Investigating Grammatical Complexity in Gulf Arabic Speaking Children with Specific Language Impairment (SLI)

Saleh Shaalan

Thesis submitted in partial fulfilment of the requirements for the

Degree of Doctor of Philosophy

University College London

May 2010

I, Saleh Shaalan confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

This is the first investigation of Specific Language Impairment (SLI) in Gulf-Arabic (GA) speaking children. The thesis consists of two main sections, in the first one, I discuss the definitions of SLI and the various theories put forward to account for the deficits seen in this population. I also discuss the importance of cross-linguistic investigations of SLI and why studying SLI in GA may prove useful in testing the accounts of SLI that argue for a general processing deficit vs. those that argue for a domain specific account of SLI. The remaining section of the first part is dedicated to describing the various language tests developed to identify children with SLI in GA. These tests were conducted with approximately 88 typically developing children and 26 children with SLI between the age of 4;6 and 9;4 years old. In the second part of the thesis, I report on two experiments investigating syntactic and phonological complexity in GA speaking children with SLI. The first experiment investigates the comprehension of three types of word orders: a canonical SVO, and two word orders that involve fronting of the direct object (OSV and OVS). Results showed that children with SLI differed from the TD groups on the sentences with fronted NP's, but not on the canonical word order. The second experiment involves a nonword repetition test where syllable length and consonant clusters are systematically controlled to contrast the influence of both phonological short-term memory and phonological complexity. The results are consistent with accounts that argue for a significant role of phonological complexity in NWR and question the "centrality" of phonological capacity in nonword repetition. The final chapter summarises the findings of the thesis and its contribution to theories of SLI in general, and to the study of SLI in Arabic in particular.

Acknowledgments

My first thanks are to all the children who participated in my project in Qatar and their families who allowed me to collect data and conduct experiments at schools and houses. I would like to thank all the schools that gave me access and helped in organising visits and obtaining consents. In particular I would like to thank the administration and staff of Omar Bin Al-Khattab primary school and Al-Arqam Academy for Girls, where I conducted most of my experiments. I would like to thank the many people who helped in data collection in Qatar, especially Mrs. Nadia Al-Hitmi of Omar School, Mrs. Wafaa Sameer of Al Arqam Academy and Mrs. Yusra Medhat. I would like also to thank Ms Amal Al-Kuwari for her help in recording the stimuli.

I would like to thank my supervisors Professor Chris Donlan and Professor Ad Neeleman for their support, help, and encouragement. I sincerely acknowledge the supervision I have received from Professor Heather van der Lely and the useful comments I have got from the following people (in alphabetical order): Professor Renée Béland, Professor Elabbas Benmamoun, Dr. Maria Black, Dr. Sami Boudelaa, Professor Shula Chiat, Dr. Nichola Gallon, Professor John Harris, Dr. Chloe Marshall, Professor JF Prunet and Dr. Jyrki Tuomainen. I enjoyed discussing this work during the viva with my examiners: Professor Neil Smith and Dr. Theo Marinis and I thank them for their invaluable insights. All errors, however, remain mine.

I would like to thank the staff and students of the (former) department of Human Communication Sciences and the UCL Centre for Developmental Language Disorders and Cognitive Neuroscience: Flavia Adani, Mike Coleman, Faisal Haque, Cecile Monjauze, Steve Newton, Angela Pozzuto, Marisa Silveira, Athina Skordi and Outi Tuomainen.

Thank you Mariam, Bashir and Layla for putting up with me throughout my PhD. My brothers and sisters have always been supportive and encouraging and they deserve my sincere thanks.

Finally, no words will suffice to express my gratitude to my mother, to whom I dedicate this work.

Table of Contents

List of Tables

List of Figures

1. Introduction to Specific Language Impairment (SLI)

This thesis investigates one of the most common developmental language disorders, namely, specific language impairment (SLI) in Gulf Arabic speaking children aged between 6 and 9 years old. The aim of this introductory chapter is to provide a definition of SLI based on best clinical and research practices. The criteria adopted for SLI in this project are elucidated and justified, as the project endeavours to ensure that the sample of children is representative of the population of children with SLI typically studied in other languages. This is of crucial importance due to the well-attested heterogeneity of SLI, which is attributed to both selection criteria and underlying phenotypic differences in the linguistic and non-linguistic profiles of children with SLI. This will be followed by an overview of syntactic, morphological, phonological, lexical and pragmatic deficits of children with SLI, with reference to languages such as English and Hebrew, which shares many characteristics with Arabic. The following section deals with the genetic nature of SLI and the many attempts being pursued to provide a clearer link between phenotypic and genotypic profiles of SLI. The chapter concludes with a discussion of the importance of studying SLI with reference to theories of language and cognition in order to show how investigating SLI in a non Indo-European language can be of great importance to both the theory and clinical practice of SLI.

