The relationship between nonword repetition, root and pattern effects, and vocabulary in Gulf Arabic speaking children

Mariam Abdel Khaleq Yousef Khater

This thesis is submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at City, University of London

Department of Language and Communication Science

City, University of London

October, 2016

Table of Contents

Tables

Figures

Acknowledgements

I would like to express my sincere gratitude to the people who supported me during my research journey. I could not have done my doctoral research without their generous support.

First, I would like to express my sincere gratitude to my supervisor Professor Shula Chiat who supervised my research for many years and who always provided her insight and valuable comments that opened new horizons to me and guide me marvellously throughout my PhD

I would like to extend my heartfelt thanks and gratitude to my supervisor Dr. Rachael-Anne Knight. Her unconditional support, patience, motivation, immense knowledge and above all, belief in me were invaluable. I could not have imagined having a better advisor and mentor for my PhD.

I would like also to thank Dr. Chloe Marshal for her guidance and help in the early stage of my project. My friend and colleague Dr. Outi Tuomainen deserves special thanks for her continuous help and support throughout different parts of the projects.

I would love to express my great gratitude to my mother and father, who taught me to have a dream and belief in myself; without their support, I would not be who I am now. A great acknowledgment is due to my mother in law and my family and friends around the world.

Finally, I would love to thank my husband Saleh Shaalan for his endless love and support and to my children who always filled my heart with happiness: my son Bashir and my daughters Layla and Alyaa. To whom I dedicate this thesis.

Declaration

I, Mariam Khater, grant powers of discretion to the University Librarian to allow this thesis to be copied in whole or in part without further reference to the author. This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgement.

Abstract

Nonword repetition has received great attention in the last three decades due to its ability to distinguish between the performance of children with language impairment and their typically developing peers and due to its correlation with variety of language abilities, especially vocabulary skills.

This study investigates early phonological skills, as represented by nonword repetition (NWR), in TD Gulf Arabic speaking children and those with language impairment and tries to examine findings in relation to two important NWR hypotheses, namely the phonological short term memory account (PSTM, Gathercole& Baddeley, 1990a) and the linguistic account of Snowling, Chiat & Hulme (1991).

In the first experiment, a new Arabic word and nonword test (WNRep) was developed and conducted with 44 TD children and a clinical group (CL) that consisted of 15 children with language impairment. The participants' ages were between two and four years old. The results show that the TD group scored significantly higher than the CL group on the WNRep and across one, two and three syllable words/nonwords and that NWR scores correlated significantly with receptive and expressive vocabulary tests. Apart from its ability to differentiate between TD and those with language impairment, NWR results revealed significant differences in groups' performance even on one syllable word and nonwords, which differs from findings in other languages.

These results raise questions about whether these findings relate to the characteristic root and pattern morphology in Arabic. Therefore, the second experiment in chapter 5 was conducted to investigate the effects of roots and patterns on TD children's repetition skills and their relation to receptive and expressive vocabulary tests. A root and pattern nonword repetition test (RAP-NWR) was developed to measure this effect. The RAP-NWR consisted of three different types of root and pattern combinations (real root and nonpattern nonwords, real pattern and nonroot nonwords and nonpattern and nonroot nonwords). All 89 participants were TD Gulf Arabic speaking children aged two to seven years old and divided into six age bands. Results showed that these children's repetitions were sensitive to the presence of roots but not patterns and that RAP-NWR scores were significantly correlated with both vocabulary tests.

Findings from both studies show that while phonological storage may explain some of the results of children's performance on NWR, there are a myriad of phonological and morphological factors that could have significant effects on NWR, such as the effects of roots and patterns, and it seems that roots more important role to play as it roots awareness emerges earlier than pattern awareness. Based on these findings, clinical utility of root and pattern NWR tests is discussed and further investigations of effects of roots and patterns on NWR are recommended.

Abbreviations

AEVT	Arabic Expressive Vocabulary test
APVT	Arabic Picture Vocabulary Test
ASD	Autistic Spectrum Disorder
CL	Children with Language Impairment
CNRep	The Children's Test of Nonword Repetition
DSM-5	Diagnostic and Statistical Manual of Mental Disorders
DV	Dependent Variable
EL	Expressive Language
EOWPVT	Expressive One Word Picture Vocabulary Test
IQ	Intelligent Quotient
IV	Independent Variable
LTM	Long Term Memory
MSA	Modern Standard Arabic
NR-NP	Non Root –Non Pattern
NR-P	Non Root-Pattern
NRT	Nonword Repetition Test
NTM	Nonconcatenating Templatic Morphology
NWR	Nonword Repetition
OCP	Obligatory Contour Principle
OME	Otitis Media with Effusion
OT	Optimality Theory
PLS-3	Preschool Language Scale-3
PPC	Percentage of Phonemes Correct
PPVT	Peabody Picture Vocabulary Test
PSRep	PreSchool Repetition Test

PSTM	Phonological Short Term Memory
RAP-NWR	Root And Pattern Non Word Repetition
R-NP	Root Non-Pattern
SC	Sentence Comprehension
SLI	Specific Language Impairment
SNHL	Sensory Neural Hearing Loss
SR	Sentence Repetition
STM	Short Term Memory
TD	Typically Developing
TDAM	Typically Developing Age Match
TDY	Typically Developing Younger children
WM	Working Memory
WNRep	Word Nonword Repetition
WWC	Whole Word Correct

International Phonetic Alphabet (IPA) symbols for Arabic consonants

Arabic letter	IPA symbols
ء	ʔ
ب	b
ت	t
ث	θ
ج	ʤ
ح	ħ
خ	x
د	d
ذ	ð
ر	r
ز	z
س	s
ش	ʃ
ص	sˤ
ض	dˤ
ط	tˤ
ظ	ðˤ
ع	ʕ
غ	ʁ

ف	f
ق	q
ك	k
ل	l
م	m
ن	n
ه	h
و	w
ي	j

1. Introduction

Interest in nonword repetition as an assessment tool has developed considerably over the last three decades. This interest stems from Gathercole and Baddeley's (1989, 1990a) findings that children's performance on nonword repetition (NWR) correlates with various language abilities. The ability to repeat nonwords is considered an important predictor of language learning, especially during the early stages of language development (Gathercole, 2006).

A large number of studies that investigated nonword repetition in typically developing children and in children with language impairment showed that the latter group have a significant difficulty with this task (Archibald & Gathercole, 2006; Bishop et al., 1996; Chiat & Roy, 2007; Conti-Ramsden, 2003; Dollaghan & Campbell, 1998; Edwards & Lahey, 1998; Ellis Weismer et al., 2000; Gathercole, 2006; Gathercole & Baddeley, 1990; Gray, 2003; Montgomery, 1995, 2002; Munson, Edwards, & Beckman, 2005; Oetting & Cleveland, 2006; Snowling et al., 1991). It has been found that children with language impairment had more difficulty repeating nonsense words compared with typically developing children. Furthermore, children's performance on NWR could predict later vocabulary development (Gathercole & Baddeley, 1989). Gathercole and Baddeley (1990a) demonstrated that children with language impairment had proportionally more difficultly repeating longer nonwords than shorter ones. Moreover, children with low performance on NWR showed difficulty in measures of productivity of grammatical abilities (Botting& Conti-Ramsden, 2001) and poor sentence comprehension (Montgomery, 1995). These findings encouraged researchers to investigate the nature and underlying processes of nonword repetition.

The phonological short term memory (PSTM) account of nonword repetition (Gathercole & Baddeley, 1989; 1990a; Gathercole et al., 1994) suggests that there is a high correlation between NWR and novel word learning because both are constrained by phonological storage. The PSTM account claims that storage of phonological information in the phonological loop (which is a key component in the working memory model by Baddeley and Hitch, 1974) is critical for learning words (Gathercole et al., 1997; Gathercole & Baddeley, 1990). For example, there is evidence that the phonological loop is responsible for sorting and processing sound combinations and that any deficit in this part would affect word learning. The PSTM account added that nonword repetition is a pure measure of phonological short term memory and they considered the NWR task as a clinical marker of language impairment. However, Snowling, Chiat & Hulme (1991) considered the phonological loop as only the initial process in NWR and there are various other processes that contribute to NWR, such as speech perception, phonological awareness and output processes. Beside these processes, there are other factors that play a role in nonword repetition, such as lexical knowledge and phonotactic probability.

Although there is a vast number of studies that explored phonological short term memory, as measured by nonword repetition and other language measures, the nature of that link is unclear. For example, the link between PSTM and long term memory (as in lexical acquisition) is also unclear. Furthermore, in the case of children with language impairment, there is a debate about whether the deficit in PSTM will cause this impairment or the poor performance on the PSTM task is a result of that impairment (Gathercole, 2006). Nevertheless, NWR tests have become a common tool to measure the ability of repeating novel words and they are used with different populations: children and adults, typically developing individuals and those with language impairment, and across different languages.

This current study attempts to investigate early phonological skills in young Gulf Arabic speaking children with both typical and atypical language development, with special focus on nonword repetition in the light of the unique phonological and morphological characteristics of the Arabic language, especially the presence of root and pattern morphology in Arabic. Findings will be discussed in relation to current theories of NWR, especially phonological short term memory (Gathercole &Baddeley, 1990a) and phonological processing theory (Snowling, Hulme & Chiat, 1991) and word formation in Semitic languages.

Many theories try to explore the underlying processes of word formation in Arabic as a Semitic language. The first account is the morphological processing account, which is based on generative theory (McCarthy, 1979). This account considers the root as an essential morpheme in word formation. The second account is the whole word processing account, which is based on optimality theory (Prince and Smolensky, 1993). The whole word processing account is an output oriented framework that regards the templates, not the root, as the essential part in word formation. The third account proposed by Caramazza et al. (1988) supported mixed or dual route processing. This account is a combination of the whole word and the morphological processing accounts.

The thesis includes two studies; the first one explores NWR in Gulf Arabic speaking children in relation to vocabulary skills and the second study investigates in depth NWR in the light of the Arabic and root and pattern morphology.

Study 1: *Nonword repetition in correlation with receptive and expressive vocabulary skills in Gulf Arabic speaking children*. This study will provide some essential information regarding the role and nature of phonological storage and phonological processing and vocabulary size in Arabic speaking children with typical and atypical language development. This study targets young children aged from 2;0 (years; months) to 4;0 years old as there is no study that has investigated this relationship in Gulf Arabic speaking children at this age. There are also few studies that examined the relationship between phonological skills and expressive vocabulary in children with and without language disorders.

This study will add cross-linguistic evidence from Arabic to the existing literature on the role of NWR and phonological storage in typical and atypical language development. Some findings of non-Indo-European languages (e.g., Cantonese) did not find NWR as a clinical marker in children with specific language impairment (SLI) (Stokes et al. 2006). Therefore, findings from Arabic will shed some light on this relationship from another non-Indo-European language. Moreover, the current study will involve developing some new assessment material that are necessary to evaluate phonological and vocabulary skills of children with and without language impairment in Arabic.

The children will be tested using the following:

A word and nonword repetition test (WNRep): this test was developed by this author to be used with Gulf Arabic preschool children and it is based on the methods used to create The Preschool Language Test (Roy & Chiat, 2004).

The Arabic Picture Vocabulary Test (APVT) (Shaalan, 2010): It is a receptive vocabulary test developed to be used with Gulf Arabic speaking children. It is based on the methods used to develop the Peabody Picture Vocabulary test (PPVT) (Dunn & Dunn, 1997).

The Arabic Expressive Vocabulary Test (AEVT): This test was developed during this project and is based on the methods used to create the Expressive One Word Picture Vocabulary Test (EOWPVT) (Martins & Brownell, 2000).

Study 2: Root and Pattern Nonword Repetition test (RAP-NWR). This study will provide information about the effects of roots and patterns on children's performance on NWR. In addition, it will explore the relationship between children's performance on the root and pattern NWR and receptive and expressive vocabulary tests. This study will shed some light on children's underlying processes and their recognition of Arabic roots and patterns in the light of different phonological and morphological theories.

This study involves developing The Root and Pattern Nonword Repetition test (RAP-NWR) that consists of nonwords that have different combinations of root and pattern conditions in Arabic (e.g., root-nonpattern, nonroot-pattern, nonroot-nonpattern).

1.1 Thesis Structure

This thesis investigates early phonological skills in TD Gulf Arabic speaking children and those with language impairment and discusses findings in relation to common theories of NWR. Moreover, we examine the effects of roots and patterns in Arabic and their relation to both theories of NWR and word formation in Semitic languages.

The rest of the thesis is organized as follows:

Chapter 2 will review NWR and its relation to vocabulary development in TD children and those with language impairment. We will examine the main theories of NWR, with special emphasis on two main theories that try to explain NWR deficits in children with language impairment, namely the phonological short term memory theory (Gathercole & Baddeley, 1990a) and the phonological processing theory (Snowling, Hulme, & Chiat, 1991).

Chapter 3 is dedicated to introducing the properties of the phonological and morphological system of Gulf Arabic and its prominent feature of the presence of roots and patterns. Word formation theories in Semitic languages are explored and we discuss how root and pattern morphology develops in these languages.

In *Chapter 4*, we present the first study where we developed a Word and Nonword Repetition test (WNRep) to investigate phonological storage and wordlikeness effects and examine if this test succeeds in differencing the performance of two group of Gulf Arabic speaking children. The first group (n=44) consists of TD developing children (average age is 3;2) and the second group (n=15) consists of children with language impairment (Clinical group), aged (average age is 3;7). The relationship between NWR performance and receptive and expressive vocabulary are examined. Moreover, we compare the effects of two scoring methods on the interpretation of these NWR results.

Chapter 5 involves conducting another experiment that looks into the effects of roots and patterns using the RAP-NWR on one hand and word length on the other hand on the performance of 89 TD Gulf Arabic speaking children aged between 2;0 to 7;0 (average age is 4;0). Results are discussed in reference to main theories of NWR and roots and patterns in Semitic languages.

Chapter 6 is dedicated to discussing the theoretical and clinical implications of the findings of this study. In this chapter, we argue that the findings of this thesis inform the current theories of NWR and word formations in Semitic languages. We also discuss the potential clinical utility of the tests developed in this thesis. Finally, limitations and directions for future research in the field of NWR in Arabic are discussed.

2. Literature Review

2.1 Nonword Repetition

Children in the early stages of language development try to mimic adult words. These words are initially novel for children, however, with time and development children will be able to repeat more difficult words and learn their meanings. Baddeley et al. (1998) proposed that children's ability to repeat novel multisyllabic words is a predictor of their overall language learning ability. Therefore, there has been an increased interest in investigating children's ability to repeat unfamiliar words in order to know more about the underlying processes that may predict language abilities in typical and atypical language development. Nonword repetition has become a common tool to measure the ability of repeating novel words, as it was used with different populations: children and adults, typically developing individuals and those with language impairments, and across different languages. Nonword repetition stimuli were manipulated in different ways to develop different tasks in order to measure different aspects of the language, such as phonology, morphology, semantic (see section 2.8 for more details).

2.1.1 Relationship between NWR and vocabulary in TD children

Studies of NWR in typically developing (TD) children have found significant correlations between NWR and receptive vocabulary (Gathercole et al., 1991; 1992, Briscoe et al., 2001; Coady & Evans, 2008). These studies explained this correlation differently, based on two main accounts that are explained in detail in section 0and 0. The first account is the phonological short term memory (PSTM) account of nonword repetition (Gathercole & Baddeley, 1989; 1990a; Gathercole et al., 1994). The second account is the phonological processing account (Snowling, Chiat& Hulme, 1991). The significant correlation between NWR and vocabulary is not surprising as it is argued that NWR mimics child's task when learning new words, as learning a new word involves attending to novel acoustic information that is used to create phonological representations. Therefore, TD children with better vocabulary tend to score better on NWR tasks when compared with children with lower vocabulary scores (Bowey, 1996; 2001; Gathercole and Baddeley 1989; Metsala, 1999). Though this correlation was found to be significant at age of four (Gathercole and Baddeley 1989; Gathercole et al. 1991), it was no longer significant by the age of five (Gathercole et al., 1992). Therefore, PSTM stops influencing vocabulary growth by this age and it is vocabulary that seems to affect NWR after this age as children have bigger vocabularies that they employ to facilitate NWR by using lexical and sublexical information. Gathercole (1992) explained that between the age of four and five years NWR is influenced by vocabulary development, while before the age of four it is NWR that has stronger influence on vocabulary development. She did not explain, however, how the direction of influence between nonword repetition and vocabulary reverses with age. Furthermore, Gathercole (2005) acknowledges other effects that may influence NWR, such as phonological complexity as evident in the difference in performance on nonwords with and without consonant clusters, and frequency of consonant sequences (Edwards, Beckman & Munson, 2004). These all can be included under the general umbrella of phonotactics.

The link between NWR and vocabulary seems to hold for receptive vocabulary, but not for expressive vocabulary. Briscoe et al. (2001) argue that this is one potentially confusing fact about the link between nonword repetition ability and vocabulary. When Briscoe et al. (2001) examined the performance of 35 TD children with a mean age of 8;6 and compared them with groups of children with specific language impairment (SLI) or with mild-to-moderate sensorineural hearing loss (SNHL), they found that receptive, rather than expressive, vocabulary tended to correlate more strongly with NWR in the SNHL and TD groups. No explanation was provided about why receptive, but not expressive, vocabulary accounted for this correlation between NWR and vocabulary development in TD children.

2.1.2 Relationship between NWR and vocabulary in children with SLI

It is important before discussing the relationship between NWR and vocabulary in children with specific language impairment (SLI) that we define SLI. Leonard (1998) defined SLI as the presence of significant receptive and or expressive language impairments in the absence of cognitive, sensorimotor, social-emotional and environmental deficits. Therefore, SLI is a disorder that is diagnosed by exclusion (Bishop, 1997). The prevalence of SLI according to Tomblin et al., (1997) for kindergarten children was 7.4%. Currently, there is a debate about the nature of SLI and whether it really exists in isolation of any other deficits and more terms are introduced to replace SLI, such as primary language impairment (Bishop, 2014; Kohnert, Windsor, & Ebert, 2009), Developmental Language Disorder (Bishop, Snowling, Thomspon, Greenhalgh & The CATALISE Consortium, 2016) or Language Learning Impairment (LLI), however the debate about the nature of SLI is beyond the scope of this thesis and SLI remains the mostly widely used term to describe this population (Bishop; 2014). For more information about the debate on SLI, please see Bishop, 2014; Conti-Ramsden, 2014; Gallagher, 2014; Lauchlan & Boyle, 2014; Leonard, 2014; Norbury, 2014; Reilly, Bishop & Tomblin, 2014b; and Reilly et al., 2014a).

Since Gathercole and Baddeley's (1990) paper, NWR has received a greater amount of attention in the study of SLI due to NWR tasks' potential use as a major tool to identify children with SLI (Botting & Conti-Ramsden, 2001; Conti-Ramsden & Hesketh, 2003; Dollaghan & Campbell, 1998; Ellis Weismer et al., 2000; Gathercole & Baddeley, 1990a; Gray, 2003).

Gathercole and Baddeley (1990a) reported that children with SLI repeated significantly fewer nonwords correctly when compared with TD children of a similar age who were matched on nonverbal intelligence, as well as a younger language-matched group. The

NWR skills of children with SLI (aged 8;6) were compared with those of typically developing age-matched children (TDAM) and typically developing younger (TDY) children. Children with SLI performed significantly more poorly than their TDAM and TDY peers on the three- and four-syllable nonwords. The mean performance of children with SLI was approximately 4 years below their chronological age. Gathercole and Baddeley (1990a) attributed this deficit to limitations in the phonological short term memory of children with SLI. Children with SLI demonstrated proportionally more difficultly in repeating longer nonwords than shorter ones, indicating according to the authors, that they had limited phonological capacity. According to them, SLI is essentially a disorder of phonological short-term memory (Baddeley et al.,1998; Gathercole & Baddeley, 1990a).

Many studies showed more evidence supporting Gathercole & Baddeley's (1990) claim that lower performance on tests of working memory in children with SLI is mostly captured by an explanation of poor storage and processing of phonological information (Bishop et al., 1996; Conti-Ramsden & Hesketh, 2003; Dollaghan & Campbell, 1998; Ellis Weismer et al., 2000). Conti-Ramsden and Hesketh (2003) compared the performance of 5 year old children with SLI to that of typically developing language matched (TDLM) peers (aged 3;0) on four possible clinical markers: (a) a past-tense task, (b) a noun plural task, (c) a NWR task, and (d) digit span. Children with SLI performed significantly below the TDLM children on digit span and NWR. Therefore, Bishop et al. (1996) proposed that nonword repetition could be considered a primary behavioral marker of SLI and that deficits in these children's ability to retain phonological representations over time could be the underlying cause of some of the syntactic deficits in children with SLI.

2.2 Theories of NWR

Many studies have found a positive relationship between performance on nonword repetition and vocabulary size in typical and atypical children learning different languages (e.g., Briscoe et al., 2001; Botting & Conti-Ramsden, 2001; Coady & Evans, 2008; Conti-Ramsden & Hesketh, 2003; Dollaghan & Campbell, 1998; Ellis Weismer et al., 2000; Gathercole & Baddeley, 1990a; Gathercole et al., 1991; 1992; Gray, 2003).

There is a growing body of research that examines the relationship between word learning and NWR. These studies argue that learning words is supported by the phonological storage that underlies verbal working memory (Gathercole & Baddeley, 1989; 1990a; 1993). Based on Baddeley and Hitch's (1974) working memory model, researchers (e.g., Gathercole et.al, 1995; Gathercole, Hitch, Adams, & Martin, 1999) found that phonological working memory mediates children's vocabulary learning, at least before 5 years of age. That is, the phonological representation of novel words is stored in the phonological loop temporarily until novel words find a place in long term memory (the mental lexicon). The following section explores the underlying processes that are associated with a novel word's journey to be a mental lexicon entry through the Psycholinguistic model of single word processing (Stackhouse and Wells, 1997) and working memory model of Baddeley and Hitch (1974) and Baddeley (2003).

2.2.1 Psycholinguistic model of single word processing (Stackhouse and Wells, 1997)

This model seeks to characterise the processes involved in naming, word and nonword repetition, based on results of a series of assessments in a group of children with typical and atypical language development and some single case studies. According to this model, discrimination of the phonetic and lexical (semantics) level is an essential mechanism that is required before processing different word types (word or nonword). The accuracy of discrimination will guarantee a better phonological and semantic representation and therefore a better word output. Different components of the Stackhouse and Wells (1997) speech processing model are shown in the following figure.

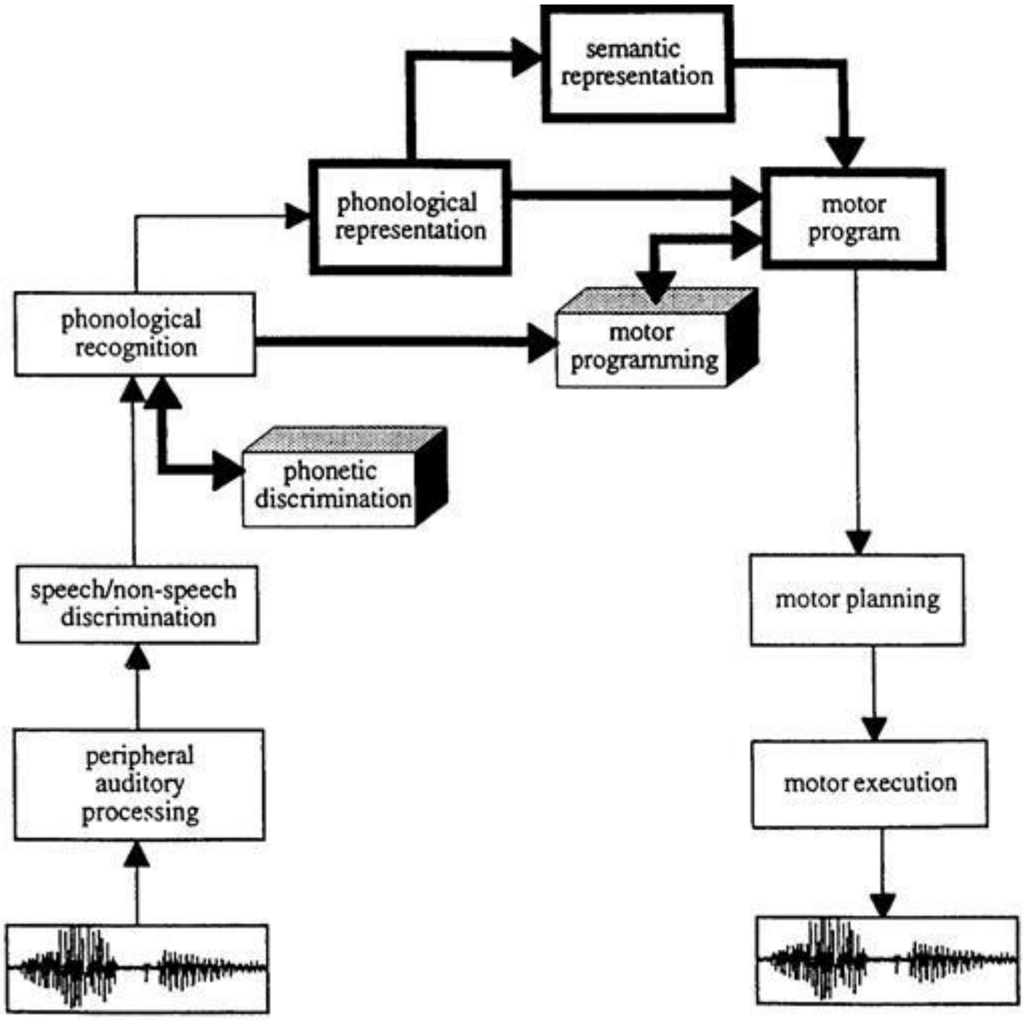

Figure 1: The Speech Processing Model from Stackhouse and Wells (1997), p 166.

The naming task, according to this model, requires long-term lexical activation of the elicited words, which in turn activates the semantic representation, this is followed by an activation of the stored phonological representation. Finally, the motor processes related to that word would be activated for word production. Incorrect production of words, according to this model, could be due to incorrect storage of that word at any of these levels of activations (semantic, or phonological).

The process activated during repeating real or familiar words could be similar to the process involved in naming words. If the long-term lexical representation of that word is activated by the input, then the semantic and phonological representations will be involved. This kind of processing is called deep processing. If the child is not familiar with that word or has fuzzy semantic or phonological representation of that word then a shallow level of processing would occur. Shallow processing would depend on the child's phonological perception and phonological short term memory storage, which in Stackhouse and Well's model is represented as phonological recognition.

According to Stackhouse and Wells' model, repetition of nonwords does not require semantic and phonological representations as the nonword cannot be retrieved from the lexicon. Therefore, a new motor program is created by selecting and combining a stored phonological unit (if the nonword has parts that correspond to existing phonological units, such as real syllables) to form an accurate repetition for the nonword. In addition. The last stage before nonword execution is the motor planning stage.

Stackhouse and Wells' (1997) model gives a detailed outline of the sequence of processes involved in producing word-size phonological units, whether naming, repeating real word, or repeating nonwords. However, there are a few studies that examined this model, in comparisons to the other models of nonword repetition (e.g., the working memory model). This might be due to the way in which this model was originated. The main aim of this model was to identify which processes are impaired in the path to word production in different children, hence the focus on single case studies. Furthermore, this model treated auditory processing, phonological processing and motor output processes as integral parts of this model, however it fails to address how long term phonological representations are established and therefore did not comment on how these representations contributed to nonword repetition. Furthermore, this model does not examine how familiar features of lexical phonology (such as wordlikeness or phonotactic probability) can affect repetition accuracy.

Since the focus of the current study is examining the relations between long-term knowledge the nonword repetition (and not studying the different stages of processing), we will not pursue models of the production processes, such as Stackhouse & Wells (1997).

2.2.2 The working memory model

The original Baddeley and Hitch (1974) model of working memory consisted of three components: A central control system called the central executive and two subsidiary systems called the visuo-spatial sketchpad and the phonological loop. See Figure 2.

Figure 2: The three component model of working memory as proposed by Baddeley and Hitch (1974) reproduced from Baddeley (2003, p.191)

As proposed by Baddeley and Hitch (1974), the central executive is responsible for the control of working memory. It is a pool with limited capacity of general processing resources and it controls the interaction between the other components. The visuo-spatial sketchpad is the second component and it is responsible for integrating and processing visual, spatial and nonverbal information.

The third component is the phonological loop, which is specialized for the storage of verbal material. It consists of two components, phonological storage which holds speech-based information for 1-2 seconds, and the articulatory rehearsal component, which converts words into articulatory or spoken words before entering phonological storage. According to Baddeley (2003), the phonological loop is responsible for sorting and processing sound combinations, so any impairment in this part of working memory will cause deficits in the phonological representation, thus affecting the process of learning new words (Archibald & Gathercole, 2006; Baddeley et al., 1998; Gathercole & Baddeley, 1990a; Gathercole et al., 1999). Baddeley et al. (1998) proposed that 'the function of the phonological loop is to provide temporary storage of unfamiliar phonological forms while more permanent memory representations are being constructed' (1998, p. 159).

Baddeley (2003) added the episodic buffer to this model of working memory in order to understand the process by which information from various subsystems was combined into a temporary representation. Therefore, the buffer oversees integrating and temporarily storing visual and auditory information into a single episode. It has also connections with long-term and semantic memories. Baddeley (2003) added that the episodic buffer carries out the essential function of feeding information into and retrieving information from long term memory, under the direction of the central executive.

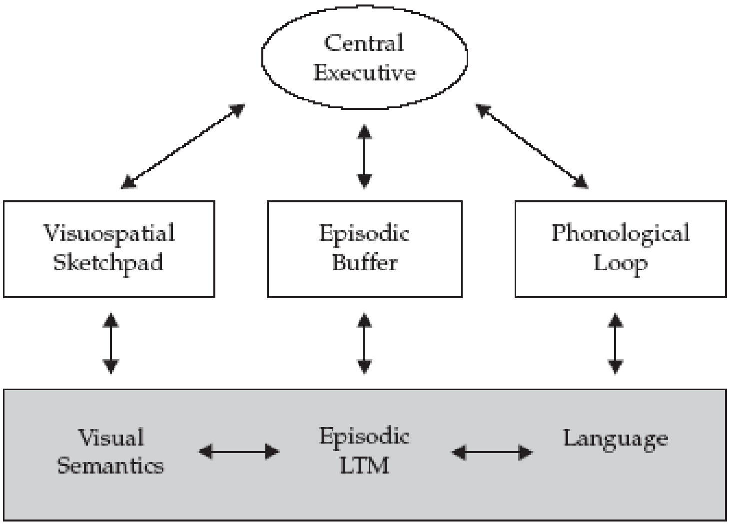

Figure 3: The revised Baddeley and Hitch Working Memory Model (Baddeley, 2003 p.196)

The working memory model (Baddeley, 2003) might succeed in explaining the different factors that are involved in working memory, however it is not clear how these factors work together. For example, there is evidence that the phonological loop is responsible for sorting and processing sound combinations and any deficit in this part would affect word learning (Archibald & Gathercole, 2006; Baddeley et al., 1998; Gathercole & Baddeley, 1990a; Gathercole et al., 1999). However, it is not clear yet if the deficit will affect the phonological loop only or if it will combine with other components. Furthermore, the WM model does not explain how semantic, syntactic and lexical components interact in NWR (Baddeley, et al.1998). The link between verbal short term memory (STM), which is another term for the PSTM, and long term memory (LTM) is also questioned by Cowan (2001), who found that Baddeley's (2003) WM model failed to explain adequately how verbal STM is linked to LTM.

In light of this WM model, there are two different accounts that try to investigate this model in relation to NWR by exploring the mechanisms underlying nonword repetition in the light of the working memory model. They also examine if the verbal short term memory represented by NWR will be assisted by the long-term memory represented by linguistic knowledge or by phonological memory only. The first one is the phonological term memory account by Gathercole & Baddeley (1989; 1990) and Gathercole et al. (1994). The other account is the linguistic account or the phonological processing account by Snowling, Chiat & Hulme (1991). Although both accounts agree that the phonological loop is a key component in the WM model, they differ on identifying the main processes that underlie performance on NWR. The following section will discuss these accounts in more details.

2.2.3 The phonological short term memory (PSTM) account of nonword repetition

The phonological short term memory (PSTM) account of nonword repetition (Gathercole & Baddeley, 1989; 1990; Gathercole et al., 1994) proposed that a person's ability to repeat what they hear helps in learning new words. Gathercole (2006) suggested that there is a high correlation between nonword repetition and novel word learning because both are constrained by phonological storage. That is, children's ability to repeat an unfamiliar sequence of phonemes can determine their ability to store and learn a sequence of phonemes pertaining to a novel word. Since every word is a novel word when it is first introduced to the child, both nonword repetition and vocabulary acquisition may have shared cognitive and neural mechanisms (Gupta & MacWhinney, 1997). The proponents of the PSTM have claimed that storage of phonological information in the phonological loop is critical for learning words (Gathercole et al., 1997; Gathercole & Baddeley, 1990). Gathercole (2006) also proposed that 'although this is not the only route by which new phonological structures can be acquired (lexically mediated learning is one alternative), it is a primitive learning mechanism that is particularly important in the early stages of acquiring a language' (p. 251).

According to the PSMT account, the phonological loop is responsible for storing temporary phonological information which influences nonword repetition. Therefore, NWR is considered a pure measure of PSTM, as it presents nonlexical material that allows the elimination of any familiarity effect. Furthermore, the phonological structure of nonwords does not require a long term lexical memory; the accurate repetition of the nonwords requires only a temporary storage for sound sequence in the phonological loop (Baddeley, Gathercole & Papagno, 1998; Dollaghan & Campbell, 1998). The PSTM account explains that word learning (long term lexical) difficulty in children with SLI is due to difficulties with phonological short term memory (phonological loop) (e.g., Archibald and Gathercole, 2006; Bishop et al., 1996; Gathercole and Baddeley, 1990a).

Although the PSTM account acknowledges the contributions of the different phonological processes that link between nonword repetition and vocabulary knowledge as proposed by the phonological processing account of Snowling, Chiat & Hume (1991), Gathercole et al. (1991) claim that PSTM is the most significant contributor to children’s subsequent vocabulary development, and that the phonological processes offer little in predicting future vocabulary size.

2.2.4 The linguistic account of nonword repetition

This account is also known as the linguistic account or the phonological processing account. This account attempts to provide an additional explanation for the nature of the relationship between NWR and vocabulary acquisition (Snowling, Chiat& Hulme, 1991). This account considers the phonological loop as only the initial process in nonword repetition, and there are various other processes that participate in the same task, such as speech perception, phonological awareness, and output processes. Beside these processes there are other factors that play a role in nonword repetition, such as lexical knowledge and phonotactic probability. Gathercole et al. (1991) proposed that lexical knowledge can be used to support nonword repetition as the similarity to the sublexical units within nonwords increases, which in turn will influence accurate nonword repetition. The linguistic account, challenges the PSTM account and argues that nonword repetition is not a pure measure of phonological short term memory, as there is a myriad of other factors that are involved in the processing of nonwords, some are influenced by PSTM, while others are not. These various factors that influence NWR are discussed in detail in the following section with reference to these two competing accounts of NWR.

2.3 Factors Affecting Nonword Accuracy

2.3.1 Age effects

Many studies have found an increase in nonword repetition accuracy with age (Edwards et al., 2004; Gathercole & Baddeley, 1989; Gathercole et al., 1991). Both the competence of temporary storage and articulation skills increase with age and support the subvocal rehearsal function of the phonological loop (Hoff et al., 2008). Subvocal rehearsal helps in the maintenance of phonological memory traces in the store (Gathercole & Baddeley, 1989). According to the PSTM account, the presence of the phonological loop would increase children’s accuracy in nonword repetition. In response to this, the phonological processing account argues that the phonological loop supports the accuracy of repetition up to the age of five years, while lexical knowledge continues to support children's repetition across ages.

2.3.2 Vocabulary size

Another variable influencing nonword repetition is vocabulary size, or the estimate of how many words a child has in his/her mental lexicon. Children with typical language development who obtain high scores on vocabulary measures are more likely to have better nonword repetition performance in contrast to children with low vocabulary scores (Bowey, 1996; Edwards et al., 2004; Gathercole & Baddeley, 1989; 1990a; Gathercole et al., 1991).

The association between vocabulary knowledge and nonword repetition skill becomes stronger as children's vocabularies increase in size. That is, the connection between phonological working memory and vocabulary size changes throughout development. Gathercole and Baddeley (1990a) proposed that children who performed better on nonword repetition exhibited faster learning for new words than the children who were less skillful at nonword repetition. Gathercole (1995) and Gathercole and Baddeley (1991) found that children with low repetition accuracy showed low scores on receptive vocabulary knowledge. Correspondingly, children who showed better performance on nonword repetition achieved high vocabulary scores on standardized vocabulary tests (Gathercole, 1995; Gathercole & Baddeley, 1991). These results indicate the bidirectional nature of the relationship between working memory and long term knowledge of words and word parts. The first part of these results supported the PSTM account, where the phonological loop supports learning and repeating novel words in children before the age of five, while lexical knowledge, according to the linguistic account, supports word learning. On the other hand, Horohov and Oetting (2004) argued that the amount of exposure to new words influences word learning more than lexical knowledge; the more the child is exposed to a new word the more he/she will be able to maintain it as a mental lexicon entry. Furthermore, Coady and Evans (2008) stated that children with language impairment succeeded in learning new words but they needed more exposure over time to the words.

2.3.3 Nonword length

Studies have found a strong effect of nonword length (number of syllables) on repetition performance (Gathercole & Baddeley, 1989; 1990; Gathercole et al., 1994), which is the main support for PSTM. In these studies, longer nonwords have typically resulted in more repetition errors than shorter nonwords. Baddeley (2002) explained the effect of word length as the longer the word is the more likely it will be forgotten because repeating longer words mean there are more demands on memory, articulation, and duration. Santos, Bueno & Gathercole (2006) explored nonword repetition skills in typically developing children from four to ten years old and found that errors increased with longer nonwords (of four and five syllables) and that accuracy increased with age. Furthermore, children with language impairment have consistently displayed significant deficits when repeating long nonwords of three or more syllables (Bishop et al., 1996;

Dollaghan & Campbell, 1998; Gathercole & Baddeley, 1990a) when their performance was compared with TD children, or when compared with their own performance on shorter nonwords (one and two syllables vs three and four syllables). Results from both typically and atypically developing children supported the PSTM account, which claimed that nonword repetition is a pure measure of PSTM and the phonological loop is an essential component for temporary phonological storage. While the phonological processing account acknowledges the effects of word length, it also acknowledges the contributions of other factors that are discussed in this section.

2.3.4 Wordlikeness

Wordlikeness is defined as the degree of likeness between a nonword and the phonological form of words stored in an individual's lexicon and it has been shown to influence NWR results (Dollaghan, Biber & Campbell, 1995; Edwards et al., 2004; Gathercole, 1995). Many studies (e.g., Gathercole, 1995; Gathercole & Adams, 1994; Gathercole et al., 1999) suggested that the more wordlike the nonword, the more likely that it will be repeated accurately. According to Baddeley (1975), judging wordlikeness is based on two main factors: wordlikeness can be either through phonological similarity where the nonword consists of a known or familiar phoneme structure, or lexico-semantic similarity where the nonword contains a real word, such as "under" as in "*under*brantuand" (CNRep, Gathercole & Baddeley, 1996) or a real morpheme, such as "ing" as in "blonterstap*ing*" (CNRep, Gathercole & Baddeley, 1996).

Wordlikeness was investigated in many NWR studies by manipulating the degree of wordlikeness. Gathercole (1991) investigated the repetition accuracy of 70 typically developing children at 4 and 5 years of age. The nonword task used in this study was rated by 20 adults on a five-point scale, from one which was very unlikely to be a real word, to five which was very likely to be a real word. Children from both groups repeated high wordlikeness nonwords more accurately than low wordlikeness nonwords. The correlation between the rated wordlikeness and the children's repetition accuracy was significant at ages 4 and 5 years old. Moreover, Archibald and Gathercole (2006) tested wordlikeness effects on groups of children aged between 7-11 years old who belonged to three groups: children with SLI, age matched and language matched children. All groups repeated high wordlike items better than the low wordlike items, although the difference was minimal for the TD age matched group.

Roy and Chiat (2004) reported significant contributions of wordlikeness effects in 66 TD children between the ages of two to four years old as they performed significantly better on words than nonwords. These results were replicated by Chiat & Roy (2007) with 315 children. These proponents of the phonological processing account predicted that children with language impairment would show less sensitivity to wordlikeness effects due to their overall deficits in phonological skills.

The findings of these studies indicate that repetition of high wordlikeness items is facilitated by long-term lexical knowledge and is therefore less sensitive to phonological memory constraints, whereas nonword repetition for low wordlikeness nonwords is dependent on phonological memory. Furthermore, typically developing children benefit from wordlikeness effects while children with language impairment were less sensitive to these effects.

2.3.5 Phonotactic probability

A closely related concept to wordlikeness is phonotactic probability. Phonotactic probability is the likelihood of the occurrence of a specific sound or sound combinations in a given language (MacRoy-Higgin& Dalton, 2015; Storkel, Armbruster & Hogan, 2006). Phonotactics refers to the sequences of phonemes that could be common or rare in that language and phonotactic probability is the frequency with which phonological segments or sequences of phonological segments occur in words in a given language (Vitevitch & Luce, 2005). So, while wordlikeness is related to how a nonword resembles an existing real word, phonotactic probability refers to the frequency of the distribution of phonemes and phoneme sequences in a particular language. For example, the phonemes and phoneme combinations in the nonword /hesələm/ are of high phonotactic probability, while the nonword /gufegɛd/ has a low phonotactic probability because its constituents have low frequency of occurrence in the English language

(Munson et. al, 2005). However, the two are related in that nonwords consisting of sound combinations of high phonotactic probability are judged to be more wordlike (Frisch, Large, & Pison, 2000). A growing body of research provided evidence that phonotactic probability influences nonword repetition in TD children. Beckman and Edwards (1999) reported that children aged three to five years old repeated common sound sequences more accurately when compared with rare sound sequences. Similar results were also reported with older children aged seven to eight years old (Gathercole, Frankish, Pickering, & Peaker, 1999). More evidence about the influence of phonotactic probability in nonword repetition in children with SLI was provided by Munson et al. (2005), where they compared the performance of three different groups: an SLI group, a TD aged matched group and a language matched group. A main effect of phonotactic probability was found as all groups tended to repeat nonwords with high-frequency sound sequences more accurately than low-frequency ones. Investigating the influence of the phonotactic probability on nonword repetition on one hand, and the influence of nonword repetition and word learning on the other hand, was explored by some studies that tried to investigate the effect of phonotactic probability on learning novel words. For example, Storkel and Rogers (2000) examined the effects of phonotactic probability on learning novel words in three groups of children aged 7, 10 and 13 years old. The three groups were exposed to nonwords paired with unfamiliar object referents; half of the nonwords had high-frequency sound sequences and the other half had low-frequency sound sequences. The two older groups identified more referents of common than rare sound sequences, while the youngest group showed no consistent effects of phonotactic probability. Storkel and Rogers (2000) explained that the lack of effects of the phonotactic probability in the youngest group was due to their phonological representation which was not like adults. The limited lexical knowledge at this age did not support children's sensitivity to high-frequency sounds. It is expected that the older the children are the better their phonotactic knowledge will be. So, as vocabulary size increases with age, children's knowledge of high-frequency sounds sequences will increase.

Overall, all these studies show that phonotactic probability influences nonword repetition, however children from different age groups are affected differentially as the sensitivity to phonotactic probability develops with age.

2.3.6 Neighbourhood density

Neighbourhood density is defined as "the number of words that sound similar to a given word" (Storkel, Armbruster & Hogan, 2006, p.1176). It is highly correlated to phonotactic probability (Vitevitch& Luce, 1999), however, the difference between the two is that neighbourhood density refers to the number of words that sound similar to a given word. This is unlike phonotactic probability, where the similarity is measured by the frequency of occurrence of individual sounds or a sequence of combined sounds.

Furthermore, the difference between neighbourhood density and wordlikeness is that the sequence of the sounds in wordlikeness is identical to a real word or a part of it, whereas in neighbourhood density one sounds or more is substituted with different sounds. There are two properties related to neighbourhood density that were investigated in research. First, the size of neighbourhood density where words differ from one another by the number of neighbours they have. For example, the word *cat* has 36 neighbours (e.g., *sat, hat, can)*, while the word *void* has only six neighbours (e.g., droid, voice). The second property is the neighbourhood frequency effect where neighbourhood words are varied in their frequency or familiarity, so a given word could be similar to a frequent word and infrequent word at the same time. However, highly frequent neighbours are more likely to be related to a given word (Vitevitch & Luce, 1999). Roodenrys and Hulme (2002) investigated the effects of neighbourhood density on short term memory using immediate recall tasks where the stimuli were selected according to their neighbourhood size and word frequency. Results showed that memory span is greater for high-frequent words versus low frequent ones. Furthermore, memory performance was also greater for the larger size of neighbourhood density versus the smaller size. According to these findings, neighbourhood density could influence nonword repetition accuracy and nonwords with high frequent neighbourhoods could be easier to repeat compared to less frequent ones.

2.3.7 Stress patterns

There are a few studies that investigated the effects of stress on nonword repetition. Dollaghan et al. (1995) investigated children's performance on nonword repetition by manipulating effects of stress using two patterns. The first pattern consisted of a nonsense syllable stress, i.e., a stress pattern that does not exist or very atypical of English words, where nonwords did not contain any weak stress and only tense vowels were used. The other stress pattern was typical of English stress pattern (a weak syllable and strong syllable stress pattern). Results showed that children were more likely to accurately repeat nonwords with stressed syllables corresponding to real words than nonwords with nonsense syllable stress. Furthermore, Roy and Chiat (2004) investigated stress effects on the type of nonword repetition errors, and found that children aged 2 to 4 years were sensitive to stress when repeating words and nonwords. Children were more likely to omit the unstressed syllables than the stressed ones. The PSTM account did not consider stress patterns as a factor that would contribute significantly to nonword repetition accuracy. For example, in the design of their NWR test (CNRep, 1996) Gathercole & Baddeley did not control for stress effects.

In the previous section, factors affecting nonword repetition were discussed, e.g., age, vocabulary size, nonword length, wordlikeness, phonotactic probability, neighbourhood density and stress pattern. Most of these factors showed that apart from effects of phonological storage, there are important effects that are not accounted by the PSTM and it seems that the linguistic account of NWR addresses some of these factors. In the following section, we examine some of the processes involved in NWR, which mediate between nonword repetition and language learning.

