presented in Table 37 with asterisks according to cut-offs commonly used in the literature: three asterisks represent a score equal to or in excess of 1.5 standard deviations from the mean (a severe deficit), two asterisks represent a score 1.0 to 1.4 standard deviations from the mean (a moderate deficit), and one asterisk represents a score .8 to .9 standard deviations from the mean (a mild deficit). In Appendix 5 it can be seen that for measures involving accuracy (i.e., word and nonword reading accuracy, vocabulary, and PA) z scores indicating a deficit are negative, and for measures involving time (i.e., word and nonword reading times, RAN) z scores are positive. Table 37 includes results for both accuracy and reading times, separated by a forward slash for each measure.

[4] There were no apparent associations with number of years of residence in Turkey or dialect of Arabic spoken.

Table 37

Z Scores for Word and Nonword Reading Accuracy and Reading Times for Grade 4 and 5 PR children

	PR1G4	PR2G4	PR3G4	PR4G4	PR5G5	PR6G5	PR7G5	PR8G5	PR9G5	PR10G5	PR11G5	PR12G5
English regular	**/***	-/-	***/-	**/-	***/-	***/***	***/*	-/-	**/***	-/-	***/-	-/**
English exception	***/***	-/**	-/**	***/-	***/-	***/**	***/-	***/-	**/***	-/***	*/-	-/***
English nonword	**/***	*/-	**/-	***/***	***/**	**/***	***/-	-/-	***/-	***/***	**/-	***/***
Arabic word	-/-	-/-	-/-	-/-	-/-	**/**	-/-	-/-	-/***	-/-	-/-	-/***
Arabic nonword	-/**	***/-	-/-	**/-	-/-	***/-	-/-	***/-	-/***	-/-	-/-	***/**
Turkish word	*/***	-/-	**/***	**/-	***/***	**/**	**/-	-/-	**/***	*/***	***/*	-/***
Turkish nonword	-/***	-/-	-/***	***/-	***/***	-/***	-/-	-/**	**/**	-/***	***/-	**/***

Note: *=.8-.9, **=1.0-1.4, ***=1.5+

Deficits in word reading and in nonword reading were found across all languages. No child showed a deficit in only one language, most (11/12) showed a deficit in English and Turkish reading, and seven of these children showed an additional deficit in Arabic reading (in four cases Arabic nonword reading only, in the other three cases Arabic word and nonword reading).

Cases PR2 and PR8 showed the mildest deficits (PR2: just in exception word reading time and nonword reading accuracy for English and nonword reading accuracy for Arabic, PR8: English exception word and Arabic nonword reading accuracy and Turkish nonword reading times). For most of the other ten children a deficit in word reading was accompanied by a deficit in nonword reading for English and Turkish (which showed most deficits), except in the case of PR7 who showed a deficit for Turkish words but not nonwords, and PR8 who showed a deficit for English exception words but not nonwords. A summary of deficits found in different types of words and nonwords across languages is presented in table 38.

Table 38

Summary of Deficits in Different Types of Words and Nonwords Across Languages

	Words*		Nonwords	
	Accuracy	Time	Accuracy	Time
Eng. regular	8	4	11	6
Arabic	1	3	5	3
Turkish	7	8	4	8

*English exception word totals: accuracy=8 time=7

In terms of accuracy, there were more children showing weakness in English than Turkish and Arabic. The least deficits were observed for Arabic words, which is perhaps unsurprising as it was children's L1. In terms of times there were more children showing weakness in Turkish than English and Arabic. The reading time deficits for Turkish is most probably because Turkish is children's L3. For nonwords in terms of accuracy, the largest number of deficits were recorded for English. This is likely due to the complex and inconsistent nature of the GPCs in English, while the fewest deficits were recorded for Turkish, which has the advantage of being transparent and does not have the visual complexities of Arabic. However, there were most deficits recorded for Turkish for nonword reading speed. This is in line with the findings reported in the literature review for transparent writing systems where poor reading was manifested in slow reading speed rather than inaccurate reading.

It is noted that for English and Arabic there are more deficits for nonwords than words - English opaqueness and Arabic complexities may encourage the use of lexical processes – and the children could be more proficient in these because of higher exposure to these two languages before learning Turkish.

As with the results for reading, the performance of the 12 poor readers in LRVs is presented with asterisks in Table 39.

Table 39

Deficits in Vocabulary PA and RAN (Letter/Digit) for the Grade 4 and Grade 5 Poor Readers

	PR1G 4	PR2G 4	PR3G 4	PR4G 4	PR5G 5	PR6G 5	PR7G 5	PR8G 5	PR9G 5	PR10G 5	PR11G 5	PR12G 5
English Vocab.	**	-	**	-	***	-	*	-	**	-	-	***
Arabic Vocab.	-	-	-	**	-	-	-	-	-	-	-	-
Turkish Vocab.	*	-	-	*	-	-	-	-	-	-	*	*
English PA	*	-	-	***	***	*	**	***	***	-	***	*
Arabic PA	**	-	-	-	-	-	-	-	-	-	***	-
Turkish PA	*	-	-	***	***	-	-	-	-	-	**	-
English RAN	*/**	-/-	***/**	-/-	*/***	***/-	-/-	***/-	**/***	-/-	-/-	-/-
Arabic RAN	-/**	-/-	-/-	-/*	-/-	***/-	-/-	-/*	-/-	***/***	-/-	-/-
Turkish RAN	-/***	-/-	-/-	***/-	*/*	-/***	-/-	-/-	*/-	-/**	***/**	*/-

Note: *=.8-.9, **=1.0-1.4, ***=1.5+

When looking at the LRVs, most deficits in terms of vocabulary were found in English (N=6), followed by Turkish (N=4), followed by Arabic (N=1). However, Turkish deficits were all mild. Little correspondence was found across languages (i.e., there was no child with a deficit of vocabulary in all three languages, three with a deficit in two languages, and four with a deficit in one language (all English)). Deficits in two languages were: one Arabic-Turkish, two English-Turkish, no child had a deficit in English and Arabic.

Vocabulary

For vocabulary, six poor readers had deficits in English, two of these children also had mild deficits in Turkish, two additional children had mild deficits in Turkish but no deficit in English, and one of these had a moderate deficit in Arabic as well. Of the six children with a vocabulary deficit in English, all showed deficits of word and nonword reading in English. Five further children with no evidence of a vocabulary deficit in English showed deficits of reading words and nonwords in English, and one (PR8G5) showed a deficit of exception word reading.

For Arabic only PR4G4 showed a deficit in vocabulary and this child showed a weakness in Arabic nonword reading. An additional three children with no vocabulary deficit showed a weakness in Arabic nonword reading and a further three showed a weakness in Arabic word and nonword reading. For Turkish, all four children who showed a deficit in vocabulary also showed a weakness in Turkish word and nonword reading. An additional four children showed a weakness in Turkish word and nonword reading but not Turkish vocabulary, and a further two children showed a weakness in either Turkish word or nonword reading but not Turkish vocabulary.

PA

For PA, there was little correspondence across languages, however, most deficits were again in English (9/12 PR children). Two of these children also had deficits in Turkish, and two other children had deficits of PA in all three languages. Of the nine children with PA deficits in English, eight had deficits of word and nonword reading in English while one (PR8G5) showed a deficit in exception word reading accuracy only. For Arabic, two children, showed a deficit in PA, one of these children had a deficit in Arabic nonword reading time. For Turkish, all four children who showed a deficit in PA also showed a weakness in Turkish word and nonword reading.

RAN

For RAN, most deficits were in Turkish (N=8) followed by English (N= 6), then Arabic (N=5) and little correspondence across languages was found. That is, two children had a RAN deficit in all three languages, five children had a deficit in two languages - two deficits in Arabic-Turkish, one English-Turkish, and two English-Arabic. Also, three children had a RAN deficit in only one language. Of the six children with a RAN deficit in English, five showed deficits of word and nonword reading in English, and one (PR8G5) showed a deficit of exception word reading only. For Arabic, five children, showed a deficit in RAN, three of whom had deficits in Arabic nonword reading, one a deficit in Arabic word and nonword reading and one no evidence of reading deficits in Arabic. For Turkish, of the eight poor readers with RAN deficits, all showed weakness in Turkish word and nonword reading.