1.1 Defining SLI

SLI is defined by the presence of significant receptive and or expressive language impairments in the absence of cognitive, sensorimotor and social-emotional deficits (Bishop, 1997; Leonard, 1998). Despite reference to some motor deficits in children with SLI (See Hill, 2001 for an overview) and other non-linguistic cognitive tasks (see Leonard, 1998 chapter 5), the argument for the non-specific nature of SLI still lacks convincing evidence, as there is no clear understanding of the overlap between SLI and motor tasks, especially as many of the non-linguistic cognitive tasks are mediated by language skills (Leonard, 1998). As for reports of significant deficits in social-behavioural skills in children with SLI, some researchers found that these “associated socio-behavioural problems [in internalisation, attention, social problems], if they exist at all, are of less magnitude and within the non-clinical range” (Redmond & Rice, 1998, p. 696). These social-behavioural problems appear as consequences of poor communications skills of children with SLI. Some studies found that the development of social-behavioural skills in preschoolers with SLI was not qualitatively different from their TD peers, but as they grew up, behavioural problems might have appeared not only because of the impact of their poor language skills on their socialisation patterns, but it was possible that the emotional impacts of SLI had played a role in exacerbating social-behavioural problems (Goorhuis-Brouwer, Coster, Nakken, & Spelberg, 2004). Studies of attention skills of children with SLI have not come to a conclusive result about the presence of consistent attention deficits in children with SLI (see Gillam & Hoffman, 2004; Hanson & Montgomery, 2002). The few studies available on the non-linguistic nature of SLI have not been conclusive, therefore, most researchers agree that the majority of the difficulties shown by children with SLI fall within the linguistic domains.

Tomblin et al. (1997) conducted the most widely cited epidemiological study of SLI based on a sample of 7,218 5-year old monolingual English speaking children in Iowa. The estimated prevalence rate was 7.4%. The criteria Tomblin et al. (1997) used were representative of those commonly employed by speech-language pathologists to diagnose children with SLI. Tomblin et al. (1997) estimated that 60% of those diagnosed with SLI were male, and 40% were female, though this gender difference in prevalence did not reach statistical significance. Similar prevalence rates were reported by Leonard (1998) and Bishop (1997).

1.2 Criteria for SLI

Researchers have employed various methods to diagnose children with SLI. One of these methods is the use of criteria based on the discrepancy between the child's language performance and what is expected according to his/her intellectual ability, as measured by IQ tests. Researchers have different views on the magnitude of discrepancy between language performance and mental and chronological age expectations (for an overview see Tomblin, Records & Zhang, 1996). Age-language discrepancy is usually estimated based on performance on mean length of utterance (MLU) or standardised language tests. Tomblin et al. (1996) suggested that most studies use cut-off values that are between –1 SD and –1.25 SD of the mean (16^{th}-10^{th} percentile). Despite some reference to difficulties in reaching satisfactory level of congruency between clinicians and researchers in identifying SLI (Aram, Morris,

& Hall, 1993), most researchers use Stark and Tallal's (1981) criteria as a basis for identifying children with SLI. The guidelines are based on the following exclusionary criteria (Stark & Tallal, 1981):

- Children with SLI should have passed hearing screening at 25 dB across the frequencies (250-6000 Hz).
- They should not have any social-emotional problems (e.g., autistic-spectrum disorders, schizophrenia...etc).
- No history of frank neurological deficits.
- They should obtain at least a nonverbal IQ of 85 and above as measured by the Wechsler Preschool and Primary Scale of Intelligence WPPSI (Wechsler, 1963) or Primary Scale of Intelligence Scale for Children-Revised (WISC-R) (Wechsler, 1974).

Stark and Tallal (1981) posited that children with severe phonological deficits or speech-motor deficits should be excluded, a practice that is not being followed by most researchers as many children with SLI are known to have phonological disorders concomitant with deficits in other linguistic domains. Instead, researchers exclude children with oral-motor deficits, such as developmental apraxia of speech. Another exclusionary criterion that is commonly used by researchers and suggested by Aram et al. (1993) is bilingualism. Researchers tend not to include children who come from bilingual homes to avoid an extra confounding variable.

Stark and Tallal (1981) used the discrepancy criterion of an overall language age of at least 12 months below nonverbal IQ or chronological age. Similarly, the diagnostic criteria used by the World Heath Organisation (ICD-10) (1992) employ the language-chronological age discrepancy for diagnosing SLI as they require that the child scores at least 2 SD below the expected age mean on language assessment.