2.4 The Processes Involved in Nonword Repetition

According to Gathercole, (2006) there are three potential skills, besides phonological storage, that can cause a source of variance, and therefore could mediate the relationship between nonword repetition and language learning. These three potential skills are: auditory processing, phonological processing and speech-motor processing.

2.4.1 Auditory processing

It is known that interruption to the acoustic signal from the peripheral auditory system to the nerves and within the central auditory system in the brain can affect language performance. Briscoe, Bishop and Norbury (2001) compared scores on The Children's Test of Nonword Repetition (CNRep) (Gathercole & Baddeley, 1996) of groups of children with either mild–moderate sensorineural hearing loss or SLI. They found both groups were significantly poorer at repeating nonwords when they were compared to age-matched typically developing children. The difference between the two clinical groups was more obvious with the increase in the length of the nonword. The hearing loss group and the SLI group were equivalent for the two- and three-syllable stimuli, however the SLI group performed more poorly when repeating four and five-syllable stimuli. These findings suggest that the auditory level played a potential role in nonword repetition performance. This study did not give clear explanation about the underlying reasons for these findings. It might be that the characteristics of deficits for each group was different and therefore each group processed nonword repetition in a different way, for example it could be that longer nonwords gave the hearing loss group more acoustic and auditory cues which could have helped with their repetition, while the longer nonwords did not help with repetition in the SLI group as the deficit was mainly in their phonological memory. Although both groups performed significantly lower on the NWR task when compared to TD children, each group showed different profiles in their performance. Results showed that the NWR task was beneficial in differentiating the performance of the two clinical groups.

Otitis media with effusion (OME), a relatively common condition in childhood, can raise the hearing threshold by the build-up of fluids in the middle ear causing a hearing loss for the duration of the OME episode. Gathercole and colleagues (2005) tested 39 children with OME. At 60 months of age, each child was also tested on the Children's Test of Nonword Repetition (Gathercole & Baddeley, 1996) and on an auditory digit span task. The OME group showed better repetition for longer nonwords compared with the shorter ones. This stands in clear contrast with the typical profile of children with SLI. The impairment in repeating the shorter nonword stimuli may be due to the reduced availability of prosodic and suprasyllabic cues to segment identity in the shorter items.

In summary, hearing impairment related to sensorineural hearing loss, OME or to central auditory processing difficulties, will directly influence important aspects of language processing and language acquisition, in turn this might cause impairments in nonword repetition.

2.4.2 Phonological processing

Phonological processing is one of the main mechanisms that govern nonword repetition. Many studies have tried to explore the nature of the phonological processes related to nonword repetition and from different perspectives. Once of the questions commonly raised is whether there is one major process underlying NWR or there are multiple processed involved in NWR. Beside the argument discussed above between the PSTM account (Gathercole & Baddeley, 1989; 1990; Gathercole et al., 1994) and the linguistic account (Snowling, Chiat & Hulme, 1991), there is another debate between the phonological hypothesis and the phonological sensitivity hypothesis that was supported by Metsala and Walley (1998). The definition of phonological sensitivity according to Adams (1990) is the global set of phonological processing abilities that display a hierarchy of sensitivity to different levels of phonological complexity in different cognitive operations. Phonological sensitivity can be measured according to Burt, Holm, and Dodd (1999) by using two tasks: an alliteration awareness task and a rhyme awareness task.

Contrasting the phonological storage hypothesis, the phonological sensitivity hypothesis by Metsala and Walley (1998) has a different point of view on the relationship between nonword repetition and vocabulary size. This view, called the lexical restructuring model, theorized that although phonological working memory, as measured by nonword repetition, has a role to play in vocabulary acquisition, it is not seen as an inherent capacity, but rather as one that is subject to development over time. Metsala (1999) and Bowey (1996; 2001) have suggested that NWR is used to measure phonological working memory, while phonological sensitivity tasks are used to measure knowledge and awareness of segmental information. The proponents of the lexical restructuring model showed some evidence that phonological sensitivity can significantly predict vocabulary acquisition (Metsala, 1999; Bowey, 1996, 2001). When multiple regression analysis was used and effects of phonological sensitivity were controlled, nonword repetition was not a significant predictor of word learning as measured by receptive vocabulary tests (Metsala, 1999; Bowey, 2001). On the other hand, Gathercole and Baddeley (1990a) argue that the phonological sensitivity account does not provide a clear explanation for the apparent deficit in NWR as the number of syllables increases in children with SLI. They explain that a more parsimonious argument is the one that attributes the deficit to phonological storage. The phonological sensitivity account argues that nonword repetition is not a "pure" measure of phonological working memory as suggested by Gathercole et al. (1997) and that there are factors besides performance on nonword repetition that can account for vocabulary development. Investigations by Bowey (1996; 2001) and Metsala (1999) have shown that when effects of age and IQ were controlled, both nonword repetition and phoneme sensitivity contributed significantly to vocabulary size, thus ruling out nonword repetition as a unique contributor.

2.4.3 Speech motor output processes

The motor representation of NWR requires a conversion of the auditory representation of a nonword to a motor sequence for repetition. This conversion requires articulately coordination at the muscular level of the articulators (for example the tongue, jaws, lips and the velum) in real time. Nonword repetition also requires planning and executing speech motor commands that link a phonological representation of a nonword and the repetition attempt (Snowling & Hulme, 1989). Four factors could interfere with the articulation accuracy in NWR: the first one is the presence of speech and/or language impairment. Children with poor scores on NWR could potentially have phonological disorders as a symptom of language impairment and/or other peripheral speech-motor disorders like dyspraxia, dysarthria or articulation disorders.

The second factor is stimuli complexity, as increasing the stimuli complexity will result in less accurate repetition. Archibald and Gathercole (2006) and Bishop et al. (1996) found that children with SLI showed more deficits in repeating stimuli that contained consonant clusters when compared with stimuli with single consonants. Similarly, Marshall and colleagues (Marshall et al. 2002; Marshall & van der Lely, 2009) reported that children with SLI were significantly less accurate in repeating nonwords with clusters. Children with SLI dropped or simplified consonant clusters and created clusters in incorrect positions. Therefore, Marshall and colleagues (Marshall et al., 2002; Marshall & van der Lely, 2009) argued that the PSTM account of Gathercole and Baddeley (1990a) failed to explain the phonological complexity deficit in children with SLI.

Further findings from Edwards and Lahey (1998) showed that children with SLI had greater difficulties in repeating lately emerged phonemes, i.e., phonemes that appear relatively late in phonological development. Finally, the production of nonwords improves with age as children's motor skills improve and they start to develop articulatory control on their speech (Smith, 2006). However, it is important to point out that NWR tests are scored differently; where some scoring takes into account articulatory accuracy, others may accept some substitutions and/or minor distortions and therefore scoring methods should be examined individually before we can compare various NWR tasks.

To sum up, nonword repetition accuracy is supported by essential processes such as auditory processing, phonological processing and speech-motor processing, and any deficit in any component would affect the repetition accuracy. In addition, there are some factors that would influence NWR accuracy, such as wordlikeness, phonotactic probability and nonword length. Furthermore, individual NWR tests differ in their design, stimuli, and the scoring methods used. In the following, we examine some of these NWR tests.

2.5 Different Types of NWR Tests

Researchers have used different tests to examine the role of nonword repetition and the different variables that may influence NWR. The Children's Test of Nonword Repetition (CNRep; Gathercole & Baddeley, 1996) is considered to be one of the first tests developed to assess nonword repetition. It consists of 40 nonwords equally divided into 2 to 5 syllable nonwords. However, researchers might argue that the CNRep is not a pure measurement of nonword repetition due to the many confounding factors included in the test. For example, it includes many wordlike nonwords and syllables that could be a word or a morpheme by themselves, such as "underbrantuand". It seems that nonwords were solely chosen according to the number of syllables. Moreover, Gathercole and Baddeley (1996) did not control for the presence of consonant clusters (e.g., */taflest/*).

The most significant finding from the CNRep was the presence of a strong effect of word length on NWR, a finding that has been replicated in most of the subsequent tests of NWR. Conti-Ramsden and colleagues used CNRep to evaluate potential clinical markers of SLI in a group of 5 years old children (Conti-Ramsden, 2003) and a group of 11 year old children with a previous history of SLI (Conti-Ramsden, Botting, & Faragher, 2001). Results indicated that CNRep provided a useful clinical marker for SLI children; however, sentence repetition was a more useful marker in the older age group.

The Nonword Repetition Test (NRT; Dollaghan & Campbell, 1998) was developed to avoid the confounding variables seen in CNRep such as the presence of wordlikeness effects and consonant clusters. The authors avoided both nonwords that are composed of lexical and sublexical elements, and nonwords that have clusters. They also significantly reduced the stimuli to 16 items instead of 42 as in the CNRep to reduce test duration. The words in the NRT were equally divided into 1-4 syllables in length. The authors included early developing consonants only to avoid any articulatory effects. All the vowels used were tense vowels and they used a non-English stress pattern to avoid the presence of any syllables with weak stress.

Archibald and Gathercole (2006) presented a comparison between these two tests and they found that both the CNRep and NRT could distinguish between children with SLI and TD children with a high level of accuracy. The CNRep test, however, had greater ability to identify children with SLI as the test items included sublexical units, grammatical morphemes and consonant clusters and children with SLI had difficulty with some of these variables. Although it is possible that these children with SLI might have benefitted from wordlikeness effects in the CNRep test (Archibald & Gathercole, 2006). Graf Estes, Evans, and Else-Quest (2007) reported that the CNRep test might succeed in differentiating the SLI group from TD groups more than the NRT. However, they argued that the NRT test is more linked to basic phonological skills and phonological memory due to its control for effects of wordlikeness, consonant clusters, and articulatory complexity.

The Preschool Repetition (PSRep) Test was developed by Roy and Chiat (2004) and it aims to measure phonological skills of children between 2 and 4 years. One of the distinguishing factors in the design of PSRep was the examination of the effects of prosody. The PSRep consists of 36 test items (18 words and 18 nonwords). Words and nonwords were equally divided into 1-3 syllables, with systematic control of stress. Therefore, half of the two syllable words and a third of the three syllable items start with a weak stress. The nonwords were created by alternating the vowel in single syllable words ("mouse" becomes /mis/ and reversing two consonants in each word to create a corresponding nonword (e.g., "dinosaur" becomes /ˈsainədɔ/ to ensure they are phonologically matched. Chiat and Roy (2007) found that this test reliably differentiated between typically developing children and children at risk of language impairment. They showed that this test was not influenced by socioeconomic status and the results showed strong effects of word length, prosody (stress) and age (Chiat & Roy, 2007).

2.6 Cross Linguistic Studies

Apart from studies that investigated nonword repetition in European Languages (e.g.: Spanish: Gibrau & Schwartz, 2007; Swedish: Kalnak, Peyrard-Janvid, Forssberg, & Sahlén (2014); Dutch: de Bree, Rispens & Gerrits, 2007, Rispens & Parigger, 2010), there are few studies that looked into NWR in non- European languages (e.g.: Cantonese: Stokes, Wong, Fletcher & Leonard, 2006; Arabic: Shaalan 2010). The importance of investigating NWR cross linguistically is to confirm that deficits found in children with SLI are not unique to the English language. Furthermore, nonword repetition tests could potentially be used to identify children at risk of language disorders in other languages.

Most of the results from European languages supported the results of English studies in nonword repetition. In Spanish, Girbau and Schwartz (2007) found strong correlation between the Spanish NWR and some standardised language measures. They found that a NWR following the phonotactic patterns of Spanish produced the same length effects as found in earlier results in English studies (e.g., Gathercole and Baddeley, 1990a)

In Dutch, de Bree, Rispens & Gerrits (2007) and Rispens & Parigger (2010) investigated nonword repetition's utility as a clinical marker for Dutch children with

SLI and dyslexia and compared the two groups to a control group of TD children. Both clinical groups performed significantly below the mean of the control group.

In Swedish, Kalnak, Peyrard-Janvid, Forssberg & Sahlén (2014) explored the accuracy of NWR as a clinical marker to distinguish between a school-aged Swedish SLI group and TD children. They found that NWR distinguished between SLI and TD groups with 90.2% sensitivity and 97.7% specificity at a cut-off level of −2 standard deviations for binary scoring of nonwords (Kalnak et al., 2014).

In Russian, Kavitskaya, Babyonyshev, Walls and Grigorenko (2011) found that, similarly to other languages, phonological memory affected children's ability to repeat words. Their results indicated that for children with SLI it was always more difficult to represent and repeat a longer nonword than a shorter nonword.

There are fewer studies that examined NWR in non-European languages when compared to European languages. For example, Stokes, Wong, Fletcher and Leonard (2006) studied NWR in Cantonese and found that there was no significant difference between children with SLI and TD children on performance on NWR. They attributed this lack of significant difference to the nature of Cantonese as a tonal language that is characterised by a very simple syllabic structure (CV only) and limited possible syllabic combinations, with no irregular stress or consonant clusters. Therefore, the unique simple syllabic structure of Cantonese did not represent a significant challenge to Cantonese speaking children with SLI. Therefore, it seems that the phonological processes underlying nonword repetition may vary according to the characteristics of a given language. For example, Gulf Arabic, another non-European language, was investigated by Shaalan (2010) who found results contrary to those in Cantonese. Shaalan (2010) investigated NWR of phonologically complex nonwords that differed from control nonwords by position and number of consonant clusters. He found that Gulf Arabic speaking children with SLI (average age 7;8 years) performed significantly less accurate than their age and language matched peers on NWR. However, there was no attempt to manipulate root and pattern effects as all his stimuli consisted of nonroots and nonpatterns, except for eight nonwords that consisted of a common existing pattern that he used as distractors/control stimuli.

2.7 The Importance of Investigating NWR in Arabic

The interest in studying nonword repetition has developed because it is relatively a simple task that does not require extensive preparation compared with other speech or language tasks. These lists of nonsense words can be used with a variety of populations and can be applied in a wide range of ages from very young children to adults. The aim of NWR can vary to measure different aspects in the language (e.g. phonology, morphology and semantics), though nonwords do not have meaning or syntactic function. By manipulating the stimuli used in NWR, some aspects of the language can be tested, for example, the semantic sensitivity can be tested by controlling the stimuli’s wordlikeness, phonological complexity, and phonotactic frequency. NWR tasks have been investigated in different languages where the NWR task was developed according to the specification of that language.

Nonword repetition was used in different developmental language impairments (e.g. SLI (Gathercole & Baddeley, 1990); Dyslexia (Catts et al., 2005); autism spectrum disorders (ASD) (Bishop, 2008; Marshall et al., 2013); Downs Syndrome (Comblain, 1999) and most of these studies found that these children had scores that were significantly lower compared with typically developing children. Furthermore, NWR helps in understanding the processes underlying the performance of the different developmental language impairment groups. For example, children with SLI performed differently on a NWR task compared a group of children with dyslexia (Marshall & van der Lely, 2009). Similar differences were also found with children with ASD group when they were compared to children with SLI (Marshall et al., 2013).

This study investigates NWR in Arabic speaking children and therefore has implications specific to this language. Research in Arabic language processing and acquisition has started to develop only recently. There are very few resources and studies on language processing and acquisition of different varieties of Arabic and language assessment tools or tests are scarce. Therefore, in these circumstances, a NWR task is an easy, flexible and powerful tool that can shed light on how TD children and children with language impairments perform on this task that has been found to be a good clinical marker in many languages. Furthermore, Arabic as a Semitic language has a unique morphology that is based on roots and patterns, hence, it is important to investigate if Arabic children perform differently on NWR. In addition to its unique morphology, Arabic is characterized by a very rich phonological system.

In the following chapter, Arabic phonology, morphology and word formation will be discussed in detail in order to understand how the specification of the Arabic language may or may not affect the underlying processes that are implicated in the NWR task.

2.8 Summary

This chapter began by defining nonword repetition and how it has become an important assessment tool as it was found consistently beneficial in differentiating children with language impairment from typically developing children and in predicting children's vocabulary size. An overview was presented about the different theories that explore the underlying processes that play a role in the performance on NWR. In light of this review, this study will provide some essential information regarding the role and nature of phonological skills, as measured by Arabic nonword repetition test, and vocabulary size in Gulf Arabic speaking children with typical and atypical language development. Examining this in a typologically different language like Arabic and comparing findings to those of European languages, might help in understanding theories of phonological processing. This study will also provide essential information about vocabulary development and phonological skills in Gulf Arabic speaking children. Gulf Arabic is the variety of Arabic language that is used in the eastern parts of the Arabian Peninsula, which includes the countries of Bahrain, Kuwait, Qatar, United Arab Emirates, and the eastern province of Saudi Arabia (Johnstone, 1967). There are many sub-dialects that are used in the Gulf area (e.g., city dweller dialects vs Bedouin dialects), however most of these differences are lexical in nature as the subdialects share many phonological and morphosyntactic features (Holes, 1989; 2004).

3. The Phonology, Morphology and Word Processing of Gulf Arabic

3.1. Introduction

Arabic belongs to the Semitic family of languages and is characterised by the presence of diglossia, whereby the spoken dialects of Arabic are used alongside Modern Standard Arabic (MSA) (Ferguson, 1959). MSA is used as the formal language (in academic curricula, newspapers, books and some programs on the television), while the spoken dialects are used in everyday situations and have different syntactic, morphological, phonological, and lexical properties from MSA.

In the following, the phonology and morphology of Gulf Arabic are briefly described.

3.2 The Phonology of Gulf Arabic

Gulf Arabic contains 30 consonant phonemes with places of articulation across the whole vocal tract from lips to glottis. Table 1 lists the consonants of Gulf Arabic.

The vowel system in Modern Standard Arabic according to Watson (2002) consists of three short vowels /a, ı, u/ and their corresponding longer vowels /a:, i:, u:/. In addition, there are two main diphthongs in MSA /aj/ and /aw/. Gulf Arabic long vowel /e:/ and /o:/ corresponding to diphthongs in MSA. Figure 4 shows the vowel system of Gulf Arabic.

Table 1: Consonant inventory of Gulf Arabic (adapted from Bukshaisha, 1985).

	Bi-labial	Labio-dental	Inter-dental	Alveo-lar	Palato-alveolar	Palatal	Velar	Uvular	Phary n-geal	glott al
Stop	b			t d tˁ			k g	q		ʔ
Nasal	m			n						
Fricative		f	θ ð ðˁ	s z sˁ	ʃ			x ʁ	ħ	h
Affricate					ʧ ʤ					
Trill				r						
Lateral				l 1ˁ						
Approx.						j	w		ʕ	

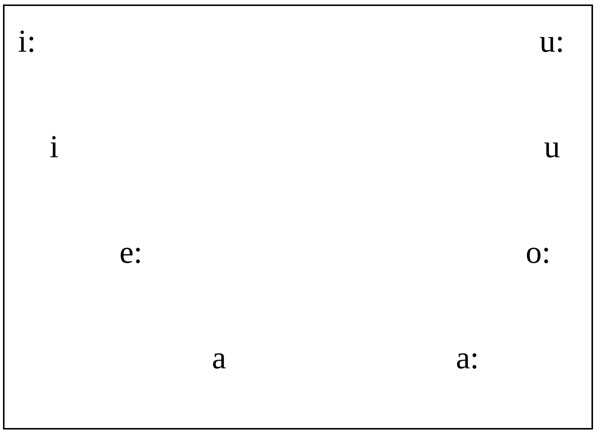

Figure 4:The vowel system of Gulf Arabic, Adapted from (Mustafawi, 2006, p.8)

Unlike standard Arabic, and some other dialects of modern Arabic, Gulf Arabic tolerates initial, medial, and final clusters and therefore there are at least 10 types of syllables in Gulf Arabic, as listed by Bukshaisha (1985). These include the following types: CV, CV:, CCV, CCV:, CVC, CV:C, CVCC, CCVC, CV:CC, and CCV:C.

Stress in Gulf Arabic is regular and depends on syllable weight, as is the case in many varieties of Arabic. The final syllable is stressed if it has a long vowel (e.g., CV:) or consonant cluster (e.g., CVCC), otherwise stress falls on the penultimate syllable (Hole, 1989).

3.3 Phonological Acquisition in Arabic

There are few studies that have investigated the acquisition of Arabic phonology and Gulf Arabic. Ayyad (2011) describes the phonological development patterns of Kuwaiti Gulf Arabic-speaking children. In this study, 80 preschool children (45-62 months) were included and divided in two different groups: a younger group (45-54 months old) and an older group (55-62 months).

More than 90% of the younger group showed mastery of the following consonants across different word positions: stops /b/, /t/, /tˤ/, /d/, /k/, /g/, /q/, /ʔ/, nasals /m/, /n/, fricatives /ð/ˤ, /ħ/, /h/, /x/, affricate /tʃ/, liquid /r/, and glides /w/, / j/. The older group acquired the same phonemes as the younger group up to 90% of mastery plus / ʁ/, / l/,

/ʃ/, /sˤ/. There were 75-89% of the younger group who acquired the following: stops /tˤ/, /q/, fricatives, /sˤ/, /ð/, /ʃ/, /ʁ/, /χ/, /ʕ/ and lateral /l/, while the older group acquired: /θ/, /ð/, /dʒ/, /ʕ/. Less than 75% of the younger group acquired fricatives /s/, /sˤ/, /θ/, /z/; voiced affricate /dʒ/, and trilled /r/, while the older group acquired /s/, /sˤ/, /r/ less than 75% of the time.

Amayreh (2003) and Amayreh and Dyson (200) studied the acquisition of consonants in Jordanian Arabic, which is not a variety of Gulf Arabic but shares many characteristics and speech sounds with Gulf Arabic.

Table 2 summarizes their findings about the development of consonants in Jordanian Arabic.

Table 2: The stages of the acquisition of Arabic consonants, adapted from Amayreh (2003).

	Babbling	12-24 ms	2-3:10 yrs	4-6:4 yrs	6:5-8 yrs
Stops	b	b, d, t, ʔ	k, q, g		tˤ, dˤ
Fricatives/affricates	h	sˤ, ʕ, ħ, h	f	s, χ, θ, ð, ʁ, sˤ, ʤ	ðˤ, z
Sonorants/liquids	m	m, n, l		r	
Glides	w, j	w, j			
Totals		13	+4	+8	+4 = 29

When comparing the results of these studies we can see that there is an agreement about the consonants acquired at the age of three to four years, as this is the age that overlaps in both studies, however some emphatic consonants in Kuwaiti children seem to be acquired earlier than seen in Jordanian children. These different findings in these two Arabic dialects could be because Kuwaiti Gulf Arabic has more frequent emphatic and later developing consonants (e.g., /θ/, /ð/, /ðˤ/) when compared to Jordanian Arabic (Ayyad, 2011). Kuwaiti children were able to acquire emphatic consonants earlier than Jordanian children and this is consistent with studies that have found that phoneme frequency in the lexicon of children can have strong effects on phonological acquisition (Edwards, Beckman & Munson, 2015).

In light of the developmental acquisition of the Arabic consonants (Amayreh, 2003; Amayreh & Dyson, 2000; Ayyad, 2011), the assessment material used in this current study in general and the nonword repetition tests in particular were developed according to this information about phonological development in children. This will be discussed in detail in chapter four and five.

3.4 The Morphology of Gulf Arabic

3.4.1 The root and pattern system in Gulf Arabic and other Semitic languages

Arabic and other Semitic languages share the same principle of derivational morphology of the root and pattern. In traditional analysis, the root of most words consists of three consonants that form the semantic abstraction. Words are then derived from this root by the superimposition of a vocalic pattern. Thus, words derived from the same root are usually related to the meaning of the consonantal roots (Holes, 2004).

shows an example of the root √k.t.b, which represents the common meaning of ‘writing’, and its derived forms (Versteegh, 1997).

Table 3: Derivatives of the Root √k.t.b (Versteegh, 1997, p. 85).

Root	Derived Form	Class/Number	Gloss
	kataba	Verb	he wrote
	yaktub	Verb	he writes
	kitaab	Noun, singular	book
k-t-b	kutub	Noun, plural	books
	katib	Noun, singular	writer
	kuttaab	Noun, plural	writers
	maktab	Noun, singular	desk/office
	makaatib	Noun, plural	offices
	maktaba	Noun, singular	library
	maktabaat	Noun, plural	libraries

shows ten derivational forms, while in the Dictionary of Modern Standard Arabic there are 32 different derivational words that belong to different lexical categories for the same root √k.t.b. These derivations have meanings related to “writing”, “letters”, or “books”. This account of Arabic morphology is called derivational morphology or root and pattern morphology. In this account the words are formed in different lexical categories by derivations from the root.

The template is another way of looking at the root and pattern in Arabic. Templates are where roots and patterns are arranged in a predetermined order or ready forms, where the consonantal root and the vocalic pattern are inserted into a skeleton to form different lexical categories, such as nouns, verbs, adjectives…etc. Nonconcatenating Templatic Morphology Languages (NTM) is a term used to describe languages that use the root, pattern, templates and affixes to form words. This account considers the template as the main unit in word formation, in contrast to derivational morphology that considers roots as the main unit.

Figure 5 shows the structure of the template according to Béland and Mimouni (2001). The most common templates that are used in Modern Arabic are CaCaCa, CaCCaCa, nCaCaCa, CtaCaCa, CiCaaC, CuCCaaC, CiCaaCii, CuCayyiC, maCCaCa, muCaaCiC , where C here represents the consonants of the roots (Truker, 2010).

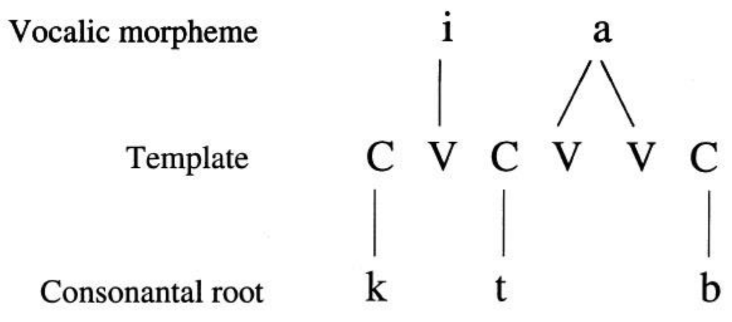

Figure 5: The root, vocalic pattern and the template (Béland and Mimouni, 2001, p.84).

These templates can be used with many roots to form the same category. For example, adding the root √k.t.b to the template maCCaCa, which represents the singular noun, gives /mak.ta.ba/ which means "library". Adding a different root, for example √d.r.s (which means "study") to the same template maCCaCa will give /mad.ra.sa/ (which means "school"). The template principle is called *Wazen* in Arabic and *Binyan* in Hebrew.

In summary, there are two different ways to look at the root in Arabic. Derivational morphology or root and pattern morphology suggests that words are derived from the root in different lexical categories. The NTM is another account that considers the template to be the main part in word formation; each template represents a specific lexical category when the root and the vocalic pattern are added to that template to form a word.

3.4.2 Word formation and models of morphology

Semitic languages in general, and Arabic and Hebrew in particular, have been studied in the morphology literature to explore if the root or the template plays a critical role in word formation. It is important to understand and determine whether word formation in Semitic languages is based on root or pattern or combination of the two in order to determine the morphological status of each one. McCarthy (1979) developed a generative account of root and patterns that considers the root as a separate morpheme that forms a word after adding a vocalic pattern. This account is supported by Marantz (1997) and Arad (2003; 2005) who assumed that root, vocalic pattern and templates are all morphemes and these three morphemes are responsible for word formation. However, some researchers (in Hebrew: Bat-El 1994, 2003; Ussishkin 2000; 2005; Arabic: McCarthy, 1993) argued that the consonantal root is not relevant in some or all types of word formation. In this account, the base for word formation is not a consonantal root, but an output form that has already been derived, as explained by the Output-Output Faithfulness constraint (Benua, 2000). Instead of an affixation process, an overwriting process occurs to form words.

Some of the studies that found that root is not necessarily needed for all word formation were based on the Optimality Theory (OT) of Prince and Smolensky (1993) and McCarthy and Prince (1993). Optimality Theory is an output oriented constraint-based framework. The structure of Optimality Theory consists of a generator, an evaluator, and constraints. The generator forms all possible output candidates for a given input, while the evaluator evaluates the candidates taking into consideration language-specific constraints. The optimal candidate is the one that does not violate the constraint ranking.

Some studies have tried to apply Optimality Theory to Semitic languages like Coptic and Arabic. Kramer (2007) concluded that the base for Coptic root and pattern morphology must be a consonantal root, and not an output form. Kramer's conclusion was based on his trial to apply the OT principles to the Coptic Language.

Tucker (2009; 2010) followed the same approach as Kramer by comparing the application of OT (the output form of the template) and generative theory (root and pattern morphology) to form a model for word formation in Egyptian Arabic morphology and verb derivation. Tucker developed a combined model for the two theories to explain the process of verb derivation in Arabic. He concluded that root morphology is essential, however the prosodic templates also play a role in forming words.

These studies found that root morphology is essential in word formation, however recent studies of modern standard Semitic languages started to accept the combined model of optimality and generative theories. In the following section, some word formation models will be discussed.

There are a number of theories that model the way words are accessed from the mental lexicon. The word as a mental lexicon has two parts, one related to the form of the word that includes phonological and morphological information and the other is related to the meaning and includes semantic and syntactic information of the word (lemma). An argument is raised between the different models that tried to explain word processing as they differ on whether the word is processed as a whole unit (whole word account) or broken down to smaller morphemes (decompositional account). Another account (mixed or dual route model) claims that the word cannot be processed as a whole word alone or as morphemes alone, instead the dual route model account suggests having mixed or dual model that includes both models of word processing.

The whole word access model was supported by Butterworth (1983), who claimed that all different forms of words are acquired as they are heard and these words are listed in the mental lexicon as a complete phonological form. The whole word route is considered as a fast mapping with one level of processing. This model does not explain for example how complex words, combined words, novel words or root and derivation-based words are processed as these types of words consist of two words or complex suffixes and prefixes. Furthermore, it does not explain how novel words link to their meanings.

On the other hand, Taft and Forster (1975) who supported the decomposed access models (morphological processing route) claimed that words are processed by isolation of the morphemes that compose the word. This is followed by access to meaning. Unlike the whole word access model, the decomposed access model of word processing slows down the time required to recognize the words while it helps to develop a meaning for novel words. The main criticism for this type of model is that it does not explain how simple words, which do not require sublevel representation, are processed in this model.

The third account by Caramazza et. al (1988) supports the mixed or dual route model. This model is a combination of the whole word access model and the decomposed access models. Caramazza et al. (1988) claimed that word processing required an activation for both models, where the whole word representation for familiar morphemes (e.g. root and affixes), and complex and novel words are processed through the decomposed access models. The parallel dual–route model (Baayen & Schreuder, 1995; 1999) added to the dual rout model another layer of word analysis: sensitivity to the frequency and familiarity of the word and morpheme. This layer of analysis will help in the decision of selecting the appropriate route for word processing. The assumption underlying these models is that the more frequent the word (morpheme, root or affixes) the more likely the word will be processed faster. Therefore, the whole word route is used with the familiar words, while novel words are processed in morphological process route and finally words that contain both familiar and unfamiliar units (e.g. familiar affixes with an unfamiliar root) will be accessed by the dual route processing. Many studies investigated the models used to process real and pseudo words in English and Hebrew. Studies of words by Burani and Caramazza (1987) and Burani and Baayen (2002) showed that access times and accuracy to suffixes and accuracy to suffixed derived words were significantly related to root familiarity. Another piece of evidence for the effects of familiarity of the suffixes was presented by Burani et al. (1997) who combined medium frequency roots with suffixes belonging to two distinct frequency ranges. The results showed that the roots combined with unfamiliar suffixes took longer to process when compared with roots with familiar suffixes. Data from Taft and Forster (1975) and Caramazza et. al (1988) showed that participants in the lexical decision task had difficulty rejecting nonword items that included a real morphemes or affixes that exist in the language, while it was easier for participants to reject nonwords that did not include any real morphemes. According to these findings, word processing cannot be explained without considering the word structure.

3.4.3 Word formation in root based languages

Each language has it is own lexical system that regulates word formation processes whereby novel words are created according to these language specific rules. Plaut and Gonnerman (2000) discuss morphological productivity, which is different in Hebrew compared to English. In English, there are many words that share the same root or stem (= root + affix(es)) but not the same meaning, for example *object, objection, objective, objectionable*. These words have similar bases but different meanings, but this is not common in root based languages. One of the main differences between Arabic and English according to Shamsan and Attayib (2015) is that in Arabic the root is consonantal and the pattern is the vowel/s affixed to the root, while in English morphemes are continuous and roots must have vowels. Affixes used to derive or inflect words from the root in Arabic may take the form of prefixes, infixes or suffixes. This contrasts with English where derivational and inflectional morphemes can only be suffixes or prefixes to the root/stem. Therefore, infixation in Arabic is very common and can disrupt the sequence of the root as in the word /ʔiq**t**atal/ “to fight each other”, where the sequence of the root √q.t.l “to fight or kill” is disrupted by the insertion of the consonant (morpheme) /t/ to indicate reflexivity. The difference in lexical rules between Semitic languages and English could cause different patterns in word formation, furthermore, the processes underlying nonword repetition could also be affected by these rules.

Deutsch and colleagues (1998) developed a model representing the Hebrew lexicon in a dual route system. The first route was the whole word level consisting of nouns and verbs, and the second route was a subword morphological level consisting of roots and patterns. They used a priming paradigm to investigate the role of roots and patterns in lexical access of *verbs* in Hebrew. They asked 96 undergraduates, all native speakers of Hebrew who were divided equally to perform two tasks: a lexical decision task and a naming task. Results showed there was a strong priming effect with root and pattern primes as participants’ reaction times were significantly faster on verbs that have roots and patterns. They concluded that while *only roots* facilitated lexical access in Hebrew nouns, roots *and* patterns played facilitatory effects in Hebrew verbs. They concluded that all words in Hebrew (nouns or verbs) are generally derived from roots and that the root is the basic morphological unit in the language. However, patterns play an important role in the verbal system of Hebrew as they are represented on the subword morphological unit along with roots. Further application of this model with adult participants was provided by Deutsch, Frost, Pelleg, Polatsek, and Rayner (2003). They used three types of written words that had different root and pattern combinations. They used an eye tracking task where fixation time and gaze duration were measured when participants were reading the stimuli that consisted of words and nonwords. Results showed that the root facilitated word recognition for verbs and nouns, while the pattern facilitated word recognition for verbs but not nouns. These differential effects for roots and patterns could be because Hebrew verbs consist of seven patterns only, while nominal patterns exceed one hundred patterns; this makes verbal patterns more salient and prominent than nominal patterns. The authors concluded that roots and patterns serve as an organizing principle of the Hebrew lexicon. They also concluded, when they applied the dual route model of Deutsch et al. (1998), that lexical processing in Hebrew may have a whole word retrieval level but it is mandatory to have the subword morphological level, as morphological decomposition is an essential part in lexical processing in Hebrew.

There are also many studies that have looked at the semantic side of the root as it plays a role in word processing in general. Berent and Shimron (1997) asked young Hebrew speaking adults to rate nonwords containing nonroot and root items, in order to examine their root structure sensitivity. The participants were asked to what extent a nonword sounded like a real word. The findings showed that participants were more likely to reject nonwords that were constructed from a nonroot. Additionally, Clark and Berman (1984) found that preschool children were able to predict the meaning of a novel word that included a real root but sounded like a nonword. For example, when children were presented with the nonword /limgof/, they responded by giving the meaning of ‘wearing boots’. Here, these children applied the meaning of /magfayi/) which means ‘boots’ in Hebrew as both the real words and the nonword shared the same root √ m.g.f.

In conclusion, it is clear from examining these studies that there are two different accounts to explain word processing; the whole word retrieval processing account versus the morphological structure account. There is also strong evidence that supports the semantic side of the root in word processing. All these factors play a role in word formation in different proportions. The typological similarities between Arabic and Hebrew as both are Semitic languages could allow us to extend these models to Arabic.

3.4.4 The development of root and pattern morphology

Understanding the development of word formation processes in the domain of the root and pattern is essential in Semitic languages. As mentioned above Semitic languages have a bounded morphology that has certain inflectional and derivational constrains on word formation. It is important to investigate if children who speak a Semitic language abide by these constraints in order to form new words. If they do, it is important to examine what developmental milestone map they follow in order to reach adult-like skills in forming new words. Furthermore, understanding the nature of the typical development and emergence of root and pattern morphology in Semitic languages will help to determine some of the underlying process related to this development. Nonword repetition tasks as discussed in section (2.1.1. and 2.1.2) help researchers and clinicians understand how children develop and acquire new words and explore the underlying process related to this development in TD children and those with language impairments. Therefore, it is essential to extend these valuable findings to languages like Arabic

Berman (2003) investigated word formation in 60 children aged from 3-10 years old and 12 adult speakers of Hebrew. The participants were required to interpret and produce innovative verbs based on familiar nouns or adjectives. The results of this study showed that children from the age of three could interpret and produce a novel verb based on the adjective or noun. Secondly, children were better at identifying a consonantal root in order to interpret a novel verb than when coining a verb from familiar adjectives or nouns. Thirdly, children used the morphological pattern (Binyan) in forming new verbs. The major difference between children and adults was the variability of their production. The adults in this study used the root more consistently and correctly than the 7 to 10 year old children, who in turn performed better than 3-4 year old children. The results of this study showed clearly that morphological skills develop with age and improve as children grow older. Results show also that these skills are required for word formation. Therefore, this study demonstrated that the process of word formation is a function of increased age and vocabulary knowledge.

Malenkey (1997) tested 100 children from kindergarten, 3rd, 6th and 9th grades who spoke Hebrew. One task tested these children's ability to isolate roots and patterns from nonwords. The other task tested their ability to use the same root and pattern in another word by analogy. Results showed that children's awareness of the root started at kindergarten, while pattern awareness did not start until the age of 10. A similar study by Karwar and Sakran (1997) tested 80 children from kindergarten and 1st, 2nd, and 4th grades and adults, all native speakers of Palestinian Arabic. Two tasks were presented to participants. The first task was a root relation task that required an identification of the root relation between two words such as kitab/maktabe "book/library"; both shared the root √k.t.b. The other task was an analogies task, which is an Arabic version of the task used by Malenkey (1997) described above. The analogies task tested participants' ability to apply a root to another word. Both tasks required some training before testing. The results showed an early awareness of the root starting from kindergarten. The same developmental trend seen in Malenkey (1997) was observed in these participants, however this Arabic study did not include investigation of the pattern. The results of these studies indicate that Hebrew and Arabic Palestinian speaking preschoolers showed an early and gradual awareness of root morphology as they recognized the root in relation to other words. By contrast, only older children were able to explicitly identify and analyze root morphology accurately and consistently.

Studies that investigated word formation in children acquiring Semitic languages (Clark & Berman 1984; Berman 1999; Ravid, 2000; Ravid, Avivi, and Levy, 2003) provided some general conclusions. Firstly, children's ability to coin new words from a familiar input increases with age. Secondly, children use the consonantal root to understand and produce new words. Thirdly, children as young as four years old were able to relate the root as a base for forming new words. Fourth, only at the age of 7 and older do children start to use the vocalic pattern knowledge to derived new words. Finally, even younger children can coin new words using one of the restricted templates available in their language.

3.5 Summary

In light of the studies that have been discussed in this chapter, four main points helped in directing aims, objectives and methods of the two experiments in this study. First, based on some of the evidence reviewed earlier, this study will follow the assumptions of the generative theory (McCarthy, 1979) which considers the root as the essential morpheme unit in word formation in Semitic languages. Second, the evidence provided by the semantic processing account showed that children were more sensitive to roots than they were to patterns (Berent and Shimron, 1997; Deutsch et al., 1998). Third, children started to recognize root at the age of the kindergarten while template awareness was only acquired at the age of ten. Finally, these results clearly showed a developmental trend for root and pattern morphological skills, where awareness of root morphology emerged significantly earlier than pattern morphology.

As the current study targets young children aged from 2-7 years old, it was preferred to investigate the root and pattern morphology but not template morphology as children's awareness about patterns and templates develop later (at age of 10 as discussed earlier, see Malenkey, 1997). In order to investigate the semantic versus the morphological aspects of root and pattern, a nonword repetition test was developed with different subtypes to understand the processes involved in root and pattern morphology and word formation. Finally, a preliminary investigation of the developmental trajectory of root and pattern morphology in Gulf Arabic speaking children at these young ages (2 to 7 years) will be investigated in this study.

4. The relationship between word and nonword repetition and receptive and expressive vocabulary skills in Gulf Arabic speaking children

4.1 Introduction

Many dimensions of nonword repetition skills have been investigated, especially those that examine its relationship with language impairment and the theory of phonological short term memory as discussed in the previous chapters. Moreover, performance on nonword repetition has been found to correlate positively with vocabulary size, particularly before the age of four (Bowey, 1996; Edwards et al., 2004; Gathercole & Baddeley, 1989, 1990; Gathercole et al., 1991) which adds to the value of NWR as a predictor or language impairment and vocabulary skills in children. Gathercole (2006) suggested that both NWR and vocabulary skills share the same process of phonological memory, which causes this positive relationship.

While most of the emphasis has been on NWR skills in children with language impairments, lately there has been an increasing interest in using real word repetition along with NWR to examine the contributions of the different processes involved in NWR. One of the main reasons to include real word repetition in some experiments, such as to help in controlling for articulatory processes since there are many young TD children who fail to complete NWR due to their limited articulation skills (e.g., Chiat& Roy, 2007; Stokes & Klee, 2009). Hoff et al. (2008) used the individualised scoring method of controlling articulation performance in nonword repetition in a different direction, as they used real word repetition along with nonwords. They argued that they used real words to control for articulation because real words place less demand on phonological memory than nonwords. Therefore, if the child was able to repeat a real word correctly, but not the nonword that consisted of the same phonemes, this indicated an error in phonological memory. However, if the child failed to repeat both the nonword and the real word, this indicated the child had limited articulation skills.

4.1.1 Real word repetition

Investigating real word repetition took another direction, in addition to its use to control for articulation in NWR tasks. When used along with the NWR (Chiat& Roy, 2007; Hoff et al., 2008; Roy & Chiat, 2004) word length was found to affect the accuracy of real word repetition in typically developing children and children with language impairment and in a differential manner: typically developing children were less affected by word length in real word repetition than NWR, as the word familiarity helped in improving their repetition accuracy (Chiat& Roy, 2007; Dispaldro et al., 2009; 2011; Roy & Chiat, 2004). However, Chiat & Roy (2007) found that children with language impairment were less affected by word familiarity, and that the interaction between length and word status (word vs. nonword) was not significant for the children with language delay. The authors attributed this to the idea that although real word repetition puts less demands on phonological short term memory compared with nonword repetition, children with language delay could not benefit from this advantage in their performance on real word repetition due to their language deficits.

Dispaldro et al. (2013) investigated the utility of real word repetition as a clinical marker in children with SLI, along with nonword repetition, in Italian children aged between 3;11-5;8. Both tasks, real word repetition and non-word repetition, succeeded in distinguishing the SLI group from the TD children. The authors explained that this was due to the common skills that are required for both of these tasks, such as PSTM, motor planning, and linguistic knowledge (e.g., phonotactics). However, the two tasks differed on which skills were more central for each task. For example, NWR relied on PSTM more than real word repetition did, while real word repetition relied more on semantic knowledge, i.e., previously learned words. And since children with language disorders are less efficient at learning words, this might have explained their poor performance on real words too (Dispaldro et al., 2013).

Dispaldro et al. (2011) found that real word repetition, but not NWR was an excellent predictor of grammatical abilities in very young TD Italian children. In their study, they administered three lists, a NWR list and two other lists that consisted of early acquired vocabulary and late acquired vocabulary. They also conducted two production tasks of grammatical structures, namely production of third person object clitic and third person plural inflections on present tense (e.g., /*mangiano*/ “they eat”). They found that while performance on NWR, but not real word repetition, correlated with word length, performance on real words (early and acquired) was a better predictor of grammatical abilities. Real word repetition showed that there were strong influences of lexical knowledge, where real words were retrieved from long term memory, and in the absence of lexical knowledge nonwords were retrieved mainly through phonological short term memory. Based on these findings, authors argued that real word repetition may provide a better reflection of children's overall linguistic abilities, due to the presence of lexical and semantic representations along with phonological representations. They explained that relying on phonological short term memory only may underestimate children's overall linguistic abilities. Therefore, they supported the use of real word repetition in clinical settings as a predictor of language abilities.

A summary of findings of studies that investigated the utility of real word repetition as a diagnostic tool showed that this task was very promising. A replication of these findings in different languages with different populations could contribute to clinical practice where real word repetition can be used instead of or along with NWR, especially for young children who might have difficulties repeating nonsense words. It would be easier for the examiner to use real words with very young children whose phonological development might not be fully matured. It is also important to investigate the predictive value of real word repetition to different language measures, such as vocabulary and morphology.

Investigators who studied real word repetition and NWR used different scoring methods when analysing the results of their tests. In the following section, the different scoring methods are discussed in detail to explore their effects on real words and NWR experiments.

4.1.2 Comparison of scoring methods

The two main scoring methods that are used in most word and nonword repetition studies are the whole word correct (WWC) method and the percentage of phonemes correct (PPC) method. The WWC method was used to score the CNRep test (Gathercole & Baddeley, 1996) where each item was scored as a whole, either entirely correct or incorrect. The PPC, on the other hand, was used to score the Nonword Repetition Test (NRT; Dollaghan & Campbell, 1998) where scoring was based on the number of phonemes produced correctly in each item. Graf Estes et al. (2007) compared these two scoring methods across 23 different English studies with TD children and children with SLI for word and NWR. Graf Estes et al. (2007) hypothesized that the WWC method could penalize children with SLI as these children might have more phonological errors compared to TD groups, while the PPC scoring method could provide a more appropriate evaluation of the two groups. The results of their investigation were contrary to the hypothesis; they showed that the difference between the TD and the SLI groups was smaller in the WWC scoring method (d=.48) than it was with the PPC method (d=1.17). The authors attributed these results to the large standard deviation when using the WWC method. Deevy et al. (2010) found that these two different scoring methods did not influence the groups' results on the NWR. The difference between the two groups based on the two scoring methods was minimal (d=2.26 vs 2.12). Contrary to previous English studies, an Italian study by Dispaldro et al. (2013), which investigated word and nonword repetition with TD children and children with SLI, using the WWC and PPC scoring methods, found that the magnitude of the group differences was greater under the WWC scoring method (d=2.57) than the PPC scoring method (d=1.38). These results showed that scoring methods may contribute to children's scores on word and nonword repetition. Both Dispaldro et al. (2013) and Deevy et al. (2010) did not allow for developmental phonological errors under the WWC scoring method, while they allowed it under the PPC scoring method. However, it was not clear why the Italian study found a significant effect for using different scoring methods for repetition tasks yet this effect was not found in the English study which used the same scoring methods. Further research is required to investigate the effects of using different scoring methods in different languages. Exploring the possible effect of using different scoring method on Arabic word and nonword repetition is an aim for this current study, especially due to the distinguished properties of Arabic phonological and morphological systems that set it apart from most European languages.