4.3 Discussion

This chapter reported the performance of twelve trilingual poor readers (PR) aged nine to eleven years in reading and spelling and literacy-related variables (LRVs) in English, Arabic, and Turkish. The performance of the PR group was first compared to the performance of a comparison group (CG) of 20 trilingual TD children. The CG was created to be matched as closely as possible to the PRs on age and vocabulary level, in addition to being from the same school as the PR group.

Individual profiles of each of the poor readers were examined. As in the coverage of the results for the TD children in Chapter 3, findings for LRVs are covered first, then those for reading, and finally for spelling.

Findings for literacy-related processes

For both groups, the vocabulary scores were highest for Arabic, followed by English and finally Turkish. These results are in line with the language background of the children for whom Arabic is the language of birth, English is the second language (main language of instruction), and Turkish is an additional language of instruction. Nevertheless, the reading scores (see Table 31) showed that for CG Arabic was read least accurately and slowest, English was read the most accurately, and fastest, and Turkish was between the other two. The reading accuracy scores for the PR were similar in English and in Arabic but the reading times for Arabic were slower, and Turkish was the least accurate and the slowest type of letter strings read by the PR.

The scores of the PR group in PA reflected worse performance than the performance of the CG. For deletion, the PR group performed better in Arabic, followed by Turkish then English, while the CG performed best in Turkish, followed by Arabic, then English. Blending Arabic words was the most accurate in both groups, then English followed by Turkish in the PR while blending English and Turkish were at the same level for the CG.

For RAN, both CG and PR groups showed the same pattern, i.e., times were fastest in English and slowest in Arabic. For Arabic and Turkish, times were faster for RAN letters than digits, while the opposite was the case for English, likely due to the correspondence of names and sounds for letters in Arabic and Turkish.

Findings for reading

The results for reading accuracy revealed a significant difference between the two groups in all types of letter strings in favour of the CG. The English regular words were read more accurately than exception words and nonwords by the CG and by the PR group. For Arabic, words were read more accurately than nonwords in the comparison and the PR group, and in Turkish, words were also read more accurately than nonwords in both groups.

Overall, English regular words were found to be the most accurately read by the CG and the PR group. However, the Arabic words were almost read as accurately as regular English words by the PR group. English exception words were the least accurately read by the CG and by the PR groups. For nonwords, the two groups showed different performance. English nonwords were the most accurate for the CG, while Turkish nonwords were the most accurate by the PR group. Still, Arabic nonwords were the least accurately read by both groups.

For reading times, English exception words were read faster than the other two types of letter string by the CG and PR groups. Regular words were read faster than nonwords by the CG, and Arabic and Turkish words were read faster than nonwords by this group. For the PR group, however, none of these differences was significant. This was interpreted as a deficit in the PR group as a whole in at least some aspect of lexical processes. The fact that the PR group also showed significantly lower levels of accuracy than the CG in reading nonwords (as well as words) across the three languages was interpreted as indicating, in addition deficits in the PR group as a whole in sublexical processing.

Findings for spelling

The results for spelling accuracy showed that for English, nonwords were spelled more accurately than real words. This was the same for PR and CG, but with more accurate performance in the CG. For Arabic, words were spelled more accurately than nonwords by both groups and for Turkish, the PR group spelled nonwords more accurately than words, while the CG spelled the words more accurately than nonwords. In English and Turkish spelling but not Arabic the CG were more accurate than the PR group.

When the performance in spelling accuracy for real words was compared across the three languages, the results revealed that for the PR group, Arabic words were the most accurately spelled, which is a finding that is in line with the relatively high level of Arabic vocabulary knowledge for this group. English exception words, on the other hand, were the least accurately spelled type of words. For the CG, Arabic words were spelled most accurately, and English regular and exception words, which were spelled at a comparable level, were the least accurately spelled. For nonwords, the PR group spelled Arabic nonwords most accurately while English nonwords were the least, while the CG spelled Turkish nonwords most accurately and Arabic the least.

Individual Poor Reader Profiles

The individual profiles of PR were examined with the aim of looking for selective deficits of lexical or sublexical processes, as well as association of these across languages, and association with deficits in the LRVs.

In terms of accuracy for words, there was a similar number of children showing weakness in English and Turkish while for nonwords there were far more children with a deficit in English than Turkish. This may be due to the transparency of Turkish which leads to encouragement to rely on sublexical processes and greater proficiency in this reading process, while English opaqueness and Arabic complexities encourage use of lexical processes. The results for reading times would seem to support this interpretation as there were more children showing deficits for Turkish than English, especially for words, indicating that the children are relying on time-consuming sublexical processes.

Turning to the association of deficits in the LRVs with reading profiles, it had been expected that poor nonword reading would be associated with a deficit in PA, while poor word reading would be associated with a deficit in vocabulary or RAN. However, there was little evidence of selective deficits, and mixed support for the expected associations of reading deficits with weaknesses in LRVs.

For English, only one PR child - PR8 - had a selective deficit, and this consisted of a severe deficit for exception word reading accuracy (the other PR children had a mix of word and nonword deficits for English). In terms of LRVs, PR8 had a severe deficit of PA, as well as of RAN. The deficit in PA would be expected to result in a weakness in nonword reading, and this was not observed. The weakness in exception word reading shown by PR8 might also be expected to be associated with slow word reading in Turkish and Arabic, on the basis of studies showing a correspondence of reading profiles across the languages of multilingual poor readers (e.g., Douklias et al., 2009). However, this was not observed for PR8, who showed poor Arabic nonword reading accuracy and slow Turkish nonword reading time.

The literature on the development of reading in Arabic indicated that there may be an association with vocabulary knowledge due to the diglossia and visual complexities of the writing system. I expected poor vocabulary may therefore be associated with poor word reading in Arabic. However, the findings showed that only one PR child – PR4 - showed a deficit in Arabic vocabulary, and this child showed a deficit in Arabic nonword reading accuracy, rather than word reading accuracy. Of the three PR children with slow Arabic word reading – PR6, PR9, PR12 – just one showed a deficit of Arabic RAN. Of the children with weak nonword reading in Arabic PR2, PR4, PR6, PR8, PR12, none had a PA deficit in Arabic.

For Turkish, the expectation was again that slow word reading time would be associated with poor vocabulary or RAN. Most of the PR group had a mix of word and nonword deficits in Turkish (N=9). Eight of the PR group showed slow word reading, and of these, three had deficits of both vocabulary and RAN, and a further four of them had a deficit of RAN only, one PR child had slow word reading but no associated LRV deficit. Five had poor nonword reading accuracy – PR4, PR5, PR9, PR11 and PR12 – of these children three had deficits of PA in Turkish and two did not.

In summary, in terms of selective lexical deficits there was one PR child for English and one (different) child for Turkish. In terms of selective sublexical deficits there were four children who showed this for Arabic and one of these children also had a selective deficit of nonword reading for Turkish. Most of the PR children had weakness in reading both words and nonwords, and deficits in the LRVs did not seem to be systematically related to the reading deficits, nor did the reading deficits seem to be consistent across languages.

The outcome is likely due to the diverse situation of the participants. Although they were all from the same school the number of years of residence in Turkey differed (from 2 to 5) so they had different levels of knowledge of Turkish, as well as English, and they were from different Arabic-speaking regions, and so spoke dialects that had different degrees of overlap with Modern Standard Arabic, which was the medium of instruction and testing. The results may also have been different if different measures had been used, or different means of analysis (for example, different cut-offs for denoting deficits in reading and LRVs).

5 Chapter 5: Discussion

This study aimed to investigate whole-word (or lexical) and sub word phonological (or sublexical) processes in reading in multilingual Arabic-English-Turkish children aged nine to eleven years. Reading processes were examined in relation to selected literacy-related variables (LRVs), namely phonological awareness, rapid automatised naming and vocabulary knowledge, and the following research questions were considered.