However, the use of both language-IQ discrepancy and age scores is currently not widely practised by researchers in child language disorders. Language-IQ discrepancy was based on the unproven assumption that there are differences between groups of children who show this discrepancy and those who do not. However, both groups seem to demonstrate similar profiles of language deficits (Leonard, 1998). Moreover, there has been no evidence of heritability for this language-IQ discrepancy, as monozygotic (MZ) twins do not show clear discrepancy for language-IQ, despite showing strong concordance for language impairment

(Bishop, 1994; 2004). Moreover, Plante (1998) questioned the use of nonverbal IQ altogether as an exclusionary criterion for SLI, as it seems that children exhibit similar language phenotypes regardless of whether they fall above or below a nonverbal IQ of 85. Bishop et al. (1995) studied the genetic basis of SLI in monozygotic and dizygotic twins and found that language performance of "unaffected" MZ twins was not significantly different from the their affected co-twins on two of four linguistic measures, but the affected co-twins were not included in the SLI group because of the lack of large discrepancy between verbal and non-verbal ability. This was not the case in DZ twins where "unaffected" twins were significantly different from affected ones in both language and non-verbal intelligence measures. However, Bishop (1994) cautions against relaxing diagnostic measures: "In our current state of knowledge, research studies may be best advised to continue to use discrepancy criteria, simply to avoid the possibility of selecting a heterogeneous mix of children with diverse aetiologies" (Bishop, 1994, p. 108).

The use of age scores for language measures (or language age) can be criticised for lack of theoretical foundation. The use of age scores can result in misrepresentation of participants' characteristics, as children of the same age are known to display a wide range of abilities. The use of language age can wrongly imply that a 1-year lag in a 2-year-old child is equivalent to 1-year delay in an 8-year old child (Paul, 1995). Therefore, matching on age may result in serious psychometric misrepresentations (see also Bishop, 1997; Botting & Conti-Ramsden, 2004; Kamhi, 1998; Lahey, 1990; Plante, 1998). Therefore, the use of standardised scores, z-scores or percentile ranks is much preferred.

Therefore, instead of resorting to language age-IQ discrepancy as a criterion for SLI, a preferred and more accurate cut off point is the one proposed by some researchers (Leonard, 1998; Leonard, 2003; Tomblin et al., 1996), namely a test score of 1.25 standard deviation (SD) below the mean for the individual's age on a comprehensive language test that covers major areas of grammatical and lexical development or -1.25 SD on two or more language subtests. Tomblin et al. (1996) proposed this criterion as it corresponded to the clinical judgements of speech-language pathologists. This also corresponds to the criterion suggested by Paul (1995), which is having a score that is below the tenth percentile.

Apart from these documented deficits in language abilities, some researchers use the presence of functional limitations on the child's ability to communicate effectively and in different social situations as another criterion in the diagnosis of SLI. According to this view, the deficit is defined not only in terms of "statistical abnormality, but [also] in terms of disability" (Bishop, 1997, p.23)

Despite these differences on which criteria should be used to diagnose SLI given its nature and the notorious difficulties in establishing reliable sub-groups within SLI (Bishop, 1994; Conti Ramsden, Crutchley & Botting, 1997; Miller, Kail, Leonard & Tomblin, 2001; Rapin & Allen, 1987; Tager-Flusberg & Cooper, 1999; Tomblin & Zhang, 1999), there is an overall agreement that SLI is a reliable and stable diagnosis (Conti-Ramsden & Botting, 1999). The heterogeneity observed in SLI reflects the underlying variability of language skills rather than being an indication of measurement errors (Bishop, 1994; Botting & Conti-Ramsden, 2004). For example, Botting and Conti-Ramsden's (2004) classification of the different cluster groups of children with SLI showed that despite children's movement between different subgroups of SLI, possibly as a function of age and performance on different tasks, their language deficits were stable. A recent attempt to find subgroups was based on the presence of concomitant phonological impairments alongside morphosyntactic impairments. It concluded that the subtype of children with language impairment only performed better than the two other subtypes of children with phonological SLI (with or without final consonant deletion) on measures of finite morpheme production and syntactic complexity (Haskill & Tyler, 2007).

Overall, despite the heterogeneity of SLI and the difficulties in finding subgroups, most researchers agree on some criteria to diagnose children with SLI based on having significant impairments in receptive and/or expressive language skills and meeting some exclusionary criteria.

1.3 An Overview of the linguistic characteristics of SLI

Children with SLI have linguistic profiles characterised by the presence of significant deficits in expressive and/or receptive verbal communication. In the following, the major deficits shown by children with SLI in the areas of syntax, morphology, phonology, lexical-semantics, and pragmatics are reviewed.

1.3.1 Syntactic and morphological deficits in children with SLI

Most studies of children with SLI report significant deficits in areas of syntax, grammatical (inflectional) morphology and syntax. These two grammatical components are undeniably the areas most investigated in the linguistic profiles of children with SLI (Bishop, 1997; Leonard, 1998).