4.1.3 Aims

This study has the following aims: First, to investigate whether real word repetition and/or NWR have the potential to be used as diagnostic assessment tools to distinguish typically developing children from children with language impairment in Gulf Arabic speaking children. Second, it will provide some essential information regarding the nature of the relationship between real word repetition and nonword repetition and receptive and expressive vocabulary skills in Arabic speaking children with typical and atypical language development. Third, it examines the predictive value of real word repetition and/or nonword repetition with receptive and expressive vocabulary. Fourth, it will investigate if using two different scoring methods (PCC and WWC) will have any impact on these three areas of investigation. Fifth, it will investigate the effects of the item length on children's repetition accuracy. Furthermore, the present study will also add cross-linguistic evidence from Arabic to the existing literature on the role of real and nonword repetition in typical and atypical language development. As discussed earlier in section 2.6, some findings of non-Indo-European languages (e.g., Cantonese) did not find nonword repetition as a clinical marker in children with SLI (Stokes et al., 2006), while a previous study in Gulf Arabic that used phonologically complex nonwords with school aged children with SLI found a significant effect for NWR (Shaalan, 2010), this study uses less phonologically complex stimuli with younger children as we avoided the use of clusters. Finally, the current study will involve developing some new assessment materials that are necessary to evaluate phonological and vocabulary skills of children with and without language impairment in Arabic. Therefore, a battery of tests was used with these children that included: a word repetition test, a nonword repetition test, a receptive vocabulary test, and an expressive vocabulary tests. These tests will be described in the following section.

4.2 Tests Developed to Assess Children's Phonological and Vocabulary Skills

The children were tested using the following:

A word and nonword repetition test (WNRep). This test was developed for this current study to be used with Gulf Arabic preschool children, and is based on the methods used to create The Preschool Language Test (Roy & Chiat, 2004).

The Arabic Picture Vocabulary Test (APVT, Shaalan, 2010). This receptive vocabulary test was previously developed to be used with Qatari Gulf Arabic speaking children. It is based on the methods used to develop the Peabody Picture Vocabulary test (PPVT) (Dunn & Dunn, 1997) and showed good psychometric properties (see Shaalan, 2010).

The Arabic expressive vocabulary test (AEVT). This test was developed during this project by this researcher, and it is based on the methods used to create the Expressive One Word Picture Vocabulary Test (EOWPVT; Martins & Brownell, 2000).

Finally, two screenings for articulation disorders and developmental verbal dyspraxia were employed (Shaalan, 2010).

4.2.1 Challenges of conducting research in Arabic

It was necessary to develop new speech and language assessment material in order to achieve the aims of the current study. However, there were some challenges that faced the investigator during the process of developing and applying these tests.

The main challenge in developing assessment material in Arabic in general, and Gulf Arabic in particular, was the limited resources and references. There are no standardized speech or language tests targeting Gulf Arabic speaking children at any age, and at this young age in particular. There are a few studies that have developed assessment material as part of PhD theses; however, none of these tests was published. Shaalan (2010) developed a battery of language tests to investigate SLI in Gulf Arabic speaking children. These tests included the following: Sentence Comprehension (SC) test, Expressive Language (EL) test, Sentence Repetition (SR) Test, and Arabic Picture Vocabulary Test (APVT). These tests were conducted with 88 typically developing children and 26 children with SLI aged between 4:6 and 9;4 years old. He also developed a nonword repetition test and used it with a group of children (children with SLI, TD age-matched children, and TD language-matched children). The general findings of Shaalan's (2010) study showed that children with SLI performed significantly worse than the typically developing children on most tasks. Another unpublished assessment test was by Ayyad (2011), who developed a single –word and object elicitation test (eliciting words by showing pictures and objects) to evaluate consonants and vowels across different word positions, different word lengths and different word structures in monolingual 4-5 year old Gulf Arabic speaking children in Kuwait. The aim of Ayyad's study was to evaluate the phonological development for TD preschoolers in Kuwait. Alqattan (2015) investigated phonological development in children acquiring Kuwaiti Arabic before age 4 by analysing speech samples, through a cross sectional study for 70 typically developing children aged 1;4 and 3;7 years.

In this current study, the invistigator used some of Shaalan's (2010) speech and language battery tests, though the current study targets differnt age groups and goals. Furthermore, the rest of the studies that were mentioned above were either targetting different age groups and aims or were applied concurrently to this study, so the researcher in this current study was not able to access other studies' assessment matierlas. Therefore, it was necessary to develop assessment materiales specifically to be used in this project.

In the following, the tests used in this experiment are described.

4.2.2 The Arabic word and nonword repetition (WNRep) test

The WNRep test design. The main objective in developing the Arabic Word-Nonword repetition test (WNRep) was to assess two to four year old Gulf Arabic speaking children’s early phonological skills as measured by word length in syllables and to examine its utility in distinguishing children with language impairment from typically developing children. Furthermore, the investigator wanted to examine if word repetition or/and nonword repetition can predict receptive and expressive vocabulary size in both groups.

The WNRep was modeled following the Preschool Repetition Test (PSRep, Roy & Chiat, 2004). As mentioned earlier, the PSRep consists of 36 test items; 18 words followed by 18 nonwords. Words and nonwords are made up of equal numbers of 1-3 syllable items, with systematic manipulation of stress, so that half the words have strong/weak stress (SW), while the other half have weak/strong (WS) stress. The words and nonwords were phonologically matched, with nonwords created by altering the vowel in single syllable words and reversing two consonants in each word to create a corresponding nonword (Roy& Chiat, 2004).

To develop the WNRep test many factors were taken into consideration to control for variables that might influence this task. The variables that were considered were: articulatory complexity, word length, wordlikeness, language specific phonotactic rules, and word familiarity.

Articulatory complexity. Qatari Gulf Arabic has 30 consonants, however only 10 consonants were used in the WNRep, and most of them were early acquired sounds. The sounds used were: /b/, /d/, /t/, /k/, /f/, /s/, /m/, /n/, /l/ and /r/. The selection of these sounds was based on Ayyad's (2011) study as she found that 90+% of the younger children aged (45-54) months acquired the following sounds: /b/, /t/, /t/, /d/, /k/, /g/, /q/ /ʔ/, /m/, /n/, /w/, / j/ /r/,/ðˤ/, /ħ/,/h/, /tʃ/, and /χ/. Amayreh (2003) and Amayreh & Dyson (2000) found that children by age 3;10 mastered most of these sounds. Although /s/ and /r/ are not early acquired sounds, they were included as it was difficult to find a range of common words that do not contain one of these consonants. According to Bukshaisha (1985), all these consonants can occur in any position in Gulf Arabic words. Clusters were also avoided in this study as we tested young children who may not have mastered the production of clusters yet. Ragheb and Davis (2010) showed that a monolingual child acquiring Cairene Arabic and aged 2 years and 8 months had difficulties producing final clusters, which were commonly substituted with geminates (e.g., /bint/ "girl", was substituted with [bitt]. Ayyad (2011) found that only one consonant cluster was acquired in word-final position by the age of 4;0 in Kuwaiti Gulf Arabic. Therefore, clusters in this current study were avoided as the participants were aged between two and four years old and may not have mastered the production of clusters at this age. On the other hand, gemination which is defined as a cluster of two identical consonants, the first consonant occupies a syllable coda and the second consonant represents the onset of the following syllable (Delattre, 1971), was included as Alqattan (2015) found that 16% of words in Kuwaiti Gulf Arabic have geminate consonants. Therefore, geminate consonants, but not clusters, were included in this current study.

All Arabic long and short vowels according to those mentioned in an earlier section (see Figure 4) were included in the WNRep test: /a/, /a: /, /u/, /u:/, /i/, /i:/, /e:/, /o:/ (Mustafawi, 2006). No diphthongs were included in the test as Modern Standard Arabic (MSA) diphthongs (e.g., /ai/ and /aw/) are almost always realised as long vowels /ɛː/ and /oː/ (Ayyad; 2011). Furthermore, there are no studies that have shown the age of acquisition of diphthongs in Arabic.

Word length. The WNRep test consists of 48 items equally divided into one, two and three syllable items. Both the list of words and nonwords consisted of 8 one syllable items, 8 two syllables items and 8 three syllable items. The items on each list were ordered in a semi-random fashion.

Language-specific phonotactic rules. In order to respect the phonotactic rules of Arabic, vowel length, shape of syllables and stress of the words were kept the same for the nonwords. Therefore, all the nonwords kept the same word structure and consonants of the words from which they were created, however, although vowel length was kept the same, vowels were changed in all nonword items. The word structures used in both lists were as follow: for one syllable items the syllable structures used were CVC. For the two word length words and nonwords they were: CV.CVC and CVC.CVC, for three syllable words and nonwords, they were: CVC.CV.CV, CV.CV.CVC, CV.CV.CV, and CV.CVC.CVC. Four items in the WNRep test with the structures CVC.CVC and CV.CVC.CVC included a geminate which as mentioned earlier is a cluster of two identical consonants.

Nonwords were developed while respecting Arabic phonological constraints. Therefore the nonwords developed for the current experiment were phonotactically legal in Arabic, i.e., they respected phonological constraints in Arabic. One of the most important constraints in Arabic is called the Obligatory Contour Principle on place of articulation (OCP-Place), where consonants produced in the same place of articulation are disfavoured within the same root (Frisch, Pierrehumbert & Broe, 2004; Frisch & Zawaydeh, 2001). For example, roots with labial consonants, such as /b/, /f/, and /m/ are very rare in Arabic. For example, there are no words in Arabic that has the root √f.b.m, and therefore no nonwords were formed in this experiment based on these illegal combinations. Examples of consonantal roots that are phonotactically legal but do not exist in Gulf Arabic are √k.d.f or √s.b.d (Shaalan, 2010).

Wordlikeness. As mentioned earlier the WNRep test has two lists: a word list and nonword list. The word list has words that were selected from speech samples of Qatari children aged two to six years old, (which will be explained further in the next paragraph), therefore the words were common and familiar to the children. In contrast, the converted nonword list did not consist of any familiar roots and or patterns, therefore the nonword list was designed to avoid any wordlikeness.

The original words used to form the WNRep were all nouns; no verbs or adjectives or adverbs were used to create this test. As the test is designed for very young children, the words were familiar and common for the target age group. All the items used in the word repetition test were selected from speech samples that were collected in a previous study by this author with 56 Qatari children aged from two to six years old (Khater & Shaalan, 2007). The words and nonwords are phonologically matched, and nonwords were created by alternating the vowel in all syllables of the words taking in consideration to avoid any real vocalic pattern that is common Gulf Arabic. The vocalic patterns that were used with nonwords were as follow:

1. For the two syllabic nonwords: (i-u) (e.g., /bituk/), (u-u) (e.g., /sukkub/), (a-u) (e.g., /lakus/), (o-i) (e.g, /moril/) and (u-a) (e.g., /nujam/.
2. For the three syllable nonwords the vocalic patterns used were as follow: (a-a-i), (a- i) (e.g., /jasari/), (u-i), (e.g.,/kusimɑ/), (u-a-ɔ) (e.g., /lufanɔ/), (u-a-u) (e.g., /nufatul/) and (u-a-ə) (e.g.,/fumajjək/).

Furthermore, none of the nonwords that were transformed from the real words contain any real roots in order to avoid any morphological similarity; all the nonwords consist of nonroots. The nonexistence of these nonwords was also checked in an Arabic dictionary *Mu'jam Al-Waseet* (Mustafa et al., 2004) to ensure that no real root was used as a nonroot item. However, investigating the root and pattern was not an aim for this current study. Though the one syllable nonwords were not real words in Gulf Arabic on their own, it is possible to find those syllables as parts of real multisyllabic words. Moreover, there was no attempt to control for stress patterns in WNRep items, as items were randomly given strong/weak (SW) and weak/strong (WS) stress patterns. Table 4 shows some examples of how the words were transformed to create the nonwords. To see the complete list of words used in the WNRep test see Appendix A and Appendix B.

All the WNRep test items (48 words and nonwords) were recorded by a female native speaker of Qatari Gulf Arabic. Recording of stimuli was conducted in a soundproof room at the City University London Phonetics Laboratory.

Table 4: Examples of words and nonwords used in the WNRep test.

Word length	**Word structure**	**Words**	**Nonwords**
One syllable	CVC	/kis/ “bag”	/sɑːk/
Two syllables	CV.CVC	/seː.kəl/ “bicycle”	/lɑ.kus/
	CVC.CVC	/laimuːn/ “lemon”	/nul.jɑːm/
Three syllables	CVC.CV.CV	/tɑn.nuːrə/ “skirt”	/nat.ta.rɔ/
	CV.CV.CVC	/ti.li.fɔn/ “telephone”	/nu.fɑ.tul/
	CV.CVC.CVC	/mu.kaj.jəf/ “air-conditioner”	/fu.maj.jək/
	CV.CV.CV	/sa.ma.ka/ “fish”	/ku.si.ma/

Note: Full stop indicates syllable boundaries.

Word and nonword familiarity. After completing the WNRep, a familiarity task was conducted with five Gulf Arabic speaking adults (three males and two females) to rule out having any non-familiar words or familiar nonwords in the WNRep test. The researcher met the participants individually in a quiet room and they listened to the words and nonwords live. Each participant was asked to state if the word/nonword was familiar or non-familiar. None of the participants did identify any of the words as an unfamiliar, nor did they identify any of the nonwords as a familiar word.

4.2.3 Developing the Arabic Expressive Vocabulary Test (AEVT)

The Arabic Expressive Vocabulary (AEVT) test was developed to be used in this current study with Gulf Arabic speaking children, as there is currently no Arabic expressive vocabulary test available to be used with Gulf Arabic speaking children. This test was designed to be administered with young children aged from two to four years old. The Arabic Expressive Vocabulary test was developed following the methodology used in the Expressive One-Word Picture Vocabulary Test (EOWPVT, Brownell 2000). The EOWPVT is a standardized test that provides assessment for the verbal expressive vocabulary for English individuals aged 2;0 to over 80 years old. The Arabic version followed the same principles of the picture display, scoring and organizing the stimuli in groups according to age bands. The bands were 11 months for each group and there were 8 groups. Each group consisted of 8 pictures (with a total of 64 pictures for the whole test). There were also four trial items to enable the children to understand the test procedures and to be familiar with the examiner and the test itself. A booklet was made that consisted of 68 pages (4 pages for practice items and 64 for test items). Each page contained one coloured picture taken from non-copyrighted materials (Windows clip art). All pictures were taken from the same source to make sure that all pictures had the same characteristics.

Selecting AEVT items was based on item type or category and item difficulty. So the items belonged to different groups, (e.g., actions or verbs, single and plurals nouns that belonged to different categories (households, animals, food, clothes, toys), Secondly, the items were arranged according to their difficulty from easy to difficult to follow the normal expressive vocabulary development in Gulf Arabic speaking children. To order the test items according to their difficulty, two methods were used to determine the words' difficulty level to children. The first method was using representative speech samples that were collected in a previous study by Khater & Shaalan (2007), from 56 Qatari children aged 2;0- 6;0 years old. The items used in the test for each age band were selected from the speech samples at the same age group. The other method used was a familiarity rating collected from 24 Qatari Gulf-Arabic speaking adults for 600 words (Shaalan, 2010) that was used in order to develop the Arabic Picture Vocabulary test (APVT) which is also used in this current study. Table 5 shows some examples of the words selected in the AEV. For the complete list of AEVT words see Appendix C.

Table 5: Examples of some of the words used in the AEVT test.

Age	2:02:11	3:0-3:11	4:0-4:11	5:0-5:11	6:0-6:11	7:0-7:11	8:0-8:11
Word	ku:ra	qami:sˤ	hadijja	yasgi:	malʕab	ħaʃara:t	ʔustˤuwani:
gloss	“ball”	“shirt”	“gift”	“watering”	“stadium”	“insects”	“cylinder”
Word	sajjara	kursi	murabbaʕ	ʕankabu:t	ʤisir	kawkab	mintˤa:d
gloss	“car”	“chair”	“square”	“spider”	“bridge”	“planet”	“air balloon”
Word	ta:kil	ʃaʤara	ðibbana	mastˤtara	quful	timΘa:l	ʕadasa
gloss	“she eats”	“tree”	“fly”	“ruler”	“lock”	“statue”	“lens”

4.2.4 The Arabic Picture Vocabulary Test (APVT) (Shaalan, 2010)

This test was developed to be used with Qatari Gulf Arabic speaking children by Shaalan (2010). The researcher in this current study modified the order of the APVT items according to the item analysis done by Shaalan (2010) following testing 107 children (4;6 to 9;6 years old) in order to be used in this study. Furthermore, two pictures were also modified or changed by the current author as they were found to be difficult to recognize according to Shaalan (2010). The APVT consists of 132 words and organised into 11 groups of 12 words ranked according to their difficulty based on a familiarity rating (Shaalan, 2010) of each item. The criteria for choosing these words were similar to those used in the British Picture Vocabulary test (BPVT, Dunn et al, 1997). It is worth noting that the distractors used in both BPVT and APVT did not follow a certain pattern as no attempt was made to arrange distractors in a systematic way.

4.3 Hypotheses and Predictions of the Study

The WNRep test, together with the receptive and expressive vocabulary tests, were administered to a group of typically developing children (TD) and a clinical group of children with language impairment (CL) to evaluate the following hypotheses:

1. The TD group will have significantly higher scores than the CL group on all measures (repetition of 1-, 2- and 3-syllable words and nonwords, as measured by percent phonemes correct (PPC) and whole words correct (WWC); receptive vocabulary and expressive vocabulary.
2. Item length will affect repetition performance in both groups, with a significant decrease in scores as length increases.
3. Word type will affect repetition performance in both groups, with significantly higher scores for words than nonwords.
4. Scores for WNRep will be significantly correlated with scores on the APVT and AEVT in both groups, and using both scoring methods.

As indicated in the above hypotheses, it was furthermore hypothesised that using two different scoring methods (Percentage of Phonemes Correct (PPC) and Percentage of

Whole Words Correct (WWC)) would not affect outcomes of this current study, i.e. the two scoring methods would yield the same effects of group and item factors.

4.4 Methods

4.4.1 Participants

Forty four typically developing children (TD) and 15 clinical children (CL) were tested. The mean age of the TD children was 38.4 months or 3;2 years. The mean age of the children in the CL group was 43 months or 3;5 years. A summary of participants' characteristics is shown in Table 6 below.

Table 6: Summary of the characteristics of participants

Participants	TD Group	Clinical Group
Number of participants (Male: Female)	44 (21:23)	15 (10:5)
Mean age in months (years)	38 (3;2)	43 (3;7)
Range in months (years)	27-47(2;3- 3;11)	33-57 (2;9-4;9)

The TD children were recruited from two kindergartens whose managers were willing to distribute invitation letter to parents. Ethical approval was obtained from City University (see Appendix J) and was submitted to the hospital in Qatar where some of the participants had been enrolling in speech therapy services.

Friends and acquaintances helped also in recruiting more. Only Qatari children who speak the Qatari dialect as a first language were included in this study. All typically developing children who were included in this study had no history of speech or language delay or impairment and did not have any neurological, developmental or cognitive impairments, such as congenital malformation, hearing loss, or autism. This information was obtained from the consent forms that were completed by the parents of all children. (See appendices D and E and for parent information sheet and consent forms).

The children in the clinical group (CL) were recruited from the Speech and Language Department at Hamad Medical Corporation, the main government hospital in Doha, Qatar. Children who were referred to the clinic with impaired language not combined with any history of congenital abnormalities, cognitive disabilities, hearing loss, oral – motor difficulties or autism were included in this study. The criteria of selecting the CL group in this experiment might be very similar to the diagnostic criteria of SLI according to SLI definitions by Bishop (1997) and Leonard (1998), however it was not possible to label the CL group in this study as an SLI group, due to the insufficient investigations by the referral sources (e.g., no IQ tests were conducted in the speech clinic). Furthermore, the debate about labeling children with specific language impairment (SLI) increased after the recent decision to not include SLI in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) (American Psychiatric

Association, 2013). For this reason, the *International Journal of Language and Communication Disorders* dedicated an issue in 2014 (Ebbels, 2014) to discuss the advantages and disadvantages of continuing to use the SLI term to diagnose children with unexplained language disorder. Most of the commentators favoured dropping SLI as a term, as research has not provided strong evidence that supports the continued use of the current definition and the exclusionary criteria (Reilly et al., 2014). But others argued that changing the label risked breaking the link with past research (Gallagher, 2014; Rice 2014; and Taylor, 2014).

Initially the researched attempted to have divide children equally into 6-month age bands, but it was difficult to find an adequate number of Qatari 2-3 year old children from nurseries, as Qatari families prefer not to send their children at this age to nurseries. Therefore, it was not possible to have *age groups* within the CL or TD with equal numbers of children in each age band. Therefore, participants were divided into CL and TD groups with no specific age bands within each group.

4.4.2 Procedures and scores

All children completed the battery of tests (APVT, EVT, and WNRep) in the same session, which ranged between 45 and 90 minutes depending on child's ability to tolerate the tasks. All children received an articulation screener and a developmental verbal dyspraxia screener (Shaalan, 2010) to rule out any severe phonological disorders.

Children who needed a break were given 10 to 15 minutes to play with some toys. Finally, the WNRep task was then administered, with the word repetition task followed by the NWR.

The APVT was typically used first with the children as it is less demanding compared with the other tests that were used in this study. Children were given the following instructions “Together we will see a picture book. I will name one of the pictures and I want you to point to the picture I am talking about. Let’s try a couple of pages”. This was followed by two practice items (“shoe” and “fish”). For each item children were presented with four pictures and they were required to point to the correct response. None of the children had any difficulties with the instructions. All responses were recorded on a score sheet and children got 1 for a correct answer and 0 for an incorrect answer. The test stopped after eight errors in one group. The total score is the total number of correct answers.

The Expressive Vocabulary Test (EVT). Children were given the following instructions in Arabic “you will see some pictures and I want you to name what you see in the picture. Let’s try some pictures…”. This was followed by four trial items. None of the children had difficulty with the instructions. All children’s responses were written on a score sheet and were audio taped through a microphone attached to an Olympus VN-5500PC DNS Digital Voice Recorder. Children got 1 for a correct answer and 0 for an incorrect answer. The responses with incorrect pronunciations were considered correct if they were intelligible for the examiner. When the child was unable to correctly name six consecutive illustrations, testing was discontinued. The total raw score was computed by subtracting the number of errors the child made from the last ceiling item. For example, a child who stopped at item number 30 and had 10 errors would have a raw score of 20. The scoring method used was adapted from the one used in the Expressive One-Word Picture Vocabulary Test (EOWPVT, (Martins & Brownell, 2000). The total time for administration and scoring was 15- 20 minutes. The children were rewarded after completing this test with stickers or small toys.

The WNRep test. All testing was conducted in a quiet room. The instruction for each child was the equivalent of the following (in Arabic) "You will listen to funny and mixed up words and I want you to repeat them the way you hear them. Now let's try this". This was followed by three trial items. Stimuli were presented from a laptop through a pair of external speakers. Stimuli were repeated when the child did not pay attention to the first production of the stimuli. This happened more often with the younger children. Children were never presented with the stimuli more than twice. No response was recorded as zero. Few self corrections were noted and accepted as correct responses. Children's productions were audiotaped through a microphone attached to an Olympus VN-5500PC DNS Digital Voice Recorder and analysed later. All children's responses were scored using two different methods. First, for the whole words correct

(WWC) method, each repetition was scored either correct (1) or incorrect (0). Minor misarticulations (especially distortion of /r/, and /s/ or substituting /l/ for /r/ or / θ / for /s/) were counted as correct. As mentioned earlier these sounds may not be mastered at this young age. Second, the children's responses were analysed and scored using the Percentage of Phoneme Correct (PPC) method, where the total numbers of correct phonemes in each item of the word and nonword were added and divided by the total number of phonemes to give a percentage that constituted the PPC. The total number of phonemes in both repetition tasks were 122.

Reliability. Reliability is defined as the consistency of performing a test (DeVellis, 2012) and it is measured to ensure that changes in test scores should be only due to the changes in the variable being measured. Inter-rater reliability is a method used to assess reliability and it measures the correlation between the scores of two different examiners or raters. In both experiments in chapters 4 and 5, a second examiner who is a Gulf Arabic speaking speech-language therapist and who has experience in scoring and administrating repetition tasks was asked to score 10% of the children's scores using the PPC scoring method in WNRep and Arabic Expressive Vocabulary test (EAVT). The inter-rater agreement between the two examiners was (α =1.0) for the EAVT, (α =.90) for WNRep test. These results indicate a higher level of inter-rater agreement.

4.5 Results and Analysis

The first question that this study tried to address was whether the performance of Typically Developing (TD) and Clinical (CL) groups differed on the WNRep test that consisted of one, two, and three syllable long words and nonwords. Therefore, a repeated measures ANOVA was conducted to investigate the main effect and the interactions of the variables. The second question was to investigate the correlation between the TD and CL scores on the WNRep test and the two receptive and expressive vocabulary tests used in this study. This was performed to evaluate the predictive value of the vocabulary size based on the repetition skills. Finally, this study investigated if using two different scoring methods, namely Percentage of Phonemes Correct (PPC) and Percentage of Whole Words Correct (WWC) will affect the interpretation of the outcomes. Therefore, all analyses were conducted using both scoring methods. All analyses were conducted using the SPSS statistical package, Version 18 (SPSS Inc., 2009). The first section presents the descriptive and inferential statistics of the PPC scoring method followed by the WWC scoring methods results. Gender analysis was conducted to explore the effects of gender on children's performance from both TD and CL groups across tasks. Results showed there was no significant difference between the performance of male and female participants in both TD and CL groups and across different repetition and vocabulary tasks (see Appendix K).

4.5.1 Results of PPC scoring method

Descriptive statistics of WNRep. The descriptive statistics of the children's performance on word and nonword repetition for both CL and TD groups, using the PPC scoring method, are presented in Table 7.

The descriptive statistics of the TD and CL group scores on different word and nonword lengths (one, two and three syllable) and for each word type are presented in Table 8.

Table 7: Scores of for word and nonword repetition tests for both typically developing children (TD) and clinical children (CL) Percentage of phonemes correct (PPC) method

	Group	Age (months)	Word repetition	Nonword repetition
TD	M	38.39 (3:3)	96.7	79.2
N=44	SD	6	6.45	11.62
	Range	26-47	91-122	73-122
CL	M	43.3	85.36	69.9
N=15	SD	8.09	15.93	20.13
	Range	33-57	66-122	55-110

Table 8: Children's mean scores (and standard deviations, SDs) on the different word lengths for the words and nonword repetition test, based on the PPC (Percentage of Phoneme Correct) scoring method. TD = typically developing children, CL= clinical group.

Repetition Type	Group Type	Mean	SD	N
1 syllable words	TD	99.71	5.78	44
	CL	93.88	7.53	15
2 syllable words	TD	97.09	5.33	44
	CL	87.28	12.11	15
3 syllable words	TD	95.12	8.32	44
	CL	80.12	18.66	15
1 syllable nonwords	TD	95.64	6.72	44
	CL	86.66	14.45	15
2 syllable nonwords	TD	88.68	11.24	44
	CL	81.08	14.03	15
3 syllable nonwords	TD	85.78	14.29	44
	CL	69.21	23.17	15

The first question in this experiment was whether the performance of the TD and CL group differed across different word lengths and word types. It was hypothesized that the TD group would score higher than the CL group across different word types and lengths. Furthermore, it was expected that both groups would perform better on words vs nonwords and shorter words/nonwords vs longer words/nonwords. To answer these questions, a repeated measures ANOVA was conducted to investigate the effects of word length (3: one, two and three syllables' length), and word type (2: words and nonwords) as within subject factors, and group (2: TD vs. CL) as a between subject factor. There was a significant main effect of group (F (1,57) =20.59, p=<.001, η^2=.265) as the TD group had significantly higher PPC scores on word/nonword repetition (M=87.95, SD=9.035) than the CL group (M=77.63, SD=18.03. There was also a significant main effect of word type (F (1,57) =42.4, p<.001, η^2=.4.27) as children's performance was better for words (M=93.8, SD=9.3) than nonwords (M=76.9, SD=7.6). A significant main effect of word length was found (with Greenhouse-Geisser correction which was conducted due to violation of spehcericity) ε = 0.785, F(1.51, 86.2) = 4.61, p < .05. Children's PPC scores decreased as the word length increased (1 syllable: M =85.19, SD=10.24; 2 syllables: M=69.17, SD=14.78, 3 syllables: M=59.75, SD=20.79). The interaction between word length and word type was significant (ε = 0.869), F(1.73, 99.0) = 2.54, p < .05.

In addition to these significant main effects, there were also significant interactions between group and word length (F (1,57) =5.15, p=.027, η^2=.083), and between word type and length (F (1,57) =1.37, p<.001, η^2=.002). The interaction between group and word type, on the other hand, was not significant (F = (1,57) =.126, p=.723, η^2=.002), and nor was the three-way interaction between group, word type and length (F (1,57) = .137, p=.713, η^2=.002). Figure 5 illustrate the effects of group and word length and the interaction between these factors across the words types respectively.

To follow up the word length main effect a Bonferroni corrected post hoc test revealed a significant difference in children's performance from both groups on different word and nonword lengths (one, two and three syllables). The longer the word and the nonword was, the less accurate the repetition was. Children's scores were significantly higher on one syllabic words/nonwords when compared to two syllable words and nonwords ($p<.05$). They also scored higher on one syllable words and nonwords when compared to three syllable words and nonwords ($p<.05$). Similarly, their scores on the two syllabic words/nonwords were significantly higher than on the three syllable words and nonwords ($p<.05$). To follow up the interaction between word length and groups, an independent t-test was conducted to compare each group (TD and CL) at each one, two and three syllables word/nonword length. Results from the statistical comparisons (see Table 9) showed that there was a significant difference between the groups at one and two and three syllables (words and nonwords). The CL group performed significantly less accurately than the TD group at every length of word and nonword (see Table 9 for means and SDs for words and nonwords separately and Figure 7 for means across the word type).

To investigate if the word length is significant within the TD group performance on one, two and three word/nonword repetition, a Paired Sample t-test showed that one syllable vs two syllable is significant ($t(43) = 3.5$, $p=.001$), two syllable vs. three syllable is significant ($t(43) = 4.14$, $p<.001$) and one syllable vs. three syllable was also significant (t (43) =2.52, p=.015). Likewise, in the CL group there was a significant difference between the different word lengths: one syllable vs two syllable was significant (t (14) =2.36, p=.033), two syllable vs. three syllable was significant (t (14) =4.17, p=.001) and one syllable vs. three syllable was also significant (t (14) =3.5, p=.004).

Therefore, the interaction between group and length must have arisen from the magnitude of the difference between groups at different lengths. Looking at the effect sizes in Table 9 and slope of the graphs in Figure 6, it is evident that the magnitude of the difference between groups increases with length, with the most marked difference between groups occurring in the three syllable items. Hence, the CL group were more affected by word/nonword length than the TD group.

Table 9: Summary of the independent samples t-test results comparing TD and CL groups at each length for word and nonword repetition using PPC scoring method.

Word Length	t - value	Significance	Effect size
One syllable	t(57) =3.24	p<.001	η^2=0.304
Two syllable	t(57) =3.36	p=.001	η_2=.0.431
Three syllable	t(57) =3.27	p<.001	η^2=0.400

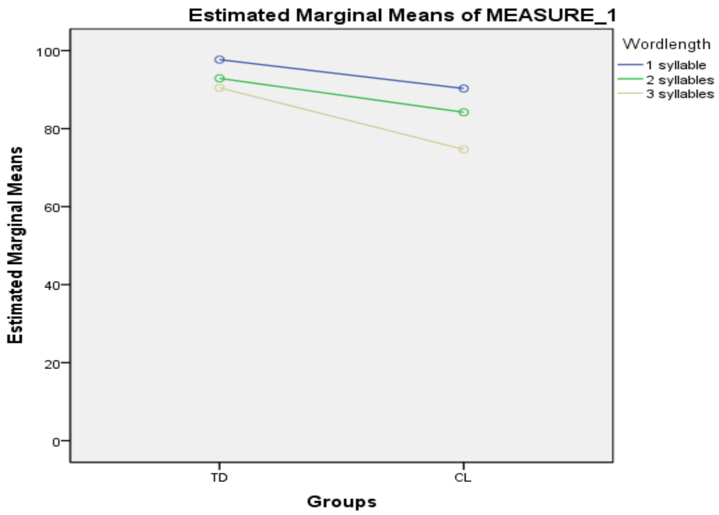

Figure 6: Performance of TD and CL groups on WNRep test (one, two and three syllables length) using PPC scoring method.

Next, to follow up the significant interaction between word length and word type, paired-samples t-tests were conducted to compare the combined groups' performance on word and nonword repetition at each length. The results showed significant differences between words and nonwords at every word length: one syllable (t (58) = 4.3, p<.001), two syllables (t (58) =5.9, p<.001), and three syllables (t (58) =5.4, p=.001). There were also significant differences when the one syllable nonwords were compared with the two syllable nonwords (t (58) = 3.2, p=.002), and the two syllable nonwords with the three syllable nonwords (t (58) = 3,32, p<.001). Thus, the accuracy of nonword repetition increased when the nonword length decreased. Similarly, one syllable words were repeated significantly better than two syllable words (t (58) =3.2, p=.002), one syllable words than three syllable words (t (58) = 4.1, p<.001) and two syllable words than three syllable words (t (58) =3.3, p=.002. Therefore, the interaction between word type and word length must have arisen, again, from the magnitude of the difference between the word type at different lengths. Looking at the slope of the graph in Figure 7 is evident that the magnitude of the difference between word type increases with length, with the most marked difference between word type occurring in the three syllable items. Hence, the nonwords repetition are more affected by increasing of the length than the word repetition.

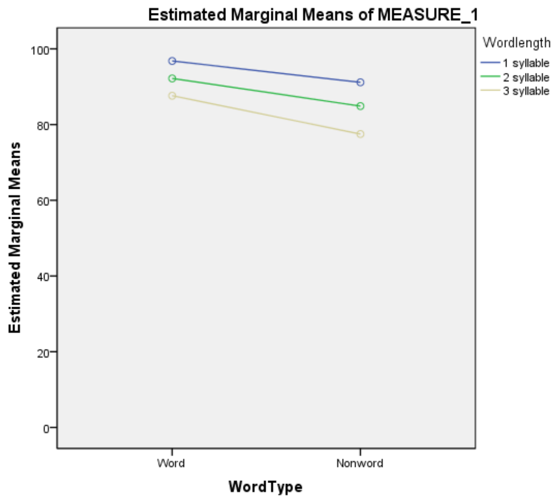

Figure 7: Children's performance on word and nonword repetition test (one, two and three syllables length) using PPC scoring method

The results of the main ANOVA and the follow up analysis are consistent with the hypothesis raised in section 4.3. The TD group scored higher than the CL group across different word types and lengths. Moreover, both groups had superior scores in repeating words vs nonword and higher scores in repeating shorter words/nonwords vs longer words/nonwords.

4.5.2 Results of WWC scoring method

Descriptive statistics of WNRep. The descriptive statistics of the children's performance on word and nonword repetition for both CL and the TD groups, using the WWC scoring method are presented in Table 10.

Table 10: Scores of for word and nonword repetition tests for both typically developing children (TD) and clinical children (CL) based on whole word correct (WWC) method

	Group	Age (months)	Word repetition	Nonword repetition
TD	M	38.39 (3:3)	92.5	75.08
N=44	SD	6	2.44	3.49
	Range	26-47	16-24	11-23
CL	M	43.3	60.8	49.38
N=15	SD	8.09	6.09	5.36
	Range	33-57	6-24	2-18

The descriptive statistics of the TD and CL group scores on different word and nonword length (one, two and three syllable) are presented in Table 11 including mean and standard deviation in each word and syllable type.

Table 11: Children's performance on the different syllable lengths for words and nonword repetition test based on the (WWC) scoring method.

Repetition Type	Group	Mean	Std. Deviation	N
One syllabic words	TD	97.72	5.57	44
	CL	86.66	15.28	15
Two syllabic words	TD	91.19	11.62	44
	CL	62.50	26.72	15
Three syllabic words	TD	88.35	16.03	44
	CL	51.66	37.16	15
One syllabic nonwords	TD	87.21	14.38	44
	CL	69.16	24.94	15
Two syllabic nonwords	TD	72.15	18.45	44
	CL	50.83	28.13	15
Three syllabic nonwords	TD	67.32	23.68	44
	CL	31.66	21.58	15

The first question in this experiment was whether the performance of the TD and CL groups was different across different word lengths and word types. It was hypothesized that the TD group would score higher than the CL group across different word types and lengths. Furthermore, it was expected that both groups would perform better on words vs nonwords and shorter words/nonwords vs longer words/nonwords. A repeated measures ANOVA was conducted to investigate the effects of the word length (3: one, two and three syllables' length), and word type (2: words and nonwords) as within subject factors, and group (2: TD vs. CL) as a between subject factor. There was a significant effect of group $F(1,57)=20.38$, $p=<001$, $\eta^2=.253$) as the TD group had significantly higher WWC scores on word/nonword repetition (M=83.7, SD=2.9) than the CL group (M=55.09, SD=5.7). There was also a significant main effect of word type ($F(1,57)=42.46$, $p<.01$, $\eta^2=.427$) as children's performance was better for words

(M=85.24, SD=19.75) than nonwords (M=68.71, SD=19.75), and significant main effect of word length ($F (1,57) =19.93$, $p<.001$, $\eta^2=.416$) was also found as the WWC score decreased as word length increased (1 syllable: M=95.8, SD=7.37; 2 syllables: M=90.6, SD=9.48, 3 syllables: M=79.94, SD=13.76). Mauchly's test indicated that the assumption of sphericity was violated ($\chi2(2) =$, $p < .05$), therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity. Main effects of word length was significant ($\varepsilon = 0.757$), $F(1.51, 86.24) = 29.54$, $p < .05$. The interaction between word length and word type was also significant ($\varepsilon = 0.869$), $F(1.73, 99.0) = 2.54$, $p < .05$.

In addition to the significant main effects, there were also significant interactions between group and word length ($F (1, 57) =4.72$, $p=.012$, $\eta^2=.146$). The interaction between group and word type, on the other hand, was not significant ($F (1, 57) = .424$, $p=.518$, $\eta^2=.098$), nor was the three-way interaction between group, word type and length ($F (1, 57) = 1.34$, $p=.264$, $\eta^2=.023$). Figure 8 illustrates the effects of group and word length and the interaction between these factors across word types respectively.

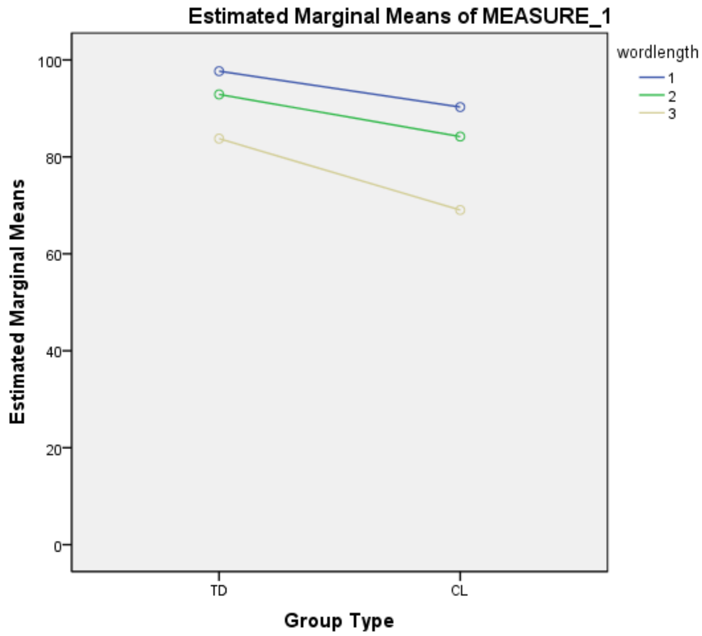

Figure 8: Performance of TD and CL groups on (one, two and three syllables length) NWRep using WWC scoring method.

To follow up the main effect of word length a Bonferroni corrected post hoc tests revealed a significant difference in children's performance from both groups on different word and nonword lengths (one, two and three syllables). The longer the word and the nonword the harder it was to repeat. Children's scores were significantly higher on one syllabic words /nonwords when compared to two syllable words and nonwords ($p<.05$). They also scored higher on one syllable words and nonwords when compared to three syllable words and nonwords ($p<.05$). Similarly, their scores on the two syllabic words/nonwords were significantly higher than on the three syllable words and nonwords ($p<.05$). To follow up the interaction between the groups and word length, an independent samples t-test was conducted to compare groups (TD and CL) at each one, two and three syllables word/nonword length. Results in Table 12 showed that there was a significant difference between the groups at one and two and three syllables (words and nonwords). The CL group performed significantly less accurately than the TD group at every length of word and nonword (see Table 11 for means and SDs, and Figure 8).

To investigate if the word length is significant within the TD group performance in one, two and three word/nonword repetition, Paired Samples t-tests showed that one syllable vs two syllable is significant (t (43) =1.9, p=.005), two syllables vs. three syllables is significant (t (43) =4.4, p<.001) and one syllable vs. three syllables was also significant (t (43) =4.1, p<.001). The paired samples t-tests for the CL group showed significant differences between one syllable vs three syllable (t (14) =3.51, p=.003), two syllables vs. three syllables is significant (t (14) =3.79, p=002) but one syllable vs. two syllables was not significant (t (14) =0.75, p=.465).

Therefore, the interaction between group and length must have arisen from the magnitude of the difference between groups at different lengths. Looking at the effect sizes in Table 12 and slope of the graph in Figure 8, it is evident that the magnitude of the difference between groups increases with length, with the most marked difference between groups occurring in the three syllable items. Hence, the CL group are more affected by word/nonword length than the TD group and the magnitude for the CL group between one and two syllables were not significant.

Table 12: Summary of the independent samples t-test results for group performance at each length for word and nonword repetition using WWC scoring method.

Word Length	t - value	Significance	Effect size
One syllable	t(57) =3.74	p<.001	η^2=0.432
Two syllable	t(57) =3.36	p<.001	η^2=.0.400
Three syllable	t(57) =4.121	p<.001	η^2=0.511

Next, to follow up the interaction between word length and word type, a paired-samples t-test was conducted comparing the combined groups' performance on word and nonword repetition at each of the three word lengths. The results showed a significant difference between words and nonwords at: one syllable length (t (58) =4.32, p <.001, two syllable length (t (58) =5.9, p<.001), however there was no significant difference between the three syllable words and nonwords length (t (58) =1.74, p=.087). There was also a significant difference when the one syllable nonwords were compared with the two syllable nonwords (t (58) = 3.2, p=.002), and the two syllable nonwords with the three syllable nonwords (t (58) = 3.32, p<.001) one syllable nonword with three syllable nonwords (t(58)=5.6, p<.001). Similarly, one syllable word length was significantly better than two syllable words repetition (t (58) =5.2, p =.002). And one syllable words vs. three syllable words (t (58) = 13.4, p<.001). Children also scored better in two syllabic words repetition than three syllabic words (t (58) =19.9, p<.001). Therefore, the interaction between word type and word length have probably arisen from the magnitude of the difference between the word type at one and two length but not the three word/nonword length.

The results of the main ANOVA and the follow-up analysis are consistent with the hypotheses put forward in section 4.3. The TD group scored higher than the CL group across different word types and lengths. Moreover, both groups scored better when repeating words vs nonwords and better when repeating shorter words/nonwords vs longer words/nonwords.

Groups' performance on receptive and expressive vocabulary tests

To investigate the children's performance on the receptive vocabulary test (APVT), an independent sample t-test- was conducted for the TD group and CL group. It was hypothesised that the CL group will score significantly less than the TD group in APVT. Results showed that the difference was not significant between the TD group (M= 23.4, SD= 7.2) and CL group (M= 20.8, SD= 6.9) in APVT (t (57) = 1.21, p = .228). This result was contrary to what was hypothesised. Figure 9 shows the scores of the TD and CL group on the APVT.

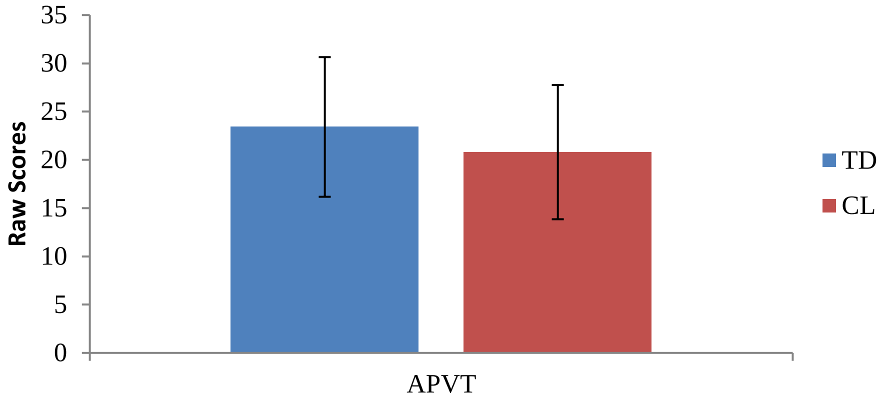

Figure 9: TD and CL performance in Arabic Picture Vocabulary test (APVT).

Another independent sample t-test was also conducted for TD group and CL group to investigate children's performance on the expressive vocabulary test (AEVT). It was also hypothesised that the CL group will score significantly less than the TD group on AEVT. Results showed that the TD group (M=18.84, SD=5.63) scored marginally higher than the CL group (M=15.73, SD=5.42), however the difference between the two groups just failed statistical significance ($t(57) = 1.85$, $p = .069$). This result was contrary to what was hypothesised. Figure 10 shows the group's performance on AEVT.