1) What are the interrelationships in reading processes across the three languages?

2) What are the patterns of association of LRVs with reading accuracy and reading time for words and nonwords considering the varying degree of transparency between English, Turkish and Arabic? And would the patterns of association be different across the languages?

3) Are the deficits in literacy and reading-related variables in the PR group across their three languages predictable from patterns of association of LRVs found in typically developing children?

4) Would reading profiles observed across the different languages of Arabic-English-Turkish poor readers be comparable to what would be found using the profiling approach used previously for English speaking poor readers.

To address the aims and to answer the research questions the study involved typically developing (TD) children as well as poor readers (PR). The findings from the TD group (N=38) were used to investigate the first two research questions. The findings from the PR group (N=12) were used to address the third and fourth research questions. Examination of the reading and LRV scores for the PR group and TD groups revealed that the scores for the former group were substantially lower for the English BPVS assessment. A subsample of the TD group was therefore selected so that English vocabulary scores were not significantly different from those of the PR group. Accordingly, a smaller comparison group CG (N=20) was chosen from the TD group where the participants were selected to match the PR group based on age and vocabulary scores.

For the first question, I investigated reading accuracy and reading time for words and nonwords in the children's three languages and analysed the relationships among the lexical and sublexical processes. I predicted that there would be a relationship among all three languages for accuracy for nonwords since the predominant process for reading nonwords is sublexical. For real words, I predicted significant correlations between English regular words on one hand and Arabic and Turkish words on the other hand because they should recruit sublexical processes due to writing system transparency and are likely to recruit lexical processes as well, as seen from studies of transparent writing systems showing lexical involvement. Reading regular words was also expected to be associated with reading nonwords across all three languages due to the recruitment of sublexical processes for these. Finally, English exception words were not expected to show correlations with Turkish and Arabic word reading as reading English exception words should tap predominantly lexical processes (although real word reading is thought to involve an element of sublexical processing as well as lexical), and the relation with English regular word reading should be significant due to the shared lexical component of regular and exception word reading.

Findings from the TD group suggested that word and nonword reading processes were associated for English and Turkish and for Arabic and Turkish, but that processes for reading words and nonwords in English and Arabic were, to an extent, independent. This overlap between English and Turkish and Arabic and Turkish might be attributed, as predicted, to the shared features. That is, the Arabic influence on the Turkish vocabulary still exists in the sense that around 30% of the Turkish vocabulary, if not more, are loanwords with an Arabic origin and with a somehow Arabic pronunciation, even though they are written in Latin script and share no similarity in grammatical rules. English, on the other hand, shares features with the Turkish writing system, such as being written in the Latin script, being read from left to right, having an upper- and lower-case version of each letter, and that it can be written in a cursive as well as print form (more details are in section 1.1.4).

Regarding spelling, scores from the TD children revealed that, for English, the findings were the opposite to the findings for reading, in the sense that, English nonwords were the most accurately spelled type of letter string and exception words were the least accurately spelled. However, for Arabic and Turkish, findings were similar to the findings from reading accuracy - words were spelled more accurately than nonwords. When the performance in spelling was compared across all three languages, and unlike in reading, spelling of Arabic words was the most accurate, followed by spelling Turkish words. English words were the least accurately spelled words, although regular words were spelled slightly more accurately than exception words, which was the least accurately spelled type of real word across all languages. These findings seem to be in line with findings from the Arabic study conducted by Saiegh-Haddad and Taha (2017). They argued that the accurate early spelling of their TD readers was a reflection of the transparent mapping between phonemes and graphemes. Also, in another Arabic study, Taha's (2016b) findings from the sixth grade suggested that spelling performance exceeded reading. However, findings from spelling nonwords across the three languages in the present study were like those from reading, in that English nonwords were the most accurately spelled and Arabic nonwords were the least accurately spelled.

For the second research question, I expected different patterns of associations of LRVs with reading accuracy and reading time for words and nonwords across the children's three languages. PA (as a sublexical marker) should be associated with letter strings according to their transparency, and hence I did not expect PA to be related to reading accuracy for exception words in English. On the other hand, I expected vocabulary knowledge, (as a lexical marker) to be strongly associated with reading accuracy and reading times for both English exception words and with Arabic word reading. Vocabulary knowledge was expected to be least strongly associated with nonword reading in all three languages, and with Turkish word reading accuracy, since the script is the most transparent of the three target languages. As for RAN, I expected it to be associated with reading times for all letter strings across the three languages and that it would, in particular, be associated with reading accuracy for English exception words and Arabic words, but less so with English regular words and Turkish words. This prediction was based on findings from previous studies in Greek (a transparent orthography) where RAN seemed to be not related to accuracy but rather to reading fluency (Papadopoulos, Spanoudis & Georgiou, 2016). Furthermore, I did not expect RAN to have a relation with accuracy for nonwords in any of the three languages.

The results from the TD children showed that for Turkish, word reading was associated with vocabulary, as well as PA and RAN, while for reading nonwords, PA was the only significantly associated variable. My prediction that English nonword reading would be strongly associated with the sublexical marker PA was borne out. Moreover, nonword accuracy had significant associations with vocabulary and RAN.

The significant association with vocabulary seemed to be attributed to the use of lexical processes, as well as sublexical, when it comes to English reading (due to the presence of exception words). Furthermore, the fastest reading times were observed for exception words, and the correlation analyses revealed that the strongest association for this letter string type was with vocabulary, a lexically-related variable. Reading English exception words is likely to draw on semantics because of the poor reliability of sublexical processes for providing correct responses, as noted earlier. Reading English exception words was also significantly associated with RAN, in line with the predictions, and this association was the strongest of all the reading-RAN pairings. Additionally, findings that RAN was associated with word and nonword reading in English and Arabic, and with word reading but not nonword reading in Turkish, are in line with findings from studies which found that RAN was a significant predictor of word reading ability, both in terms of accuracy and fluency (Powell and Atkinson, 2020). Rahman (2020) recently reported a study with Turkish speaking children where RAN was associated with both word and nonword reading times, as in different writing systems including German (Moll & Landerl, 2009) and Spanish (Onochie-Quintanilla et al., 2017). Rahman suggests that RAN may be involved both in sublexical processes as well as providing rapid access to the lexicon.

Regular word reading, which draws on both lexical and sublexical processes, showed an association with all three LRVs. Furthermore, for reading nonwords in English and Arabic, vocabulary was significantly associated (as well as PA and RAN). The association of nonword reading and vocabulary in English and Arabic could be due to habitual higher activation of lexical processes in these two languages, due to language-specific features (such as diglossia in Arabic, lack of consistency of GPCs in English), but it might also be due to easier activation of semantics, due to higher levels of receptive vocabulary than in Turkish, or else to higher degrees of word-likeness of the nonwords in Arabic and English compared to Turkish.

Regarding the third research question, about whether the deficits in literacy and LRVs in the PR group across their three languages would be predictable from patterns of association from TD children, the expectations were that for English, which is opaque with a mixture of decodable and non-decodable printed words, we might expect both RAN and vocabulary deficits to be associated with poor reading of exception words, while a PA deficit would be associated with poor nonword reading. In Arabic and Turkish on the other hand, we might expect a lexical reading deficit to be manifest in slow reading speed for words and associated with a vocabulary or RAN deficit, while poor reading of nonwords might be associated with a deficit of PA.

Findings from the PR and CG groups showed that Arabic vocabulary knowledge scores were the highest for the PR, then English and finally Turkish. These results were in line with the language background of the children for whom Arabic was the language of birth, English was the second language (main language of instruction), and Turkish was an additional language of instruction. However, the reading scores showed that for the CG Arabic was read least accurately and slowest, English was read the most accurately, and fastest, and Turkish was between the other two. The reading accuracy scores for the PR were similar in English and in Arabic but the reading times for Arabic were the slowest and Turkish was the least accurately read and the slowest type of letter string for the PR.

The scores of the PR group in PA reflected worse performance than the performance of CG. For RAN, both PR and CG showed the same pattern, i.e., times were fastest in English and slowest in Arabic. Reading English exception words was significantly associated with RAN.