The lack of significant difference between the CL and the TD group could be due to the small CL group sample (N=15), and it also could be due to three outlier children from the CL group who performed well in all tests though they met the criteria of the CL group selection. The children's scores in both vocabulary tests the APVT and AEVT were recalculated while excluding the three outlier children (n=12). The results of the revised independent sample T-test showed that the TD group performed significantly better than the CL group on both receptive and expressive vocabulary tests. On the receptive vocabulary test (APVT), the TD group performed significantly better (M=23.4, SD= 7.2) than the CL group (M=18.17, SD= 4.8), t (54) = 2.36, p = .022). On the expressive vocabulary test (AEVT), the TD group had significantly better raw scores (M=18.8, SD= 5.6) than the CL group (M=13.92, SD= 4.3), t (54) = 1.85, p = .007).

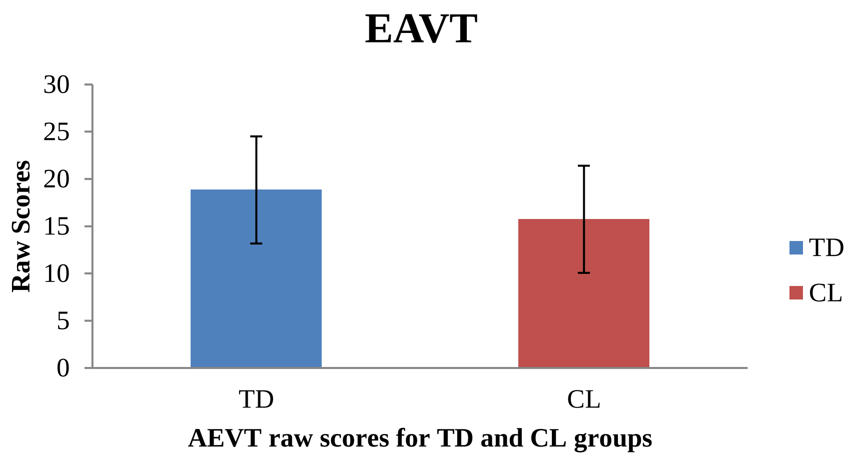

Figure 10: TD and CL groups' performance on Expressive Arabic Vocabulary test (AEVT).

4.5.3 Correlation between word and nonword repetition and receptive and expressive vocabulary in TD children

A Pearson product-moment correlation coefficient was calculated to measure the correlation between word and nonword repetition and the Arabic Picture Vocabulary Test (APVT) and the Expressive Arabic Vocabulary Test (AEVT) skills in TD children (n=44), using the WWC scoring method. It was expected that word and nonword repetition scores would significantly correlate with receptive and expressive vocabulary test on one hand and the word repetition scores would also significantly correlate with nonword repetition on the other hand. Results of the various correlations are shown in Table 13. The reference of labelling the correlation strength in this section are based on the following r, 0-0.19 is regarded as very weak, 0.2-0.39 as weak, 0.40-0.59 as moderate, 0.6-0.79 as strong and 0.8-1 as very strong correlation

For WWC scoring method Results showed that word repetition test significantly correlated with age (r=.51, p<.000), the APVT (r=.49, p<.001), the AEVT (r=.65, p<.001), and the nonword repetition test (r=.56, p<.001). In addition, the nonword repetition test significantly correlated with age in months (r=.38, p<.011), the APVT (r=.37, p<.013), the AEVT and the word repetition test (r=.47, p<.001). In addition, the APVT and AEVT were correlated with each other (r=.612, p<.000). Both vocabulary tests were correlated with age in months (age and APVT, r=.491, p<.001 and age and

AEVT, r=.806, p<.001).The results of the correlation using WWC scoring method are consistent with what was expected.

Table 13: Correlations between word and nonword repetition scores of the typically developing (TD) children and the Arabic picture vocabulary test (APVT) and the Expressive vocabulary test (AEVT), using the WWC scoring method (*n*=44).

		Age in Months	APVT	AEVT	Word	Nonword
Age Months	Pearson Correlation	1	.491**	.806**	.518**	.382*
	Sig.(2-tailed)		.>001	>001	>001	.<.05
APVT	Pearson Correlation	.491**	1	.612**	.497**	.370*
	Sig.(2-tailed)	.001		>001	>001	<.05
AEVT	Pearson Correlation	.806**	.612**	1	.651**	.471**
	Sig.(2-tailed)	.000	>001		>001	>001
Word	Pearson Correlation	.518**	.497**	.651**	1	.562**
	Sig.(2-tailed)	<.001	>001	.>001		>001
Nonword	Pearson Correlation	.382*	.370*	.471**	.562**	1
	Sig.(2-tailed)	<.05	<.05	>001	>001	

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

For the PPC scoring method, Pearson product-moment correlations coefficients were also calculated to measure the correlation between word and nonword repetition and receptive and expressive vocabulary in typically developing children (TD) (n=44). Results of the various correlations are shown in Table 14. Results showed that the word repetition test significantly correlated with age (r=.453, p<.01), the APVT (r=.331, p<.05), the AEVT (r=.560, p<.001), and the nonword repetition test (r=.602, p<.001). In addition, the nonword repetition test significantly correlated with age in months (r=.322, p<.05), AEVT (r=.376, p<.05) and the Arabic Picture Vocabulary test APVT (r=.331, p<.05).

Table 14: Correlations between word and nonword repetition scores of the typically developing children and the Arabic picture vocabulary test (APVT) and the Expressive vocabulary test (AEVT) using the PPC scoring method (*n*=44)

		Age Months	APVT	AEVT	Word	Nonword
Age in Months	Pearson Correlation	1	.491**	.806**	.453**	.322*
	p value		<.001	. <.001	. <.01	<.05
APVT	Pearson Correlation	.491**	1	.612**	.331*	.331*
	Sig. (2-tailed)	.001		<.001	.028	.034
AEVT	Pearson Correlation	.806**	.612**	1	.560**	.376*
	Sig. (2-tailed)	<.001	<.001		<.001	<.05
Word	Pearson Correlation	.453**	.331*	.560**	1	.602**
	Sig. (2-tailed)	<.05	<.05	<.001		<.001
Nonword	Pearson Correlation	.322*	.331*	.376*	.602**	1
	Sig. (2-tailed)	<.05	<.05	<.05	<.001	

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

In addition, the APVT and AEVT were strongly correlated with each other (r=.612, p<.001). Both vocabulary tests were correlated strongly with age in months (age and APVT, r=.491, p<.001 and age and the AEVT, r=.806, p<.001). The results of the correlation using PPC scoring method were consistent with what was expected.

The Pearson product-moment correlation coefficient was also calculated to measure the correlation between word and nonword repetition and receptive APVT and expressive AEVT vocabulary skills in the clinical group (CL) (n=15). Results of the various correlations are shown in Table 15. Results showed that the word repetition test significantly correlated with the nonword repetition test (r=.786, p<.001). However, the word repetition test did not correlate significantly with other variables e.g., age (r=.113, p<.0688), the APVT (r=.110, p<.697), or the AEVT (r=.155, p<.582). Nonword repetition correlated significantly with word repetition, however the correlation with the other factors was not significant (age in months, APVT, and the AEVT). The APVT and AEVT were strongly correlated (r=.692, p<.000). However, neither vocabulary tests nor word and nonword repetition tests did correlate with age in months. The results of the correlation using the WWC scoring method were partially contrary to what was expected as the word/nonword repetition scores did not correlate with the receptive or expressive scores. On the other hand, the significant correlation between word and nonword repetition was consistent with what was hypothesised.

Table 15: Correlations between word and nonword repetition scores of the Clinical (CL) Group and the Arabic picture vocabulary test (APVT) and the Expressive vocabulary test (AEVT), Using WWC scoring methods(*n*=15).

		Age in Months	APVT	AEVT	Word	Nonword
Age in months	Pearson Correlation	1	.416	.201	.113	.099
	Sig.(2-tailed)		.123	.473	.688	.724
APVT	Pearson Correlation	.416	1	.692**	.110	.121
	Sig.(2-tailed)	.123		.004	.697	.668
AEVT	Pearson Correlation	.201	.692**	1	.155	.419
	Sig.(2-tailed)	.473	.<.01		.582	.120
Word	Pearson Correlation	.113	.110	.155	1	.786**
	Sig.(2-tailed)	.688	.697	.582		<001
Nonword	Pearson Correlation	.099	.121	.419	.786**	1
	Sig.(2-tailed)	.724	.668	.120	<001	

**. Correlation is significant at the 0.01 level(2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

For the PPC scoring method, the correlation between word and nonword repetition and receptive and expressive vocabulary in the clinical group (CL) (n=15), as shown in

Table 16, showed that the nonword repetition test significantly correlated only with the word repetition test (r=.733, p<.002). However, the word repetition test did not correlate significantly with age (r=.032, p<.910), the APVT (r=.131, p<.634), or the AEVT (r=.215, p<.443) Furthermore, the nonword repetition test did not correlate with the other variables (age (r=-.160, p=.569), APVT (r=-.020, p.943) and the AEVT (r=.411, p=128). The APVT and AEVT were strongly correlated (r=.692, p<.000). None of the vocabulary tests did correlate significantly with age in months, or the word and nonword repetition test. These results of the correlation using the PPC scoring method were partially contrary to what was expected as the word/nonword repetition scores did not correlate with the receptive or expressive vocabulary tests. On the other hand, the significant correlation between word and nonword repetition was consistent with what was hypothesised. One possible reason for this nonsignificant correlation was the inclusion of three children who scored within normal range in both vocabulary tests who had received speech therapy for more than three months of speech therapy while the other participant in the CL group had either no speech therapy or they were just referred to the speech therapy unit at the time of the study. This information about length of speech therapy services was available after the individual results for each participant were analysed. Reanalysis for the correlation was conducted after the three participants were excluded and the results showed significant correlation between vocabulary tests and word nonword repetition (see Appendix I). Figure 11 and Figure 12 show sscatterplots of the correlation between (APVT), (AEVT), word and nonword repetition tests for all children (TD and CL) using PPC scoring method and the WWC scoring method respectively.

Table 16: Correlations between word and nonword repetition scores of the Clinical Group and the Arabic picture vocabulary test (APVT) and the Expressive vocabulary test (AEVT), using PPC scoring method (*n*=15).

		Age in months	APVT	EVAT	Word	Nonword
Age in Months	Pearson Correlation	1	.416	.201	.032	-.160
	Sig. (2-tailed)		.123	.473	.910	.569
APVT	Pearson Correlation	.416	1	.692**	.131	-.020
	Sig. (2-tailed)	.123		.004	.643	.943
APVT	Pearson Correlation	.201	.692**	1	.215	.411
	Sig. (2-tailed)	.473	.004		.443	.128
Word	Pearson Correlation	.032	.131	.215	1	.733**
	Sig. (2-tailed)	.910	.643	.443		.002
Nonword	Pearson Correlation	-.160	-.020	.411	.733**	1
	Sig. (2-tailed)	.569	.943	.128	.002	

**. Correlation is significant at the 0.01 level (2-tailed).

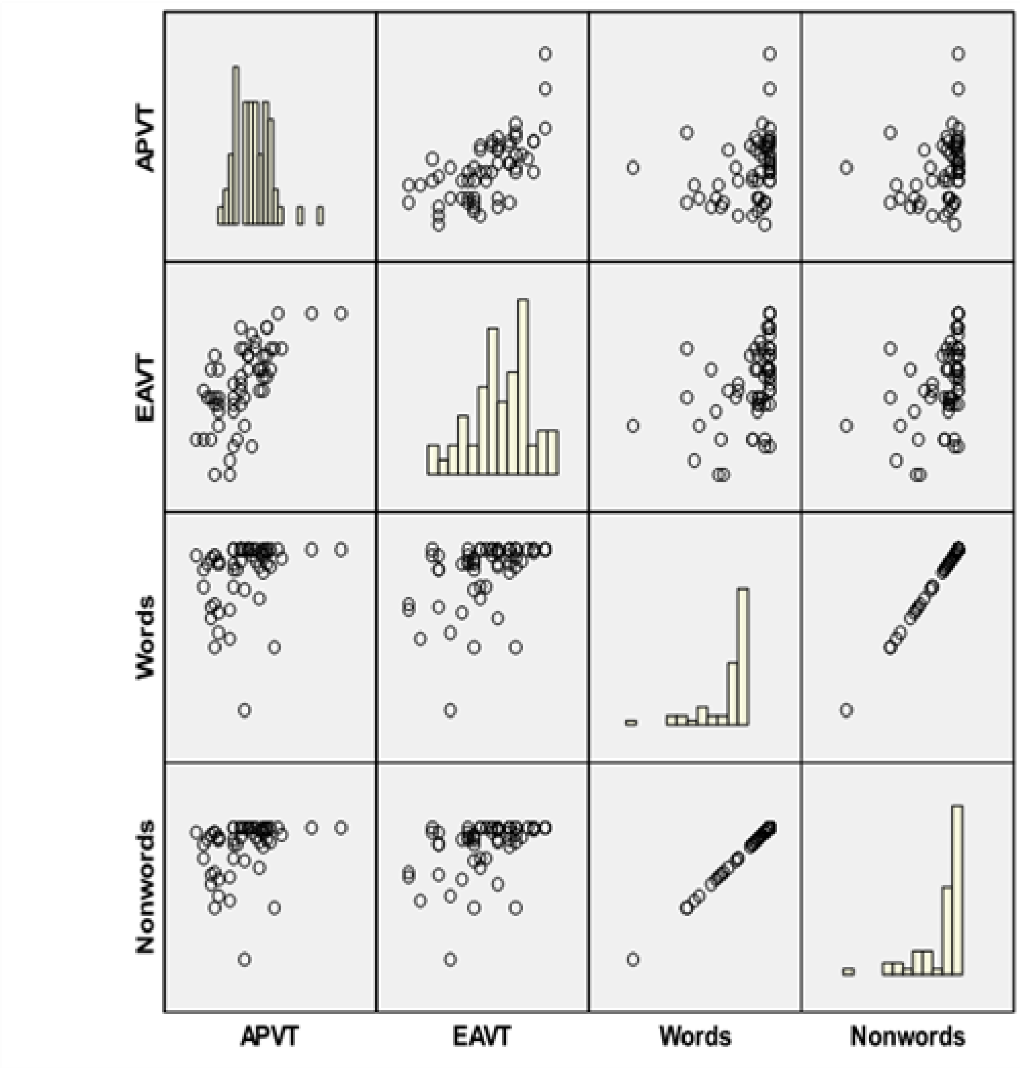

Figure 11: A scatterplot of the correlation between Arabic Picture Vocabulary Test (APVT), the Arabic Expressive Vocabulary Test (AEVT), word and nonword repetition tests for all children (TD and CL) using PPC scoring method.

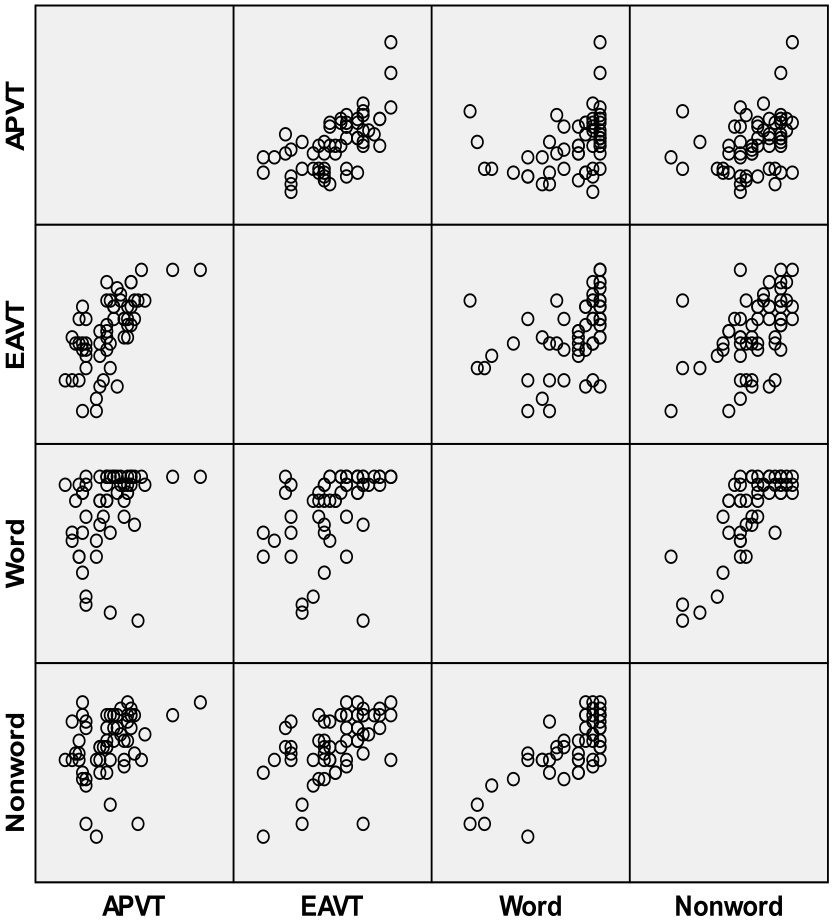

Figure 12: A scatterplot of the correlation between Arabic Picture Vocabulary Test (APVT), the Arabic Expressive Vocabulary Test (AEVT), word and nonword repetition tests for all children (TD and CL) using WWC scoring method.

Regression Analysis. A multiple regression analysis was conducted using a stepwise method as a second step after conducting the correlation analyses, to evaluate the predictive value of the vocabulary size based on the repetition skills. Two sets of regression analyses were conducted to explore if the scoring methods WWC and PPC for word and nonword repetition would make any difference in predicting the expressive and receptive vocabulary skills in both TD children and CL children.

Regression analysis based on the PPC scoring method for TD children. A multiple linear regression was calculated to predict the receptive vocabulary skills measured by the children's scores on the APVT (DV, dependant variable) based on children's performance on word repetition and nonword repetition (IV, independent variables), to see if the children's performance on word/nonword repetition skills can predict their receptive vocabulary size. A significant regression equation was found (F (1,42) =5.168, p = .028), with an R^2 of .110. The children's performance on APVT is equal to (-20.412+.553) (nonword repetition). The children's performance on nonword repetition can predict 11% of the variance of the APVT, however word repetition was excluded as its predictive value was not significant (see Figure 13). Another set of regression was conducted with the AEVT as the DV and word and nonword repetition as the IV. A significant regression equation was found (F (1,42) = 19.236, p < .001), with an R^2 of 0.314. The AEVT predicted weight is equal to (-39.301 + .733) (nonword repetition), The children performance in nonword repetition can predict 31% of the variance of the

AEVT. Word repetition was excluded as its predictive value was not significant. The excluded predictor or variable may have significant correlation as in this current analysis, however it is possible that not all of them will be statistically significant in the same multiple linear regression model (see Figure 14).

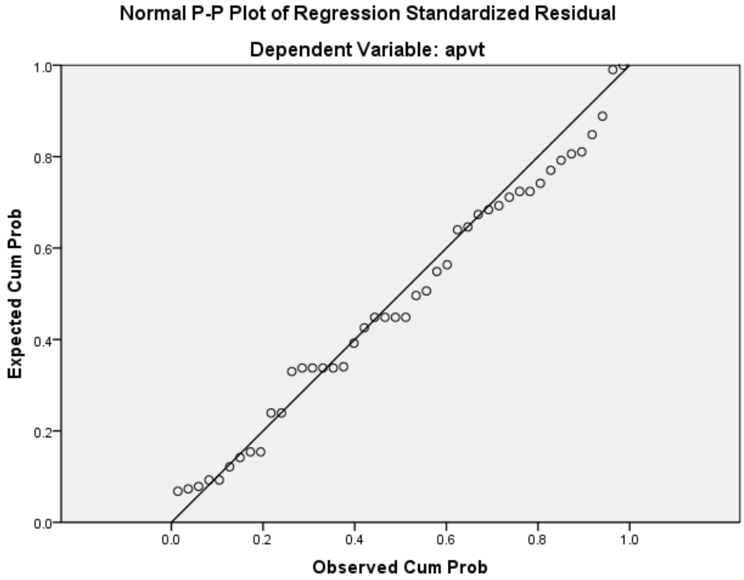

Figure 13: Regression model with word and nonword repetition as predictor of APVT scores for TD children using PPC scoring method.

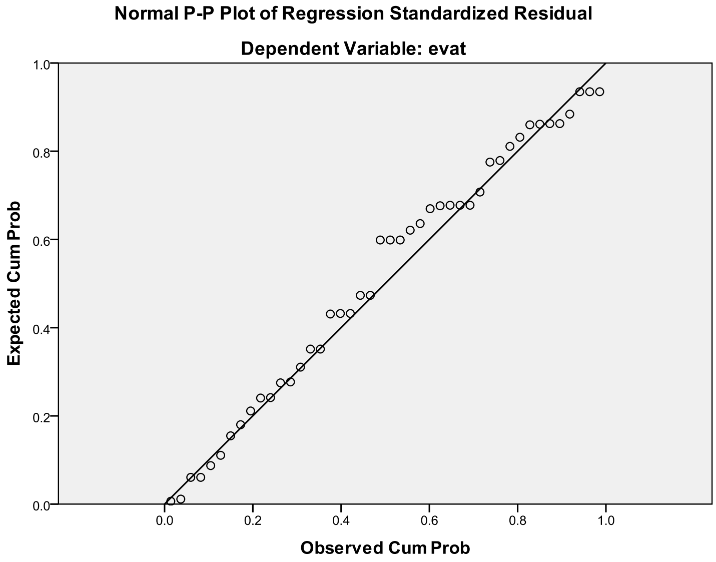

Figure 14: Regression model with word and nonword repetition as predictor of AEVT scores for TD children using PPC scoring method.

Regression analysis based on the PPC scoring method for CL children. A multiple linear regression was calculated for the clinical group this time, to predict the receptive vocabulary skills measured by the children's scores on the APVT (DV, dependant variable) based on children's performance on word repetition and nonword repetition (IV, independent variables). A non-significant regression equation was found (F (1, 13) =.226, p = .0643), with an R^2 of .017. The children's performance on APVT is equal to (14.86+ .085) (nonword repetition). The children's performance on nonword repetition can predict 1.7% of the variance of the APVT (see Figure 15). Another set of regressions was conducted with the AEVT as the DV this time and the word and nonword repetition as IV. A non-significant regression equation was found (F (1,13) =

.627, p < .443), with an R^2 of 0.046. The AEVT predicted weight is equal to (8.153+ .108) (nonword repetition), therefore children's performance on the nonword repetition can predict 4.6% of the variance of the AEVT (see Figure 16).

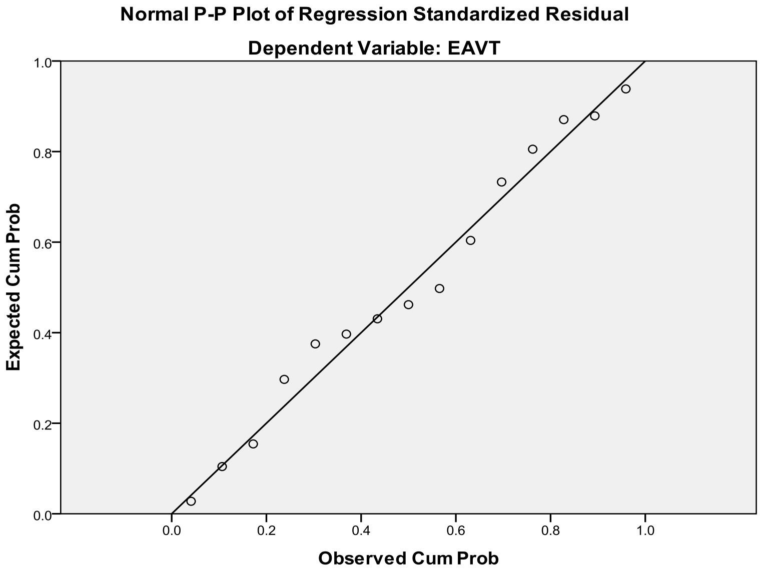

Figure 15: Regression model with word and nonword repetition as predictor of AEVT scores for CL children using PPC scoring method.

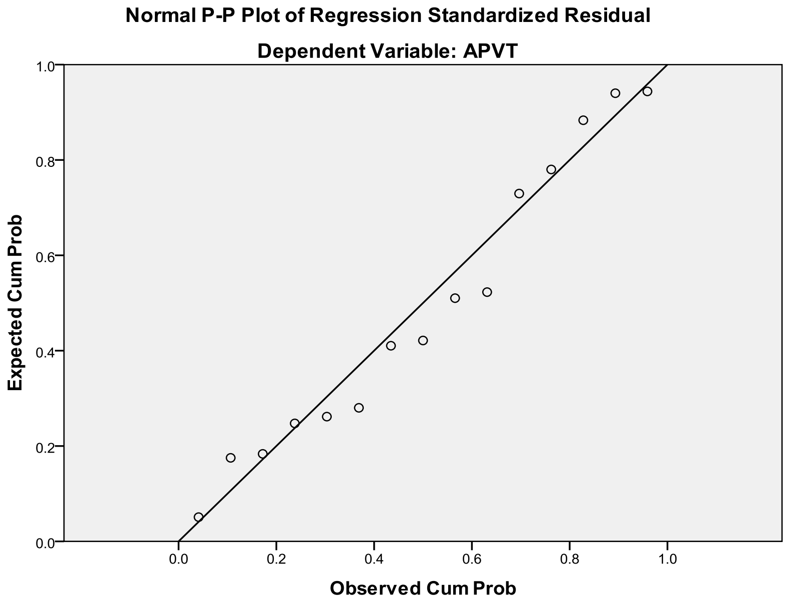

Figure 16: Regression model with word and nonword repetition as predictor of APVT scores for CL children using PPC scoring method.

Regression analysis based on the WWC scoring method for TD children. A multiple linear regression was calculated to predict the receptive vocabulary skills measured by children's scores on the APVT (as the dependant variable) based on children's performance on word repetition and nonword repetition as independent variables). A significant regression equation was found (F (2,42) =7.16, p = .002), with an R^2 of .259. Children's performance on APVT is equal to (-9.37 +.301) (word repetition) +.066 (nonword repetition). Children's performance on the word and nonword repetition together can predict 26% of the variance of the APVT (see Figure 17).

Another regression was conducted with the AEVT as the DV and word and nonword repetition as IV. A significant regression equation was found ($F(2,40) = 17.226$, $p < .000$), with an R^2 of 0.463. The AEVT predicted weight is equal to -15.675 +.331 (word repetition) + .053 (nonword repetition). Children's performance on the word and nonword repetition together can predict 46.3% of the variance of the AEVT (see Figure 18).

In both sets of regression, word repetition was stronger than nonword repetition in predicting the children's performance on both APVT and AEVT, as the beta value was higher for word repetition than nonword repetition.

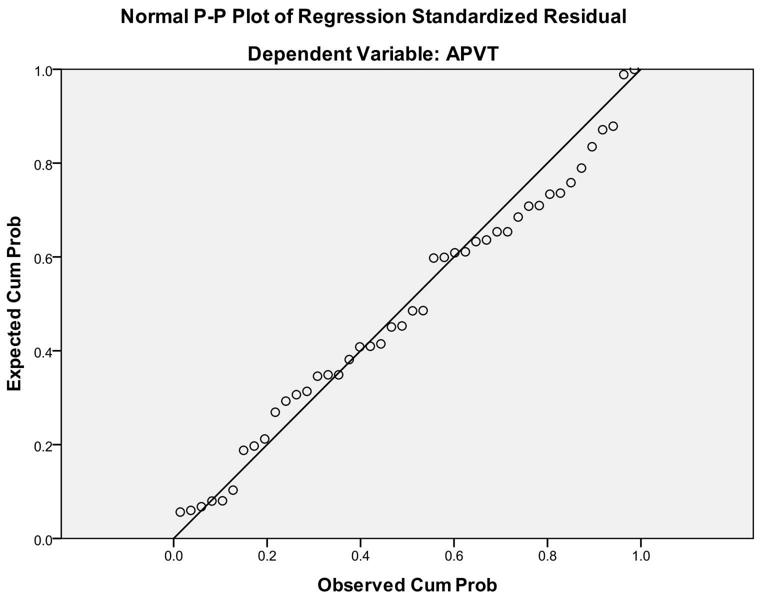

Figure 17: Regression model with word and nonword repetition as predictor of APVT scores for TD children using WWC scoring method.

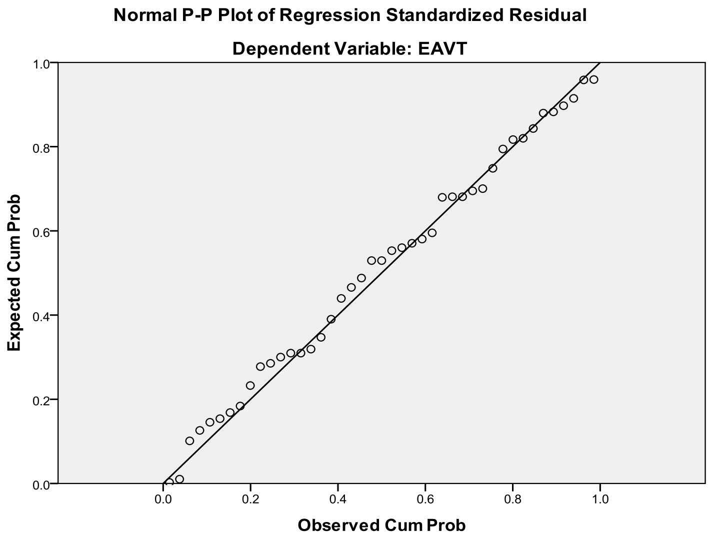

Figure 18: Regression model with word and nonword repetition as predictor of AEVT scores for TD children using WWC scoring method.

Regression analysis based on the WWC scoring method for CL children. A multiple linear regression was calculated for the clinical group this time (N=15), to predict the receptive vocabulary skills measured by the children's scores on the APVT (DV) based on children's performance in word repetition and nonword repetition (IV). A non-significant regression equation was found ($F(2,12) = .092$, $p = .912$), with an R^2 of .015) The children's performance in APVT is equal to (18.716 + .011) (word repetition) +.091 (nonword repetition). The children's performance in the both word and nonword repetition can predict 1.5% of the variance of the APVT (see Figure 19). Another set of regression was conducted with the AEVT as the DV and word and nonword repetition as IV. A non-significant regression equation was found ($F(2,12) = 2.059$, $p < .170$), with an R^2 of .0256. AEVT predicted weight is equal to (12.528 -0.97) (word repetition) +.188 (nonword repetition), children's performance in word and nonword repetition can predict 2.6 % of the variance of the AEVT (see Figure 20).

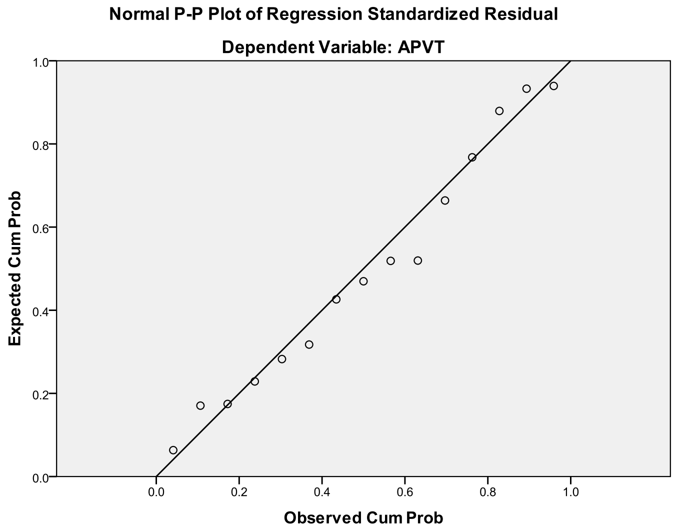

Figure 19: Regression model with word and nonword repetition as predictor to APVT for CL children using WWC scoring method.

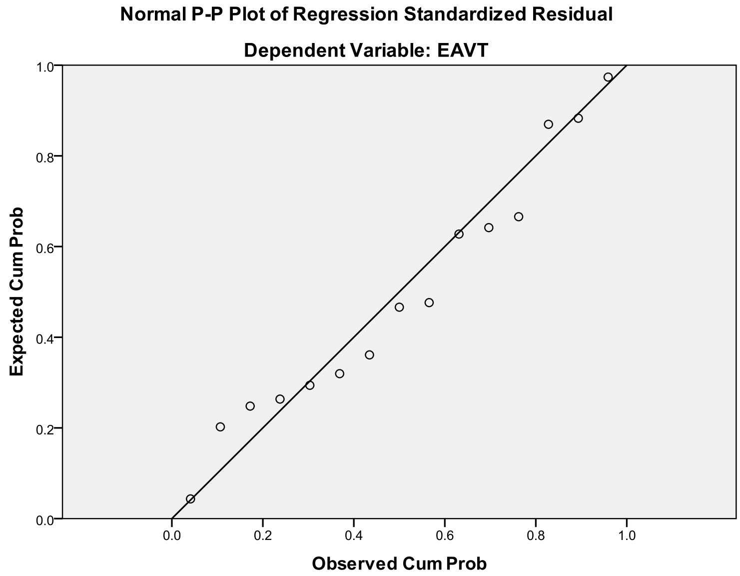

Figure 20: Regression model with word and nonword repetition as predictor to AEVT for CL children using WWC scoring method.

4.6 Discussion

This study was designed to investigate five main issues. First, if the WNRep test can be used as a useful diagnostic tool to distinguish typically developing children from those with language impairment. Second, it investigated the effects of word length (one, two and three syllables) and word type (word vs nonword) on children's repetition accuracy. Third, it aimed at providing some essential information regarding the nature of the relationship between real word repetition and nonword repetition and receptive and expressive vocabulary in Arabic speaking children with typical and atypical language development. Fourth, it evaluated the predictive value of real word repetition and or nonword repetition to receptive and expressive vocabulary. Fifth, it examined if using two different scoring methods (percentage of phonemes correct (PPC) and percentage of whole words correct (WWC)) makes a difference in evaluating the above goals. There were a few participants who did not complete the tasks. The percentage of participants who refused to complete the tasks was 14.5% from the total number of the children in both groups (typical and clinical). This percentage is high compared with Chiat and Roy (2007), which was 6%. This was the case because the majority of children who refused to complete the task had not been to any nurseries or school so they might have found the investigator and or the testing procedures unfamiliar. However, all the children who refused to continue the expressive task were able to do the receptive vocabulary test which did not require any verbal responses as they were able answer the receptive vocabulary stimuli by pointing to the pictures. These children's performance was within normal range when compared with other children who completed all receptive and expressive tasks.

4.6.1 Word and nonword repetition as a diagnostic tool

The results of this study showed that the clinical group performed significantly worse than their typically developing peers on word and nonword repetition tests, regardless of which repetition scoring methods were used. Therefore, these results extend the validity of word and nonword repetition as a potential clinical marker of language impairment in Gulf Arabic. Furthermore, the WNRep test can be clinically useful in other Arabic dialect besides Gulf Arabic as the real word items used in the WNRep items can be found in many other Arabic dialects. The stimuli used in this task consisted of early acquired sounds and familiar syllable structures that can be found in many other Arabic dialects, which may support the viability of this task in identifying children with language impairment in Gulf Arabic and other Arabic dialects. However, these are preliminary findings that should be supported by larger studies the include larger number of participants to confirm the diagnostic validity of such test.

The findings in this study were consistent with the prediction of this experiment. It was expected that TD children would perform better than the CL children on word and nonword repetition tests, which is also consistent with many studies that found nonword repetition as a clinical marker for children with language impairment (Botting & Conti-Ramsden, 2001; Conti-Ramsden & Hesketh, 2003; Dollaghan & Campbell, 1998; Ellis Weismer et al., 2000; Gathercole & Baddeley, 1990b; Gray, 2003). All these studies found that children with language impairment repeated significantly fewer nonwords correctly when compared with TD children of similar age, matched on nonverbal intelligence, as well as a younger language-matched group.

Real word repetition succeeded also in distinguishing the typically developing children from the clinical children. These findings strengthen the utility of real word repetition as a possible useful clinical tool for children with language impairment and support Dispaldro et al.'s (2013) findings who demonstrated that real word repetition can be a diagnostic assessment tool that distinguishes TD Italian children from those with SLI. Although real word repetition required different underlying processes compared with nonword repetition, both tasks succeeded in identifying the CL group from the TD group. The main skills required for word repetition are different from those required for nonword repetition, as real word repetition relies on existing phonological and semantic knowledge accumulated from learning previous vocabulary, whilst nonword repetition relies heavily on PSTM, although both tasks share some common skills, such as speech perception, oral-motor skills , PSTM, lexical and linguistic knowledge. These shared skills probably could allow real word repetition to function as a possible diagnostic tool to distinguish CL group from the TD ones in Gulf Arabic speaking children, as the CL children are also likely to have smaller vocabularies, so this may in turn affect real word repetition.

The usefulness of using real word repetition as a diagnostic tool could contribute to clinical practice, as using real word repetition could be more familiar than a nonword task, especially for young children. It will also be easier for the examiner to use it with children. However, before recommending using real word repetition as a diagnostic tool, it is important to investigate the utility of real word repetition along with nonword repetition with a larger sample from both TD and CL Gulf Arabic speaking children.

4.6.2 Implications regarding effects of word types

The Main ANOVA in both scoring methods showed that there was a significant effect of word type, and the results of the paired t-test also showed that the TD and the CL groups were significantly more accurate in word repetition than nonword repetition in both scoring methods. This shows that children in the current study were sensitive to lexical familiarity. These findings replicate the findings from a variety of studies (e.g., Gathercole, 1995; Gathercole & Adams, 1994; Gathercole et al., 1999) who suggest that the more wordlike the item, the more likely that it will be repeated correctly. These findings are consistent with Roy and Chiat (2004) who found that children aged two to fours years old performed better on words than nonwords when they conducted the Preschool Repetition test (PSRep). Furthermore, Gathercole, Willis, Emslie, and Baddeley (1991) found that repeating nonwords that sounded like real words was more accurate than nonwords that sounded less “word-like” on the CNRep test.

Consequently, this study’s finding is consistent with a phonological processing account (Snowling, Chiat & Hulme, 1991; Chiat, 2001) of nonword repetition. According to this account previous lexical knowledge supports better nonword repetition.

On the other hand, the interaction between word type and group type was not significant with either scoring methods. There was no difference between groups’ performances (TD and CL) on different word types (words and nonwords). Both groups were significantly more accurate on word repetition than nonword repetition. These findings also support the phonological processing account (Snowling et al. 1991, Chiat, 2001). Though the CL children have more limited vocabulary skills, they benefited from their lexical knowledge as they repeated words more accurately than nonwords.

4.6.3 Implications regarding effects of word length

The results of word and nonword repetition in this current study showed that there was a significant effect of word length (one, two and three syllable words/nonwords) in TD and CL groups using both PPC and WWC scoring methods. The longer the word or the nonword was, the less accurate the repetition was for both groups.

These findings support many studies from the PSTM account of nonword repetition that claimed that longer nonwords resulted typically in more repetition errors than shorter nonwords (Gathercole & Baddeley, 1989; 1990a; Gathercole et al., 1994). According to the PSMT account, the phonological loop is responsible for storing temporary phonological information which influences nonword repetition. The longer the word is the more likely it will be forgotten because repeating longer words means there are more demands on working memory, and the results from this experiment showed length effects in both words and nonwords.

Furthermore, the interaction between word length and word type was significant. Children’s scores increased when the number of the syllable in word and nonword repetition decreased, both groups scored better on word repetition than nonword repetition across all word lengths (one, two and three syllables). However, the magnitude of difference was largest for three syllable words across both groups. Children were affected more by length in nonwords vs words repetition. This finding is consistent with Chiat and Roy (2007) and Roy and Chiat's (2004) findings that word length affected the accuracy of real word repetition in TD children and children with language impairment differently; TD children were less affected by word length in real word repetition than nonword repetition. However, Chiat & Roy (2007), found that the clinical group were less affected by word familiarity as the interaction between the length and word status was not significant for children with language delay, while in this current study both groups benefitted from lexical knowledge. These findings, as mentioned earlier, support the PSTM account for NWR, as there was a significant effect of word length in this experiment. However, the finding that both groups benefited from lexical familiarity, as they scored better on word repetition than they did on nonword repetition across all word lengths, supports the phonological processing account (Snowling et al. 1991; Chiat, 2001) of nonword repetition. According to this account previous lexical knowledge supports better nonword repetition. The underlying processes required for repetition may not be explained by one account only as there are many processes which may contribute word and nonword repetition, such as speech perception, PSTM, oral-motor skills, lexical and linguistic knowledge.

4.6.4 Relationship between word and nonword repetition and vocabulary tests

The fourth aim in this study was to explore the nature of the relationship between real word repetition and nonword repetition and receptive and expressive vocabulary in Arabic speaking children with typical and atypical language development

The correlation results of the TD group and CL group are very similar across the different scoring methods. Word and NWR were significantly correlated with each other, and with age, and receptive and expressive vocabulary across the two scoring methods.

The significant correlation between nonword repetition and the receptive vocabulary used in this study (APVT, Shaalan, 2010) using PPC and WWC scoring methods is consistent with many studies in the literature (Gathercole et al., 1991; 1992; Briscoe et al., 2001; Coady & Evans, 2008). On the other hand, Shaalan (2010) found only weak correlation between the same receptive test used in this study (APVT) and the NWR test he developed. There could be multiple reasons for this finding. First, the participants in Shaalan (2010)'s study had an average age of 7;8 for the SLI group and the chronological matched group and 5;8 for the language matched group with 11 participants in each group. The average age for current participants was 3;2 years for the TD children and 3;7 for the CL group. Gathercole (2006) explained that the older the children grow, the weaker the relationship becomes between NWR and vocabulary skills. Therefore, the wider gap in ages between participants in this study and Shaalan (2010) could explain the differences in findings. Moreover, three of the participants in the CL scored within normal range on both receptive and expressive vocabulary tests; these children received at least 3 months of speech language therapy that could have influenced the results of their vocabulary scores.

The correlation between nonword repetition and expressive vocabulary (AEVT) was significant in both scoring methods. These findings were not consistent with a few studies that previously investigated this relationship. Those studies found a weak or non-significant relation between the NWR and expressive vocabulary (e.g. Briscoe et al., 2001; Stokes et al., 2013; Conti-Ramsden, 2001). Our findings can be explained as the word/nonword repetition and expressive vocabulary share the same skills of speech perception, oral-motor planning and articulation and therefore it is not surprising to find this correlation. Both tasks have joint requirements of articulating and pronouncing a series of sounds, whether words or nonwords. Therefore, children who perform well on repetition tasks might perform well on expressive vocabulary. Moreover, there are other studies that found significant correlations between the two skills, such as Kovas et al., 2005 and Krishnan et al. (2013), which found a strong evidence for the correlation between nonword repetition, oral-motor skills and articulation. Another possible explanation could be found in the typological properties of Arabic as a root and pattern language. It is possible that both tasks (NWR and expressive vocabulary) are mediated by awareness of root and patterns in Arabic. This experiment, however, was not designed to examine the contributions of roots and patterns; and this will be investigated in the next experiment.

These strong correlations could support the utility of real word repetition in predicting vocabulary skills in children, however no causal relation is suggested in this relationship. There are no other studies that investigated the relationship between real word repetition and other vocabulary tests, but there are some studies that investigated this relationship with different aspects of language. Chiat and Roy (2007) found a strong correlation between PSRep (consisting of real word and nonwords) and receptive and expressive language skills tested using Preschool Language Scale -3 (PLS-3) (which consists of auditory comprehension and expressive language tasks). Furthermore, Dispaldro et al. (2011) found that real word repetition was a good predictor of grammatical skills in Italian but not in English. They found that while NWR correlated with phonological storage, real word repetition correlated better with grammatical abilities due to the presence of lexical and semantic representations along with phonological representations in real words (see 4.1.1).

Correlations for the CL group. The correlation results were consistent for the CL group across the two different scoring methods. There was a significant correlation between word repetition and nonword repetition and between receptive and expressive vocabulary. However, the correlation between word and nonword repetition tests and the receptive and expressive tests was not significant. These findings are contrary to those of Bowey (1996; 2001), Gathercole and Baddeley (1989), and Metsala (1999). This can be due to the small number of the CL sample (n=15). When both groups were added to the same correlation analysis the correlation between word and nonword repetition and the receptive and expressive vocabulary tests was strongly significant across the two different scoring methods (see Appendix I: Correlation between word and nonword repetition scores and receptive and expressive vocabulary tests for TD and CL groups).

On the receptive (APVT) and expressive (AEVT) tests, the CL group had lower scores on these tests than the TD group, however the difference between the two groups on the vocabulary test was not statistically significant. This could be due to the small CL group sample (N=15). Furthermore, it could be because the age average of the CL group is higher than the TD (3;7 years for CL vs 3;2 for TD group) and it also could be due to three children from the CL group who performed well on all tests although they met the criteria of the CL group selection. The good performance of these three children who were all males could be because they received speech therapy sessions for longer than 3 months. When children’s scores on both vocabulary tests (the APVT and AEVT) were recalculated excluding the three outlier children who received speech therapy, the results showed a significant difference between the CL group and the TD group on both vocabulary tests. It is important to know that after excluding the three male participants the ratio between male to female participants became 7:5.

4.6.5 The predictive value of word and nonword repetition

A multiple regression analysis was conducted to evaluate the predictive value of receptive (APVT) and expressive (AEVT) vocabulary skills based on word and nonword repetition skills for the TD and CL groups across the two different scoring methods the PPC and WWC.

For the PPC scoring method, TD children’s performance on nonword repetition can predict 11% of the variance of the receptive vocabulary test (APVT) and 31% of the expressive vocabulary test (AEVT). Word repetition was excluded as its predictive value was not significant. The excluded predictor or variable may have significant correlation as in this current analysis, however it is possible that not all of them will be statistically significant in the same multiple linear regression model. The predictive value of nonword repetition performance for the CL group was not significant for APVT (1.7%) nor was it for the AEVT (4.6%).

The Regression results for the TD children using the WWC scoring method showed higher predictive value. Word and nonword repetition together can predict 25% of the variance of the receptive vocabulary test (APVT) and 46.3 % of the expressive vocabulary test (AEVT). The predictive value for the nonword repetition performance for the CL group was not significant for the APVT (1.5%) or the AEVT (2.6%). The regression finding in this study for the TD group and in the two scoring methods is consistent with other studies that found nonword repetition as a predictor of vocabulary size (Bowey 1996; 2001; Gathercole & Baddeley 1989; Gathercole et al. 1991; Metsala, 1999). Unlike the TD group, the CL group's performance on word and nonword repetition could not predict the receptive and the expressive vocabulary skills. This might be due to the small sample size (n=15) or that children with language impairment could have a different language profile. Chiat and Roy (2007) investigated the predictive power of the PSRep test for language at 4-5 years and 9-11 years and found a small predictive value for the PSRrep, especially for longer terms. Furthermore, Chiat and Roy (2007) attributed this to the presence of a proportion of children who failed the PSRep due to severe speech difficulties that affected their expressive language.

4.6.6 Effects of scoring methods

The aim of using two different scoring methods was to investigate if this would make any difference in analyzing the results of word and NWR. By comparing all the results in this study using the two scoring methods, it was found that these two different scoring methods did not make any significant contribution to the groups' results on the word and nonword repetition. Some results were significant in one scoring method but not in others, but the direction of performance across all results was the same for both groups. The main effects and interactions for different variables were very similar in both scoring methods. In correlation and regression analyses the WWC scoring method showed stronger significant correlations and higher predictor value between repetition tasks and vocabulary tests than the PPC scoring method. However, the WWC scoring method failed to detect the difference in performance between 3 syllable nonwords vs. 3 syllable words repetition. These findings were contrary to our expectation; it was expected that the PPC scoring method would be more sensitive to articulatory competence. The more the misarticulations were the less the percentage of the PCC would be, while the WWC would be less sensitive to the number of errors made. The current investigation showed that the difference between scoring methods was small. These findings are consistent with Graf Estes et al. (2007) and Deevy et al's. (2010) findings, which found no difference between using WWC and PPC scoring methods in English. On the other hand, our findings were not consistent with Dispaldro et al's. (2013) findings in Italian, that found significant effects for using different scoring methods for repetition tasks, where the WWC accounted for more significant differences between the TD group and the clinical group in comparison to the PPC. One possible explanation for lack of significant difference between the two scoring methods has to do with the criteria used to select the stimuli. Most of the sounds used were early developing sounds that could be produced correctly by most children at this age, including those with language impairment. In addition, consistent articulation errors were allowed for both groups and this factor could have reduced the difference between the two scoring methods. Moreover, all words and nonwords have no clusters and children with language impairment in Gulf Arabic were found to have significant difficulties repeating nonwords that contain different types of clusters (Shaalan, 2010).

Investigating the effects of using different scoring methods in Arabic might help make an appropriate decision in preference for one scoring method over the other. However, it would be more useful to apply these different scoring methods with different studies targeting various populations of TD and atypically developing children. Although the WWC might be easier to administer, less time consuming and potentially less prone to inter rater disagreements, this study recommends using the PPC method. Results as reported earlier showed that the PPC, but not the WWC, differentiated between the performance of the children on three syllable words vs three syllable nonwords. However, more comparisons are needed to investigate differences while considering levels of phonological complexity of the stimuli (e.g., early vs. late developing, clusters vs. non-clusters…etc.).

4.7 Summary

This study examined the viability of a word and nonword repetition test (WNRep) as a diagnostic tool to distinguish Gulf Arabic speaking children with language impairment (CL) from typically developing (TD) peers. Results of the WNRep showed that this test could be a potentially useful diagnostic tool, as the clinical group (CL) performed significantly worse than the TD group on word and nonword repetition and across different word and nonword lengths. Furthermore, both groups' performances on WNRep significantly correlated with the Arabic receptive and expressive vocabulary tests.

Analysis of the data shows that the performance of both groups on nonword repetition was partially consistent with the phonological short-term memory account (PSTM) (Gathercole, 2006; Gathercole & Baddeley, 1990). The CL group performed significantly worse than the TD group on one, two, and three syllable words and nonwords. The PSTM account explains the part of the findings related to word length, while the phonological processing skills account (Chiat, 2001; Snowling et al., 2001) presents a better explanation for real word repetition findings that are related to word familiarity and lexical knowledge. Furthermore, real word repetition results were consistent with Dispaldro et al's. (2013) study, which found word repetition a useful diagnostic tool for children with language impairment in Italian.

The correlations between word and nonword repetition and the Arabic receptive and expressive vocabulary tests were significant for the TD group in both scoring methods. This is unlike the CL group, where the correlations between the repetition tasks and the vocabulary tests were not significant. Consequently, regression findings in this study for the TD group, using the two scoring methods, are consistent with other studies that found nonword repetition as a predictor of vocabulary size (e.g., Bowey, 1996; 2001; Gathercole and Baddeley, 1989; Gathercole et al., 1991; Metsala, 1999).

In addition, comparing the results of the two scoring methods used in the present study (PPC: percentage of phonemes correct, and WWC: percentage whole word correct) showed there was no significant difference between both methods. These findings are consistent with Graf-Estes et al. (2007) and Deevy et al's. (2010) findings in English. On the other hand, our findings were not consistent with Dispaldro et al's. (2013) study, which found that scoring methods influenced results, where the magnitude of the group differences was greater under the WWC scoring method than it was under PPC.

The results in this chapter provide some initial information about the performance of TD Gulf Arabic speaking children and those with language impairment on word and nonword repetition and how their performance on these tasks correlated with receptive and expressive vocabulary tests. These results show significant effects of wordlikeness as children had better NWR scores on words vs. nonwords. However, wordlikeness effects in Arabic are influenced by roots and patterns and therefore it is important to examine carefully the effects of *both* roots and patterns in NWR and how each one of them influence NWR. In the next chapter, we will explore the impact of the typology of the phonological and morphological system as represented by root and pattern effects in Gulf Arabic on TD children's repetition skills. To do that, we developed a new nonword repetition test to explore the effects of the Arabic root and pattern morphology on children's repetition skills.

5. Root and Pattern Nonword Repetition Test

5.1 Introduction

Arabic is a Semitic language that uses root and pattern morphology (see chapter 3). A popular example of a root is √k.t.b, which means "writing". By changing the pattern and/or adding some affixes, speakers can create up to 14 different words from that root. For example, /kita:b/ means "book", /maktaba/ means "library" and /maktab/ means "desk". Therefore, many new words that an Arabic speaking child is exposed to might already partially exist in his/her phonological memory because of their links to an existing representation of the root. So, these new words are not in fact entirely new. However, there are few studies that can explain the processes that underlie the development of vocabulary in Arabic speaking children. In the previous study (chapter 4) wordlikeness effect (word vs. nonwords) was examined and found to be significant. However, wordlikeness effects need to be examined in relation to root and pattern effects in Arabic. While both wordlikeness effects and root and pattern effects refer to similarity of a nonword to an existing word (wordlikeness) or an existing root (root effects), roots have a special role in the generative theory of morphology (McCarthy, 1997) and they are known to be constrained by some phonotactic principles that are root-specific (e.g., OCP-Place). Both Frisch & Zawaydeh (2001) and Gwillimas and Marantz (2015) found an emerging evidence to separation between root effects and wordlikeness effects as both found that adult participants were more sensitive to violation of root-specific rules, such as Obligatory Contour Principle- OCP-Place (which does not allow co-occurrence of two initial consonants from the same place of articulation) than they were to wordlikeness or neighborhood density effects. Both pointed to superiority of the root as the principal unit responsible for spoken word recognition. However, the role of root and pattern in NWR in children has not been studied before and therefore this study aims to provide an initial exploration of how children use roots and pattern morphology when presented with different combinations of these units. This initial investigation is necessary before we can compare root effects with wordlikeness effects.

In this chapter, we try to investigate if root and patterns effects have a special role to play in nonword repetition. Based on current literature of NWR in Arabic, we do not know if children do in fact use their knowledge about roots to form new words and at what age they start to use this principle to help them form new words. We also examine roots contributions in comparison with pattern effects in the performance of young TD Gulf Arabic speaking children.

Many studies in the last two decades showed a strong correlation between nonword repetition and vocabulary size (e.g., Bowey, 1996; 2001; Gathercole and Baddeley 1990a, 1990b, Metsala, 1999). A child's ability to repeat nonsense words correctly is linked to his/her ability to acquire new words easily, as the new words are essentially nonsense words when the child first hears them. Then, with repeating new words in different contexts, he/she will develop semantic and phonological representations for these words. While this is the situation for learners of most languages, the Arabic speaking child might process many new words by referring to a basic root in his/her mind.

This chapter tries to reveal more information about phonological and morphological processes in Arabic speaking children. A nonword repetition test was developed based on the root and pattern principle of Arabic to see if there are any effects of root and pattern on nonword repetition itself and on vocabulary development in Arabic speaking children.

This current study tries also to explore different error types associated with nonword repetition. Error pattern analysis has frequently been used to determine the types of errors made by children in NWR tests. It also allows for the determination of simplifications and substitutions that children may make that are influenced by the length and level of phonological complexity of the target. Error analysis may reflect developmental phonological errors that decrease with age or are influenced by phonological impairment; and it may also reflect the influence of lexical phonology the child has acquired.

Edwards and Lahey (1998) assessed nonword repetition in school-aged children with SLI and TD children and found that phoneme substitution was more frequent than phoneme omission for both groups, while phoneme addition was not common. Marton and Schwartz (2003) reported that children with language impairment made more errors compared to TD children. However, the error patterns were similar; substitution was more frequent than other error types, and the percentage of errors increased with nonword length. Furthermore, Marton et. al (2003) found the LI group had a significantly higher number of multiple errors within the same nonword in comparison to the TD group.

Marshall, Harris and Van der Lely (2003) found that typically developing children and children with LI exhibited lexicalisation through changing the phoneme(s) of a nonword to create a real word, for example nonword /klɛt/ realized as "collect", indicating a tendency to convert the nonword to a common word or phonological sequence that was stored in their memory. However, they did not report the frequency with which lexicalisation took place.

In this current study, error analysis was conducted to assess the results of all participants to explore errors that 'morphized' roots or patterns in the different nonword types, and phonological errors that were related to syllabic and phonemic levels (consonants and vowels). Furthermore, we wanted to examine which sounds were more problematic and in which position; therefore, consonant by consonant and vowel by vowel analyses were conducted.

5.2 Aims

This study examines the effects of roots and patterns on children's performance in a NWR task in Gulf Arabic that involves different combinations of roots and patterns. It will also explore how these effects of root and pattern interact with effects of phonological storage by using two and three syllable nonwords. This study will shed some light on how effects of roots and pattern knowledge unfold across different age groups in Gulf Arabic speaking children. Furthermore , this study will investigate the relationship between children's performance on the root and pattern NWR and receptive and expressive vocabulary tests in Gulf Arabic and compare findings to other languages. Moreover, error analysis will be explored in this experiment in the light of root and pattern related errors. Finally, Findings of this experiment will be discussed in the light of NWR theories and root and pattern studies.

5.3 Variables Considered in The Design of the Root and Pattern Nonword Repetition Test (RAP-NWR)

The main objective of developing the root and pattern nonword repetition test (RAP-NWR) test was to examine if there are any effects of different roots and/or patterns on children's performance on NWR in Arabic.

5.3.1 RAP-NWR test

There are three possible combinations of roots and patterns in Arabic nonwords in the RAP-NWR:

Type 1: Root-nonpattern (R-NP) nonwords. These items have real Arabic roots and non-existing (or rare) patterns. For example, in the nonword /katub/ the root is √k.t.b "write" , while the pattern is (a-u) , which is not common in spoken dialects.

Type 2: Nonroot-pattern (NR-P) nonwords. These items have a non-existing root and a real pattern. For example, in the nonword /kafas/, the pattern (a-a) is very common, while √k.f.s is not a real root.

Type 3: Nonroot-nonpattern (NR-NP) nonwords. These items have non-existing roots and uncommon patterns. An example is the nonword /dafuk/, where √d.f.k is a non-existing root and the pattern (a-u) is not very common in spoken dialects.

All the roots that were used in the stimuli were checked in the Arabic dictionary *Al-Mu'jam Al-Waseet* (Mustafa et al., 2004). The words are listed alphabetically in this dictionary and follow the root system. The non-existence of the nonwords was also checked in *Al Waseet* to ensure that no real root was used as a nonroot item.

The 'real' patterns or vocalic levels used in the RAP-NWR test were common and they respected the phonotactic rules of Gulf Arabic. Examples of common or familiar patterns in Gulf Arabic were (a-a) for two syllabic nonwords as in /latas/ and for three syllabic words (a-a-a) as in /lafabad/. Examples of uncommon or unfamiliar patterns were (a-u) and (u-i) for two syllabic nonwords, as in /fasud/ and /sumif/; and for three syllabic nonwords (a-u-a) and (u-i-a), as in /danufas/ and /fusibal/. The uncommon patterns can be found in some forms of Classical Arabic which are not used in the Gulf Arabic dialect. Unlike the common patterns, the uncommon ones consisted of a series of unrepeated and different vowels, which is expected to increase the difficulty of repeating uncommon pattern compared to common patterns. The selection of the common pattern in the design of this test was limited by a lack of resources that classified the frequency of the patterns in young children acquiring Arabic. Therefore, the researcher relied on children's speech samples that were used in the first study (See section 4.2.2) to determine the familiarity and the frequency of different patterns used by children. Therefore, to control for common versus uncommon patterns, moderately frequent patterns were avoided, (e.g., (u-u), (u-u-u), (a-i) and only very common or non-common patterns were used in this study.

Further criteria were applied to control for morphological information and familiarity.

5.3.2 Considerations regarding phonological and information

To develop the RAP-NWR test only the consonants /k/, /l/, /m/, /n/, /s/, /b/, /t/, /f/, /d/ were included as these are mostly early acquired sounds in Arabic and indeed in many languages and most of them should exist in the consonant repertoire of the participants. The sound /s/ could be an exception, as it is not an early developing sound, but it was included because it is a very common sound and distortions or substitutions of /s/ were tolerated in scoring.

In order to control for morphological information, /b/ was excluded at the beginning of the nonwords as it is a preposition in Arabic that means “with” (e.g. /bi:di/ “with my hand”); /n/ was also avoided at the beginning as it is a first person plural pronoun (e.g. /na:kil/ “we eat”); and /m/ was avoided as it indicates negation when appearing initially (e.g., /ma:ni/ “I’m not”). Initial /t/ was avoided as it is used as a feminine gender marker in verbs and pronouns (e.g., /tadrus/ “she studies”) and /l/ was avoided initially as it means ‘to’ and because it is also used as a short form of the definite article /il/ “the”. Some sounds were also excluded at the end of the nonwords to avoid any morphological information. For example, final /t/, /m/, and /k/ are also pronouns in Arabic and they were excluded in most of the nonwords, except in the following items: /fulit/, /fulitak/ and /dafuk/.

5.3.3 Word length considerations

The RAP-NWR test items consisted of either two or three syllables. It was difficult to create one syllable nonwords as it was hard to respect Arabic phonotactic and morphological rules with root based items. Many studies showed that the effect of the word length starts from three syllables and above with words longer than two syllables becoming more difficult to repeat (Dollaghan & Campbell, 1998; Gathercole &

Baddeley, 1990a; Montgomery, 2004). Therefore, a decision was made to compare children's performance on nonwords that are two and three syllables in length only.

5.3.4 Considerations regarding familiarity

Familiarity rating is a measure of a person's frequency of exposure to a word. Gernsbacher (1984) concluded that familiarity rating appears to be a stronger predictor for word recognition when using pronounceable versus unpronounceable stimuli. Furthermore, familiarity rating was found to be a better predictor for low frequency words when using word versus nonword stimuli, which supports the argument that semantic components affect familiarity rating for low frequency words (e.g., Balota, Pilotti & Cortese, 2001).

Unlike the study in the previous chapter, a familiarity test was applied to provide more information on how adult participants use their lexical, phonological and morphological knowledge to rate and categorize a list of nonwords according to their familiarity.

A familiarity rating questionnaire was developed to be used with Gulf Arabic speaking adults in order to obtain their judgment on a set of words and nonwords. These ratings were used in the selection of the RAP-NWR items. Furthermore, the ratings allowed us to compare if the adults process the different types of the root and pattern differently. We expected adults to find items with real roots and real patterns more familiar compared with nonroot or nonpattern items as the root contains semantic information, while the vocal pattern carries grammatical information (morphology and syntax).

A questionnaire was designed that consisted of a list of 119 words and nonwords of two and three syllables length, with different combinations of real, and non-existing, or unfamiliar, roots and patterns (see Appendix F). Participants were asked to rate each word or nonword in terms of familiarity. The scale was from 1 to 7, where 7 is most likely to be an Arabic word and 1 is not an Arabic word. All items were developed based on the RAP-NWR subtypes mentioned above. An additional list of real words was added to this experiment as fillers in order to distract participants from listening to long lists of nonsense words.

The questionnaire was administered to six female adults. The age range for the participants was from 25-40 years old and all were native speakers of Gulf Arabic. The participants belonged to different occupations; there were three social workers, two teachers, and a secretary. A brief explanation about the general aim of this study was provided to the participants. None of the participants had participated in similar experiments before. The participants and the examiner sat in a quiet room. The examiner then read the stimuli and participants listened and wrote their ratings by circling the appropriate number (from 1 to 7) on the answer sheets in front of them. The answer sheets did not include any written stimuli to avoid any orthographic cues.

The instructions were adapted from Gilhooly & Logie (1980). In addition to the 119 task items, three trial items were conducted to make sure that the participants understood the instructions.

Table 17 below shows some descriptive statistics of the participants' performance for the different subtypes.

Table 17: Descriptive summary of the participants' performance on the familiarity rating test for the different words and nonword types.

	Real Words	Type1 (R-NP)	Type2 (NR-P)	Type3 (NR-NP)
Total no. of items	34	28	33	24
Mean	6.5	3.4	2.9	2.6
SD	.65	1.4	1	.95
Range	4.8-7.0	2.0-6.6.7	1.7 -6.8	1.5-5.7

Note. Type 1= root- nonpattern, Type 2= nonroot-pattern, Type 3= nonroot- nonpattern.

A Friedman's test was conducted to compare the difference in rating among the three different types of nonwords (Type 1= root- nonpattern, Type 2= nonroot-pattern, Type 3= nonroot- nonpattern). Results showed that there was a statistically significant difference in familiarity score between the three different nonword types ($\chi^2(2)$ = 79.61, $p< 0.01$), with a mean rank familiarity score of 3.4 for Type 1, 2.9 for Type 2 and 2.6 for Type 3. Follow-up pairwise comparisons were conducted using a Wilcoxon test and it showed there was a significant difference in familiarity scores between Type 1 and 2 nonwords (Z =-5.78 7, $p < 0.01$), Type 1 and 3 (Z = -7.869, p = 0.000) and Type 2 and 3 (Z = -2.576, p = 0.010). The generalization about the hierarchy of the familiarity is as follows:

Type 1 (R-NP)>Type 2 (NR-P) > Type 3 (NR-NP)

The results of these familiarity ratings were compatible with the hypothesis that participants were influenced by root familiarity. Participants found real words significantly more familiar than the other types. Moreover, participants found nonwords with real roots, i.e., Type 1 (R-NP) more familiar than Type 2 (NR-P) and Type 3 (NR-NP). The presence of a real root in Type 1(R-NP) explains these findings, which increases the familiarity of this type compared with the types with nonroots. Furthermore, the pattern types did make a significant difference to the participants' judgments between Type 2 (NR-P) and other Types. The items selected for inclusion in the RAP-NWR test were within the range +1 to -1 SD from the total mean of familiarity rate for each type.

5.4 Hypotheses and Predictions of the Root and Pattern Nonword Repetition Test (RAP-NWR)

The RAP-NWR test was designed to examine the effect of roots and patterns on children's nonword repetition. It was predicted that scores for items containing a real root or a real pattern would be significantly higher than scores for items containing a nonroot and nonpattern. No prediction was made regarding scores for items containing a nonroot vs items containing a nonpattern.

In addition to the experimental hypothesis, we expected performance on the RAP-NWR to reinforce findings for performance on the NWRep test: we predicted significant effects of length, with children gaining higher scores for two-syllable than three-syllable items; and scores on the RAP-NWR test would be significantly correlated with age, and with scores on the receptive and expressive vocabulary tests.

5.5 Methods

5.5.1 Participants

Eighty nine typically developing (TD) Gulf Arabic speaking children participated in this experiment. All children who participated in this study had Gulf Arabic as their first language according to parental report, and none of them participated in the previous experiment.

The participants had a mean age of 48 months. There were approximately equal numbers of children tested in six age bands between the ages of two and seven years old. Children from age 2 to 4 years were divided into six-month-age bands while the group of children aged 5 to 7 years were divided into 11-month bands. The reason behind that is that language development in younger children is expected to accelerate more rapidly when compared to older groups. The children were recruited from two kindergartens and two schools. All children had no history of hearing loss, congenital abnormalities, oral-motor difficulties or autism and all had no history of referral to speech and language therapy services by parental report. A summary of participants' characteristics is shown in Table 18.

Table 18: Characteristics of the participants in the RAP-NWR Test.

Age Groups	Typically Developing Children
Age Band 1: 2;0-2;6 years	
Number of participants (Female: Male)	14 (5:9)
Mean age in months (years)	27.5 (2;3)
Range in months (years)	24-30 (2;0-2;6)
Age Band 2: 2:7 – 3:0 years	
Number of participants	16 (8:8)
Mean age in months (years)	33.6 (2;9)
Range in months (years)	31-36 (2;7-3;0)
Age Band 3: 3;1-3;6 years	
Number of participants	13 (6:7)
Mean age in months (years)	38.6 (3;2)
Range in months (years)	37-42 (3;1-3;6)
Age Band 4: 3;7-4;0 years	
Number of participants	16 (5:11)
Mean age in months (years)	47.1 (3;9)
Range in months (years)	43-48 (3;7-4;0)
Age Band 5: 5;0-5;11 years	
Number of participants	15(6:9)
Mean age in months (years)	62.3(5;3)
Range in months (years)	60-66 (5;0-5;6)
Age Band 6: 6;0-7;0years	
Number of participants	15 (9:6)
Mean age in months (years)	78 (6;5)
Range in months (years)	72-84 (6;0-7;0)
Total number of participants	89 (39-50)
Mean age in months (years)	48.1 (4;0)
Range	24-84 (2;0-7;0)

5.5.2 Material and procedures

Developing the RAP-NWR. The selection of the items used in the RAP-NWR test was based on the familiarity rating test scores. The RAP-NWR test consists of three different types of stimuli, each has 12 items with a total of 36 items (see Appendix L for a complete list of all stimuli). The RAP-NWR consists of two and three syllables items only. The items used in the RAP-NWR were selected from the nonwords of the familiarity rating test. The three different types are shown in Table 19. See appendix G for full RAP-NWR test.

Table 19: An example of the different type of roots and patterns in two and three syllable nonwords that were used to develop the RAP-NWR.

Types	No. of syllables	Root/Nonroot	Pattern	Nonword
Type1 (R-NP)	2	√k.s.b	a-u	/ka.sub/
(root-nonpattern)	3		a-u-a	/ka.su.bad/
Type2 (NR-P)	2	√d.m.b	a-a	/da.mab/
(nonroot-pattern)	3		a-a-a	/da.ma.baf/
Type3(NR-NP)	2	√d.f.k	a-u	/da.fuk/
(nonroot-nonpattern)	3		a-u-a	/da.fu.kab/

Unintentionally, there were five items that violated the OCP-place phonotactic constraint in Arabic. The 5 items were (item no 4:/fusibal/, 21: /damabaf/, 26: /lafabad/, 31: /sumifal/, and 34: /sumif/). Therefore, results were recalculated with and without these items as will be explained in the results section.

The RAP-NWR test was presented live to each participant individually, unlike the first study where recorded stimuli were presented from a laptop through speakers. The recorded stimuli were avoided due to the difficulties that were found when conducting WNRep in the first study (see section 4.4.2) as young children were uncomfortable with stimuli being delivered through the laptop. Some of these children did not respond in the beginning of the task and they needed some time to adapt to the procedure and the task. Therefore, it was preferred in this current study to present the stimuli live to provide a spontaneous and natural atmosphere to present the stimuli. Stimuli were repeated when the child did not pay attention to the first production of the stimuli. This happened more often with the younger children. Children were never presented with the stimuli more than twice, if there was no response even after repetition it was reported as zero. Few self-corrections were noted and they were accepted as correct responses. Examples of children's responses are presented in Appendix M. All testing was conducted in a quiet room. The examiner would start the session by playing with the child to establish rapport. The instructions for each child were the equivalent of the following (in Arabic) "You will listen to funny and mixed up words and I want you to repeat them the way you hear them. Now let's try this…". This was followed by three trial items. Children's productions were audiotaped through a microphone attached to an Olympus VN-5500PC DNS Digital Voice Recorder. The children were rewarded after each task with stickers or small toys. Children were given some breaks during the session when they showed signs that they lost interest or got tired from the task. Each repetition was scored using the percentage-of-phonemes-correct (PPC) scoring method. The PPC was selected to use in this study, as it found earlier in chapter four that there was no difference between using different scoring methods. Minor misarticulations (especially distortion of /r/, and /s/ or substituting [l] for /r/ or [Ɵ] for /s/) were counted as correct.

For the error analysis, children's errors were classified into two main phonological categories. The first category of errors was related to syllabic errors: any deletion of any whole syllable was reported as a syllable deletion error. The second category related to segmental errors. This category was divided into two main types: root errors (consonants) and pattern errors (vowels), and each type was divided into three subtypes: substitution, deletion and addition. In addition, morphization, whereby children changed nonroots to real roots (by substituting consonants) or changed nonpattern to patterns (by substituting vowels) was found in children's responses and recorded.

In addition to the RAP-NWR test, The Arabic Picture Vocabulary Test APVT, a receptive vocabulary test (Shaalan, 2010) and Arabic expressive vocabulary test (AEVT) that were used in the previous experiment were also conducted for 59 children aged 2-4 years old (age group 1-4). Due to time and logistic constraints it was not possible to conduct the vocabulary test with the older groups of 5-7 year olds (age group 5 and 6).

5.6 Results

All data were analysed using repeated measures ANOVAs with nonword types (type 1 (R-NP) vs. type 2 (NR-P) vs. type 3 (NR-NP)) and word length (2-syllables, 3-syllables) as within subject factors and age band as a between subject factor. A Bonferroni correction ($\alpha' = \alpha/k$) was applied to all follow-up tests (Pairwise comparisons and t-tests). It was hypothesised that children would score better in repeating RAP-NWR types that included familiar roots or patterns vs unfamiliar types that had no familiar roots or patterns. Furthermore, it was expected that children would score better in repeating short nonwords vs. long nonwords. Results showed that there was a significant effect of nonword type ($F(2,158) = 5.052$, $p < .001$, $\eta^2 = .060$), Pairwise comparisons showed there was a significant difference in the performance of the children between Type 3 (NR-NP) and Type 1 (R-NP) ($p < .005$). However, the difference between Type 1 (R-NP) and Type 2 (NR-P) was not significant ($p = .127$), and the difference between Type 2 (NR-P) and Type 3 (NR-NP) ($p = .106$) did not reach significance either. In general, children performed more accurately on Type 1 (R-NP) when compared to Type 3 (NR-NP) (81% vs 84.5%). The difference was not large; however, it was statistically significant. Figure 21 shows children’s performance on the three different RAP-NWR types. In general children were more accurate on repeating Type 1 than they were on other types.

There was a significant effect of word length ($F(1,79) = 106.17$, $p < .001$, $\eta 2 = .573$). A pairwise comparison showed there was a significant difference in the performance of the children on two syllable vs. three syllable nonwords ($p < .001$). In general, children performed more accurately on two syllables than they did on three syllable items (76.6 % vs 88.5%). Figure 22 shows how children, in general, scored better on two syllable nonwords than three syllable ones.

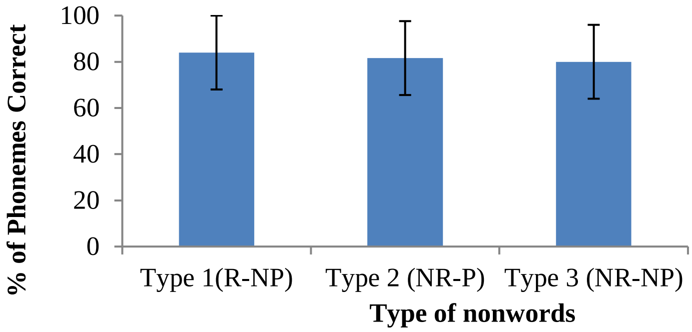

Figure 21: Children's performance on the RAP-NWR test across the three different types of nonwords.

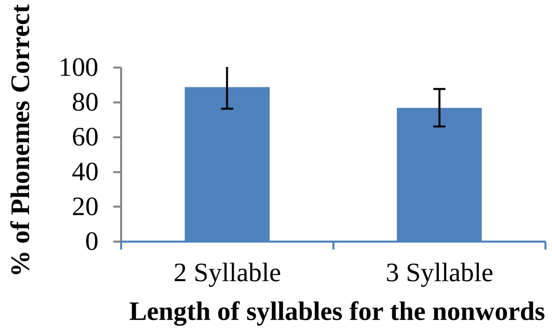

Figure 22: Percentage of correct responses on RAP-NWR test across two and three syllable nonwords for all participants.

Moreover, there was a significant age effect (F=3.64, p<.001). Pairwise comparison showed there was a significant difference between age group 1 and the following age groups: age group 3 (p<.034), age group 4 (p<.000), age group 5 (p<.000), and age group 6 (p<.000) in their nonword repetition score as measured by the percentage of correct phonemes. Age group 1 (children aged between 2;0 and 2;6 years old) performed significantly worse than all other age groups, except age group 2. There was also a significant difference between age group 2 (2;7-3;0) on one hand and age groups 4 (p<.039), age group 5 (p<.000) and age group 6 (p<.000) as age group 2 scored significantly less on the RAP-NWR test when compared to these groups. Age group 3 scored significantly better than age group 1 (p<.034), and scored significantly less when compared to age group 6 (p<.032). Age group 4 children scored significantly better than age group 1 (p<.000) and age group 2 (p<.039). In addition, age group 5 had a more significant score on the test when they were compared to age group 1 (p<.000) and age group 2 (p<.001). Finally, age group 6 scored significantly higher on the test when they were compared to age group 1 (p<.001), age group 2 (p<.001) and age group 3 (p<.032). So the following summarizes the group comparisons:

Age group 1= age group 2 but < age groups 3,4,5, and 6.

Age group 2=age group 3, but < age groups 4,5, and 6.

Age group 3=age groups 4,5, but < age group 6.

Age group 4=age group 5, 6.

Figure 23 shows how children performed on the RAP-NWR test across the six different age groups. In general, the older the children were the more accurate they performed on the RAP-NWR test (with percentage of correct responses ranging between 75% and 97.7%).

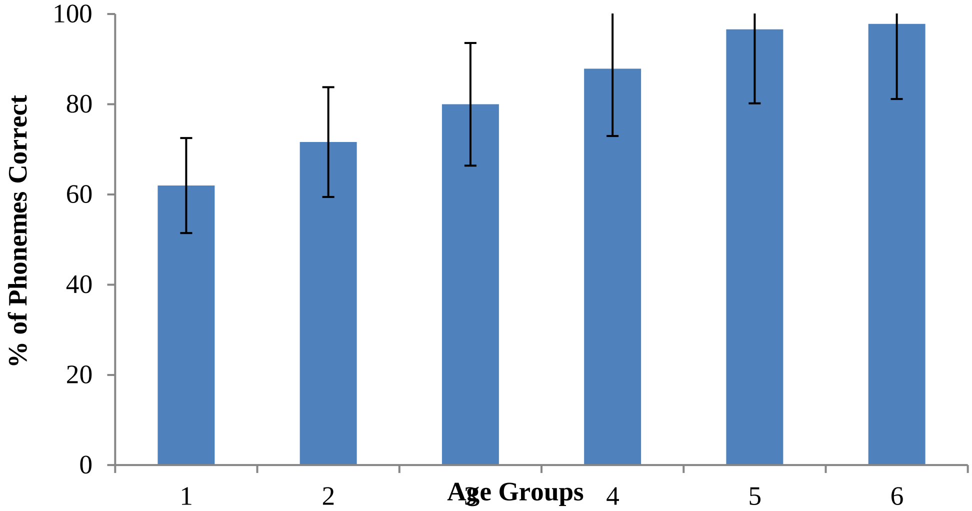

Figure 23:Performance of children on RAP-NWR based on age groups. Age group 1: 2;0-2;6, age group 2: 2;7-3;0, age group 3: 3;1-3;6, age group 4: 3;7-4;0, age group 5: 5;0-5;11, age group 6: 6;0-7;0.

All data were analysed using repeated measures ANOVAs with nonword types (type 1 (R-NP) vs. type 2 (NR-P) vs. type 3 (NR-NP)) and word length (2-syllables, 3-syllables) as within subject factors and age band as a between subject factor. Mauchly's test indicated that the assumption of sphericity had been violated ($\chi2(2)$ =, $p < .05$), therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity. Main effects of RAP-NWR Type was significant ($\varepsilon = 0.819$), $F(1.63, 886.1) = 3.50$, $p < .001$. The interaction between nonword length and RAP-NWR type was also significant ($\varepsilon = 0.795$), $F(1.59, 801.1) = 6.03$, $p < .001$.

5.6.1 The interaction between Types of nonwords and word length

Results of the main ANOVA, showed there was a significant interaction between nonword type and nonword length $F(1.59, 801.1) = 6.03$, $p < .001$. Table 20and

Figure 24 shows that children performed more accurately on two syllable nonwords than they did on three syllable nonwords across RAP-NWR types. The difference between children's performance on the three different root and pattern types (R-NP, NR-P, NR-NP) was minimal on two syllables, while on the three syllable nonwords the difference was greater.

Table 20: Descriptive statistics for the children performance on the RAP-NWR with two and three word length.

RAP-NWR Type	Word length	Mean	SD
Root-nonpattern	2 syllables	88.85	1.6
(R-NP)	3 syllable	80.19	2.1
Nonroot-pattern	2 syllable	88.26	1.56
(NR-P)	3 syllable	76.46	2.05
Nonroot-nonpattern	2 syllable	88.4	1.7
(NR-NP)	3 syllable	73.65	2.16

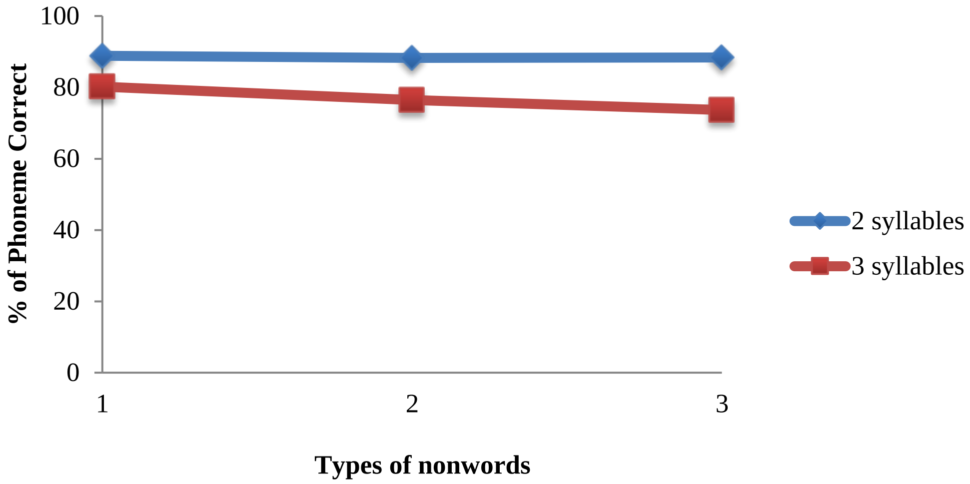

Figure 24: Performance on different nonword types and word lengths. Types of nonwords: Type 1 (R-NP), Type 2 (NR-P), Type 3 (NR-NP).

A series of paired-samples t-tests was conducted to follow up on this interaction. A comparison of children performance on the three different types of root and pattern combinations (R-NP, NR-P, NR-NP) and on the two different types of word length (two and three syllable nonwords) was conducted.

For type 1 (R-NP), there was a significant difference on two syllable nonwords (M=89.1, SD=18.1) and three syllable nonwords (M=80.3, SD=24.7), (t (88) = (5.36), p=.000) as the children repeated accurately R-NP nonwords composed of two syllables more than they did on R-NP with three syllables. Similarly, these children scored significantly better on Type 2 nonwords with two syllables (M = 88.4, SD = 17.7) than they did on Type 2 with three syllables (M=76.6, SD=24.1), (t (88) = 7.19, p= .000). In addition, these children had significantly higher scores on when repeating Type 3 nonwords with two syllables (M = 88.6, SD=19,05) when compared to Type 3 with three syllable (M=73.94, SD), (t (88) =7.17, p =.000). In general, the difference between children's performance on the three different root and pattern types was minimal on two syllable, while on the three syllable nonwords the difference was greater.

5.6.2 Interaction between age and nonword length

The interaction between nonword length and age was significant as reported earlier in the main ANOVA (F=5.571, p<.000). Table 21shows descriptive statistics of children's performance (all age groups) on different nonword lengths (two and three syllable nonwords). The results of paired-samples t-test was conducted to follow- up the interactions between the six different age groups on the RAP-NWR with two and three syllables nonwords are also presented in Figure 25.

Table 21: Descriptive statistics (M =Mean, SD= Standard Deviations) of the performance of the different age groups on the RAP-NWR different word lengths (two and three word lengths).

	2 syllable		3 syllable		Paired-samples t-test	
Age Group	M	SD	M	SD	t-value	Significance
Age group 1 (n=14)	70.00	26.59	53.5	27.15	t (13) = 4.96	p <.001
Age group 2 (n=16)	81.11	14.14	62.10	19.56	t (13) = 5.79	p <.001
Age group 3 (n=13)	87.00	13.28	72.95	21.93	t (13) = 4.97	p <.001
Age group 4 (n=16)	94.79	5.58	81.00	18.05	t (13) = 4.01	p <.01
Age group 5 (n=15)	99.11	1.12	94.12	5.83	t (13) = 3.23	p <.01
Age group 6 (n=15)	99.25	0.99	96.29	4.32	t (13) = 3.21	p <.01

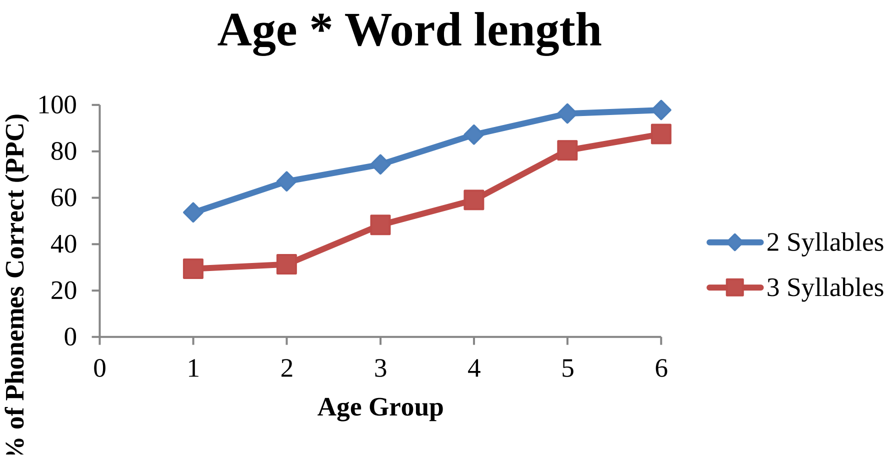

Figure 25: Performance on different word length based on different age groups. Age group 1: 2;0-2;6, age group 2: 2;7-3;0, age group 3: 3;1-3;6, age group 4: 3;7-4;0, age group 5: 5;0-5;11, age group 6: 6;0-7;0.

To investigate effects of word length (two and three syllables) across all age groups, one way ANOVA was conducted, the results showed that there is a significant difference between age group at two syllable nonwords length ($F(5,84) = 9.41$, $p < .001$, and three syllable nonwords length ($F(5,84) = 10.99$, $p < .001$), a multiple comparison with Bonferroni post hoc correction was conducted and the results were are shown in Table 22.

Table 22: Multiple comparison results of the difference between age groups on two and three syllable nonwords.

Age groups	Two syllable nonwords significance level	Three syllable nonwords significance level
Age group 1 vs. age group 2	p=.1	p=.1
Age group 1 vs. age group 3	p =.02	p=.1
Age group1 vs. age group 4	p <.001	p=.388
Age group 1 vs. age group 5	p< .001	p=.001
Age group 1vs. age group 6	p<.001	p <.001
Age group 2 vs. age group 3	p=.1	p=.1
Age group 2 vs. age group 4	p=.002	p=.024
Age group 2 vs. age group 5	p=.005	p<.001
Age group 2 vs. age group 6	p<.001	p<.001
Age group 3 vs. age group 4	p=.541	p=.1
Age group 3 vs. age group 5	p=.023	p=.015
Age group 3 vs. age group 6	p=.21	p=.005
Age group 4 vs. age group 5	p=.1	p=.425
Age group 4vs. age group 6	p=.1	p=.166
Age group 5 vs. age group 6	p=.1	p=.1

Adjustment for multiple comparisons: Bonferroni.

To summarize the results, each age group was significantly more accurate on two syllable nonwords than it was on three syllable nonwords. Furthermore, the magnitude of difference between the two and three syllable decreased when the children's ages increased.

5.6.3 The interaction between age and nonword types

The results of main ANOVA showed that there was no significant interaction between age groups and the RAP-NWR types ($f(10, 158) = .64$, $p < .77$). This might be due to the small number of participants in each age group and/or the small number of items in each RAP-NWR type. When the main effect of the nonword types was calculated (see section 5.6.1) by including all groups together (not separated) the results showed there was a significant difference between Type 1 (R-NP) and Type 3 (NR-NP). Table 23 shows some descriptive statistics about each group scores on the three different RAP-NWR types.

Table 23: Descriptive statistics (M=Mean, SD= Standard Deviations) of the performance of the different age groups on the three different types of the RAP-NWR test.

	Type 1 (R-NP)		Type 2 (NR-P)		Type 3 (NR-NP)	
Age Group	M	SD	M	SD	M	SD
Agegroup1 (n=14)	63.79	29.48	60.52	27.36	56.94	24.59
Age group 2 (n=16)	72.83	18.97	70.57	17.27	66.67	19.37
Age group 3 (n=13)	81.62	20.06	79.19	16.53	75.64	21.14
Age group 4 (n=16)	88.28	12.97	84.81	14.77	87.15	13.33
Age group 5 (n=15)	98.06	3.08	95.83	3.96	94.81	6.14
Age group 6 (n= 15)	98.24	2.61	97.41	3.04	96.94	3.77

Figure 26 shows how the children from the different six age groups performed on the three different RAP-NWR types.

The younger groups 1, 2, 3 and 4 performed differently on the three different types. Children across the first four age groups performed better on Type 1 (R-NP) when compared with Type 2 (NR-P) and Type 3 (NR-NP). The older the children were, the less the difference was in their performance across the three different RAP-NWR types.

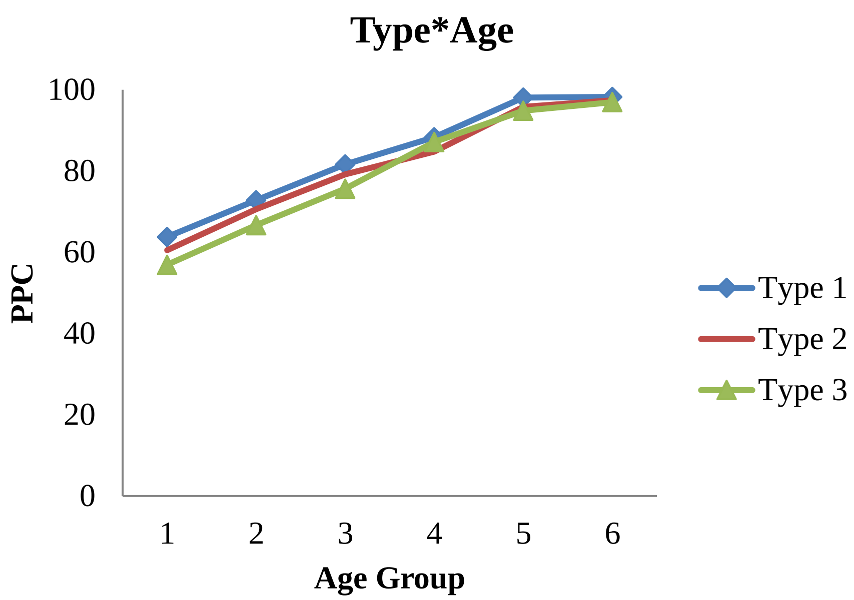

Figure 26: Performance on different nonword types (1= R-NP, 2=NR-P, 3 NR-NP) based on different age groups. Age group 1: 2;0-2;6, age group 2: 2;7-3;0, age group 3: 3;1-3;6, age group 4: 3;7-4;0, age group 5: 5;0-5;11, age group 6: 6;0-7;0.

The previous parts of analysis tried to explore children's performance, across all age groups on RAP-NWR with reference to the different nonwords type and lengths. Children performed significantly more accurately on two syllable nonword vs three syllable nonwords. Furthermore, children scored significantly more accurately on Type 1 (root-nonpattern) than Type 3 (nonroot-nonpattern). These findings support what was hypothesized earlier in section 5.4. Nevertheless: it was expected that children would score significantly better on items that included familiar patterns (Type 2) vs items that included unfamiliar patterns (Type 1 and 3). The results, however, were contrary to the hypothesis and there was no significant difference in children's performance on Type 2 (familiar patterns) vs Types 1 and 3 (unfamiliar patterns).

In the following section, the correlation between children's performance on RAP-NWR and the receptive and expressive vocabulary tests will be explored.

As mentioned earlier in this chapter, the RAP-NWR test unintentionally contained five items that violated the OCP-Place phonotactic constraint. Therefore, all the results were analysed again by excluding these items that violated the OCP-Place constraint. There were no differences found in the results with and without excluding those five items., so only results for the full set of items are included here. See

Appendix H for the results excluding the five items violating OCP-Place in the RAP-NWR test.

5.6.4 Relationship of RAP-NWR test to other vocabulary tests

The vocabulary tests in this experiment were conducted with 59 children only (age groups 1 to 4) as mentioned earlier in section 5.5. The descriptive statistics of all participants in each of the Arabic Picture Vocabulary Test (APVT), the Arabic Expressive Vocabulary Test (AEVT) and RAP-NWR test are displayed in Table 24.

A Pearson product-moment correlation coefficient was calculated to measure the correlation between RAP-NWR performance and receptive and expressive vocabulary tests. It was hypothesised that the children's scores on RAP-NWR test would correlate significantly with age, receptive and expressive vocabulary tests. Results of the various correlations are shown in

Table 25 and Figure 26. Results showed that RAP-NWT significantly correlated with age (r=.53, p<.001), the Arabic Picture Vocabulary test (APVT) (r=.62, p=.001), and the Arabic Expressive vocabulary test (AEVT) (r=.75, p<.001)

Table 24: Means (and standard deviations) of the raw scores on the RAP-NWR test.