As for the fourth research question, concerning the profiling approach used previously for English speaking poor readers, it was expected, as in the research of, for example, Douklias et al. (2009), that underlying deficits would be consistent across languages. The results from the TD children indicated that Arabic and Turkish reading share processes (due to morphology/vocabulary) while Arabic and English do not seem to, and English nonword reading is associated with Turkish word and nonword reading (due to the same writing script, etc.). Moreover, and as studies on the transfer of skills in multilingual readers with specific learning difficulties showed unexpected discrepancies between abilities across languages of the readers, the transfer of a lexical deficit might be less consistent across languages (Verhoeven, 1994).

To answer this question, two sets of analyses were conducted. One was a group comparison, where the results in reading and LRVs for the poor reader were compared with those of the typically developing readers. Findings indicated that, as a group, the poor readers had weaknesses in both lexical and sublexical reading processes across the three languages. The second set of analyses involved creating individual profiles for the poor readers in terms of deficits in reading words and nonwords and LRVs. Most children had deficits of both word and nonword reading in English and Turkish. There was one child with a selective lexical deficit in English and one with a sublexical deficit in Turkish. For Arabic, three children had deficits of both word and nonword reading and a further four children had sublexical deficits only. There did not appear to be any consistent association of deficits across the children's three languages and nor were there any consistent associations of reading deficits with those in the literacy-related variables.

Limitations of the study and future research

There were several limitations of the present study. First, the PA tasks proved to be relatively easy for the children in the present study. The high level of accuracy observed in the tasks means that the results involving PA must be viewed with some caution. It would be better to administer more complex PA tests, such as Spoonerisms, in future work. It might also be better to use PA tests that focus solely at the level of phonemes (rather than easier word- and syllable-level tasks, especially in deletion and blending). This should have two advantages: phonemes are the level at which sublexical GPCs operate; and a larger range of performance on PA tasks should make clearer the association between PA and reading skills.

A further limitation is that the number of participants was relatively small, due to the extensive nature of the data collection over three languages. Given the scarcity of studies in the area, it was considered important to collect detailed data using a multiple cases approach for the research with the poor readers. Researchers such as 'Shallice' (1988) and 'Yin' (1984) have argued for the case study approach for increasing our understanding of cognitive mechanisms. The method can be particularly useful in areas where there is little prior evidence, as in the present study, since it involves detailed data collection together with replications over time, to provide a solid foundation of knowledge. It was considered important to provide detailed data on a number of measures across languages in the current study to provide a basis for further data collection.

A further consideration relates to the measures used in the study. The aim was to employ measures that were equivalent assessments of the same construct in each language (lexical and sublexical processes, PA, RAN and vocabulary). However, it is often noted in research on multilingual populations that the issue of equivalence of assessments across languages is complex. In studies of language processes, it is difficult to ensure, for example, that words in an assessment in one language are comparable in familiarity with items in an assessment in another language, due to lack of resources such as psycholinguistic databases. In the current study attempts were made to select assessments that would be equivalent across languages, but detailed exploration of the assessments used may indicate variables that need to be taken into consideration. In addition, I focused on reading at the word level only in the present study. The associations of cognitive and linguistic skills to reading at the sentence and text level, should be pursued in future research.

It needs to be noted that the study explored a limited number of literacy-related variables. Research into cognitive and linguistic processes underlying reading indicates a range of variables that are associated with reading apart from those targeted in the present study, for example visual attention span (e.g., Valdois et al. 2011) and morphological knowledge (e.g., Kirby et al., 2012). It would have been informative to have included assessments of these, however the research reporting the association of these variables with reading development and reading difficulties was published, in many cases, after data collection for the current study had taken place (for example, the study of Asadi et al., (2017) showing a relation of reading in Arabic to morphology).

Finally, it was acknowledged that the children were diverse in terms of their experience of the different languages, and their language background, which could make detection of association of deficits particularly difficult. For future work, I would suggest that participants be recruited who are speakers of one dialect or originally from one Arabic-speaking region, rather than a mixed group from different Arabic-speaking regions. Also, for future work, a longitudinal study would be a way to assess the causal relations among different LRVs and lexical and sublexical processes in the different languages. It would also be informative to look at associations with sentence and text level reading performance.

Conclusion

The present study, to the best of the author's knowledge, is the first to examine reading and reading-related processes in trilingual children who speak English, Turkish and Arabic. The group study with the typically developing readers revealed associations across languages indicating that the underlying cognitive mechanisms are shared by, for example, Arabic and Turkish word reading, English and Turkish nonword reading, and that these mechanisms are differentially associated with literacy-related variables (LRVs).

The study with the poor readers involved two sets of analyses. One was a group comparison where the results in reading and LRVs for the PR group were compared with those of TD children. Findings indicated that, as a group, the PR children had weaknesses in both lexical and sublexical reading processes across the three languages. The second set of analyses involved creating individual profiles for the poor readers in terms of deficits in reading words and nonwords and LRVs. The results were not in line with expectations (that reading profiles might be shared across languages and that reading deficits would be consistently associated with different LRVs). Potential reasons for the failure to support the predictions were put forward, including the diverse backgrounds of the children and method employed for designating deficits.

The findings are original in addressing reading in three languages that vary across a range of variables – script, reading direction, transparency. English and Turkish share script but not transparency, and Turkish and (vowelised) Arabic share transparency but not script or writing direction. These features appeared to lead to differences in associations of lexical and sublexical processes, and association with LRVs. Findings such as these are informative for theories of reading and reading development. The more we understand how features of different writing systems affect underlying reading processes the more informed these theories will be. We also saw that the poor readers in the study showed varying levels of reading deficit and weaknesses in LRVs, with relatively lesser degrees of deficit in Arabic, the children's first language. While the results for the individual profiles did not support the predictions set out concerning associations it may still be the case that assessment of lexical and sublexical abilities for reading and of LRVs will be helpful for remediation purposes for children such as the poor readers in the study. Results in terms of severity of deficit can be used to guide decisions about which type of intervention will be appropriate. It will be important, though, to increase the scope of assessment to cover more recently identified potential underlying deficits.

Word Count: 46400

References:

Abu-Rabia, S., & Shakkour, W. (2014). Cognitive Retroactive Transfer (CRT) of Language Skills among Trilingual Arabic-Hebrew and English Learners. *Open Journal of Modern Linguistics*, *04*(01), 1–20. https://doi.org/10.4236/ojml.2014.41001

Abu-Rabia, S. (2001). The role of vowels in reading Semitic scripts: Data from Arabic and Hebrew. *Reading and Writing*, *14*(1), 39-59.

Adams, M.J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: Massachusetts Institute of Technology Press.

Alves, L. M., Siqueira, C. M., Ferreira, M. do C. M., Alves, J. F. M., Lodi, D. F., Bicalho, L., & Celeste, L. C. (2016). Rapid Naming in Brazilian Students with Dyslexia and Attention Deficit Hyperactivity Disorder. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00021

Asaad, H., & Eviatar, Z. (2013). The effects of orthographic complexity and diglossia on letter naming in Arabic: A developmental study. *Writing Systems Research*, *5*(2), 156-168.

Asadi, I. A., Khateb, A., Ibrahim, R., & Taha, H. (2017). How do different cognitive and linguistic variables contribute to reading in Arabic? A cross-sectional study from first to sixth grade. *Reading and Writing*, *30*(9), 1835-1867.

Babayiğit, S., & Stainthorp, R. (2007). Preliterate phonological awareness and early literacy skills in Turkish. *Journal of Research in Reading*, *30*(4), 394-413.

Babayiğit, S., & Stainthorp, R. (2010). Component processes of early reading, spelling, and narrative writing skills in Turkish: A longitudinal study. *Reading and Writing*, *23*(5), 539-568.

Babayiğit, S., & Stainthorp, R. (2011). Modeling the relationships between cognitive–linguistic skills and literacy skills: New insights from a transparent orthography. *Journal of Educational Psychology*, *103*(1), 169–189. https://doi.org/10.1037/a0021671

Besner, D., Twilley, L., McCann, R. S., & Seergobin, K. (1990). On the association between connectionism and data: Are a few words necessary? *Psychological Review*, *97*(3), 432–446. https://doi.org/10.1037/0033-295X.97.3.432

Bingöl, A. (2003). Ankara'daki ilkokul 2. ve 4. sınıf öğrencilerinde gelişimsel disleksi oranı. *Ankara Üniversitesi Tıp Fakültesi Mecmuası*, *56*(2).

Boets, B., Wouters, J., Van Wieringen, A., & Ghesquiere, P. (2007). Auditory processing, speech perception and phonological ability in pre-school children at high-risk for dyslexia: A longitudinal study of the auditory temporal processing theory. *Neuropsychologia*, *45*(8), 1608-1620.

Bowers, P. G., & Newby-Clark, E. (2002). The role of naming speed within a model of reading acquisition. *Reading and Writing*, *15*(1), 109-126.

Bradley, L., & Bryant, P. E. (1983). Categorizing sounds and learning to read—a causal connection. *Nature*, *301*(5899), 419-421.

Campbell, R., & Butterworth, B. (1985). Phonological dyslexia and dysgraphia in a highly literate subject: A developmental case with associated deficits of phonemic processing and awareness. *The Quarterly Journal of Experimental Psychology Section A*, *37*(3), 435-475.

Caravolas, M., Bruck, M., & Genesee, F. (2003). Similarities and differences between English-and French-speaking poor spellers. In N. Goulandris (Ed.), *Dyslexia in different languages: Cross-linguistic comparisons* (pp. 157–180). Whurr Publishers.

Caravolas, M., Lervåg, A., Defior, S., Seidlová Málková, G., & Hulme, C. (2013). Different Patterns, but Equivalent Predictors, of Growth in Reading in Consistent and Inconsistent Orthographies. *Psychological Science*, *24*(8), 1398–1407. https://doi.org/10.1177/0956797612473122

Carroll, J. M., Snowling, M. J., Stevenson, J., & Hulme, C. (2003). The development of phonological awareness in preschool children. *Developmental Psychology*, *39*(5), 913–923. https://doi.org/10.1037/0012-1649.39.5.913

Castles, A., & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition*, *91*(1), 77–111. https://doi.org/10.1016/S0010-0277(03)00164-1

Catts, H. W. (1996). Defining dyslexia as a developmental language disorder: An expanded view. *Topics in Language Disorders*, *16*(2), 14–29. https://doi.org/10.1097/00011363-199602000-00004

Chow, B. W. Y., McBride-Chang, C., & Burgess, S. (2005). Phonological processing skills and early reading abilities in Hong Kong Chinese kindergarteners learning to read English as a second language. *Journal of educational psychology*, *97*(1), 81.

Cline, T. (2000). Multilingualism and dyslexia: Challenges for research and practice. *Dyslexia*, *6*(1), 3-12.

Coltheart, M., Masterson, J., Byng, S., Prior, M., & Riddoch, J. (1983). Surface dyslexia. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *35*(3), 469–495. https://doi.org/10.1080/14640748308402483

Coltheart, M. (1978). Lexical Access in simple reading tasks. In G. Underwood (Ed.), Strategies if information processing (p.p. 151-216). San Diego, CA: Academic Press

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204–256. https://doi.org/10.1037/0033-295X.108.1.204

Comeau, L., Cormier, P., Grandmaison, É., & Lacroix, D. (1999). A longitudinal study of phonological processing skills in children learning to read in a second language. *Journal of Educational Psychology*, *91*(1), 29–43. https://doi.org/10.1037/0022-0663.91.1.29

Da Fontoura, H. A., & Siegel, L. S. (1995). Reading, syntactic, and working memory skills of bilingual Portuguese-English Canadian children. *Reading and Writing*, *7*(1), 139-153.

da Silva, P. B., Engel de Abreu, P. M. J., Laurence, P. G., Nico, M. Â. N., Simi, L. G. V., Tomás, R. C., & Macedo, E. C. (2020). Rapid Automatized Naming and Explicit Phonological Processing in Children With Developmental Dyslexia: A Study With Portuguese-Speaking Children in Brazil. *Frontiers in Psychology*, *11*, 928. https://doi.org/10.3389/fpsyg.2020.00928

Daniels, P. T., & Share, D. L. (2018). Writing System Variation and Its Consequences for

Reading and Dyslexia. *Scientific Studies of Reading*, *22*(1), 101–116. https://doi.org/10.1080/10888438.2017.1379082

de Jong, P. F., & van der Leij, A. (1999). Specific contributions of phonological abilities to early reading acquisition: Results from a Dutch latent variable longitudinal study. *Journal of Educational Psychology, 91*(3), 450–476. https://doi.org/10.1037/0022-0663.91.3.450

de Jong, P. F., & Share, D. L. (2007). Orthographic Learning During Oral and Silent Reading. *Scientific Studies of Reading*, *11*(1), 55–71. https://doi.org/10.1080/10888430709336634

Douklias, S. D., Masterson, J., & Hanley, J. R. (2009). Surface and phonological developmental dyslexia in Greek. *Cognitive Neuropsychology*, *26*(8), 705–723. https://doi.org/10.1080/02643291003691106

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. https://doi.org/10.1044/2015_JSLHR-L-13-0310

Durgunoğlu, A. Y., Nagy, W. E., & Hancin-Bhatt, B. J. (1993). Cross-language transfer of phonological awareness. *Journal of Educational Psychology*, *85*(3), 453.

Ehri, L. C. (1995). Phases of development in learning to read words by sight. *Journal of Research in Reading*, *18*(2), 116–125. https://doi.org/10.1111/j.1467-9817.1995.tb00077.x

Elbeheri, G., & Everatt, J. (2006). Literacy ability and phonological processing skills amongst dyslexic and non-dyslexic speakers of Arabic. *Reading and Writing*, *20*(3), 273–294. https://doi.org/10.1007/s11145-006-9031-0

Everatt, J., Smythe, I., Adams, E., & Ocampo, D. (2000). Dyslexia screening measures and bilingualism. *Dyslexia*, *6*(1), 42-56.

Facoetti, A., Lorusso, M. L., Paganoni, P., Cattaneo, C., Galli, R., Umiltà, C., & Mascetti, G. G. (2003). Auditory and visual automatic attention deficits in developmental dyslexia. *Cognitive Brain Research*, *16*(2), 185–191. https://doi.org/10.1016/S0926-6410(02)00270-7

Frederickson, N., & Frith, U. (1998). Identifying dyslexia in bilingual children: A phonological approach with inner London Sylheti speakers. *Dyslexia*, *4*(3), 119-131.

Friedmann, N., & Haddad-Hanna, M. (2014). Types of Developmental Dyslexia in Arabic. In E. Saiegh-Haddad & R. M. Joshi (Eds.), *Handbook of Arabic Literacy* (Vol. 9, pp. 119–151). Springer Netherlands. https://doi.org/10.1007/978-94-017-8545-7_6

Frith, U. (1985). Beneath the Surface of developmental dyslexia. In *In surface Dyslexia*. J. London. https://curve.coventry.ac.uk/cu/items/e89e56d2-be06-ab12-c817-b564d69e5ed5/1/Frith_13_0836770266.pdf

FRLL, (2012). Forum for Research in Literacy and Language, Institute of Education.

Geva, E. (2008). Facets of metalinguistic awareness related to reading development in Hebrew: Evidence from monolingual and bilingual children. In *Learning to read across languages*(pp. 166-199). Routledge.

Gillon, G. T. (2004). *Phonological awareness: From research to practice*. Guilford Press

Gleason, J. B., & Ratner, N. B. (2009). *The development of language / Jean Berko Gleason, Nan Bernstein Ratner [editors].* (7th ed). Pearson.

Gottardo, A., Yan, B., Siegel, L. S., & Wade-Woolley, L. (2001). Factors related to English reading performance in children with Chinese as a first language: More evidence of cross-language transfer of phonological processing. *Journal of educational psychology*, *93*(3), 530.