Age Groups	RAP-NWR	APVT	AEVT
Group 1 (2;0-2;6) years			
Number of participants	14	14	14
Mean Raw Score & (SD)	13.0 (56)	13.2(2.54)	8.5(3.08)
Range of scores	38-197	11-20	2-13
Group 2 (2;7-3;0 years)			
Number of participants	16	16	16
Mean Raw Score & (SD)	151(35.1)	15.3(3.24)	12.5 (2.09)
Range of scores	77-208	9-20	10-17
Group 3 (3;1-3;6) years			
Number of participants	13	13	13
Mean Raw Score & (SD)	170 (39.1)	20 (7.3)	16.2 (4.4)
Range of scores	101-208	9-37	10-25
Group 4 (3;7-4;0) years			
Number of participants	16	16	16
Mean Raw Score & (SD)	187.3 (27)	26 (1.09)	18.2 (3.0)
Range of scores	106-216	14-52	13-24

In addition, the APVT significantly correlated with age (r=.62, p=.000), the AEVT (r=.64, p<001) and the RAP-NWR test (r=.39), p<.001).). The AEVT correlated significantly with age (r=.75, p<001).), APVT (r= .64, p=.001), RAP-NWR test (r=.62, p<.001). These results of the correlation analysis were consistent with the prediction of a significant correlation between RAP-NWR test scores and APVT, AEVT and age. See Table 25 and Figure 27.

Table 25: Correlations between RAP-NWR test, age the other tests: APVT (Arabic Picture Vocabulary Test), AEVT (Expressive Arabic Vocabulary Test). (n=59).

		RAP-NWR	APVT	AEVT	Age in months
RAP-NWR	Pearson Correlation	1	.398**	.621**	.539**
	Sig. (2-tailed)		.002	.000	.000
APVT	Pearson Correlation	.398**	1	.640**	.602**
	Sig. (2-tailed)	.002		.000	.000
AEVT	Pearson Correlation	.621**	.640**	1	.755**
	Sig. (2-tailed)	.000	.000		.000
Age in Months	Pearson Correlation	.539**	.602**	.755**	1
	Sig. (2-tailed)	.000	.000	.000	

**. Correlation is significant at the 0.01 level (2-tailed).

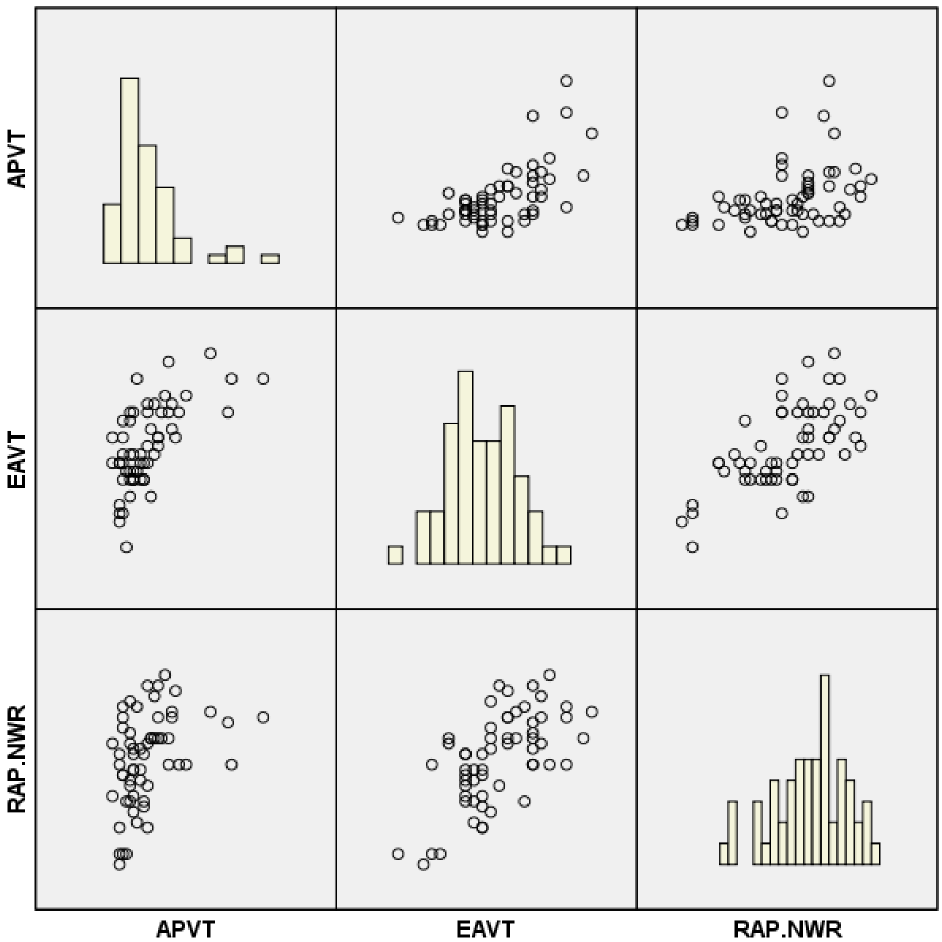

Figure 27: A scatterplot of the correlation between Arabic Picture Vocabulary Test (APVT), the Arabic Expressive Vocabulary Test (AEVT) and Root and Pattern Nonword repetition test RAP-NWR.

To measure the correlation between each RAP-NWR type and APVT and EAVT, another A Pearson product-moment correlation coefficient was calculated to measure the correlation between Type 1, 2 and 3 scores and receptive and expressive vocabulary tests Results showed that Type 1 significantly correlated with Arabic Picture

Vocabulary test (APVT) (r=.387, p=.002), and the Arabic Expressive vocabulary test (AEVT) (r=.545, p<.001). Type 2 significantly correlated with APVT (r=.345, p=.007), and AEVT (r=.595, p<.001). Type 3 significantly correlated with APVT (r=.395, p=.002), and AEVT (r=.621, p<.001). Furthermore, Type 1 correlated with Type 2 (r=.828, p<.001), Type 1 correlated with Type 3 (r=.851, p<.001). Type 2 correlated with Type 3 (r=.842, p<.001).

5.6.5 Error analysis

Error analysis was conducted to explore:

- Errors at the morpheme level, whereby non-words were replaced with real morphemes (roots or patterns); this is the morphological equivalent to lexicalisation whereby a non-form is replaced with a real lexical item, and will be termed 'morphization'.
- Errors at the phonological level, whereby syllables or phonemes (consonants and vowels) were omitted or replaced. Furthermore, we analysed children's performance on each consonant and vowel depending on their position in the nonword.
- When a syllable was deleted it was not counted as a consonant or vowel deletion. Similarly, root and vowel morphization were not counted substitiutions.

Examples of different types of errors found in the children's responses are listed in Table 26.

Table 26: Examples of the error types in responses to RAP-NWR test

	Type of error	Target nonword	Child's response
Syllabic errors	Syllable deletion	/la.fa.bad/	[la.fa]
Segmental errors	Consonant substitution	/la.mus/	[la.bus]
	Consonant deletion	/fu.lit/	[fu.li]
	Consonant addition	/fu.lit/	[fu.litt]
	Vowel substitution	/ka.tub/	[ka.tab]
	Vowel deletion	/da.fas/	[da.fs]
	Vowel addition	/sa.bu.daf/	[sa.bu.dafa]
Morphization	Non-root to root	/sa.ka.dab/	[sa.ka.bab]
	Non-pattern to pattern	/ka.tub/	[ka.tab]

This section presents descriptive statistics showing the distribution of error types according to RAP-NWR types. A series of paired t-tests was conducted to compare children's accuracy on each consonant and vowel across different RAP-NWR types. We then examined in detail the substitutions and deletions for each target consonant and vowel according to their position in the nonword. We also explored if consonant/vowel position in the nonword affects children's repetition accuracy.

Distribution of error types

Table 27 shows, for each error type, the frequency of occurrence and the percentage of all errors for all participants. As can be seen, the vast majority of errors were consonant substitutions (47.8%). Consonant deletions accounted for 26.92% with consonant additions making up less than 1% of total errors. Vowel errors were far less frequent, and vowel deletions were the most common error in patterns (8.83% of total errors),. Vowel substitutions were 4.65% and vowel additions were less than 1%. Morphization for roots and patterns made up less than 2% of the children's total errors, with almost equal numbers of real root and real pattern substitutions.

Table 27: Distribution of error types on RAP-NWR test for all children (n=89)

Type of error		Frequency	% of all errors
Syllable deletion		138	7.12
Non-root morphization		30	1.55
Non-pattern morphization		35	1.81
Consonant Errors	Substitution	926	47.81
	Deletion	522	26.94
	Addition	11	0.57
Vowel Errors	Substitution	90	4.65
	Deletion	171	8.83
	Addition	14	0.72
Total		1937	100.00

Distribution of consonant and vowel errors according to RAP-NWR type

Table 28 shows the percentage of correct consonants (PCC) and percentage of correct vowels (PCV) according to RAP-NWR type. The results in Table 28 showed that children performed more accurately on repeating consonants in type 1 (which contained real roots) when compared to types 2 and 3. Similarly, children repeated vowels in type 2 (which contained real patterns) more accurately compared to types 1 and 3. The accuracy of repeating nonwords that included familiar consonant or vowel sequences was higher compared with unfamiliar ones.

Table 28: Percentage of consonants and vowels correct according to RAP-NWR type

NW Type	PCC	PVC
Type 1	83.60	88.20
Type 2	79.24	89.10
Type 3	78.81	85.96

PCC=Percent consonants correct, PVC= Percent vowels correct

To explore if the percentage of consonants correct (PCC) was affected by RAP-NWR type, a paired t-test was conducted. This showed that the PCC in type 1 (M= 83.60, SD=20.41) was significantly higher than PCC in type 2 (M=79.24, SD= 19.25), (t (88) =4.05, $p<.001$) and type 3 (M=78.8, SD= 17.95), (t (88) =4.06, $p<.001$). The PCC in type 2 (M=79.24, SD= 19.25) was not significantly different than PCC in type 3 (M=78.8, SD= 17.95) (t (88) =.37, $p=.712$). These results suggested that children were more accurate in repeating consonant sequences that occured as real roots.

Another paired t-test was conducted to explore if the percentage of vowels correct (PVC) was influenced by different RAP-NWR types. The results showed that PVC in type 1 (M= 88.2, SD=18.4) was not significantly different from type 2 (M=89.10, SD=16.77) (t (88) = 91, p=.36). However, PVC in type 1 (M= 88.2, SD=18.4) was significantly higher than type 3 (M=85.9, SD=19.8) (t (88) =2.08, p=0.04). Furthermore, PVC in type 2 (M=89.10, SD=16.77) was significantly higher than type 3 (M=85.9, SD=19.8) (t (88) =2.82, p=.006). These results indicated that children were more accurate in repeating nonwords that consisted of real pattern (Type 2) when compared to nonwords that consisted of nonpattern -nonroot nonwords (Type 3). On the other hand, repeating Type 1(root-nonpattern) was almost as accurate as repeating Type 2 words that consisted of a real patterns.

Distribution of phoneme substitution and phoneme deletion according to target phoneme and its position

Table 29 and Table 30 show in detail the distribution of substitutions and deletions, according to children's responses for consonants and vowels respectively and based on the target consonant/vowel and its position. The frequency of occurrence for each targeted consonant and vowel are also displayed in the same table.

Table 29: Distribution of substitutions and deletions for each consonant according to consonant position in the nonword. C=consonant.

Target consonant and frequency of occurrence in RAP-NWR test	**Position of the consonant**			
	C1	**C2**	**C3**	**C4**
/k/ C1=10 C2=4 C3=1 C4=2	Substitutions n=39 /k/->[t] n=26 /k/->[s] n=4 /k/->[d] n=3 /k/->[m] n=2 /k/->[b] n=2 /k/->[l] n=1 /k/->[n] n=1 Deletions n=17	Substitutions n=45 /k/->[f] n=18 /k/->[b] n=10 /k/->[t] n=7 /k/->[s] n=4 /k/->[d] n=3 /k/->[l] n=3 Deletions n=12	Substitutions n=18 /k/->[t] n=6 /k/->[f] n=5 /k/->[d] n=4 /k/->[b] n=3 Deletions n=3	Substitutions n=19 /k/->[t] n=15 /k/->[l] n=2 /k/->[d] n=1 /k/->[f] n=1 Deletions n=5
/b/ C1=0 C2=6 C3=8 C4=6		Substitutions n=29 /b/->[f] n=16 /b/->[m] n=3 /b/->[d] n=3 /b/->[t] n=2 /b/->[m] n=2 /b/->[l] n=1 /b/->[k] n=1 /b/->[s] n=1 Deletions n= 12	Substitutions n=75 /b/->[f] n=27 /b/->[t] n=20 /b/->[d] n=15 /b/->[s] n=8 /b/->[m] n=4 /b/->[n] n=1 Deletions n=12	Substitutions n=27 /b/->[k] n=10 /b/->[f] n=8 /b/->[t] n=5 /b/->[s] n=3 /b/->[l] n=1 Deletions n=21
/s/ C1=7 C2=6 C3=10 C4=1	Substitutions n=30 /s/->[t] n=11 /s/->[k] n=7 /s/->[d] n=6 /s/->[l] n=2 /s/->[m] n=2 /s/->[n] n=1 /s/->[ʃ] n=1	Substitutions n=65 /s/->[f] n=19 /s/->[t] n=12 /s/->[d] n=11 /s/->[b] n=10 /s/->[l] n=5 /s/->[n] n=4 /s/->[k] n=4	Substitutions n=21 /s/->[f] n=13 /s/->[b] n=4 /s/->[l] n=2 /s/->[d] n=1 /s/->[b] n=1	Substitutions n=10 /s/->[f] n=5 /s/->[n] n=2 /s/->[ʃ] n=1 /s/->[m] n=1 /s/->[k] n=1

	Deletions n=19	Deletions n=2	Deletions n=10	Deletions n=9
/d/ C1=7 C2=1 C3=7 C4=4	Substitutions n=16 /d/->[s] n=5 /d/->[k] n=3 /d/->[m] n=3 /d/->[l] n=3 /d/->[n] n=1 /d/->[t] n=1 Deletions n=17	Substitutions n=8 /d/->[t] n=6 /d/->[f] n=2 Deletions=0	Substitutions n=37 /d/->[b] n=16 /d/->[t] n=10 /d/->[k] n=7 /d/->[n] n=1 /d/->[s] n=1 /d/->[f] n=1 /d/->[l] n=1 Deletions n=14	Substitutions n=27 /d/->[b] n=9 /d/->[f] n=5 /d/->[s] n=5 /d/->[l] n=3 /d/->[t] n=3 /d/->[n] n=1 /d/->[l] n=1 Deletions n=21
/f/ C1=4 C2=8 C3=5 C4=4	Substitutions n=9 /f/->[s] n=4 /f/->[b] n=2 /f/->[t] n=1 /f/->[m] n=1 /f/->[k] n=1 Deletions n=15	Substitutions n=52 /f/->[b] n=25 /f/->[s] n=9 /f/->[k] n=8 /f/->[t] n=4 /f/->[m] n=3 /f/->[d] n=2 /f/->[l] n=1 Deletions n=16	Substitutions n=56 /f/->[s] n=29 /f/->[t] n=5 /f/->[m] n=5 /f/->[l] n=4 /f/->[b] n=4 /f/->[k] n=3 /f/->[d] n=3 /f/->[n] n=2 /f/->[d] n=1 Deletions n=11	Substitutions n=47 /f/->[s] n=24 /f/->[d] n=8 /f/->[l] n=5 /f/->[k] n=5 /f/->[t] n=4 /f/->[b] n=1 Deletions n=57
/t/ C1=0 C2=2 C3=2 C4=0		Substitutions n=21 /t/->[b] n=9 /t/->[k] n=8 /t/->[s] n=3 /t/->[d] n=1 Deletions=0	Substitution n=42 /t/->[f] n=17 /t/->[k] n=10 /t/->[l] n=6 /t/->[s] n=6 /t/->[b] n=1 /t/->[d] n=1 /t/->[n] n=1 Deletions n=3	
/m/ C1=0		Substitutions n= 37 /m/->[f] n=11 /m/->[l] n=9		

C2=5 C3=0 C4=0		/m/->[b] n=7 /m/->[s] n=4 /m/->[n] n=3 /m/->[d] n=2 /m/->[r] n=1 Deletions n=5		
/n/ C1=0 C2=2 C3=0 C4=1		Substitutions n=30 /n/->[m] n=13 /n/->[b] n=8 /n/->[l] n=4 /n/->[t] n=2 /n/->[f] n=2 /n/->[s] n=1 Deletions n=3		Substitutions n=5 /n/->[l] n=2 /n/->[k] n=1 /n/->[t] n=1 /n/->[f] n=1 Deletions n=16
/l/ C1=8 C2=2 C3=1 C4=4	Substitutions n=21 /l/->[k] n=9 /l/->[t] n=4 /l/->[n] n=3 /l/->[m] n=3 /l/->[s] n=1 /l/->[f] n=1 Deletions n=17	Substitutions n=8 /l/->[k] n=5 /l/->[t] n=2 /l/->[s] n=1 Deletions n=4	Substitutions n=11 /l/->[m] n=7 /l/->[n] n=1 /l/->[k] n=1 /l/->[t] n=1 /l/->[f] n=1 Deletions n=3	Substitutions n=9 /l/->[n] n=4 /l/->[f] n=2 /l/->[b] n=1 /l/->[s] n=1 /l/->[d] n=1 Deletions n=27

Table 30: Distribution of substitutions and deletions for each vowel according to vowel positions in the nonword. V=vowel.

Target vowel and frequency of occurrence in RAP-NWR test	Positions of the vowel		
	V1	V2	V3
/a/ V1=30 V2=12 V3=23	Substitutions n=6 /a/->[u] n=5 /a/->[i] n=1 Deletions n=58	Substitutions n=4 /a/->[u] n=2 /a/->[i] n=2 Deletions n=20	Substitutions n=4 /a/->[i] n=3 /a/->[u] n=1 Deletions n=48
/u/ V1=6 V2=17 V3=0	Substitutions n=13 /u/->[a] n=9 /u/->[i] n=4 Deletions n=11	Substitutions n=49 /u/->[a] n=48 /u/->[i] n=1 Deletions n=23	
/i/ V1=0 V2=2 V3=0		Substitution n=14 /i/->[a] n=12 /i/->[u] n=2 Deletions n=11	

Distribution of consonant and vowel repetition according to position of the target

The percentage correct for each consonant and vowel position for each child is presented in Appendix N. Table 31 shows mean and standard deviation for each consonant and vowel position for all children.

Table 31: Percentage of correct consonants and vowels according to their positions in the nonword for all participants (n=89).

	C1	**C2**	**C3**	**C4**	**V1**	**V2**	**V3**
Mean	85.89	82.46	81.55	70.85	89.64	87.86	83.83
SD	20.42	19.25	17.94	25.38	16.92	17.22	21.15

To examine if consonant repetition accuracy was influenced by position in the nonwords, a paired t-test was conducted to compare the mean of each consonant (C) based on its position (e.g. 1,2,3 or 4). The results showed that C1 (M=85.9 SD =20.4) was repeated significantly more accurately than C2 (M= 82.4, SD=19.25), (t (88) = 4.9, p<.001), C3 (M=81.5, SD= 17.9), (t(88)=3.8, p<.001) and C4 (M=70.8, SD= 25.37), (t(88)= 8.5, p<.001). Correct productions of C2 (M= 82.4, SD=19.25) were not significantly different from C3 (M=81.5, SD= 17.9), (t (88)= 1.3, p=.191). However, C2 was repeated significantly more accurately than C4 (M=70.8, SD= 25.37), (t (88) =7.6, p<.001). Finally, C3 (M=81.5, SD= 17.9) was repeated significantly more accurately than C4 (M=70.8, SD= 25.37) (t(88)=7.5, p<.001).

To explore if vowel repetition accuracy was influenced by position in the nonwords, paired t-tests were conducted to compare the mean of each vowel (V) according to its position in the nonword. Results showed that V1 (M= 89.64, SD= 16.92) was repeated significantly more accurately than V2 (M= 87.86, SD= 17.2), (t (88)=3.94, p<.001) and V3 (M= 83.8, SD= 21.14), (t (88)= 5.47,p<.001). Vowel 2 (V2) (M= 87.8, SD= 17.2) was repeated significantly more accurately than V3 (M= 83.8, SD= 21.14) (t (88)= 4.38,p<.001).

5.7 Discussion

This study was conducted to achieve four main objectives: first, to examine the effects of three different types of roots and pattern combinations on children's performance on a nonword repetition test in Gulf Arabic. Secondly, it tried to assess effects of phonological storage combined with root and pattern knowledge by exploring the effects of nonword length using two and three syllable nonwords. Thirdly, this study aimed to shed some light on the development of the processing of roots and patterns across different age groups (children aged between two and seven years). Finally, the relationship between children's performance on the root and pattern NWR and receptive and expressive vocabulary tests was studied and compared with findings in Arabic, including the study in the previous chapter, and other languages.

5.7.1 Implications of the root effects

There are no studies that systematically examined the effects of roots and patterns on NWR, as there are few languages that have this linguistic feature. Most of the studies that examined nonword repetition in root and pattern languages did not look into this effect (e.g., Shaalan, 2010 in Gulf Arabic). The stimuli used in Shaalan (2010) differed from those used in the current studies in some aspects. Shaalan's (2010) nonword list consisted of 56 two and three syllable nonwords. He used 7 tri consonantal nonroots to create all two and three syllable nonwords and used 4 vocalic nonpatterns (in 48 nonwords) and 2 common vocalic patterns (in 8 control nonwords) to control for pattern effects. Shaalan (2010) found that TD children (average age 5;8 for language matched group and 7;8 for age matched group) were sensitive to pattern effects, while children with SLI (average age of 7;8) were not, however; his study did not examine effects of roots as all his stimuli were nonroots. Therefore, the results obtained here are unique and will shed light on some of the processes underlying children's abilities to repeat nonwords in root and pattern languages.

Children in this experiment showed significant differences on their performance between Type 1 (root-nonpattern) and Type 3 (nonroot-nonpattern). The children were significantly more accurate in repeating Type 1 nonwords (root-nonpattern) than they were in repeating Type 3 nonwords (nonroot-nonpattern). The presence of an existing, familiar root makes root-nonpattern (R-NP) nonwords easier to repeat, i.e., children were more likely to repeat a nonword accurately if it had an existing root. On the other hand, they were less likely to repeat a nonword accurately when it had a non-existing root and a non-existing pattern (NR-NP). Poor performance on (NR-NP) nonwords could be due to lack of semantic cues that were available in Type 1 (R-NP).

These root effects could differ from wordlikeness as defined in other languages due to the fact that Semitic roots consist of noncontiguous consonants that carry semantic information and are intertwined with vocalic patterns to form words. These semantic cues present in the root seem to account for the performance of these children at this age more than the vocalic pattern that carries morphological and syntactic information. Children in this current study were sensitive to consonantal roots that were embedded within the nonwords even though these consonants were separated by nonpatterns. This differs from wordlikeness in that children have to pay attention to consonants that are spread throughout a nonword and do not rely on a specific sequence of sounds as in wordlikeness effects. Children's ability to identify the root reflect that the phonological processes underlying the repetition is influenced by the root morphology.

These findings met the prediction of this study, which was that the children responded more accurately to nonword items with real roots compared with the other types. The root effect is related to semantic and wordlikeness effects, as the presence of the root in the nonword adds semantic and sublexical cues that would influence the repetition, however it differs from wordlikeness effects as explained above. Findings from a variety of studies (e.g., Gathercole, 1995; Gathercole & Adams, 1994; Gathercole et al., 1999) suggest that the more wordlike the item is, the more likely that it will be repeated correctly. Findings of this study are also consistent with results from other studies that controlled for wordlikeness effects, for example, Roy and Chiat (2004) who tested 66 children using the Preschool Repetition test (PSRep), found that children aged two to four years old performed better on repeating words than nonwords.

Accuracy on NWR improves when nonwords include either syllables that are themselves lexical items (Dollaghan, Biber & Campbell, 1993) or segments with high phonotactic frequencies (Munson, 2001). Consequently, the findings of this study are consistent with the phonological processing account of nonword repetition (Snowling et al. 1991; Chiat, 2001). According to this account, previous lexical knowledge supports better nonword repetition as familiarity will increase when there are sublexical units implanted within nonwords (Gathercole et al., 1991). Children's skills in repeating nonwords improve with age as their vocabulary developed (Masoura & Gathercole, 2005; Metsala & Walley, 1998; Munson, Edwards, & Beckman, 2005)

On the other hand, children in this study were less likely to repeat accurately Type 3 (NR-NP) than Type 1 (R-NP), as Type 3 (NR-NP) had weaker root and pattern effects. Type 3 (NR-NP) in this study provides a pure and sensitive measure for phonological memory in isolation of any lexical knowledge. Type 3 (NR-NP) nonwords consisted of non-existing roots and non-existing patterns and therefore it is the least wordlike in the RAP-NWT. The design of Type 3 (NR-NP) nonwords was similar to that used in Shaalan (2010) study where it proved critical in differentiating the performance of children with LS from age and language matched groups.

Furthermore, the current findings are consistent with studies in Hebrew which shares Arabic language the same root and pattern principle. Berent and Shimron (1997) found adult participants were more likely to reject nonwords that were constructed from a nonroot. Additionally, Clark and Berman (1984) found that preschool children were able to predict the meaning of a novel word that included a real root but sounded like a nonword.

According to the current findings, the effect of the root on children's repetition can not be isolated from children's lexical knowledge, while at the same time they are not the same skill. The RAP-NWR test was designed to measure children's sensitivity to root and pattern skills, however, children's sensitivity to the root and pattern might be a combination of both skills (root and pattern and lexical knowledge). The knowledge of root and pattern principle in children who speak Sematic languages will lead to more complex morphosyntactic skills that are needed to acquire vocabulary that belongs to different classes (e.g. nouns, verbs, adverbs, adjectives) but based on the same root, which is a different skill than lexical knowledge. Findings of RAP-NWR showed children's sensitivity to root and pattern that may combine with their lexical knowledge or familiarity. To explore how Arabic speaking children used the morphosyntactic skills that are based on roots and patterns, different tasks might be more useful like analog tasks, where the child is asked to build new word classes from the same root.

5.7.2 Implication of effects of vocalic patterns

Results of the present study showed that different types of vocalic patterns did not significantly influence children's performance on the RAP-NWR test. Children numerically performed better on Type 2 (NR-P), which consisted of very common Arabic patterns (with repeated vowels (a-a) and (a-a-a)), than they did on Type 3 (NR-NP), which consisted from uncommon patterns with different vowels (e.g.,(a-u), (a-u-a), (u-u), (u-u-a)). However, the difference was not significant (82.3% on Type 2 vs 81% on Type 3). These findings were not consistent with the prediction of this experiment. It was expected that children would find the common patterns easier to repeat than uncommon ones. The results were not consistent with Shaalan's (2010) study who used a nonword repetition test with vocalic non patterns similar to those used in the current study with older typically developing children and children with SLI. Shaalan (2010) found that the differences in accuracy on low vs. high frequency pattern was not significant in the SLI group. However, the TD group found the high frequency patterns (real patterns) (e.g., (a-a)) significantly easier than the low frequency patterns (or nonpatterns) ($p < .001$) (Shaalan, 2010). The current findings of nonsignificant effects of patterns can possibly be attributed to two main factors. First, the number of items used in each pattern type was not controlled as they were not equally distributed. Second, the participants' average age in this study was 4 years, while it was 7;8 in Shaalan's (2010) study. Malenkey (1997) reported that children's awareness to vocalic patterns was mastered by the age of 10. Moreover, the current study was not designed to test the articulatory processing in nonword repetition for roots or patterns. Therefore, the increased articulatory complexity in the uncommon patterns compared to the common patterns may have confounded the results of patterns. As mentioned earlier, the common patterns consisted of one repeated vowel (a-a) compared to three different vowels (e.g.,(a-u-a)) in the uncommon ones, which might make these uncommon patterns more difficult to articulate. The articulatory complexity could have been controlled more tightly if the numbers in all different subtypes were even. The total number of the different patterns used in RAP-NWR test was six. These different common and uncommon patterns were used randomly and unequally through the different nonword types of the RAP-NWR test. For example, the (a-u) pattern was used ten times throughout the test while the (u-i) pattern was used only twice. In addition, (a-u-a) was used eight times and (u-i-a) was used four times only. Finally, (a-a) and (a-a-a) were used six times each. The uneven number in each subtype of uncommon patterns made it difficult to control and rendered them not amenable to statistical analysis. This might be considered as a limitation in the stimuli and the design of this study and should be avoided in future study by considering an even number of items in each pattern type.

5.7.3 Implications of word length

The findings of this study showed significant differences between two and three word length conditions as the children showed better performance on two syllable nonwords when compared to three syllable nonwords. This finding might support the phonological short term memory (PSTM) account of nonword repetition (Gathercole & Baddeley, 1989; 1990; Gathercole et al., 1994). In these studies, as mentioned earlier in chapter 2, longer nonwords resulted typically in more repetition errors than shorter nonwords. Therefore, word length is considered to be the main measure of phonological short term memory (PSTM) (Baddeley, Gathercole & Papagno 1998; Dollaghan & Campbell, 1998).

Furthermore, the interaction between word length and the age showed a significant developmental trend in this study. The accuracy of children's performance on longer syllables increased with age. Children's were more accurate in repeating two syllable nonwords vs three syllable nonwords and the gap between both nonwords length scores decreased with age. These findings are consistent with studies such as Adams & Gathercole (1995). Children's articulation, memory, and phonological skills improved with age and resulted in better repetition across different word lengths.

5.7.4 The developmental trend of the RAP-NWR

The results of this study showed that in general the phonological development of Gulf-Arabic speaking children is in line with other studies of the development of phonological skills. For example, Munson (2001) found that children aged three to eight years old were less accurate in repeating nonwords when compared with adults' performance. The results of the percentage rates of the RAP-NWR test (Table 21) showed that older children outperformed their younger peers on NWR. The reason that older children scored better on NWR can be explained by many factors. First, children's skills in repeating nonwords improve with age as their vocabulary developed (Masoura & Gathercole, 2005; Metsala & Walley, 1998; Munson, Edwards, & Beckman, 2005). Second, oral motor skills improved with age (Kovas et al., 2005) as the performance of the children on nonword repetition was predicted by their oral motor skills. Krishnan et al. (2013) found that accuracy of nonword repetition was highly associated with articulation performance in preschool typically developing children. The third factor is that phonological short term memory develops with age according to the PSTM account, and the better the phonological memory is the better the nonword repetition would be. Under the age of 6 years old, PSTM will support nonword repetition, however, a shift in the developmental direction occurs with older children, as they rely on their vocabulary development to support their PSTM when repeating unfamiliar words (Gathercole et al., 1992). Finally, Children's knowledge about root and pattern in Semitic languages improves with age (Clark& Berman, 1984).

This study revealed there was a strong interaction between word length and age groups, which has been reported in studies in other languages (Baddeley, Gathercole & Papagno, 1998; Dollaghan & Campbell, 1998). Although there was a main effect of the RAP-NWR types, the interaction between age group and different RAP-NWR types was not significant. TD children at these young ages (between 2 to 7 years) were sensitive similarly to effects of roots and patterns with no clear interaction between age and type of nonwords. It is possible that this group of participants were similarly sensitive to the presence or absence of root effects as studies of Semitic languages found that children as young as three years old were sensitive to root effects (Clark & Berman, 1984). On the other hand, pattern effects lag behind those of roots as studies reported that awareness of pattern effects develops around the age of ten (Malenkey, 1997). Therefore, children in this current study were mostly sensitive to presence of roots and were mostly not influenced by types of patterns and this may explain the lack of interaction. Another possible reason for lack of interaction was the small number of stimuli in each RAP-NWR type (12 items in each nonword type) or the small number of participants in each age group (average of 15 participants in each age group).

5.7.5 RAP-NWR in correlation with other vocabulary tests

The correlation of this nonword test with expressive and receptive vocabulary tests is consistent with other findings in the literature for English and other languages. Table 25 shows significant correlations between the RAP-NWR test, age and receptive and expressive vocabulary tests.

The strong correlation between the RAP-NWR and the Arabic Expressive Vocabulary Test (AEVT) is not consistent with the findings of Briscoe et al. (2001), who showed that links between nonword repetition and vocabulary in typically developing children only held for measures of receptive vocabulary, not for measures of expressive vocabulary. In addition, Stokes et al., (2013) found only a weak correlation between NWR and expressive vocabulary. The receptive vocabulary accounted for 14.9% of the variance in NWR scores, while expressive vocabulary accounted for 5.8% (Stokes et al., 2013). Conti-Ramsden (2001) also found no correlation between nonword repetition and either receptive or expressive vocabulary size for participants aged from 7-11 years old. However, this study is consistent with the results of Shaalan (2010) in Gulf Arabic, although he reported weaker correlations due to small number of participants and the fact that he tested older children (aged five to nine years).

There are few studies that investigated the relationship between NWR and expressive vocabulary compared with numerous studies that targeted the relationship between receptive vocabulary and NWR. However, both expressive vocabulary and NWR require articulation output skills to produce the stimuli. This is not required with receptive vocabulary. Both RAP-NWR and AEVT have joint requirements of articulating and pronouncing a series of sounds, whether words or nonwords. Some studies investigated the relationship between nonword repetition and articulation and oral motor skills (Kovas et al., 2005, Krishnan et al., 2013) and found a strong evidence for the correlation between the two skills.

The strong correlation between RAP-NWR test and AEVT demonstrates that the RAP-NWR could be potentially used as a successful tool to predict the expressive vocabulary of Gulf Arabic speaking children. It may be worth in the future expanding this test to be used with both typically and atypically developing children.

The strong correlation between the receptive vocabulary test (APVT, Shaalan, 2010) and RAP-NWR is consistent with most of the studies that found a strong correlation between receptive vocabulary and nonword repetition. Consequently, NWR could become a predictor for vocabulary knowledge in both typically and atypically developing children (Briscoe et al., 2001; Coady & Evans, 2008; Gathercole et al., 1991; 1992).

In conclusion, the Root and Pattern Nonword Repetition test (RAP-NWR test) correlated with the Arabic Picture Vocabulary Test (APVT, Shaalan, 2010), a result that was reported in most of the studies in different languages. There are few studies that investigated the relationship between nonword repetition and expressive vocabulary. Our findings showed a strong correlation between the Arabic Expressive vocabulary test (AEVT) and RAP-NWR. These finding are preliminary and further research would increase our understanding of the underlying processes that link different tasks with each other.

5.7.6 Error Analysis

The errors and distribution of error types presented in this study are consistent with most of the studies that explored error patterns in nonword repetition tasks (e.g.

Edwards and Lahey, 1998; Marton and Schwartz, 2003; Marshall, Harris and Van der Lely, 2003). Most of the errors were consonant/vowel substitutions followed by consonant and vowel deletions then syllable deletions. Morphization (or root/pattern lexicalisation) accounted for less than 2% for each. Less than 1% of total errors were consonant/vowel additions.

Due to the design of the RAP-NWR task in the current study, errors were analysed in terms of root and pattern status, with PCC and PVC calculated separately for RAP-NWR types. The PCC was significantly higher in type 1 nonwords that consisted of real roots vs type 2 and 3 which contained non-root nonwords. There was no significant difference between types 2 and 3, where both types consisted of non-root nonwords. Thus, familiarity of the consonant sequence (root) influenced the accuracy of children's repetition. Likewise, PVC was significantly higher when children repeated items belonging to type 2 (that contained real patterns) when compared to type 3 (that contained non-patterns). On the other hand, PVC in type 2 (that contained real patterns and non-roots) vs type 1 (that consisted of non-patterns and real roots) was not significant. Therefore, it seems that the presence of real roots influenced the accuracy in repeating non-patterns in type 1. The familiarity of the root gave the children more chances to focus on the difficult part which in this case was the nonpattern. Furthermore, the fact that PVC of type 1 was significantly higher than type 3, which consisted of non-roots and nonpatterns, might support this conclusion.

The accuracy of repeating consonants in RAP-NWR test appears to be influenced by their position in the nonwords. Initial consonant (C1) was produced significantly more accurately than C2, C3 and C4. Furthermore, C2 was repeated more accurately than C3 and C4. Results showed the C4 was the most problematic consonant compared with other consonants in different positions. C4 is most vulnerable and this occurs only in the longer items (3 syllable) where it is also the final consonant –both factors may contribute to its vulnerability. This finding supports the previous studies that showed the longer the nonword is the more likely to have more errors. Similar findings were found in vowel analysis; PVC was significantly higher for V1 followed by V2 then V3.

There were no clear pattern errors specially for consonants in the current error analysis, but there were some common error patterns that were observed:

- Substitutions typically involved substitution with a consonant one feature different from the target: /k/ to [t], /s/ to [t], /s/ to [f], /d/ to [b], /d/ to [t], /f/ to [s], /t/ to [k], /n/ to [m].
- A few substitutions occurring more than 10 times were two or three features different from the target: /k/ to [f], /b/ to [k], /t/ to [f], /b/ to [f], /f/ to [b].
- [f] seems to be a popular substitution. This could be due to the high frequency of /f/ in the test. Furthermore, consonant harmony was one of the common errors found in this analysis. Children tended to harmonise consonants, i.e., replacing one consonant in a CVC syllable to make both

consonants the same. However, the current analysis did not allow us to fully investigate the harmonisation.

- Most vowel substitutions were from /u/ or /i/ to [a]. This could be due to children's tendency to harmonise the vowels (make the vowels the same) and/or because /a/ is used in the most common patterns in Gulf Arabic (e.g, (a-a)).
- The majority of vowel and consonant deletions were due to whole syllable deletion.

5.8 Summary

This experiment addressed four main issues, first, the effects of roots and patterns on children's performance in a RAP-NWR test, which involved three types of root and pattern combinations. Secondly, it assessed phonological storage by exploring the effects of nonword length by using two and three syllable nonwords in the RAP-NWR test. Thirdly, it explored the development of the performance of different age groups in the RAP-NWR test. Finally, the relationship between children's performance on the root and pattern NWR and receptive and expressive vocabulary was examined.

The results showed that children were significantly more accurate in repeating Type 1 (root-nonpattern) than Type 3 (nonroot-nonpattern). There was no significant difference between Type 2 (nonroot-pattern) and other types. The presence of the root in the nonword items adds semantic and sublexical cues that influenced the repetition of nonwords. Children were more sensitive to root effects when compared to pattern effects and it seems that root effects may have a more important role to play in NWR than phonological storage because even on two syllable words children performed better on root nonwords than they did on nonroot nonwords. Similarly, in study 1 children with language impairment scored significantly worse on single syllable words and nonwords. Root and pattern effects were explained and distinguished from wordlikeness effects reported in other languages.

These findings were consistent with the phonological processing skills account (e.g., Snowling et al., 2001; Chiat, 2001) and with Hebrew studies (Berent and Shimron, 1997; Clark and Berman, 1984). Finally, correlation analysis showed that there was a strong correlation between the RAP-NWR and the Arabic Expressive Vocabulary Test (AEVT), which is not consistent with the findings of Briscoe et al., (2001), Stokes et al., (2013) and Conti-Ramsden (2001). Moreover, there was a significant correlation between RAP-NWR and Arabic Picture Vocabulary Test (APVT) which is consistent with previous studies in other languages (Gathercole et al., 1991; 1992; Briscoe et al., 2001; Coady & Evans, 2008).

The final chapter of this thesis will summarise findings and theoretical and clinical implications. Directions for future research will also be discussed.

6. Discussion

6.1 Introduction

This chapter will summarise the findings of the two studies, discuss theoretical and clinical implications of the results and suggest directions for future research.

6.2 Summary of findings

Two main studies were conducted. The first one investigated early phonological skills in Gulf Arabic speaking children using a word and nonword repetition test (WNRep) with a TD group (n=44) and a clinical (CL) group (n=15). It also investigated the relationship between the size of receptive and expressive vocabulary and the phonological skills of Gulf Arabic speaking children. An expressive vocabulary test was created by the current researcher to examine the vocabulary skills of TD and children with language impairment.

The second study followed study one to examine the effects of different combinations of Arabic roots and patterns and their interaction with word length. The study involved creating a NWR test (RAP-WNR) that includes different roots and patterns combinations and it was conducted with 89 TD children.

Summaries of the findings of the two studies are below.

6.2.1 Summary of findings of study 1

The first study, presented in chapter four, examined the phonological skills and verbal working memory skills of these children using the Word Nonword Repetition tests (WNRep). The WNRep consisted of a list of real words and a list of nonwords where item length was manipulated to tap into phonological storage. The utility of this test in identifying children at risk of language impairment was evaluated. The real word repetition task in this experiment was designed to control for articulation skills and to investigate its utility as a clinical marker for language impairment along with NWR. All real words selected for this task were familiar nouns for these young children. The results showed that the TD group (*n*=44, mean of age=38 moths) performed significantly better on word and nonword repetition when compared with the CL group (*n* =15, mean of age =43 months). Both word and NWR are potentially useful to differentiate typically developing children from the clinical ones. Furthermore, the TD and CL groups did significantly better on word repetition than they did on nonword repetition, which suggests that both groups were sensitive to lexical familiarity. The interaction between the group type (TD and CL) and the word type (word vs. nonword) was not significant; which shows that both groups were similarly sensitive to the different word types.

The word and nonword repetition items in the study consisted of one, two and three syllables. The results showed that there was a significant effect of word length in TD and CL groups in both word and nonword repetition; the longer the word or the nonword was the more difficult it was to repeat for both groups of children.

The interaction between word type and word length was significant. Accuracy of nonword repetition decreased as the number of syllables increased and the gap between the word vs nonword scores was significant at one syllable word/nonword, two syllable word/nonword and three syllable word nonword.

When comparing the results obtained from using two different scoring methods (the whole word correct (WWC) method and the percentage of phoneme correct (PPC) method, results showed almost identical results for the two methods, however correlations and regression reached higher levels of significance and with good predictive value for receptive and expressive vocabulary performance in the WWC scoring methods. In general, there was a medium to strong correlation between word and nonword repetition tests and the APVT and AEVT in TD children and across different scoring methods. As for the CL group, there was a strong correlation between APVT and AEVT, and between word and nonwords, but there was no significant correlation between either of the two repetition tasks and the two vocabulary tests using either of the two different scoring methods. Word and nonword repetition together succeeded in predicting 25% of the receptive vocabulary size and 46.3% of the expressive skills in the TD group using the WWC scoring method. With the PPC scoring method, the predictive value of nonword repetition only was 11% for receptive vocabulary and 31% for expressive vocabulary. For the CL group the predictive value ranged between 1.5% and 4.6% for both vocabulary tasks and across the two different scoring methods. The difference in correlation and regression results across the two different scoring methods in this study was minimal and did not alter the outcomes. However, it is important to take into consideration that the selection of the scoring method may influence results, though it did not in this current study.

Overall, the performance of the TD and CL groups in the first study showed a similar profile to results seen in other languages ((Botting & Conti-Ramsden, 2001 Conti-Ramsden & Hesketh, 2003; Dollaghan & Campbell, 1998; Ellis Weismer et al., 2000; Gathercole & Baddeley, 1990b; Gray, 2003) as the TD group scored better on word repetition than they did on nonword repetition, however the difference between the performance of the two groups was not significant on the vocabulary task. However, this was possibly due to the inclusion of some students who had scores in the normal range despite being in the clinical group. When these students were excluded, analysis showed stronger correlations between the CL group's performance on the WNRep and receptive and expressive vocabulary. The predictive value of both word and nonword repetition for both vocabulary task was high in TD group but not in the CL, possibly due to the great variability in the performance of the children in the CL group.

While the first study provided preliminary results about the performance of Gulf Arabic speaking children on word and nonword repetition and how it is correlated with their performance in expressive and receptive vocabulary, the second study explored the impact of the typology of the phonological and morphological system in Gulf Arabic represented by root and pattern morphology on children's repetition skills.

6.2.2 Summary of findings of study 2

Results of study 1 showed that children with language impairment (CL group) performed significantly worse than TD children across all word types (words and nonwords) and length types (one syllable, two syllable, and three syllable). Study 1 showed a significant effect of wordlikeness, however Arabic has distinct root and pattern effects that could have influenced the performance of the children. Therefore, investigating the role of roots and patterns in Arabic NWR was needed to understand if this specific characteristic of Arabic would make any difference in how children processed nonword repetition, which was the aim of the second study, presented in chapter five. A root and pattern nonword repetition test (RAP-NWR) was developed to examine if root and pattern knowledge would affect the performance of TD Gulf Arabic speaking children. The RAP-NWR consisted of three different types of root and pattern combinations: the first set of items consisted of real root and nonpattern nonwords, the second set consisted of real pattern and nonroot nonwords, and the third set consisted of nonpattern and nonroot nonwords. Eighty nine TD Gulf Arabic speaking children participated in this experiment and they were between two and seven years old and divided into 6 age bands. Results showed that TD Gulf Arabic speaking children were sensitive to the presence of the root as they performed significantly better in repeating type one which consisted of real root versus type three that contained nonroots. Moreover, there was some indication that these children might be sensitive to the presence of patterns as they performed slightly better on items containing real patterns than they did on the nonpattern items, however the difference was not statistically significant.

The second aim in this study was to explore the effects of nonword length for two versus three syllabic nonword repetition in Gulf Arabic. Results showed that TD children performed significantly better on repeating two syllabic nonwords when compared with the three syllabic nonwords, which replicated the findings of study 1 and showed that phonological storage might be an important consideration when examining performance of TD children and those with language impairment in Arabic, a finding that was reported in many languages.

The third aim of this study was to investigate the age at which TD Gulf Arabic speaking children start to develop their sensitivity to roots and patterns and how they performed across different age groups. Results showed that children were sensitive to root effects starting from the age of 2. The interaction between age and the different nonword type was not significant, which suggested that different age groups performed similarly on different nonword types. As in the first study, the interaction between different age groups and nonword length was significant. The younger the children were, the wider was the gap between their performance on two vs. three syllable nonword repetition.

The last aim was to investigate the relationship between the performance of TD children on the root and pattern NWR and receptive and expressive vocabulary tests. Results showed significant relationship between RAP-NWR and the receptive and expressive vocabulary tests.

Furthermore, the PCC analysis for different RAP-NWR types showed that PCC was significantly higher in type 1 vs type 2 and 3. There was no difference in PCC scores between type 2 vs 3. This can be explained by familiarity of the consonant sequence (root) that influenced the accuracy of children's repetition. Analysis of PVC showed that scores of nonwords belonging to type 2 were significantly better than type 3. This could be attributed to pattern familiarity that influenced repetition. On the other hand, there was no significant difference between Type 1 PVC scores and type 2, though type 1 consisted of nonpattern vs. real patterns in type 2. This could be due to the familiarity of the root, which gave the children more chances to focus on the novel part which in this case was the nonpattern.

Error analysis related to roots and patterns was also investigated in this current study. Results showed that consonant/vowel substitutions followed by deletions were the most common error patterns, which is consistent with most of the studies that looked into error patterns in nonword repetition tests (e.g., Edwards and Lahey (1998); Marshall, Harris and Van der Lely (2003); Marton and Schwartz (2003)). Morphization (or root/pattern lexicalisation) accounted for less than 2% for each and less than 1% of errors were consonant/vowel additions. The accuracy of consonants and vowels appears to be influenced by their position in RAP-NWR items. The initial vowels/ consonants were produced more accurately than the middle and final ones. This finding supports the previous studies that showed the longer the nonword is the more likely it is to have more errors. There was no consistent pattern of errors specially for consonants in the current error analysis, but there were some general error patterns that were observed. Substitutions typically involved substitution with a consonant one, two and three features different from the target consonant. Though the current analysis did not allow full investigation of harmonisation, the children in this analysis tended to harmonise consonants. Moreover, most vowel substitutions were from /u/ or /i/ to [a], possibly due to the fact that the most common pattern in Gulf Arabic is (a-a).

6.3 Theoretical Interpretations and Implications

This section will discuss the current studies' findings in light of nonword repetition theories, taking into consideration the implication of the performance on word and nonword repetition, nonword repetition in relation to receptive and expressive vocabulary, and the effect of root and pattern structure on the processes underlying nonword repetition.

6.3.1 Implications regarding word and nonword repetition

Processes involved in nonword repetition. The nonword repetition task in the WNRep test was designed to measure effects of phonological short term memory (Gathercole and Baddeley, 1990a) by increasing the length of words and nonwords from one, to two and three syllables, while also looking at effects of wordlikeness (word vs nonwords). This design was used to examine if the phonological loop as presented in the original working memory model of Baddeley and Hitch (1974) and the revised model by Baddeley (2003) is responsible for providing temporary storage of unfamiliar phonological forms as well as sorting and processing sound combinations. Therefore, any impairment in this part of working memory will cause a deficit in the phonological representation, thus affecting the process of learning new words. The results of this study were partially consistent with the PSTM theory as the longer the word or nonword was the more difficult it was for children with language impairment to repeat and children who performed poorly on nonword repetition had lower vocabulary sizes when compared with other children who did well on nonword repetition. However, the fact that children with language impairment performed significantly worse than TD children even on one and two syllable nonwords does not seem to be compatible with PSTM as most studies in English showed that differences in performance on NWR usually start at three syllables and above. This poor performance on two and even on one syllable nonwords indicates that there might be other more important considerations or processes involved in NWR in Arabic, such as effects of roots and patterns, which are closely related to wordlikeness effects. It is noted that Shaalan (2010) found similar results in that children with SLI performed significantly less accurate on NWR on two syllable nonwords. This latter interpretation is more compatible with phonological processing account (Snowling & Chiat, 2001).

The results of real word repetition in this current study raise questions regarding the processes underlying the repetition skills based on the working memory model. For example, whether the phonological loop in the working memory model is mainly responsible for providing temporary storage of unfamiliar phonological forms only, or whether deficits in the phonological loop only cause a deficit it learning new words or there are possibly other deficits in different components of the model. The working memory model failed to answer the above questions, especially when we consider the real word repetition. The model also failed to explain how the semantic, syntactic and lexical components interact with NWR, or how the short term memory is linked to long term memory. It is not possible to counter or explain the underlying processes involved in nonword repetition in one model like the working memory model, because the nonword repetition test was found to be a flexible measure that was designed to measure different language aspects (e.g., phonology, morphology) by manipulating the stimuli in order to investigate these language aspects.

Word and nonword repetition as clinical markers: The nonword repetition task in the first study succeeded in distinguishing children with language impairment from typically developing ones; the WNRep was sensitive to children's repetition skills, which in turn is related to their language skills as shown in many studies (e.g., Botting & Conti-Ramsden, 2001; Conti-Ramsden & Hesketh, 2003; Dollaghan & Campbell, 1998; Ellis Weismer et al., 2000; Gathercole & Baddeley, 1990b; Gray, 2003). Therefore, WNRep might be a potential clinical marker for language impairment in Gulf Arabic. This finding however, cannot be generalized, as the CL group in the current study may not be representative of all children at risk of language impairment in Arabic. Most of the time, and due to the lack of standardised assessment tools in Arabic, only children with moderate to severe language impairment are referred to assessment or discovered by their parents to have language deficits. Therefore, it is recommended that the utility of the WNRep test in identifying children at risk of language impairment should warrant further investigation with a larger sample of children. It is recommended to investigate the diagnostic accuracy of the WNRep in a population study that includes a full range of ages. The preliminary results of WNRep might be consistent with the previous studies that investigated NWR as a clinical marker for children with language impairment, however it is still not clear what are the factors that make some nonword tests succeed in discriminating children with language impairment better than others. Some factors might affect the utility of nonword repetition as a clinical marker, such as test characteristics and design (e.g., wordlikeness, nonword length and nonword complexity, numbers of participants, age and type of language impairment). For example, Graf Estes et al. (2007) reported that the CNRep test (Gathercole & Baddeley, 1996) was more successful in differentiating children with SLI from TD children when it was compared to another NWR test, namely the NRT (Dollaghan & Compbell, 1998).

On the other hand, the real word repetition task in the WNRep showed similar results to those found in NWR although the items in this task were familiar and real words. Just like nonword repetition, it seems that real word repetition could be a potential clinical marker in children with language impairment in Gulf Arabic as these children scored significantly lower than the typically developing children on real word repetition. Similarly, the longer the real word was, the less accurately it was repeated by both groups. The finding that real word repetition task in this current study was found to be a potential clinical marker may support the findings reported by Dispaldro et al. (2013) who argued that real word repetition can potentially be a useful diagnostic assessment tool to distinguish Italian TD children from children with SLI, despite some fundamental differences in the two tasks. Dispaldro et al. (2013) argued that the main skills required for word repetition differ from those required for nonword repetition. Real word repetition relies on existing phonological and semantic knowledge accumulated from learning previous vocabulary, while nonword repetition relies mostly on phonological short term memory (PSTM). However, both tasks share some skills, such as speech perception, oral-motor skills, use of PSTM and lexical and linguistic knowledge. These shared skills probably allowed real word repetition to succeed as a diagnostic tool that distinguished the CL group from the TD group in Gulf Arabic speaking children. The results of real word repetition could also be attributed to typological differences in Arabic. Arabic is a root and pattern language, and to form new words children may use their knowledge about root derivation, where one root is manipulated extensively to produce various items that are semantically related to the root. In Arabic new words are strongly related to children's lexical knowledge, this may explain why real word repetition could be effectively used as a clinical marker that could predict vocabulary knowledge.

Relations with receptive and expressive vocabulary. The significant correlation between the Arabic Picture Vocabulary Test (APVT) and nonword repetition reported in typically developing children in this current study is consistent with many studies (Briscoe et al., 2001; Coady & Evans, 2008; Gathercole et al., 1991; 1992). However, some other studies found no correlation between expressive vocabulary and nonword repetition (e.g. Briscoe et al., 2001; Conti-Ramsden, 2001; Stokes et al., 2013) unlike the current study, which found a strong correlation between WNRep test and expressive vocabulary.

The CL group showed weak correlations between performance on word and nonword repetition and their scores on the APVT and AEVT. When reviewing the CL children's profiles there were three children from the CL group who performed in line with average scores of the TD children on receptive and expressive vocabulary. These three children's performances on the WNRep were higher than the average score of the CL group but still lower than the TD children's average score. The three children met the criteria for children with language impairment. One possible explanation was that these children received speech therapy for a longer time (three to six months of therapy) compared to other CL participants. Another possible explanation was that these children had scores within the normal range on receptive language skills but not on expressive language, which allowed them to perform within normal limits on the receptive vocabulary test. Furthermore, the scoring method used in the expressive test allowed articulation errors since the word was still intelligible for the examiner, which might have helped these children to perform within the normal range on the expressive test, but not the WNRep as the scoring criteria were stricter and they penalised articulation errors. When the correlation analysis was conducted again excluding the three children, results showed significant correlations between word and nonword repetition and the APVT and the AEVT.

The small clinical (CL) group sample could also be a reason why the correlation between the repetition tasks and the vocabulary tasks was not significant. Having a larger sample may help overcome the individual differences among participants, especially that the performance of each child with language impairment may vary across different tests. Though the criteria of selecting the CL group in this study was strict and similar to the criteria used in many studies (e.g., Chiat & Roy, 2007; Shaalan, 2010), the duration of speech therapy received was not considered in the criteria.

6.3.2 Implications regarding roots and patterns effects

The second study in this project was conducted to examine the effects of roots and patterns on repetition skills in typically developing Gulf Arabic speaking children. There are no studies that have examined the effect of roots and patterns on NWR. Most of the studies that examined nonword repetition in root and pattern languages (e.g., Arabic and Hebrew) did not investigate directly effects of roots and patterns on NWR. Results showed that TD children performed significantly better on repeating items that consisted of real roots versus nonroot items. However, these children were less sensitive to the presence of real patterns versus nonpattern items. It is argued that there are two possible accounts for these effects of roots and patterns, namely the linguistic or phonological processing account and the root and pattern account.

The linguistic account or the phonological processing account (Snowling et al. 1991, Chiat, 2001) argues that previous lexical knowledge supports better nonword repetition; so familiar sublexical units that are implanted within nonwords increase accuracy (Gathercole et al., 1991). Nonword repetition is also improved when nonwords include either syllables that are themselves lexical items (Dollaghan, Biber & Campbell, 1993) or segments with high phonotactic frequencies (Munson, 2001).

The effects of the root in nonword repetition in the current study were significant; the participants were more accurate in repeating nonwords that contained real roots vs nonroot nonwords. However, the effect of the root might be confounded with wordlikeness effects. The main difference between wordlikeness effects (found in some English NWR test, such as CNRep, Gathercole & Baddely,1996) and root and pattern effects (as found in the current RAP-NWR test) is that the root in Arabic is a non-contiguous morpheme consisting typically of three consonants that are intertwined with a vocalic pattern. Therefore, it is not present as a word consisting of contiguous sounds like the example mentioned earlier in CNRep ('under' as in "*under*brantuand" or a real morpheme (such as 'ing' as in "blonterstap*in"*. However, despite this nonlinear property of roots, participants showed awareness of the effects of internal structure of the roots when they were embedded in nonwords.

On the other hand, the linguistic account failed to explain why children were not sensitive to real pattern nonwords versus the nonpattern items, though the real pattern used in this study were very common patterns (e.g. (a-a) for two syllabic nonwords and (a-a-a) for three syllabic nonwords). However, age could have played an important role for lack of pattern effects in NWR in this study, because Arabic and Hebrew speaking children do not develop pattern awareness at this age (see Malenky, 1997).

The second account that might explain effects of roots and patterns is the “root and pattern account” of Semitic morphology (Berent & Shimron, 1997; Plaut & Gonnerman, 2000). This account claims that root-based languages have different morphological processing when they are compared to English. The root and pattern knowledge is accumulated with age and so according to Berman (2003) Hebrew speaking children as young as three years old can recognize roots in novel nonwords, and by the age of seven to ten years old they can use roots to derive new word categories. As these children grow older, the root and pattern knowledge will be used consistently and correctly. Malenkey (1997) and Karwa & Sakran (1997) claimed that Hebrew and Palestinian Arabic speaking children’s awareness of the root started at kindergarten, while pattern awareness did not start until the age of ten. Findings of this current study support the root and pattern account. Gulf Arabic speaking children aged from two to seven years old were sensitive to roots more than they were to patterns. However, when it comes to patterns, children performed better on repeating real patterns versus the nonpattern items, but the difference was not significant. So, it seems that pattern knowledge is yet to be mastered at this young age.

6.3.3 Implication for word formation in Semitic languages

The results of the RAP-NWR in the light of the root morphology supports Kramer’s (2007) and Tucker’s (2009; 2010) view that the root is an essential morpheme in word processing in Semitic languages. Children in this study performed differently on the three different types of the RAP-NWR test depending on the presence of the root, which suggests that roots are processed in a different level. The current results, on the other hand, do not support the whole word processing account by Butterworth (1983), who claimed that the word is processed in the mental lexicon as a complete phonological entity (see section 3.4.2). Furthermore, the current results do not support Taft and Forster (1975) who argued for the decomposed access models (morphological processing route) where words are processed by isolation of the morphemes that compose the word. Participants' scores in this current study varied for each nonword type which means they processed each type differently. The mixed or the dual route model by Caramazza et. al (1988) could explain the current results better than the whole word processing model or the decomposed model. The dual route model required an activation for both models, where the familiar words or nonwords that consist of familiar morphemes (root, affixes) will be processed through the whole word processing model, while the complex and novel words are processed through the decomposed access models. The current findings showed that participants were better at repeating nonwords that consisted of real roots (e.g. Type 1 in RAP-NWR) while they scored lower for the nonwords that did not consist of any familiar morphemes (e.g., Type 3 in RAP-NW). Baayen and Schreude (1995; 1999) added another layer to the dual rout model of word analysis to help in the decision of selecting the appropriate route for word processing. This layer is the sensitivity to the frequency and familiarity of the word and morpheme. Participants showed a significant sensitivity and familiarity to root vs nonroot nonwords but they were not significantly sensitive to pattern vs nonpattern nonwords however they scored better on Type 2 vs Type 3. Our finding can be explained by Burani and Caramazza (1987) and Burani and Baayen (2002) results whose their studies showed that access times and accuracy to suffixes and accuracy to suffixed derived words were significantly related to root familiarity. The lexical decision of pseudowords were more accurate and faster when the low frequency suffixes were combined with high frequency roots. In our case Type 2 in RAP-NWR consisted of nonroots combined with high frequency patterns, therefore the current findings could support Burani and Caramazza (1987) and Burani and Baayen (2002). However, the main difference between the findings is in the stimuli used. In Arabic, the root is a non-contiguous morpheme and the pattern is a string of vowels that can be added in any position, not only suffixes that attach to the roots as in Burani and Caramazza (1987).

6.4 Clinical Implications

This thesis has some clinical implications that could help in assessing and identifying children with language impairment, especially in the domain of word and nonword repetition and vocabulary skills. This study also evaluates the utility and efficiency of using different scoring methods for NWR and the relationship between the two NWR tasks and expressive and receptive vocabulary tests in Gulf Arabic. Examination of the development of phonological skills as represented by NWR across different ages in Gulf Arabic shows that they are in line with those reported in other languages.

The NWR tests as a clinical tool in Gulf Arabic. One of the main aims of this thesis was to develop speech and language assessment material that can help clinicians and researchers working with Gulf Arabic speaking children identify children at risk of language impairment. The tests developed for this project were the Word and Nonword Repetition Test (WNRep), the Root and Pattern Nonword Repetition Test (RAP-NWR), and the Arabic Expressive Vocabulary Test (AEVT). Results showed that the WNRep test differentiated between TD children and children with language impairment and scores were correlated with expressive and receptive vocabulary tests. The WNRep test that was conducted in the first study with children with typical and atypical language development succeeded in discriminating between these two groups. The test was conducted with children aged from two to four years old, which is a critical age for identifying children at risk of language impairment. The WNRep is a short and easy assessment tool that can be used with young children with less effort and time, especially when it is compared to other speech and language assessment tests. Furthermore, the WNRep could be useful as a screening tool that can be used in schools and kindergartens by teachers. The WNRep test consists of two tasks. The first one is real word repetition, which could help children to move smoothly to the second task, which is nonword repetition. Moreover, the WNRep showed good predictive values for receptive and expressive vocabulary for the TD group and it is expected to have similar predictive value for the CL if it is conducted with larger group. However, it is important that these tests are used with a larger population to assess their utility and more measures are needed to examine the reliability and validity of these tests.

Furthermore, one should be cautious when interpreting these results and more investigation is required to replicate these findings and examine the utility of this test. One reason for this is that the children included in the clinical group (CL group) are not representative of all children at risk of language impairment. These children tend to have severe difficulties and their language deficits are more observable by parents and clinicians. However, we know that children with language impairment have varying levels of difficulties and some may even perform within normal range on some tests of nonword repetition. Some studies that examined larger numbers of children (some of whom met the criteria of SLI) found that some of these children performed within normal range on nonword repetition and some of the TD children were found to do poorly on nonword repetition (see Ellis Weismer et al., 2000; Gathercole, 2006 for a review and discussion of this). Therefore, while NWR is a good tool to identify children at risk of SLI or language impairment, we cannot conclude that it is sufficient to rule language impairment in our out. However, it is one of the tools that should be combined with other assessment tools.

The RAP-NWR test used in the second study was efficient in measuring the children's root and pattern sensitivity and predicting vocabulary skills. However, the utility of the RAP-NWR test as a diagnostic assessment tool that can distinguish children with language impairment from others was not explored in this study as all participants in the RAP-NWR study were TD children. However, its correlation with receptive and expressive vocabulary test and its sensitivity to performance based on root and pattern combinations could be further extended in children with language impairment in the future.

Scoring methods. Two scoring methods were used in this study, namely the whole word correct method (WWC) and the percentage of phonemes correct (PPC). The aim of using these two different scoring methods was to investigate if this would make any difference in analyzing the output of the word and NWR tests. Results showed there was no significant contribution of the scoring methods employed to the groups' results on the word and NWR test. Therefore, it is recommended in clinical settings to use WWC scoring method as it is easier to apply and less time consuming. On the other hand, using the PPC scoring method could be useful for research purposes as it is more informative about details of children's response to nonword repetition.

Correlation between performance on root and pattern nonwords and receptive and expressive vocabulary tests. This project involved developing an expressive vocabulary test, namely the Arabic Expressive Vocabulary Test (AEVT). The AEVT showed some potentially good indicators of appropriate psychometric properties in both studies, as children with language impairment performed significantly lower than the TD group on AEVT. This test can be used with children up to 9;11 years old, but in both studies it was used only with children aged between two to five years old. Therefore, it is recommended to include older children in later uses of this test to investigate their performance on the AEVT. Furthermore, a larger sample is needed to examine the reliability and validity of this assessment tool. The strong correlation between the RAP-NWR test and APVT and AEVT demonstrates that the RAP-NWR could be potentially used as a successful tool to predict the receptive and expressive vocabulary in Gulf Arabic speaking children. The three different types of RAP-NWR item were also significantly correlated with each other. Furthermore, each type correlated significantly with the receptive and expressive vocabulary skills. The significant correlation between RAP-NWR and receptive vocabulary is supported with many studies in the literature unlike expressive vocabulary (Briscoe et al., 2001; Coady & Evans, 2008; Gathercole et al., 1991; 1992).

Developmental trends in phonological skills. The results in this study showed that the developmental nature of the phonological skills of Gulf-Arabic speaking children is in line with other studies of the development of phonological skills in other languages. The gradual chronological development of nonword repetition skill can be explained by several factors. First, vocabulary growth helps in improving repetition skills (Masoura

& Gathercole, 2005; Metsala & Walley, 1998; Munson, Edwards, & Beckman, 2005). Second, articulation and oral motor skills improve with age (Kovas et al., 2005; Krishnan et al., 2013). The third factor is that phonological short term memory develops with age (Gathercole et al., 1992). And fourth, the awareness of root and pattern knowledge improves with age (Karwa & Sakran, 1997; Malenkey, 1997).

6.5 Limitations and directions for future research

Findings of this current study are limited by some factors. When developing the stimuli for nonword repetition in the first study, there were a few items that violated the OCP-place phonotactic constraint in Arabic. However, the results were analysed with the violated items excluded and there were no differences in the results with or without those items (See Appendix H for analysis after excluding these items). This finding may add further evidence to Frisch and Zawaydeh's (2001) study that gave some evidence that violation of OCP-Place in Arabic is allowed in certain words.

Another limitation in the first study was the selection of the clinical group. Though the criteria for selecting the CL group in this study were similar to the criteria used in other studies (e.g., Chiat & Roy 2007; Shaalan, 2010) the duration of speech therapy received was not considered in these criteria. This might have caused children who received speech and language therapy to perform as well as typically developing children. Reanalysis of the results after excluding the CL children with high performance showed stronger correlations between the CL group's performance on the WNRep and receptive and expressive vocabulary. Therefore, future studies should try to control for the effects of therapy in clinical groups.

The small CL group size could be another limitation in this study. In this study we had only 15 children in the CL group and it is always better to generalize the findings with bigger samples, especially since some of the children in the CL received speech therapy services while others did not. However, this could be extended in the future by administering these tasks with more participants belonging to different age groups.

The findings regarding word and nonword repetition and vocabulary in the first study raised several issues that warrant further investigation. Furthermore, replicating this study with larger samples from both TD and CL group and with older and younger children would help to understand and generalize these findings and add more reliability to WNRep as an assessment tool. Lack of the validity tests in the first study and reliability and validity tests in the second study were one of the main limitation in this project.

The second study showed that the RAP-NWR test, which was used with TD children only, was a useful predictor for typically developing children's vocabulary size. Therefore, it would be valuable to investigate the utility of the RAP-NWR test as a diagnostic tool to distinguish children with language impairment from others. The RAP-NWR test revealed that Gulf Arabic speaking children were sensitive to roots and patterns from very young age, however, it is not clear how Gulf Arabic speaking children process their root and pattern knowledge in order to form novel words. Further investigation is required to explore this part using different tasks, such as analogue tasks as in the studies by Karawa and Sakran (1997) and Malenkey (1997) where children were asked to develop new words based on familiar roots. Furthermore, it is important to investigate the link between root and pattern repetition and grammatical abilities (e.g., morphosyntactic tasks, understanding of grammatical structures) to assess the contribution of these structures.

Moreover, the current findings regarding stronger effects for the root could be investigated further by comparing the effects of specific phonotactic constraints on co-occurrence of root consonants (OCP-Place). This could be explored by manipulating NWR stimuli based on violation of OCP-Place and assess whether children of different ages and abilities are sensitive to this rule. Moreover, wordlikeness effects can be manipulated in the same NWR task (low vs high wordlike) in order to compare effects of wordlikeness with root effects. Investigating the effect of harmonization is also important in order to explore in depth children's error patterns.

Finally, further investigation is needed to understand the role of root and pattern knowledge in language development and its role in NWR repetition in both TD children and children with language impairment in Arabic in particular and Semitic languages in general. Both this study and that of Shaalan (2010) showed that children with language impairments have difficulties repeating nonwords (with varying levels of root and pattern combinations, especially when containing nonroots (both studies) or consonant clusters (Shaalan, 2010). Even in one syllable nonwords, as shown in the first study, children with language deficits performed significantly worse than their TD peers, which indicates that effects of roots and patterns in particular, and phonological skills in general may have stronger contributions to NWR when compared with phonological short term memory. Therefore, conducting a test like the RAP-NWR with children with language impairment may help to shed more lights on the role played by patterns and effects in NWR in children with language impairment.

6.6 Conclusion

Investigations of word and nonword repetition skills in TD children and children with language impairment in Gulf Arabic revealed the following findings. The first study showed that the more the repetition items contained lexical and sub lexical syllables, the easier it was for the children from both TD and CL groups to repeat; which supported a phonological processing account (Snowling et al., 1991) of nonword repetition, especially with the finding that children with language impairment scored significantly less on NWR even on single syllable words and nonwords. The first study showed an interaction between word length and group, a finding that has been reported in many studies where children with language impairment performed worse than their age matched peers on word and nonword repetition. However, while the first study showed effects of both word length (PSTM) and phonological processing skills (wordlikeness), the second study which was conducted with TD children tried to examine effects of roots and patterns and word length. The results of the Root and Pattern NWR test showed that although there was a length effect as children's nonword repetition across ages decreased as word length increased, root and pattern effects were also strong as children's performance was differentiated across all ages even on two syllable nonwords. These results of both experiments strongly support the importance of considering effects of roots and patters, which could be considered a type of wordlikeness effects, when designing NWR tests in Semitic languages and more studies are needed to tease apart word length effects from root and pattern effects. It is expected that such NWR tests will be of great utility in clinical practice as well as in examining the different theories of NWR.

Finally, all word and nonword repetition tasks in this study correlated significantly with receptive and expressive vocabulary test. This supports the validity of these tools to be used as a screening or assessment tools for children with language impairment in Gulf Arabic. However, replication of the results of NWR tests and vocabulary tests and the relation among them with a larger number of participants is needed to confirm these preliminary findings.

References

Al-Akeel, A. (1998). *The acquisition of Arabic language comprehension by Saudi children*. Unpublished Ph.D. thesis, University of Newcastle upon Tyne, UK.

Aljenaie, K. (2001). *The emergence of tense and agreement in Kuwaiti Arabic children.* Unpublished Ph.D dissertation, University of Reading.

Alqattan, S. (2015). *Early Phonological Acquisition by Kuwaiti Arabic Children*. Unpublished Ph.D thesis, University of Newcastle, UK.

Amayreh, M. M. (2003). Completion of the consonant inventory of Arabic. *Journal of Speech, Language, and Hearing Research*, 46, 517-529.

Amayreh, M. M., & Dyson, A. T. (2000). Phonetic inventories of young Arabic-speaking children. *Clinical Linguistics & Phonetics*, *14*, 193-215.

Archibald, L., & Gathercole, S. (2006). Nonword repetition: A comparison of tests. *Journal of Speech, Language, and Hearing Research*, *49*, 970-983.

Ayyad, H. (2011). *Phonological development of typically developing Kuwaiti-Arabic speaking pre-schoolers*. PhD dissertation, University of British Columbian, Canada.

Baayen, H., & Schreuder, R. (1999). War and peace: Morphemes and full forms in a noninteractive activation parallel dual-route model. *Brain and language, 68*, 27-32.

Baddeley, A. D. (2003). Working memory and language: An overview. *Journal of Communication Disorders, 36*, 189-208.

Baddeley, A. D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*, 158-173.

Baddeley, A. D., & Hitch, G. (1974). Working memory. *The Psychology of Learning and Motivation, 8*, 47-89.

Balota, D. A., Pilotti, M., & Cortese, M. J. (2001). Subjective frequency estimates for 2,938 monosyllabic words. *Memory & Cognition, 29*, 639-647.

Bat-El, O. (1994). Stem modification and cluster transfer in Modern Hebrew. *Natural Language & Linguistic Theory, 12*, 571-596.

Bat-El, O. (2003). The fate of the consonantal root and the binyan in Optimality Theory. *Recherches linguistiques de Vincennes, 32*, 31-60.

Béland, R., & Mimouni, Z. (2001). Deep dyslexia in the two languages of an Arabic/French bilingual patient. *Cognition, 82*, 77-126.

Benua, L. (2000). *Phonological relations between words*. New York: Garland Publishing.

Berent, I., & Shimron, J. (1997). The representation of Hebrew words: Evidence from the obligatory contour principle. *Cognition, 64*, 39-72.

Berman, R. A. (1985). The Acquisition of Hebrew. *The Crosslinguistic Study of Language Acquisition*, vol. 1, ed. D. I. Slobin, 255-371. Hillsdale, NJ: Erlbaum

Berman, R. A. (1999). Children's innovative verbs versus nouns: Structured elicitations and spontaneous coinage. *Methods in Studying Language Production*, 69-93.

Berman, R. A. (2003). Children's lexical innovations. *Language Acquisition and Language Disorders, 28*, 243-292.

Berman, R.A. & D. Ravid (2000). Acquisition of Israeli Hebrew and Palestinian Arabic: A review of current research. *Hebrew Studies*, 41, 7-22.

Bishop, D. V. M. (1997). *Uncommon understanding: Comprehension in specific language impairment*. Hove: Psychology Press.

Bishop, D. V. (2008). Specific language impairment, dyslexia, and autism: Using genetics to unravel their relationship. In C.F. Norbury, J.B. Tomblin and D.V.M. Bishop (Eds.), *Understanding Developmental Language Disorders: From Theory to Practice*. London: Psychological Press.

Bishop, D. V. M. (2014). Ten questions about terminology for children with unexplained language problems. *International Journal of Language and Communication Disorders*, *49*, 381–415.

Bishop, D. V. M., Adams, C., & Norbury, C. F. (2006). Distinct genetic influences on grammar and phonological short-term memory deficits: Evidence from 6-year-old twins. *Genes, Brain and Behaviour*, *5*, 158-169.

Bishop, D.V.M., North, T., & Donlan, C. (1996). Nonword repetition as a behavioural marker for inherited language impairment: Evidence from a twin study. *Journal of Child Psychology and Psychiatry, 37, 391–403.*

Bishop, D. V. M., Snowling, M.J., Thomspon, P.A., Greenhalgh, T., & The CATALISE Consortium. (2016). CATALISE: A multinational and multidisciplinary Delphi consensus study: identifying language impairment in children. *PLOS One, 11*, 7.

Botting, N., & Conti-Ramsden, G. (2001). Nonword repetition and language development in children with specific language impairment (SLI). *International Journal of Language and Communication Disorders, 36*, 421-432.

Bowey, J. A. (1996). On the association between phonological memory and receptive vocabulary in five-year-olds. *Journal of Experimental Child Psychology, 63,* 44-78

Bowey, J. A. (2001). Nonword repetition and young children's receptive vocabulary: A longitudinal study. *Applied Psycholinguistics, 22*, 441-469.

Briscoe, J. M., Bishop, D. V. M. & Norbury, C. (2001). Phonological processing, language, and literacy: a comparison of children with mild-to-moderate sensorineural hearing loss and those with specific language impairment. *Journal of Child Psychology and Psychiatry, 4*2, 329- 340.

Bukshaisha, F. A. (1985). *An experimental phonetic study of some aspects of Qatari Arabic*. Doctoral dissertation, University of Edinburgh.

Burani, C., & Caramazza, A. (1987). Representation and processing of derived words. *Language and Cognitive Processes, 2*, 217-227.

Burani, C., Dovetto, F. M., Thornton, A. M., & Laudanna, A. (1997). Accessing and naming suffixed pseudo-words. In: Booij G., van Marle J. (eds) *Yearbook of morphology* 1996. Springer, Dordrecht.

Butterworth, B. (1983). *Lexical representation in language production II: development, writing and Other Language Processes*. London: Academic Press.

Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition, 28*, 297-332.

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

Chiat, S., & Roy, P. (2007). The Preschool Repetition Test: An evaluation of performance in typically developing and clinically referred children. *Journal of Speech, Language and Hearing Research*, *50*, 429-443.

Clark, E. V., & Berman, R. A. (1984). Structure and use in the acquisition of word formation. *Language, 60,* 542-590.

Coady, J. A., & Evans, J. L. (2008). Uses and interpretations of non-word repetition tasks in children with and without specific language impairments (SLI). *International Journal of Language & Communication Disorders, 43,* 1-40.

Comblain, A. (1999). The relevance of a nonword repetition task to assess phonological short-term memory in individuals with Down syndrome. *Down Syndrome Research and Practice, 6,* 76-84.

Conti-Ramsden, G. (2003). Processing and linguistic markers in young children with specific language impairment (SLI). *Journal of Speech, Language, and Hearing Research*, *46*, 1029-1037.

Conti-Ramsden, G. (2014). What should we call children who struggle to talk? Taking a developmental, global perspective of diagnostic labels—reflections on Bishop (2014). Commentary on Bishop, D. V. M., Ten questions on terminology for children with unexplained language problems. *International Journal of Language and Communication Disorders, 49*, 381–415.

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

Conti-Ramsden, G., & Hesketh, A. (2003). Risk markers for SLI: a study of young language-learning children. *International Journal of Language & Communication Disorders, 38,* 251-263.

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in Brain Research, 169,* 323-38

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

Deevy, P., Weil, L. W., Leonard, L. B., & Goffman, L. (2010). Extending use of the NRT to preschool-age children with and without specific language impairment. *Language, Speech, and Hearing Services in Schools, 41*, 277-288.

Deutsch, A., & Frost, R., Pelleg, Pollatsek, & Pollatsek, & Rayner, K. (2003). Early morphological reading: Evidence from parafoveal previous benefit in Hebrew. *Psychologic Bulletin & Review*, *10*, 415-422.

Deutsch, A., Frost, R., & Forster, K. I. (1998). Verbs and nouns are organized and accessed differently in the mental lexicon: evidence from Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1238.

DeVellis, R. F. (2012). *Scale development: Theory and applications*. Los Angeles: Sage Publications.

Dispaldro, M., Benelli, B., Marcolini, S., & Stella, G. (2009). Real-word repetition as a predictor of grammatical competence in Italian children with typical language development. *International Journal of Language & Communication Disorders*, *44*, 941-961.

Dispaldro, M., Deevy, P., Altoé, G., Benelli, B., & Leonard, L. B. (2011). A cross-linguistic study of real-word and non-word repetition as predictors of grammatical competence in children with typical language development. *International Journal of Language & Communication Disorders*, *46*, 564-578.

Dispaldro, M., Leonard, L. B., & Deevy, P. (2013). Real-word and nonword repetition in Italian-speaking children with specific language impairment: A study of diagnostic accuracy. *Journal of Speech, Language, and Hearing Research, 56,* 323-336.

Dollaghan, C., Biber, M & Campbell, T. (1993). Constituent syllable effects in a nonsense-word repetition task. *Journal of Speech and Hearing Research, 36*, 1051-1054

Dollaghan, C. A., Biber, M. E., & Campbell, T. F. (1995). Lexical influences on nonword repetition. *Applied Psycholinguistics, 16*, 211-222.

Dollaghan, C., & Campbell, T. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research*, *41*, 1136-1146.

Dunn, L. M., & Dunn, L. M. (1997). *PPVT-III: Peabody picture vocabulary test*. Circle Pines, MN: American Guidance Service.

Ebbels, S. (2014). Introducing the SLI debate. *International Journal of Language and Communication Disorders*, *49,* 337–380.

Edwards, J., Beckman, M. E., & Munson, B. (2004). Frequency effects in phonological acquisition. *Journal of Child Language, 42*, 306-311

Edwards, J., Beckman, M. E., & Munson, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of speech, language, and hearing research, 47,* 421-436

Edwards, J., & Lahey, M. (1998). Nonword repetitions of children with specific language impairment: Exploration of some explanations for their inaccuracies. *Applied Psycholinguistics*, *19*, 279-309.

Ellis Weismer, S., Tomblin, J. B., Zhang, X., Buckwalter, P., Chynoweth, J. G., & Jones, M. (2000). Nonword repetition performance in school-age children with and without language impairment. *Journal of Speech, Language, and Hearing Research, 43*, 865–878.

Feghali, H. J. (2008). *Gulf Arabic: The Dialects of Kuwait, Bahrain, Qatar, UAE, and Oman: Grammar, Dialogues, and Lexicon*. London: Dunwoody Press.

Fenson, L., Dale, P.S., Reznick, J.S., Thal, D., Bates, E., Hartung, J.P., Pethick, S., & Reilly, J.S. (1993). *The MacArthur Communicative Development Inventories: User's Guide and Technical Manual*. Baltimore : Paul H. Brokes Publishing Co.

Ferguson, C. (1959). Diglossia. *Word*, *15*, 325-340.

Frisch, S.A, Large, N.R, Pison, D.B. (2000). Perception of wordlikeness: Effects of segment probability and length on the processing of nonwords. *Journal of Memory and Language, 42*, 481-496.

Gallagher, A. (2014). What's in a name? Some thoughts on Reilly *et al*. (2014). Commentary on Reilly, S., Tomblin, B., Law, J., McKean, C., Mensah, F. K., Morgan, A., Goldfeld, S., Nicholson, J. M. and Wake, M., 2014, Specific language impairment: a convenient label for whom? *International Journal of Language and Communication Disorders, 49,* 416–451

Gathercole, S. E. (1995). Is nonword repetition a test of phonological memory or long-term knowledge? It all depends on the nonwords. *Memory and Cognition, 23*, 83-94.

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics, 27*, 513-543.

Gathercole, S., & Adams, A. (1993). Phonological working memory in very young children. *Developmental Psychology, 29*, 770–778.

Gathercole, S. E., & Adams, A. M. (1994). Children′ s phonological working memory: Contributions of long-term knowledge and rehearsal. *Journal of Memory and Language, 33*, 672-688.

Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language, 28*, 200-213.

Gathercole, S. E. & Baddeley, A. D. (1990a). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language, 29*, 336-360.

Gathercole, S. E., & Baddeley, A. D. (1990b). The role of phonological memory in vocabulary acquisition: A study of young children learning new names. *British Journal of Psychology, 81*, 439-454.

Gathercole, S. E., & Baddeley, A. D. (1993). Phonological working memory: A critical building block for reading development and vocabulary acquisition? *European Journal of Psychology of Education, 8*, 259-272.

Gathercole, S. E. & Baddeley, A. D. (1994). *Working memory and language*. Mahwah, NJ: Lawrence Erlbaum Associates.

Gathercole, S. E., & Baddeley, A. D. (1996). *The Children's Test of Nonword Repetition*. London: The Psychological Corporation UK.

Gathercole, S.E.; Hitch, G..J.; Service, E; & Martin, A. J. (1997). Phonological short-term memory and new word learning in children. *Developmental Psychology, 33,* 966-979.

Gathercole, S. E., Service, E., Hitch, G., Adams, A.-M., & Martin, A. (1999). Phonological short-term memory and vocabulary development: Further evidence on the nature of the relationship. *Applied Cognitive Psychology, 13*, 65-77.

Gathercole, S. E., Tiffany, C., Briscoe, J., Thorn, A., & The ALSPAC Team. (2005). Developmental consequences of poor phonological short-term memory function in childhood: A longitudinal study. *Journal of Child Psychology and Psychiatry, 46*, 598-611.

Gathercole, S. E., Willis, C., & Baddeley, A. D. (1991). Nonword repetition, phonological memory, and vocabulary- a reply to Snowling, Chiat and Hulme. *Applied Psycholinguistics, 12*, 375-379.

Gathercole, S. E., Willis, C., Emslie, H., & Baddeley, A. D. (1991). The influences of number of syllables and wordlikeness on children's repetition of nonwords. *Applied Psycholinguistics, 12*, 349-367.

Gathercole, S.E. , Willis, C., Baddeley, A. D. & Emslie, H. (1994). The children's test of nonword repetition: a test of phonological working memory. *Memory, 2,* 103-127.

Gathercole, S. E., Willis, C., Emslie, H., & Baddeley, A. D. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental Psychology, 28*, 887-898.

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General, 113*, 256-281.

Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behaviour Research Methods & Instrumentation, 12,* 395-427.

Girbau, D., & Schwartz, R. G. (2007). Nonword repetition in Spanish-speaking children with Specific Language Impairment (SLI). *International Journal of Language & Communication Disorders, 42*, 59-75.

Graf Estes, K., Evans, J., & Else-Quest, N. (2007). Differences in nonword repetitions performance of children with and without specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 50*, 177–195.

Gray, S. (2003). Diagnostic accuracy and test-retest reliability of nonword repetition and digit span tasks administered to preschool children with specific language impairment. *Journal of Communication Disorders, 36*, 129-151.

Gupta, P. & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: Computational and neural bases. *Brain and Language, 59,* 267-333.

Gwilliams, L., & Marantz, A. (2015). Non-linear processing of a linear speech stream: The influence of morphological structure on the recognition of spoken Arabic words. *Brain & Language, 1476,* 1-13.

Hoff, E., Core, C., & Bridges, K. (2008). Nonword repetition assesses phonological memory and is related to vocabulary development in 20-to 24-month-olds. *Journal of Child Language, 35*, 903-916.

Holes, C. (1989). *Gulf Arabic*. London: Routledge.

Holes, C. (2004). *Modern Arabic: Structures, functions, and varieties*. Georgetown University Press.

Johnstone, T. M. (1967). Aspects of syllabication in the spoken Arabic of 'Anaiza. *Bulletin of the School of Oriental and African Studies, 30*, 1-16.

Kara, R. A. (1976). *The problems encountered by English speakers in learning Arabic*. University of California, Berkeley.

Karwar, C., and Sakran, M. (1998). *Developmental aspects of phonological and morphological awareness in Palestinian Arabi*c. Seminar paper, Department of Communication Disorders, Tel Aviv University.

Kavitskaya, D., Babyonyshev, M., Walls, T., and Grigorenko, E. (2011). Investigating the effects of syllable complexity in Russian-speaking children with SLI. *Journal of Child Language, 38*, 979-989.

Khater, M., & Shaalan, S. (2007). Reporting norms for mean length of utterance (MLU) in words and morphemes for Qatari Speaking Children. Paper presented at *Linguistics in the Gulf Conference*, University of Qatar, Doha, Qatar.

Kohnert, K., Windsor, J., & Ebert, K. D. (2009). Primary or "specific" language impairment and children learning a second language. *Brain and language,109,* 101-111.

Kovas, Y., Hayiou-Thomas, M. E., Oliver, B., Dale, P. S., Bishop, D. V., & Plomin, R. (2005). Genetic influences in different aspects of language development: the etiology of language skills in 4.5 year- old twins. *Child Development, 76,* 632-651.

Kramer, R. (2007). Nonconcatenative morphology in Coptic. *Phonology at Santa Cruz series.* California: University of California at Santa Cruz.

Krishnan, S., Alcock, K. J., Mercure, E., Leech, R., Barker, E., Karmiloff-Smith, A., & Dick, F. (2013). Articulating novel words: Children's oromotor skills predict nonword repetition abilities. *Journal of Speech, Language, and Hearing Research, 56,* 1800-1812.

Lauchlan, F. and Boyle, C. (2014).To label or not to label: Is this the question? Commentary on Bishop, D. V. M., Ten questions about terminology for children with unexplained language problems. *International Journal of Language and Communication Disorders, 49,* 381–415

Leonard, L. (2014). Replacing one imperfect term with another. Commentary on Reilly, S., Tomblin, B., Law, J., McKean, C., Mensah, F. K., Morgan, A., Goldfeld, S., Nicholson, J. M. and Wake, M., 2014, Specific language impairment: A convenient label for whom? *International Journal of Language and Communication Disorders, 49,* 416–451.

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

MacRoy-Higgins, M. & Dalton, K.P. (2015). The influence of phonotactic probability on nonword repletion and fast mapping in 3-year-olds with a history of expressive language delay. *Journal of Speech, Language, and Hearing Research, 58*, 1773-1779

Malenky, A. (1997). *The development of morphological awareness in Hebrew speakers.* Master's thesis, School of Education, Tel Aviv University.

Marshall, C., Harris, J., & van der Lely, H. K. (2003). The nature of phonological representations in children with Grammatical-Specific Language Impairment (G-SLI). In D. Hall, T. Markopoulos, A. Salamoura, & S. Skoufaki (Ed.), The University of Cambridge First Postgraduate Conference in Language Research (pp. 511-517). Cambridge University, Cambridge: Cambridge Institute of Language Research.

Marton, K., & Schwartz, R. G. (2003). Working memory capacity and language processes in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *46*, 1138-1153.

Martin, N. A., & Brownell, R. (2000). *Expressive one-word picture vocabulary test-4*. Novato, CA: Academic Therapy Publications.

Masoura, E. V., & Gathercole, S. E. (2005). Phonological short-term memory skills and new word learning in young Greek children. *Memory, 13,* 422-429.

McCarthy, J. J. (1979). On stress and syllabification. *Linguistic inquiry, 10*, 443-465.

McCarthy, J. (1986). OCP effects: gemination and antigemination. *Linguistic Inquiry, 17*, 207-263.

McCarthy, J. J. (1993). A case of surface constraint violation. *Constraint-based theories in multilinear phonology, special issue of Canadian Journal of Linguistics*, *3*8, 169-195.

McCarthy, J. D., & Prince, A. (1993). *Prosodic Morphology I: Constrain Interaction and Satisfaction*. Centre for Cognitive Science.

Melby-Lervåg, M., Lervåg, A., Lyster, S. A. H., Klem, M., Hagtvet, B., & Hulme, C. (2012). Nonword-repetition ability does not appear to be a causal influence on children's vocabulary development. *Psychological Science*, *10*, 1092-1098.

Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology*, *9*1, 3-19.

Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. Metsala, J. L. (Ed); Ehri, L. C. (Ed), *Word recognition in beginning literacy,* (pp. 89-120). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Montgomery, J. W. (1995). Examination of phonological working memory in specifically language-impaired children. *Applied Psycholinguistics, 16,* 355-378.

Montgomery, J. W. (2002). Examining the nature of lexical processing in children with Specific Language Impairment: Temporal processing or processing capacity deficit. *Applied Psycholinguistics*, *23*, 447-470.

Montgomery, J. W. (2004). Sentence comprehension in children with specific language impairment: Effects of input rate and phonological working memory. *International Journal of Language & Communication Disorders, 39*, 115-133.

Munson, B. (2001). Phonological pattern frequency and speech production in adults and children. *Journal of Speech, Language, and Hearing Research, 44*, 778-792.

Munson, B., Edwards, J., & Beckman, M. E. (2005). Relationships between nonword repetition accuracy and other measures of linguistic development in children with phonological disorders. *Journal of Speech, Language and Hearing Research*, *48*, 61-78.

Mustafa, I., AlZayaat, A., Hamed, A. Q., & AlNajaar, M. (2004). *AL Mu'jam AL Waseet*. Cairo: Maktabat al-Shorouq al-Dawliyah.

Mustafawi, E. M. (2006). *An optimality theoretic approach to variable consonant alternations in Qatari Arabic.* Unpublished doctoral dissertation, University of Ottawa, Ottawa.

Norbury, C. F. (2014). Language Impairment: where do we draw the line? Commentary on Reilly, S., Tomblin, B., Law, J., McKean, C., Mensah, F. K., Morgan, A., Goldfeld, S., Nicholson, J. M. and Wake, M., 2014, Specific language impairment: A convenient label for whom? *International Journal of Language and Communication Disorders , 49*, 416–451.

Oetting, J. B., & Cleveland, L. H. (2006). The clinical utility of nonword repetition for children living in the rural south of the US. *Clinical Linguistics & Phonetics, 20,* 553 - 561.

Plaut, D. C., & Gonnerman, L. M. (2000). Are non-semantic morphological effects incompatible with a distributed connectionist approach to lexical processing? *Language and Cognitive Processes, 15,* 445-485.

Prince, A., & Smolensky, Al. (1993). *Optimality Theory: Constraint interaction in generative grammar*. Tech. Rep. 2, Rutgers University Centre for Cognitive Science.