Goswami, U., Ziegler, J. C., Dalton, L., & Schneider, W. (2001). Pseudohomophone Effects and Phonological Recoding Procedures in Reading Development in English and German. *Journal of Memory and Language*, *45*(4), 648–664. https://doi.org/10.1006/jmla.2001.2790

Goswami, U. (2010). A psycholinguistic grain size view of reading acquisition across languages. In *Reading and dyslexia in different orthographies* (pp. 41-60). Psychology Press.

Grosjean, F. (2010). *Bilingual*. Harvard university press.

Güven, S., & Friedmann, N. (2019). Developmental Letter Position Dyslexia in Turkish, a Morphologically Rich and Orthographically Transparent Language. *Frontiers in Psychology*, *10*, 2401. https://doi.org/10.3389/fpsyg.2019.02401

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review*, *106*(3), 491–528. https://doi.org/10.1037/0033-295X.106.3.491

Harmer, J. (1991). *The practice of English language teaching*. New York: Longman.

Holmes, J. C. (1973). Site selection by parasitic helminths: interspecific interactions, site segregation, and their importance to the development of helminth communities. *Canadian Journal of Zoology*, *51*(3), 333-347.

Hulme, C., & Snowling, M. J. (2009). *Developmental disorders of language learning and cognition*. Wiley-Blackwell.

Jacobs, K. (2010). Defining Reading Disabilities 1 Defining Reading Disabilities: Multiple Deficits Behind One Problem.

Jiménez González, J. E., & Hernández Valle, I. (2000). Word Identification and Reading Disorders in the Spanish Language. *Journal of Learning Disabilities*, *33*(1), 44–60. https://doi.org/10.1177/002221940003300108

Joanisse, M. F., Manis, F. R., Keating, P., & Seidenberg, M. S. (2000). Language deficits in dyslexic children: Speech perception, phonology, and morphology. *Journal of experimental child psychology*, *77*(1), 30-60.

Karageorgos, P., Müller, B., & Richter, T. (2019). Modelling the relationship of accurate and fluent word recognition in primary school. *Learning and Individual Differences*, *76*, 101779. https://doi.org/10.1016/j.lindif.2019.101779

Katz, L., & Frost, R. (1992). Chapter 4 The Reading Process is Different for Different Orthographies: The Orthographic Depth Hypothesis. In *Advances in Psychology* (Vol. 94, pp. 67–84). Elsevier. https://doi.org/10.1016/S0166-4115(08)62789-2

Kinsbourne, M. (1983). Deep dyslexia. M. Coltheart, K. Patterson, & JC Marshall (Eds.). London: Routledge & Kegan Paul, 1980. Pp. xi+ 444. *Applied Psycholinguistics*, *4*(4), 377-384.

Kirby, J. R., Deacon, S. H., Bowers, P. N., Izenberg, L., Wade-Woolley, L., & Parrila, R. (2012). Children's morphological awareness and reading ability. *Reading and Writing*, *25*(2), 389–410. https://doi.org/10.1007/s11145-010-9276-5

Knoepke, J., Richter, T., Isberner, M.-B., Naumann, J., & Neeb, Y. (2014). Phonological recoding, orthographic decoding, and comprehension skills during reading acquisition. *Zeitschrift Für Erziehungswissenschaft*, *17*(3), 447–471. https://doi.org/10.1007/s11618-014-0560-z

Lafrance, A., & Gottardo, A. (2005). A longitudinal study of phonological processing skills and reading in bilingual children. *Applied Psycholinguistics*, *26*(4), 559–578. https://doi.org/10.1017/S0142716405050307

Landerl, K., & Wimmer, H. (2000). Deficits in phoneme segmentation are not the core problem of dyslexia: Evidence from German and English children. *Applied psycholinguistics*, *21*(2), 243-262.

Landerl, K., & Wimmer, H. (2008). Development of word reading fluency and spelling in a consistent orthography: An 8-year follow-up. *Journal of Educational Psychology*, *100*(1), 150–161. https://doi.org/10.1037/0022-0663.100.1.150

Landerl, Karin, Freudenthaler, H. Harald, Heene, Moritz, De Jong, Peter F, Desrochers, Alain, Manolitsis, George, . . . Georgiou, George K. (2019). Phonological Awareness and Rapid Automatized Naming as Longitudinal Predictors of Reading in Five Alphabetic Orthographies with Varying Degrees of Consistency. *Scientific Studies of Reading, 23*(3), 220-234.

Lervåg, A., Bråten, I., & Hulme, C. (2009). The cognitive and linguistic foundations of early reading development: A Norwegian latent variable longitudinal study. *Developmental Psychology*, *45*(3), 764–781. https://doi.org/10.1037/a0014132

Liberman, I. Y., & Liberman, A. M. (1990). Whole language vs. code emphasis: Underlying assumptions and their implications for reading instruction. *Annals of Dyslexia*, *40*(1), 51-76.

Lundberg, I., & H∅ ien, T. (1990). Patterns of information processing skills and word recognition strategies in developmental dyslexia. *Scandinavian Journal of Educational Research*, *34*(3), 231-240.

Lundberg, I. (2002). Second language learning and reading with the additional load of dyslexia. *Annals of dyslexia*, *52*(1), 165-187.

Manis, F. R., Seidenberg, M. S., & Doi, L. M. (1999). See Dick RAN: Rapid Naming and the Longitudinal Prediction of Reading Subskills in First and Second Graders. *Scientific Studies of Reading*, *3*(2), 129–157. https://doi.org/10.1207/s1532799xssr0302_3

Marshall, J. C., & Newcombe, F. (1973). Patterns of paralexia: A psycholinguistic approach. *Journal of psycholinguistic research*, *2*(3), 175-199.

McCloskey, M., & Caramazza, A. (1988). Theory and methodology in cognitive neuropsychology: A response to our critics. *Cognitive Neuropsychology*, *5*(5), 583-623.

McDowell, K., Carroll, J., & Ziolkowski, R. (2013). *Relations between Phonological Awareness and Vocabulary in Preschoolers*. *1*(2), 10.

Melby-Lervåg, M., Lyster, S.-A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin*, *138*(2), 322–352. https://doi.org/10.1037/a0026744

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. https://doi.org/10.1037/0022-0663.91.1.3

Miller-Guron, L., & Lundberg, I. (2000). Dyslexia and second language reading: A second bite at the apple?. *Reading and Writing*, *12*(1), 41-61.

Mirza, A., Gottardo, A., & Chen, X. (2017). Reading in multilingual learners of Urdu (L1), English (L2) and Arabic (L3). *Reading and Writing*, *30*(1), 187-207.

Moll, K., & Landerl, K. (2009). Double Dissociation Between Reading and Spelling Deficits. *Scientific Studies of Reading*, *13*(5), 359–382. https://doi.org/10.1080/10888430903162878

Moll, K., Ramus, F., Bartling, J., Bruder, J., Kunze, S., Neuhoff, N., Streiftau, S., Lyytinen, H., Leppänen, P. H. T., Lohvansuu, K., Tóth, D., Honbolygó, F., Csépe, V., Bogliotti, C., Iannuzzi, S., Démonet, J.-F., Longeras, E., Valdois, S., George, F., … Landerl, K. (2014). Cognitive mechanisms underlying reading and spelling development in five European orthographies. *Learning and Instruction*, *29*, 65–77. https://doi.org/10.1016/j.learninstruc.2013.09.003

Morfidi, E., Van Der Leij, A., De Jong, P. F., Scheltinga, F., & Bekebrede, J. (2007). Reading in two orthographies: A cross-linguistic study of Dutch average and poor readers who learn English as a second language. *Reading and writing*, *20*(8), 753-784.

Morton, J., & Patterson, K. (1980). A new attempt at an interpretation, or, an attempt at a new interpretation. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), Deep dyslexia. London: Routledge & Kegan Paul. Pp. 91-118. (a)

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. https://doi.org/10.1037/0012-1649.40.5.665

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. https://doi.org/10.1111/j.1467-9817.2004.00238.x

Niolaki, G. Z., Terzopoulos, A. R., & Masterson, J. (2014). Varieties of developmental dyslexia in Greek children. *Writing Systems Research*, *6*(2), 230–256.

https://doi.org/10.1080/17586801.2014.893862

Oktay, A., & Aktan, E. (2002). A Cross-linguistic Comparison of Phonological Awareness and Word Recognition in Turkish and English Comparaison entre la reconnaissance phonologique et la reconnaissance des mots en turc et en anglais Una comparación lingüística cruzada de la Conciencia Fonológica y del reconocimiento de palabras en turco e inglés. *International Journal of Early Years Education*, *10*(1), 37–48. https://doi.org/10.1080/09669760220114836

Onochie-Quintanilla, E., Defior, S., & Simpson, I. C. (2017). Visual multi-element processing as a pre-reading predictor of decoding skill. *Journal of Memory and Language*, *94*, 134–148. https://doi.org/10.1016/j.jml.2016.11.003

Öney, B., & Durgunoğlu, A. Y. (1997). Beginning to read in Turkish: A phonologically transparent orthography. *Applied Psycholinguistics*, *18*(01), 1. https://doi.org/10.1017/S014271640000984X

Ouellette, G., & Beers, A. (2010). *A not-so-simple view of reading: How oral vocabulary and visual-word recognition complicate the story*. 20.

Papadopoulos, T. C., Spanoudis, G. C., & Georgiou, G. K. (2016). How Is RAN Related to Reading Fluency? A Comprehensive Examination of the Prominent Theoretical Accounts. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01217

Patel, T. K., Snowling, M. J., & de Jong, P. F. (2004). A Cross-Linguistic Comparison of Children Learning to Read in English and Dutch. *Journal of Educational Psychology, 96*(4), 785–797. https://doi.org/10.1037/0022-0663.96.4.785

Peterson, R. L., Arnett, A. B., Pennington, B. F., Byrne, B., Samuelsson, S., & Olson, R. K. (2018). Literacy acquisition influences children's rapid automatized naming. *Developmental Science*, *21*(3), e12589. https://doi.org/10.1111/desc.12589

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*(1), 56–115. https://doi.org/10.1037/0033-295X.103.1.56

Powell, D., & Atkinson, L. (2020). Unraveling the links between rapid automatized naming (RAN), phonological awareness, and reading. *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000625

Pratt, A. C., & Brady, S. (1988). Relation of phonological awareness to reading disability in children and adults. *Journal of Educational Psychology*, *80*(3), 319–323. https://doi.org/10.1037/0022-0663.80.3.319

Rack, J., Hulme, C., Snowling, M., & Wightman, J. (1994). The Role of Phonology in Young Children Learning to Read Words: The Direct-Mapping Hypothesis. *Journal of Experimental Child Psychology*, *57*(1), 42–71. https://doi.org/10.1006/jecp.1994.1003

Rahbari, N., & LeFevre, J.-A. (2015). *Effects of Orthographic Depth on Developmental Dyslexia in Different Languages*. Unpublished. https://doi.org/10.13140/rg.2.1.1709.4565

Rahman, E. (2020) Unpublished PhD thesis. Brunel University

Ram-Tsur, R., Faust, M., & Zivotofsky, A. Z. (2008). Poor Performance on Serial Visual Tasks in Persons With Reading Disabilities: Impaired Working Memory? *Journal of Learning Disabilities*, *41*(5), 437–450. https://doi.org/10.1177/0022219408321141

Rayner, K., & Reichle, E. D. (2010). Models of the reading process. *Wiley Interdisciplinary Reviews. Cognitive Science*, *1*(6), 787–799. https://doi.org/10.1002/wcs.68

Ricketts, J., Nation, K., & Bishop, D. V. M. (2007). Vocabulary Is Important for Some, but Not All Reading Skills. *Scientific Studies of Reading*, *11*(3), 235–257. https://doi.org/10.1080/10888430701344306

Rowse, H. J., & Wilshire, C. E. (2007). Comparison of phonological and whole-word treatments for two contrasting cases of developmental dyslexia. *Cognitive Neuropsychology*, *24*(8), 817–842. https://doi.org/10.1080/02643290701764207

Saiegh-Haddad, E., & Henkin-Roitfarb, R. (2014). The structure of Arabic language and orthography. In *Handbook of Arabic literacy* (pp. 3-28). Springer, Dordrecht.

Saiegh-Haddad, E., & Geva, E. (2008). Morphological awareness, phonological awareness, and reading in English–Arabic bilingual children. *Reading and Writing*, *21*(5), 481–504. https://doi.org/10.1007/s11145-007-9074-x

Saiegh-Haddad, E., & Taha, H. (2017). The Role of Morphological and Phonological Awareness in the Early Development of Word Spelling and Reading in Typically Developing and Disabled Arabic Readers: Arabic Reading and Spelling in Typical and Disabled Readers. *Dyslexia*, *23*(4), 345–371. https://doi.org/10.1002/dys.1572

Savage, R., Pillay, V., & Melidona, S. (2007). Deconstructing rapid automatized naming: Component processes and the prediction of reading difficulties. *Learning and Individual Differences*, *17*(2), 129–146. https://doi.org/10.1016/j.lindif.2007.04.001

Savage, R., Pillay, V., & Melidona, S. (2008). Rapid serial naming is a unique predictor of spelling in children. *Journal of learning disabilities*, *41*(3), 235-250.

Schmalz, X., Robidoux, S., Castles, A., Coltheart, M., & Marinus, E. (2017). German and English Bodies: No Evidence for Cross-Linguistic Differences in Preferred Orthographic Grain Size. *Collabra: Psychology*, *3*(1), 5. https://doi.org/10.1525/collabra.72

Schwartz, M., Geva, E., Share, D. L., & Leikin, M. (2007). Learning to read in English as third language: The cross-linguistic transfer of phonological processing skills. *Written Language & Literacy*, *10*(1), 25-52.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*(4), 523–568. https://doi.org/10.1037/0033-295X.96.4.523

Silvén M., Poskiparta, E., Niemi, P., & Voeten, M. (2007). Precursors of reading skill from infancy to first grade in Finnish: Continuity and change in a highly inflected language. *Journal of Educational Psychology*, *99*(3), 516–531. https://doi.org/10.1037/0022-0663.99.3.516

Serrano, F., & Defior, S. (2008). Dyslexia speed problems in a transparent orthography. *Annals of Dyslexia*, *58*(1), 81–95. https://doi.org/10.1007/s11881-008-0013-6

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*(2), 143–174. https://doi.org/10.1348/000712603321661859

Shallice, T. (1988). From neuropsychology to mental structure. Cambridge, UK: Cambridge University Press

Shallice, T. (1981). Phonological agraphia and the lexical route in writing. *Brain: a journal of neurology*, *104*(3), 413-429.

Share, D. L. (2008). On the Anglocentricities of current reading research and practice: The perils of overreliance on an ‘outlier’ orthography. *Psychological Bulletin*, *134*(4), 584–615. https://doi.org/10.1037/0033-2909.134.4.584

Shatil, E., & Share, D. L. (2003). Cognitive antecedents of early reading ability: A test of the modularity hypothesis. *Journal of Experimental Child Psychology*, *86*(1), 1–31. https://doi.org/10.1016/S0022-0965(03)00106-1

Smythe, I., Salter, R., & Everatt, J. (2003). *The international book of dyslexia*. Wiley. http://www.vlebooks.com/vleweb/product/openreader?id=none&isbn=9780470862865

Snowling, M. J. (2001). From language to reading and dyslexia. *Dyslexia*, *7*(1), 37–46.

https://doi.org/10.1002/dys.185

Snowling, M. J., Gallagher, A., & Frith, U. (2003). Family Risk of Dyslexia Is Continuous: Individual Differences in the Precursors of Reading Skill. *Child Development*, *74*(2), 358–373. https://doi.org/10.1111/1467-8624.7402003

Sotiropoulos, A., & Hanley, J. R. (2017). Developmental surface and phonological dyslexia in both Greek and English. *Cognition*, *168*, 205–216. https://doi.org/10.1016/j.cognition.2017.06.024

Stanovich, K. E., & Siegel, L. S. (1994). Phenotypic performance profile of children with reading disabilities: A regression-based test of the phonological-core variable-difference model. *Journal of Educational Psychology*, *86*(1), 24–53. https://doi.org/10.1037/0022-0663.86.1.24

Stuart, M., & Howard, D. (1995). KJ: A developmental deep dyslexic. *Cognitive Neuropsychology*, *12*(8), 793-824.