Ragheb, M., & Davis, S. (2010). The acquisition of word-final clusters in Cairense Arabic: An OT analysis. Paper presented at *the Arabic Linguistics Symposium*, Austin, Texas, April 2010.

Ravid, D. (2003). A developmental perspective on root perception in Hebrew and Palestinian Arabic. *Language Acquisition and Language Disorders, 28*, 293-320.

Ravid, D., Avivi, G., & Levy, R. (2000). Derivational Morphology in SLI Children. In Michael Perkins and Sara Howard (Eds.), *New directions in language development and disorders,* 39-49. Springer, US.

Ravid, D., H. Landau & M. (2003). Lovetski. Passive morphology in the school years: An example of later language acquisition. *Script 5-6*, 129-158.

Reilly, S., Bishop, D. V. M and Tomblin, B. (2014). Terminological debate over language impairment in children: forward movement and sticking points. *International Journal of Language and Communication Disorders, 49*, 452–562.

Reilly, S., Tomblin, B., Law, J., McKean, C., Mensah, F., Morgan, A., Goldfeld, S., Nicholson, J. and Wake, M. (2014). SLI: A convenient label for whom? *International Journal of Language and Communication Disorders, 49*, 416–451.

Rispens, J., & Parigger, E. (2010). Non-word repetition in Dutch-speaking children with specific language impairment with and without reading problems. *British Journal of Developmental Psychology, 28*, 177-188.

Roy, P., & Chiat, S. (2004). A prosodically controlled word and nonword repetition task for 2-to 4-year-olds: Evidence from typically developing children. *Journal of Speech, Language and Hearing Research, 47*, 223-234.

Sahlen, B., Reuterskiold-Wagner, C., Nettelbladt, U., & Radeborg, K. (1999). Nonword repetition in children with language impairment-pitfalls and possibilities. *International Journal of Language & Communication Disorders, 34*, 337-352.

Shaalan, S. (2010). *Investigating grammatical complexity in Gulf Arabic speaking children with Specific Language Impairment (SLI).* Ph.D. Dissertation, University College London.

Shaalan, S., & Khater, M. (2006). A comparison of two measures of assessing spontaneous language samples in Arabic speaking children. Poster presented at *Child Language Seminar*, University of Newcastle upon-Tyne, U.K, 19-21 July 2006.

Shamsan, M.A., & Attayib, A. (2015). Inflectional morphology in Arabic and English: A contrastive study. *Journal of English Linguistics, 5*, 139-150.

Schreuder, R., & Baayen, R. H. (1995). Modelling morphological processing. In L. B. Feldman (Ed.), *Morphological Aspects of Language Processing* (pp. 131-154). Hillsdale, New Jersey: Lawrence Erlbaum.

Smith, B. (2006). Precautions regarding nonword repetition tasks. *Applied Psycholinguistics, 27*, 584-587.

Snowling, M., Chiat, S., & Hulme, C. (1991). Words, nonwords and phonological processes: Some comments on Gathercole, Willis, Emslie and Baddeley. *Applied Psycholinguistics, 12*, 369-373.

SPSS Inc. (2009). *PASW Statistics for Windows*, Version 18.0. Chicago: SPSS Inc.

Stackhouse, J. & Wells, B. (1997). *Children's Speech and Literacy Difficulties: A Psycholinguistic Framework.* London: Whurr Publishers.

Storkel, H. L., Armbruster, J., & Hogan, T. P. (2006). Differentiating phonotactic probability and neighborhood density in adult word learning. *Journal of Speech, Language, and Hearing Research*, *49*, 1175-1192.

Stokes, S. F., Wong, A. M., Fletcher, P., & Leonard, L. B. (2006). Nonword repetition and sentence repetition as clinical markers of specific language impairment: The case of Cantonese. *Journal of Speech, Language, and Hearing Research, 4*9, 219-236.

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behaviour, 14*, 638-647.

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech and Hearing Research, 40*, 1245-1260.

Torrington Eaton, C., Newman, R. S., Ratner, N. B., & Rowe, M. L. (2015). Nonword repetition in 2-year-olds: Replication of an adapted paradigm and a useful methodological extension. *Clinical Linguistics & Phonetics, 29*, 523-535.

Tucker, M. A. (2009). The root-and-prosody approach to Arabic verbs. *Qualifying Paper*, UCSC, 1-43.

Tucker, M. A. (2010). Iraqi Arabic verbs: The need for roots and prosody. In *Proceedings of the 28th West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla Proceedings Project.

Ussishkin, A. P. (2000). *The emergence of fixed prosody*. Doctoral dissertation, University of California Santa Cruz.

Ussishkin, A. (2005). A fixed prosodic theory of nonconcatenative templatic morphology. *Natural Language & Linguistic Theory,* 23, 169-218.

Versteegh, K. (1997). *The Arabic language*. Edinburgh: Edinburgh University Press

Watson, Janet C.E. (2002). *The Phonology and Morphology of Arabic*. Oxford University Press, London.

Williams, D., Payne, H., & Marshall, C. (2013). Nonword repetition impairment in autism and specific language impairment: Evidence for distinct underlying cognitive causes*. Journal of Autism and Developmental Disorders, 43,* 404-417.

Appendix A: Arabic Word Repetition Test

Child name: …………………… Date: …………………………….

D.O.B: …………………………. Examiner: ……………………….

Trials :
1. /mo :l / mall
2. /ka.nab/ sofa
3. /sik.ki:.na/ knife

Note: word repetition list with English translation.

Word		Word	
1. /ki:s/	Bag	13. /ra.mil/	Sand
2. /le:t/	Light	14. /wa.lad/	Boy
3. /ʔa.na.na:s/	Pineapple	15. /ti.li.fo:n/	Telephone
4. /saj.ja:.ra/	Car	16. mu.kaj.jef	Air conditioner
5. /se :f/	Sword	17. /fa:.ni:.la/	T-shirt
6. /sa.ma.ka/	Fish	18. /dik.ka:n/	Supermarket
7. /ro:b/	Yoghurt	19. /se :.kal/	Bicycle
8. /lay.mu:n/	Lemon	20. /tan.nu:.ra/	Skirt
9. /na : s/	People	21. /ki.ta :b/	Book
10. /damm/	Blood	22. /ni.ser/	Eagle
11. /lo:n/	Colour	23. /dab.ba:.sa/	Stapler
12. /fi :l/	Elephant	24. /kab.bu :s/	Cap

Appendix B: Arabic Nonword Repetition Test

Child name: ……………………   Date: …………………………….

D.O.B: ………………………….   Examiner: ……………………….

Trials :

1. /le:m/
2. /ba.kin/
3. /nus.su:.ka/

Nonwords		**Nonwords**	
1. /lu:f/		13. /ku.si.ma/	
2. /nuk.ki:d/		14. /nat.ta:.ro/	
3. /sa:k/		15. /no:l/	
4. /ro.nis/		16. /mo:d/	
5. /su:n/		17. /lo.wid/	
6. /fu.maj.jek/		18. /bi.tu:k/	
7. /sad.du:.ba/		19. /nu.fa.tu:l/	
8. /nul.ja:m/		20. /yas.sa:.ri/	
9. /fi:s/		21. /nuʔ.sa.nos/	
10. /bu:r/		22. /mo.ril./	
11. /lu:.fa:.no/		23. /tu:l/	
12. /suk.ki:b/		24. /la:kus/	

Appendix C: Arabic Expressive Vocabulary Test (AEVT)

Record Form

Name…………………………...   Gender: M / F

School/Nursery……………………   Grade………………

Examiner……………….

	Year	Month	Day
Test Date			
Birth Date			
Chronological Age			

إبدأ دائماً عند رقم 1 في المجموعة الأولى، توقف في حالة وجود 6 أخطاء متتابعة. استخدم اللهجة الدارجة محلياً عند قراءة الكلمات الدارجة. ضع(\ ) اذا كانت الأجابة خاطئة , سجل لفظ الطفل في حالة عدم نطق الكلمة بشكل صحيح في حالة كانت الأجابة صحيحة اترك الخانة فارغة.

Ceiling Item	
Minus Errors	-
Raw Score	=

التدريب:

في البداية تودد إلى الطفل واجعله مرتاحاً من خلال محادثة بسيطة، ثم اشرح له الهدف " راح نشوف كتاب فيه صور، وأبيك تقولي اسم كل صورة راح أشر عليها ". ثم ابدأ بصور التدريب(أ, ب , ج , د). شنو هذا و اشر على....(قطوة)..إذا لم يؤشر الطفل شجعه، صححه إذا كانت الإجابة غير صحيحة (فقط أثناء التدريب). إذا لم يعرف الطفل الإجابة أو تردد، اطلب منه أن يحاول، "حاول، حتى لو ما تعرف موب مشكلة". دائماً امدحه على المشاركة،

تدريب أ : (شنو هذي؟ ...."قطوة"

تدريب ب : (شنو هذي؟...." كيكة"

تدريب ج : (شنو يسوي الرجال؟ ......"يقرأ"

تدريب د : (شنو نسمي هذيلة كلهم؟..."العاب"

أ. قطوه..............

ب. كيكه...............

ج. شنو يسوي الرجال؟

يقرأ.............

د. شنو نسمي كل هذيله ؟

ألعاب ....

2:11-2:00

1. كوره.............
2. جوتي............
3. موز ....
4. يد ....
5. مفتاح............
6. سياره...........
7. عصفور ....
8. شنو تسوي البنت ؟

.تاكل...........

3:11-3:00

9. قميص............
10. قلم................
11. فرشاة أسنان........
12. طياره.............
13. كرسي............
14. شنطة.............
15. حصان...........
16. شجرة...........

4:11-4:00

17. هدية............
18. مسجد............
19. أسنان............
20. سفينة............
21. ثلاجة............
22. مربع............
23. مظلة............
24. ذبانة.............

5:11-5:00

25. شنو يسوي الرجال؟

يسقي..............

26. لمبة\ليت.........
27. شنونسمي هذيله كلهم؟

خضروات.........

28. ستارة...........
29. بركة...........
30. ركبة...........
31. عنكبوت.........
32. مسطرة..........

6:11-6:00

33. قفل.............
34. ملعب\استاذ.......
35. بطريق..........
36. شنو يسوي الرجال؟

يصبغ............

37. خريطة........
38. جسر ....
39. سمك القرش....
40. دباسة.........

7:11-7:00

41. شنو نسمي هذيله كلهم؟

حشرات..........

42. مطرقة.........
43. نعامة..........
44. كوكب.........
45. باذنجان.......
46. رصيف........
47. شلال..........
48. تمثال..........

8:11-8:00

49. منطاد..........
50. شاحنة.........
51. شنو يسوي الرجال؟

يجدف.........

52. عدسة مكبرة...
53. غيم\غيوم....
54. كرفتة.......
55. شنو هذا الشكل؟

اسطواني.....

56. شنونسمي هذيله كلهم؟

عائلة..........

9:10-9:00

57. فرس البحر .....
58. منظار\دربيل....
59. شطرنج........
60. شنو هذا الشكل؟

خماسي..........

61. كمان\كمنجة....
62. شنو نسمي هذيله كلهم؟

زواحف ....

63. شنو تسوي البنت ؟

تخيط \ تنسج......

64. قرنبيط.........

CITY UNIVERSITY LONDON

Appendix D: Parent Information Sheet

Project title: Early phonological skills as a predictor of receptive and expressive vocabulary size in Gulf Arabic speaking children.

Investigators: Ms Mariam Khater, Professor Shula Chiat, & Dr Rachael-Anne Knight
Department of Language and Communication Science
City University, Northampton Square
London, EC1V 0HB
Telephone: [redacted]
Email: [redacted]

I am a doctoral student in Language and Communication Science at City University London. As part of my studies, I am carrying out a research project to investigate children's speech and language skills. This study will provide teachers and speech language therapists with some essential information regarding phonological processing and word learning in Arabic as there are currently only limited resources for this. Developing new assessment materials is therefore essential in order to identify and assess children with language impairment.

To carry out my project I need to see 30 typically developing children aged between 5:0 – 7:0 years old at the time of testing.

The children will be asked to carry out the following test:

- A root nonword repetition test: will repeat 12 words, 12 root-non pattern nonwords, 12 non root pattern nonwords and 12 nonroot nonpattern nonwords. All these words and nonwords consist of two or three syllables. The total is 36 items.

I understand the importance of providing confidentiality to research participants. The information that will be collected will not to be discussed or communicated outside of research meetings with the Principal Investigator, Supervisors or others specifically identified by the Investigators.

I understand the importance of providing confidentiality to research participants. The information that will be collected will not to be discussed or communicated outside of research meetings with the Principal Investigators, Supervisors or others specifically identified by the Investigators.

If you kindly agree for your child to participate in this study, could you please fill in the attached consent form and give it to the principal investigator or your child's teacher or speech therapist. Your child participation is voluntary, and you have the right to withdraw at any time without giving a reason. If you have any further questions or concern please do not hesitate to contact my supervisors Professor Shula Chiat [redacted] and Dr Rachael-Anne Knight [redacted]

Thank you for your consideration.

Yours sincerely,
Mariam Khater
Doctoral student in Language and Communication Science at City University London.

If there is an aspect of the study which concerns you, you may make a complaint by contacting the Secretary to the Research Ethics Committee by phone (004420 7040 8106), or by e-mail to Alison Welton [redacted], or by writing to: **Alison Welton**,Research Governance Officer, School of Health Sciences, City University,20 Bartholomew Close, West Smithfield,London EC1A 7QN,Tel: [redacted]
Email: [redacted]

Informed Consent Form

Title of Project: Early phonological skills as a predictor of receptive and expressive vocabulary size in Gulf Arabic speaking children. **Investigators:** Ms Mariam Khater , Professor Shula Chiat, & Dr Rachael-Anne Knight

	YES	**NO**
Have you read the Parent Information Sheet?		
Have you had the opportunity to ask questions and discuss the study?		
Have you received satisfactory answers to all your questions?		
Have you received enough information about the study?		
Do you agree to your child participating in this study?		
Do you give permission to audio record the testing with your child and keep the recording until the end of the study (October 2014)?		
Do you understand that you are free to withdraw your child from the study without penalty at any stage?		
Do you give permission for any assessments of your child to be made available to your child's teacher or speech and language therapist?		

Participant's Name: ....(please print)

Participant's Date of Birth: ....

Parent's/Guardian's Name: ....

Your relationship to participant: ....

Contact Numbers: …………………………………………………………

Email: …………………………………………………………

Signature of Parent/Guardian: ....Date:..........

Appendix E: Invitation for your Schools to participate in a research study

Project title: Early phonological skills as a predictor of receptive and expressive vocabulary size in Gulf Arabic speaking children.

Investigators: Ms Mariam Khater, Professor Shula Chiat, & Dr Rachael-Anne Knight
Department of Language and Communication Science
City University, Northampton Square
London, EC1V 0HB
Telephone: [redacted]
Email: [redacted]

Secretary of Ethics Committee: Alison Welton
Research Governance Officer, City University
Telephone: [redacted]
E-mail: [redacted]

Dear Nursery Manager

I am a doctoral student in Language and Communication Science at City University London. As part of my studies, I am carrying out a research project to investigate children's speech and language skills. This study will provide teachers and speech language therapists with some essential information regarding phonological processing and word learning in Arabic as there are currently only limited resources for this. Developing new assessment materials is therefore essential in order to identify and assess children with language impairment.

To carry out my project I need to see 30 typically developing children aged between 5:0 – 7:0 years old at the time of testing..

The children will be asked to carry out the following test:

- A root nonword repetition test: will repeat 12 words, 12 root-non pattern nonwords, 12 non root pattern nonwords and 12 nonroot nonpatternnonwords. All these words and nonwords consist of two or three syllables. The total is 48 items.

I understand the importance of providing confidentiality to research participants. The information that will be collected will not to be discussed or communicated outside of research meetings with the Principal Investigator, Supervisors or others specifically identified by the Investigators.

If you kindly agree to participate in this project, please refer to the principal investigator children who come to your nursery/clinic and meet the following criteria:

- Their age is between 5 year 0 months and and 7 years 0 months.
- Child does not present or have a history of speech and or language delay/problem.
- No history of congenital abnormalities, hearing loss, oral-motor difficulties or autism.

If you have any further questions or concern please do not hesitate to contact my supervisors Professor Shula Chiat [redacted] and Dr Rachael-Anne Knight [redacted]

Many thanks for giving this your consideration.

Yours sincerely,

Mariam Khater
Doctoral student in Language and Communication Science at City University London

If there is an aspect of the study which concerns you, you may make a complaint by contacting the Secretary to the Research Ethics Committee by phone [redacted]), or by e-mail to Alison Welton [redacted]), or by writing to: **Alison Welton**,Research Governance Officer, School of Health Sciences, City University,20 Bartholomew Close, West Smithfield, London EC1A 7QN,Tel: [redacted], Email: [redacted]

Appendix F: Familiarity test

Please circle the following items using a scale that goes from 1 to 7. Use (1) to describe words that you have never heard before and are most likely not to be Arabic words; circle (2) for words that are not familiar but are more likely than 1 to be Arabic words, and so on until 7 which refers to words that sound like familiar Arabic words.

Do not be bothered if you cannot tell the meaning of words as some of them may not be real words. Please judge words by their familiarity to Arabic regardless of knowing or not knowing their meanings.

Note: The participants form didn't contain any written stimuli, the examiner read the following list to the participants.

Items	**1**	**2**	**3**	**4**	**5**	**6**	**7**
/nabat/	1	2	3	4	5	6	7
/kasataf/	1	2	3	4	5	6	7
/sakadab/	1	2	3	4	5	6	7
/dauuf/	1	2	3	4	5	6	7
/salud/	1	2	3	4	5	6	7
/damabaf/	1	2	3	4	5	6	7
/labusaf/	1	2	3	4	5	6	7
/fulitak/	1	2	3	4	5	6	7
/basamat/	1	2	3	4	5	6	7
/lamus/	1	2	3	4	5	6	7
/samak/	1	2	3	4	5	6	7
/kamulan/	1	2	3	4	5	6	7
/kalam/	1	2	3	4	5	6	7
/lakub/	1	2	3	4	5	6	7
/tasubaf/	1	2	3	4	5	6	7

/fusib/	1	2	3	4	5	6	7
/talasab/	1	2	3	4	5	6	7
/saluk/	1	2	3	4	5	6	7
/masak/	1	2	3	4	5	6	7
/kanub/	1	2	3	4	5	6	7
/sanal/	1	2	3	4	5	6	7
/kanab/	1	2	3	4	5	6	7
/kubidaf/	1	2	3	4	5	6	7
/fasud/	1	2	3	4	5	6	7
/kamasal/	1	2	3	4	5	6	7
/salukad/	1	2	3	4	5	6	7
/latas/	1	2	3	4	5	6	7
/nafas/	1	2	3	4	5	6	7
/kabaf/	1	2	3	4	5	6	7
/damab/	1	2	3	4	5	6	7
/sabudaf/	1	2	3	4	5	6	7
/tamuf/	1	2	3	4	5	6	7
/lafusab/	1	2	3	4	5	6	7
/sabud/	1	2	3	4	5	6	7
/sumifal/	1	2	3	4	5	6	7
/lafab/	1	2	3	4	5	6	7
/kadufab/	1	2	3	4	5	6	7
/sabadal/	1	2	3	4	5	6	7
/fulit/	1	2	3	4	5	6	7
/kafal/	1	2	3	4	5	6	7
/dafusal/	1	2	3	4	5	6	7
/lakafad/	1	2	3	4	5	6	7
/katab/	1	2	3	4	5	6	7
/kamas/	1	2	3	4	5	6	7
/tasub/	1	2	3	4	5	6	7

/kamafas/	1	2	3	4	5	6	7
/talas/	1	2	3	4	5	6	7
/kasubad/	1	2	3	4	5	6	7
/katubaf/	1	2	3	4	5	6	7
/kubis/	1	2	3	4	5	6	7
/kabus/	1	2	3	4	5	6	7
/kalimat/	1	2	3	4	5	6	7
/kubid/	1	2	3	4	5	6	7
/kafulab/	1	2	3	4	5	6	7
/lamak/	1	2	3	4	5	6	7
/kalad/	1	2	3	4	5	6	7
/kabafas/	1	2	3	4	5	6	7
/dasafal/	1	2	3	4	5	6	7
/saludaf/	1	2	3	4	5	6	7
/fusibal/	1	2	3	4	5	6	7
/sumif/	1	2	3	4	5	6	7
/lamakaf/	1	2	3	4	5	6	7
/kabusad/	1	2	3	4	5	6	7
/biladi/	1	2	3	4	5	6	7
/kanubaf/	1	2	3	4	5	6	7
/banat/	1	2	3	4	5	6	7
/dafan/	1	2	3	4	5	6	7
/sakub/	1	2	3	4	5	6	7
/dubisal/	1	2	3	4	5	6	7
/saluk/	1	2	3	4	5	6	7
/kaful/	1	2	3	4	5	6	7
/sakabat/	1	2	3	4	5	6	7
/kabas/	1	2	3	4	5	6	7
/lamusad/	1	2	3	4	5	6	7
/kanubaf/	1	2	3	4	5	6	7
/sakubal/	1	2	3	4	5	6	7

/fasudab/	1	2	3	4	5	6	7
/banatak/	1	2	3	4	5	6	7
/dafanat/	1	2	3	4	5	6	7
/kamul/	1	2	3	4	5	6	7
/samakat/	1	2	3	4	5	6	7
/labus/	1	2	3	4	5	6	7
/lakafad/	1	2	3	4	5	6	7
/kitabat/	1	2	3	4	5	6	7
/kaladas/	1	2	3	4	5	6	7
/dasaf/	1	2	3	4	5	6	7
/latasak/	1	2	3	4	5	6	7
/nasam/	1	2	3	4	5	6	7
/kasat/	1	2	3	4	5	6	7
/kaduf/	1	2	3	4	5	6	7
/lafabad/	1	2	3	4	5	6	7
/dafus/	1	2	3	4	5	6	7
/nabatat/	1	2	3	4	5	6	7
/lafus/	1	2	3	4	5	6	7
/lamas/	1	2	3	4	5	6	7
/lakubad/	1	2	3	4	5	6	7
/tamufal/	1	2	3	4	5	6	7
/danufas/	1	2	3	4	5	6	7
/sakab/	1	2	3	4	5	6	7
/balad/	1	2	3	4	5	6	7
/kafas/	1	2	3	4	5	6	7
/sanalab/	1	2	3	4	5	6	7
/kafasal/	1	2	3	4	5	6	7
/dafukab/	1	2	3	4	5	6	7
/kanabat/	1	2	3	4	5	6	7
/kamaf/	1	2	3	4	5	6	7

/dafuk/	1	2	3	4	5	6	7
/basma/	1	2	3	4	5	6	7
/banat/	1	2	3	4	5	6	7
/kasub/	1	2	3	4	5	6	7
/lakaf/	1	2	3	4	5	6	7
/sakad/	1	2	3	4	5	6	7
/kabasat/	1	2	3	4	5	6	7
/sabad/	1	2	3	4	5	6	7
/nasamat/	1	2	3	4	5	6	7
/katub/	1	2	3	4	5	6	7
/kafalat/	1	2	3	4	5	6	7
/nafasak/	1	2	3	4	5	6	7
/masakat/	1	2	3	4	5	6	7
/lamasat/	1	2	3	4	5	6	7
/salamat/	1	2	3	4	5	6	7

Appendix G: Root and Pattern Nonword Repetition test (RAP-NWR)

Child name: …………………… Date: …………………………….

D.O.B: …………………………. Examiner: ……………………….

<u>Trials :</u>

Saluk
Kamasal
kadufab

Items		Items	
/kabusad/		/dafuuk/	
/fasud/		/sakadab/	
/fusibal/		/damabaf/	
/kasub/		/kasubad/	
/latas/		/kafas/	
/danufas/		/lakaf/	
/lamus/		/katub/	
/dafus/		/lafabad/	
/lafus/		/kaduf/	
/kafasal/		/kamulan/	

/fulitak/		/kasataf/	
/sabudaf/		/lakafad/	
/fulit/		/sumifal/	
/sabud/		/lafusab/	
/dafukab/		/dubisal/	
/kanuub/		/sumif/	
/sabad/		/labusaf/	
/dasaf/		/sakad/	

Appendix H: Results of the RAP-NWR experiment excluding the 5 items violating the (OCP-Place)

All data was analysed using repeated measures ANOVAs with nonword Types (type 1 (R-NP) vs type 2 (NR-P) vs type 3 (NR-NP)) and Word length (2-syllables, 3-syllables) as within subject factors and age band as a between subject factor. A Bonferroni correction was applied to all follow-up tests (Pairwise comparisons and t-tests).

Results showed that there was a significant effect of nonword type (F (2,166) =3.968, p<.000, η^2=.046), and number of syllables (F (1,83) =94.57, p<.000, η^2=.533). The children had significantly better scores on repeating Type1 (R-NP) when compared with types 2 and 3. In addition, their scores on repeating two syllable nonwords was significantly better when compared with three syllabic words (see Figure 28). Moreover, there was a significant word length*age group interaction, (F=4.627, p<.000), and type*word length interaction, (F=4.715, p<.000). However, there was no interaction of age*group*type.

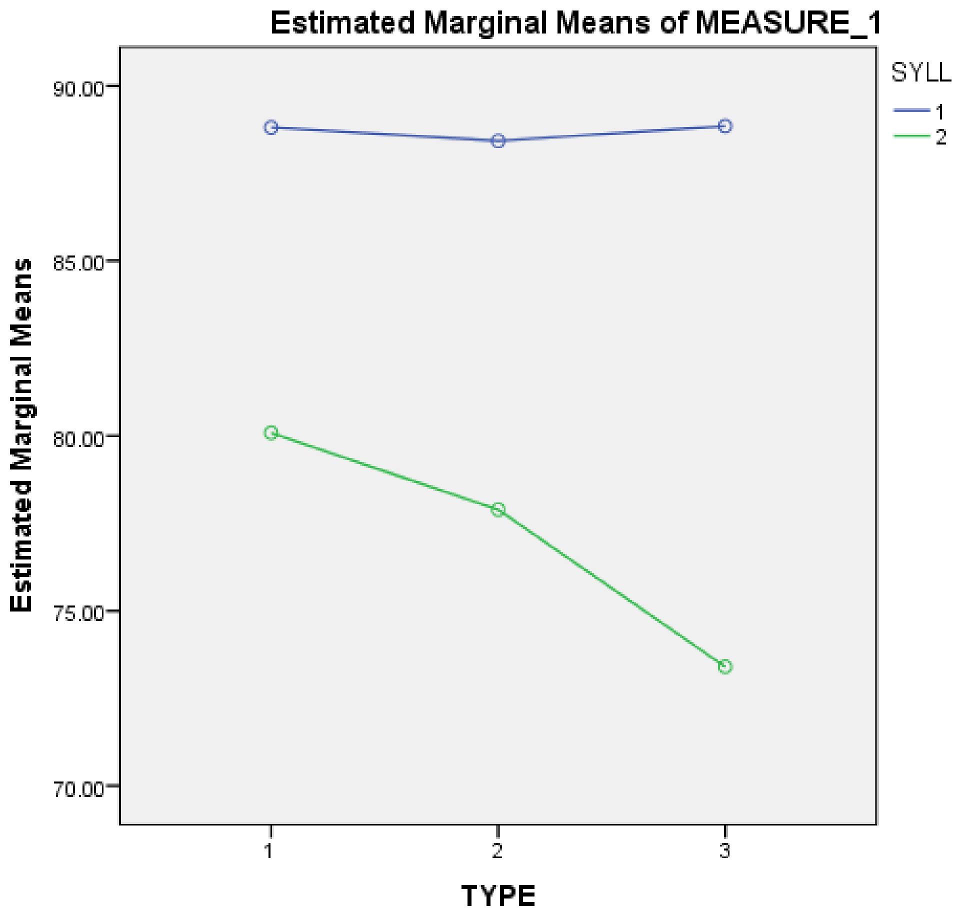

Figure 28: Performance on different nonword types and word length. Types of nonwords: Type 1 (R-NP), Type 2 (NR-P), Type 3 (NR-NP), excluding the 5 items violating the (OCP-Place).

Pairwise comparisons showed there was a significant difference in the performance of the children between Type 3 and Type 1($p<.005$). Children's performance on type 1 nonwords was significantly higher than on type 3 nonwords. However, the difference between Type 1

(R-NP) and Type 2, and Type 2 and Type 3 did not reach significance. In general, children performed more accurately on Type 1 when compared with Type 3 (81% vs 84.5%). The difference is not high; however, it is statistically significant.

Main effects of word length

Pairwise comparison showed there was a significant difference in the performance of the children in two syllable vs. 3syllable items ($p<.001$). In general children performed more accurately on two syllables than on three syllables items (77.1 % vs 88.6%).

Pairwise comparison with Bonferroni correction showed there was a significant difference between age group 1 and the following age groups: age group 3($p<.002$), age group 4 ($p<.000$), age group 5($p<.000$), and age group 6 ($p<.000$) in their nonword repetition score as measured by the percentage of correct phonemes. Age group 1 (children aged between 2:0 and 2:6 years old) performed significantly below when they are compared with all other age groups, except age group 2. There is also significant difference between age group 2 (2:7-3:00) on one hand and age groups 4 ($p<.004$), age group 5($p<.000$) and age group 6($p<.000$) on another hand. Age group 3 showed significant difference from age group 1 ($p<.002$), age group 5 ($p<.008$) and age group 6($p<.004$). Age group 4 children showed significant difference comparing with group age 1($p<.000$) and age group 2 ($p<.004$). In addition, age group 5 showed significant difference comparing with age group 1($p<.000$), age group 2 ($p<.001$). Finally, age 6 showed significant difference with age group 1 ($p<.000$), age group 2($p<.000$) and age group 3($p<.004$). So the following summarizes the group comparisons:

Age group 1= age group 2 but < age groups 3,4,5, and 6.

Age group 2=age group 3, but < age groups 4,5, and 6.

Age group 3=age groups 4,5, but < age group 6.

Age group 4=age group 5,6

Table 32 shows how the children performed on the RAP-NWR test across the 6 different age groups. In general, the older the children are the more accurate they perform on the RAP-NWR test (with percentage of correct responses ranging between 62.4% and 97.9%).

Table 32: Percentage of correct phonemes RAP-NWR for all children in different age groups, n=89.

Age Group	Percentage of correct phonemes
1 (2:0-2:6) years	62.4%
2 (2:7-3:0) years	71.9%
3 (3:1-3:6) years	81.1%
4 (3:7-4:0) years	87.6%
5 (5:0-5:11) years	96.4%
6 (6:0-7:0) years	97.9%

Appendix I: Correlation between word and nonword repetition scores and receptive and expressive vocabulary tests for TD and CL groups

		APVT	AEVT	Word repetition	Nonword repetition
APVT	Pearson Correlation	1	.643**	.268*	.268*
	Sig. (2-tailed)		.000	.040	.040
	N	59	59	59	59
AEVT	Pearson Correlation	.643**	1	.430**	.430**
	Sig. (2-tailed)	.000		.001	.001
	N	59	59	59	59
Word repetition	Pearson Correlation	.268*	.430**	1	1.000**
	Sig. (2-tailed)	.040	.001		.000
	N	59	59	59	59
Nonword repetition	Pearson Correlation	.268*	.430**	1.000**	1
	Sig. (2-tailed)	.040	.001	.000	
	N	59	59	59	59

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Appendix J: Ethical Approval

School of Health Sciences

CITY UNIVERSITY LONDON

Research Office
Northampton Square
London EC1V 0HB

Tel: [redacted]

www.city.ac.uk

Ref: PhD/12-13/12

09 April 2013

Dear Mariam / Shula / Rachel-Anne

Re: Early phonological skills as a predictor of receptive and expressive vocabulary size in Gulf Arabic speaking children

Thank you for forwarding amendments and clarifications regarding your project. These have now been reviewed **and approved** by the Chair of the School Research Ethics Committee.

Please find attached, details of the full indemnity cover for your study.

Under the School Research Governance guidelines you are requested to contact myself once the project has been completed, and may be asked to complete a brief progress report six months after registering the project with the School.

If you have any queries please do not hesitate to contact me as below.

Yours sincerely

[redacted]

Alison Welton
Research Governance Officer

[redacted]

Appendix K: Analysis of results by gender of participants for word repetition, nonword repetition, APVT and EAVT (Study 1)

Gender

Gender has been found to have little effect on children's language performance, especially in very young children. Chiat and Roy (2006) found a small effect of gender on children's performance on different language tasks. To explore the effect of gender in this current study, descriptive statistics (see Table 33) shows mean values and standard deviations of the children from both TD and clinical groups (male, female) in word repetition, nonword repetition, APVT and EAVT.

Table 33: Descriptive statistics of children's (male: female) scores on word repetition, nonword repetition, APVT, and EAVT.

	Gender	N	Mean	St. Deviation
Word Repetition	Male	33	93.49	10.97
	Female	26	94.26	6.84
	Total	59	93.83	9.31
Nonword Repetition	Male	33	76.63	8.99
	Female	26	77.26	5.60
	Total	59	76.91	7.63
APVT	Male	33	23.21	7.39
	Female	26	22.15	7.03
	Total	59	22.75	7.19
EAVT	Male	33	19.06	5.01
	Female	26	16.77	6.37
	Total	59	18.05	5.72

To investigate the effect of gender on the children's performance on word, nonword repetition, APVT and EAVT, one way ANOVA was conducted with gender as independent variable. Results showed that there was no significant difference between different male and female participants and across all measures: word repetition ($F(1,57) = .098$, $p = .756$), nonword repetition ($F(1,57) = .098$, $p = .756$, different nonword lengths ($F(5,84) = 10.99$, $p < .001$), APVT ($F(1,57) = .311$, $p = .579$) and EAVT($F(1,57) = 2.38$, $p = .128$). These findings are in line with the findings of Chiat and Roy (2006) that found a minimum effect of gender on the children's performance in different tasks.

Pearson product-moment correlations coefficients were also calculated to measure the correlation between word and nonword repetition and receptive and expressive vocabulary of male participants from both TD and CL groups (n=33). Results of the various correlations are shown in Table 34. Results showed that the word repetition test significantly correlated with AEVT ($r = .396$, $p = .034$), and the nonword repetition test ($r = 1.00$, $p < .001$). Nonword repetition is significantly correlated with AEVT ($r = .396$, $p = 034$). Furthermore, AEVT is significantly correlated with APVT ($r = .658$. $p < .001$). On the other hand, none of the repetition scores correlate significantly with APVT.

Table 34: Correlations between scores of word and nonword repetition and Arabic Picture Vocabulary Test (APVT) and Arabic Expressive Vocabulary Test (AEVT) in males from TD and CL groups (*n*=33)

		APVT	EAVT	Word	Nonword
APVT	Pearson Correlation	1	.658**	.256	.256
	Sig. (2-tailed)		<.001	.150	.150
AEVT	Pearson Correlation	.658**	1	.369*	.369*
	Sig. (2-tailed)	<.001		.034	.034
Word	Pearson Correlation	.256	.369*	1	1.000**
	Sig. (2-tailed)	.150	.034		<.001
Nonword	Pearson Correlation	.256	.369*	1.000**	1
	Sig. (2-tailed)	.150	.034	<.001	

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Pearson product-moment correlations coefficients were also calculated to measure the correlation between word and nonword repetition and receptive and expressive vocabulary of female participants from both TD and CL groups (n=26). Results of the various correlations (see Table 35) were in general similar to correlations of males scores. Results showed that the word repetition test significantly correlated with AEVT (r=.645, p<.001), and the nonword repetition test (r=1.00, p<.001). Nonword repetition was significantly correlated with AEVT (r=.645, p<.001), Furthermore, AEVT was significantly correlated with APVT (r=.640. p<.001). On the other hand, none of the repetition scores correlate significantly with APVT.

According to the series of analysis of the gender effect on the children's performance from both groups TD and CL, we conclude that gender difference has no significant effect on children's scores.

Table 35: Correlations between word and nonword repetition scores and Arabic Picture Vocabulary Test (APVT) and Arabic Expressive Vocabulary Test (AEVT) in females from TD and CL groups (*n*=26)

		APVT	EAVT	Word	Nonword
APVT	Pearson Correlation	1	.640**	.323	.323
	Sig. (2-tailed)		<.001	.108	.108
AEVT	Pearson Correlation	.640**	1	.645**	.645**
	Sig. (2-tailed)	<.001		<.001	<.001
Word	Pearson Correlation	.323	.645**	1	1.000**
	Sig. (2-tailed)	.108	<.001		<.001
Nonword	Pearson Correlation	.323	.645**	1.000**	1
	Sig. (2-tailed)	.108	<.001	<.001	

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Appendix L: RAP-NWR test classified according to types

Root-Nonpattern (Type 1)	Nonroot-Pattern (Type 2)	Nonroot-Nonpattern (Type 3)
/ka.bu.sad/	/la.tas/	/fu.si.bal/
/fa.sud/	/ka.fa.sal/	/da.nu.fas/
/ka.sub/	/sa.bad/	/da.fus/
/la.mus/	/da.saf/	/la.fus/
/fu.li.tak/	/sa.ka.dab/	/sa.bu.daf/
/fu.lit/	/da.ma.baf/	/sa.bud/
/ka.nub/	/ka.fas/	/da.fu.kab/
/ka.su.bad/	/la.kaf/	/da.fuk/
/ka.tub/	/la.fa.bad/	/ka.duf/
/ka.mu.lan/	/ka.sa.taf/	/su.mi.fal/
/du.bi.sal/	/la.ka.fad/	/la.fu.sab/
/la.bu.saf/	/sa.kad/	/su.mif/

Appendix M: Examples of children's responses in RAP-NWR test.

Nonwords	NW Type	Child 1	Child 2	Child 3	Child4
/kabusad/	1	1	0	1	[kabusat]
/fasud/	1	1	[masjad]	1	1
/kasuub/	1	1	1	[kafus]	1
/lamus/	1	[lamu]	[labus]	1	1
/fulitak/	1	1	[falala]	[fuli]	1
/fulit/	1	1	[fuli]	[fukil]	1
/kanub/	1	[kamut]	[kabab]	[kamat]	1
/kasubad/	1	1	[kubas]	[kabubed]	1
/katub/	1	[tatub]	[katab]	[kabut]	1
/kamulan/	1	1	[lamalan]	[kamuman]	1
/dubisal/	1	1	0	[dubufal]	1
/labusaf/	1	[tusosaf]	[lalubas]	1	1
/latas/	2	1	1	1	1
/kafasal/	2	[tasasal]	[kalafas]	1	1
/sabad/	2	1	[sabab]	[tabad]	[sabat]
/dasaf/	2	[dasat]	1	1	1
/sakadab/	2	[satatab]	[sababak]	[sakabab]	1
/damabaf/	2	[damadaf]	0	1	[nanamaf]
/kafas/	2	[tasat]	1	1	1
/lakaf/	2	1	[labaf]	1	1
/lafabad/	2	1	[lafa]	1	[lafadab]
/kasataf/	2	[tasasat]	[kas]	[kafafaf]	1
/lakafad/	2	[kalafad]	0	[lalafad]	1
/sakad/	2	[sadad]	0	[abufad]	[sakab]
/fusibal/	3	[subibal]	1	1	1
/danufas/	3	[danufat]	1	[damuhas]	1
/dafus/	3	1	[dadab]	1	1

/lafus/	3	1	[lalas]	1	1
/sabudaf/	3	[sadubas]	0	[kaboba]	[sabutaf]
/sabud/	3	1	1	[sakod]	1
/dafukab/	3	[tatudab]	[sabak]	1	1
/dafuk/	3	0	[sabut]	1	1
/kaduf/	3	1	1	[kafod]	1
/sumifal/	3	[sumisal]	0	0	1
/lafusab/	3	[lafusaf]	[ladubas]	1	[lasusab]
/sumif/	3	1	0	[sumi]	1

*Scores: 1= correct answer, 0= no response.

Appendix N: Percentage of correct repetitions of consonants and vowels according to their positions in the nonword.

Table 36: Percentage of correct repetitions of consonants and vowels according to their positions in the nonword for all participants (n=89).

Child	C1	C2	C3	C4	V1	V2	V3
1	100.00	86.11	88.89	55.56	97.22	97.22	88.89
2	66.67	44.44	44.44	16.67	72.22	50.00	44.44
3	91.67	69.44	58.33	77.78	97.22	88.89	83.33
4	97.22	94.44	86.11	94.44	100.00	100.00	100.00
5	97.22	91.67	91.67	94.44	97.22	94.44	77.78
6	83.33	72.22	72.22	77.78	94.44	94.44	100.00
7	100.00	94.44	88.89	66.67	100.00	100.00	94.44
8	83.33	83.33	86.11	72.22	94.44	94.44	83.33
9	100.00	94.44	86.11	83.33	100.00	100.00	100.00
10	100.00	91.67	88.89	72.22	100.00	100.00	94.44
11	97.22	83.33	75.00	83.33	94.44	94.44	100.00
12	80.56	72.22	69.44	72.22	88.89	88.89	94.44
13	80.56	72.22	75.00	55.56	86.11	83.33	77.78
14	100.00	100.00	100.00	100.00	100.00	100.00	100.00
15	80.56	72.22	75.00	55.56	86.11	83.33	77.78
16	80.56	72.22	75.00	55.56	86.11	83.33	77.78
17	88.89	86.11	77.78	72.22	88.89	86.11	72.22

18	77.78	72.22	86.11	88.89	72.22	75.00	88.89
19	88.89	83.33	77.78	66.67	91.67	88.89	72.22
20	11.11	36.11	55.56	33.33	33.33	36.11	33.33
21	97.22	88.89	86.11	94.44	97.22	97.22	100.00
22	16.67	22.22	33.33	22.22	36.11	38.89	27.78
23	100.00	91.67	86.11	83.33	97.22	100.00	100.00
24	80.56	88.89	94.44	72.22	100.00	100.00	94.44
25	80.56	80.56	80.56	61.11	91.67	86.11	94.44
26	75.00	66.67	69.44	55.56	75.00	77.78	77.78
27	100.00	94.44	86.11	88.89	100.00	100.00	100.00
28	97.22	94.44	88.89	61.11	100.00	100.00	94.44
29	94.44	91.67	75.00	66.67	97.22	94.44	94.44
30	97.22	88.89	86.11	61.11	97.22	97.22	88.89
31	88.89	77.78	69.44	66.67	86.11	86.11	83.33
32	94.44	88.89	91.67	83.33	100.00	100.00	100.00
33	77.78	66.67	72.22	50.00	88.89	77.78	61.11
34	61.11	50.00	44.44	11.11	69.44	55.56	44.44
35	88.89	77.78	69.44	55.56	94.44	91.67	77.78
36	5.56	11.11	11.11	5.56	25.00	27.78	22.22
37	63.89	66.67	63.89	33.33	88.89	80.56	72.22
38	97.22	94.44	94.44	94.44	97.22	97.22	100.00
39	75.00	72.22	75.00	44.44	83.33	77.78	50.00

40	69.44	66.67	77.78	33.33	88.89	88.89	77.78
41	100.00	94.44	77.78	33.33	97.22	97.22	77.78
42	69.44	69.44	55.56	44.44	83.33	86.11	66.67
43	77.78	75.00	72.22	38.89	77.78	77.78	50.00
44	80.56	77.78	80.56	55.56	86.11	80.56	72.22
45	94.44	86.11	83.33	55.56	91.67	91.67	88.89
46	61.11	55.56	52.78	38.89	77.78	69.44	66.67
47	97.22	97.22	91.67	94.44	97.22	97.22	100.00
48	44.44	38.89	33.33	16.67	44.44	36.11	16.67
49	83.33	72.22	66.67	38.89	86.11	83.33	77.78
50	94.44	91.67	88.89	77.78	97.22	94.44	94.44
51	44.44	47.22	66.67	44.44	52.78	58.33	44.44
52	61.11	52.78	58.33	33.33	69.44	63.89	50.00
53	97.22	97.22	97.22	88.89	100.00	100.00	100.00
54	36.11	50.00	50.00	27.78	36.11	44.44	33.33
55	44.44	47.22	61.11	72.22	50.00	50.00	72.22
56	94.44	94.44	94.44	88.89	94.44	94.44	88.89
57	100.00	100.00	97.22	94.44	100.00	100.00	100.00
58	94.44	88.89	91.67	66.67	97.22	94.44	83.33
59	69.44	58.33	58.33	33.33	72.22	63.89	38.89
60	100.00	100.00	100.00	100.00	100.00	100.00	100.00
61	100.00	100.00	100.00	100.00	100.00	100.00	100.00

62	100.00	100.00	97.22	94.44	100.00	100.00	94.44
63	100.00	100.00	100.00	94.44	100.00	94.44	100.00
64	94.44	88.89	97.22	94.44	100.00	100.00	100.00
65	94.44	86.11	94.44	83.33	94.44	83.33	88.89
66	97.22	94.44	94.44	100.00	97.22	94.44	100.00
67	97.22	100.00	86.11	77.78	100.00	100.00	83.33
68	100.00	100.00	100.00	94.44	100.00	100.00	100.00
69	97.22	100.00	100.00	100.00	97.22	100.00	100.00
70	100.00	100.00	100.00	100.00	100.00	100.00	100.00
71	91.67	97.22	94.44	66.67	94.44	94.44	83.33
72	100.00	97.22	97.22	100.00	100.00	100.00	100.00
73	91.67	91.67	91.67	72.22	97.22	97.22	88.89
74	100.00	97.22	94.44	88.89	100.00	97.22	100.00
75	86.11	86.11	86.11	77.78	97.22	97.22	100.00
76	100.00	100.00	94.44	100.00	100.00	97.22	100.00
77	100.00	94.44	97.22	94.44	100.00	91.67	100.00
78	100.00	100.00	100.00	100.00	100.00	97.22	100.00
79	100.00	100.00	97.22	100.00	100.00	100.00	100.00
80	100.00	88.89	86.11	77.78	100.00	100.00	94.44
81	94.44	88.89	86.11	72.22	94.44	94.44	88.89
82	100.00	94.44	97.22	88.89	100.00	91.67	100.00
83	94.44	97.22	100.00	100.00	100.00	97.22	100.00

84	100.00	97.22	88.89	55.56	100.00	97.22	94.44
85	100.00	100.00	100.00	100.00	100.00	100.00	100.00
86	94.44	91.67	86.11	88.89	100.00	97.22	100.00
87	100.00	100.00	100.00	100.00	100.00	100.00	100.00
88	100.00	100.00	100.00	100.00	100.00	100.00	100.00
89	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Mean	**85.89**	**82.46**	**81.55**	**70.85**	**89.64**	**87.86**	**83.83**