Swanson, H. L., Trainin, G., Necoechea, D. M., & Hammill, D. D. (2003). Rapid Naming, Phonological Awareness, and Reading: A Meta-Analysis of the Correlation Evidence. *Review of Educational Research*, *73*(4), 407–440. https://doi.org/10.3102/00346543073004407

Taha, H. (2016a). Deep and shallow in Arabic orthography: New evidence from reading performance of elementary school native Arab readers. *Writing Systems Research*, *8*(2), 133–142. https://doi.org/10.1080/17586801.2015.1114910

Taha, H. (2016b). The Development of Reading and Spelling in Arabic Orthography: Two Parallel Processes? *Reading Psychology*, *37*(8), 1149–1161. https://doi.org/10.1080/02702711.2016.1193580

Tahan, S., Cline, T., & Messaoud-Galusi, S. (2011). The relationship between language dominance and pre-reading skills in young bilingual children in Egypt. *Reading and Writing*, *24*(9), 1061-1087.

Tainturier, M.-J., Roberts, J., & Charles Leek, E. (2011). Do reading processes differ in transparent versus opaque orthographies? A study of acquired dyslexia in Welsh/English bilinguals. *Cognitive Neuropsychology*, *28*(8), 546–563. https://doi.org/10.1080/02643294.2012.698986

Temple, C. M., & Marshall, J. C. (1983). A case study of developmental phonological dyslexia. *British Journal of psychology*, *74*(4), 517-533.

Tibi, S., & Kirby, J. R. (2018). Investigating Phonological Awareness and Naming Speed as Predictors of Reading in Arabic. *Scientific Studies of Reading*, *22*(1), 70–84. https://doi.org/10.1080/10888438.2017.1340948

Tobia, V., & Marzocchi, G. M. (2014). Predictors of reading fluency in Italian orthography: Evidence from a cross-sectional study of primary school students. *Child Neuropsychology*, *20*(4), 449–469. https://doi.org/10.1080/09297049.2013.814768

Treiman, R., Mullennix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology: General*, *124*(2), 107.

Valdés, G., & Figueroa, R. A. (1994). *Bilingualism and testing: A special case of bias*. Ablex Publishing.

Valdois, S., Bidet-Ildei, C., Lassus-Sangosse, D., Reilhac, C., N'guyen- Morel, M.-A., Guinet, E., & Orliaguet, J.-P. (2011). A visual processing but no phonological disorder in a child with mixed dyslexia. Cortex, 47, 1197–1218.

Van den Bos, K. P. (1998). IQ, phonological awareness and continuous-naming speed related to Dutch poor decoding children's performance on two word identification tests. *Dyslexia*, *4*(2), 73-89.

Vaessen, A., Bertrand, D., Tóth, D., Csépe, V., Faísca, L., Reis, A., & Blomert, L. (2010). Cognitive development of fluent word reading does not qualitatively differ between transparent and opaque orthographies. *Journal of Educational Psychology, 102*(4), 827–842. https://doi.org/10.1037/a0019465

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry*, *45*, 2–40. https://doi.org/10.1046/j.0021-9630.2003.00305.x.

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. https://doi.org/10.1080/10888430709336632

Verhoeven, L. T. (1994). Transfer in Bilingual Development: The Linguistic Interdependence Hypothesis Revisited. *Language Learning*, *44*(3), 381–415. https://doi.org/10.1111/j.1467-1770.1994.tb01112.x

Verhoeven, L. (2007). Early bilingualism, language transfer, and phonological awareness. *Applied Psycholinguistics*, *28*(3), 425-439.

Wade-Woolley, L., & Geva, E. (2000). Processing novel phonemic contrasts in the acquisition of L2 word reading. *Scientific studies of reading*, *4*(4), 295-311.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, *101*(2), 192–212. https://doi.org/10.1037/0033-2909.101.2.192

Wagner, R. K., Torgesen, J. K., Laughon, P., Simmons, K., & et al. (1993). Development of young readers' phonological processing abilities. *Journal of Educational Psychology*, *85*(1), 83–103. https://doi.org/10.1037/0022-0663.85.1.83

Wallach, M., & Wallach, L. (1976). *Teaching all children to read / Michael A. Wallach, Lise Wallach.* Chicago; London: University of Chicago Press.

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*(1/2), 5–20. https://doi.org/10.1023/A:1021789804977

Wimmer, H. (1993). Characteristics of developmental dyslexia in a regular writing system. *Applied psycholinguistics*, *14*(1), 1-33.

Wimmer, H., Landerl, K., Linortner, R., & Hummer, P. (1991). The relationship of phonemic awareness to reading acquisition: More consequence than precondition but still important. *Cognition*, *40*(3), 219–249. https://doi.org/10.1016/0010-0277(91)90026-Z

Wile, T. L., & Borowsky, R. (2004). What does rapid automatized naming measure? A new RAN task compared to naming and lexical decision. *Brain and Language*, *90*(1–3), 47–62. https://doi.org/10.1016/S0093-934X(03)00419-X

Wimmer, H., Mayringer, H., & Landerl, K. (2000). The double-deficit hypothesis and difficulties in learning to read a regular orthography. *Journal of educational psychology*, *92*(4), 668.

Winskel, H. (2011). Orthographic and phonological parafoveal processing of consonants, vowels, and tones when reading Thai. *Applied Psycholinguistics, 32*(4), 739–759. https://doi.org/10.1017/S014271641100004X

Wolf, M., & Katzir-Cohen, T. (2001). Reading fluency and its intervention. *Scientific Studies of Reading*, *5*(3), 211–239. Scopus.

Yin, R.K., (1984). Case Study Research: Design and Methods. Beverly Hills, Calif: Sage Publications.

Zeguers, M., Boer, M., Snellings, P., & Jong, P. (2018). Universal and Language-Specific Predictors of Early Word Reading in a Foreign Language: An Analysis of the Skills That Underlie Reading Acquisition in Three Different Orthographies. *Developmental Psychology*, *54*. https://doi.org/10.1037/dev0000606

Ziegler, J. C., Bertrand, D., Tóth, D., Csépe, V., Reis, A., Faísca, L., Saine, N., Lyytinen, H., Vaessen, A., & Blomert, L. (2010). Orthographic Depth and Its Impact on Universal Predictors of Reading: A Cross-Language Investigation. *Psychological Science*, *21*(4), 551–559. https://doi.org/10.1177/0956797610363406

Ziegler, J. C., & Goswami, U. (2005). Reading Acquisition, Developmental Dyslexia, and Skilled Reading Across Languages: A Psycholinguistic Grain Size Theory. *Psychological Bulletin*, *131*(1), 3–29. https://doi.org/10.1037/0033-2909.131.1.3

Ziegler, J. C., Perry, C., & Coltheart, M. (2000). The DRC model of visual word recognition and reading aloud: An extension to German. *European Journal of Cognitive Psychology*, *12*(3), 413–430. https://doi.org/10.1080/09541440050114570

Ziegler, J. C., Perry, C., Ma-Wyatt, A., Ladner, D., & Schulte-Körne, G. (2003). Developmental dyslexia in different languages: Language-specific or universal?. *Journal of experimental child psychology*, *86*(3), 169-193.

Zoccolotti, P., de Luca, M., di Pace, E., Judica, A., Orlandi, M., & Spinelli, D. (1999). Markers of developmental surface dyslexia in a language (Italian) with high grapheme–phoneme correspondence. *Applied Psycholinguistics, 20*(2), 191–216. https://doi.org/10.1017/S0142716499002027