Many syntactic structures have been implicated in the linguistic profiles of children with SLI. Children with SLI have well-documented problems in the use and comprehension of wh-questions (Deevy & Leonard, 2004; van der Lely, 1998; van der Lely & Battell, 2003); comprehension of relative clauses (Adams, 1990; Friedmann & Novogrodsky, 2004; Stavrakaki, 2001) and comprehension of passives (Bishop, 1979; 1997; van der Lely, 1996; van der Lely & Harris, 1990). There is, however, disagreement as to whether children with SLI use different syntactic categories (e.g., nouns, verbs, adjectives) in a distribution similar to those of typically developing children (Leonard 1998) or they have more significant problems with verb use.

Children with SLI have shown extensive deficits in various areas of grammatical morphology. Many studies have reported a host of problems in the use of past tense (Gopnik & Crago, 1991; Rice, Wexler & Cleave, 1995; van der Lely & Ullman, 1996), copula and auxiliary be (Leonard, 1992; Rice & Wexler, 1995) and third person singular –s (Leonard, 2003; Rice & Oetting, 1993; Rice & Wexler, 1995). Rice and colleagues have argued that deficits in tense marking constitute the main clinical marker of SLI in English (Rice, 2007; Rice & Wexler, 1995, 1996a, 1996b). Rice and colleagues have found that morphemes marking tense (e.g., third person singular –s, past tense –ed, copula *be*, and auxiliaries *be* and *do*) constitute the core of the morphosyntactic deficits in English speaking children with SLI, while morphemes such as plural –s do not. The performance of children with SLI on these tense and agreement markers is significantly worse than both chronologically matched and language matched typically developing children and across different tasks (production and grammaticality judgment) (Rice, 2007; Rice & Wexler, 1996b;

Rice, Wexler, & Redmond, 1999). The findings of this account, known as the Extended Optional Infinitive (EOI), have been replicated in many other languages (see section 2.2.3 on EOI). In Dutch speaking children, de Jong (1999; 2003) reported significant deficits in tense and agreement in Dutch children with SLI, though he found that commission (substitution) errors were more common than omission errors in structures that required tense and agreement. Clahsen (1989) reported that German speaking children with SLI displayed significant deficits in using finite verbs in German due to difficulties in establishing agreement relations between subject and verb. Moreover, he studied 15 English and German children with SLI and concluded that these children showed more deficits in subject-verb agreement. Deficits in tense and agreement have been among the most widely implicated impairments in SLI and studies across many languages have found that children with SLI have difficulties using inflectional morphology to mark tense and agreement.

Studies of morphosyntactic deficits in children with SLI in Arabic and Hebrew provide divergent results on whether these children have significant deficits with inflectional morphology. In Hebrew, Dromi and her colleagues (Dromi, Leonard, & Shteiman; 1993; Dromi, Leonard & Blass, 2003) found that most inflectional morphemes did not pose special difficulty for Hebrew-speaking children with SLI, while in Arabic children with SLI showed deficits in tense and agreement markers (Abdalla, 2002).

In Hebrew, Dromi and her colleagues examined the production of different verb, noun and adjective inflections (e.g., present/past tense inflections, noun plurals, adjective-noun agreement) in 15 children with SLI and compared them to both age matched and MLU matched children (Dromi et al., 1993). They showed that children with SLI performed as high as the language control children on person, gender and tense inflections. In another study, Dromi and colleagues (Dromi et al., 2003) used two methods of assessing verb forms in Hebrew speaking children with SLI (HSLI), namely, spontaneous language samples and elicitation of verb forms. Both method showed that HSLI children performed similarly to the MPU-matched control (children matched on Morpheme Per Unit: an adaptation of the MLU developed by Dromi and Berman (1982) for Hebrew language samples) on a wide range of finite and infinitive verb forms. However, they had difficulty with some complex verb patterns (Binyamin). Both methods showed that children did not find difficulty producing appropriate agreement in present tense, while they had difficulties in agreement in past tense. Therefore, Dromi et al (2003) concluded that both naturalistic and elicited data showed that HSLI children do not have general deficits in verb morphology despite having problems with some complex forms that involve phonological or semantic complexity.

In Arabic, Abdalla (2002) examined spontaneous language samples of ten Saudi Arabic speaking children with SLI aged between 4;3 and 5;0 years old and found that they had difficulties using tensed verbs (past and present) in comparison to both age and language(MLU) matched groups of typically developing children. Moreover, she examined the production of subject-verb agreement markers in the spontaneous speech of these children and compared their performance to typically developing age and MLU matched groups. In Arabic, subject agrees with verb in person, number, and gender. Due to lack of enough tokens in spontaneous speech, all these forms were collapsed into one agreement factor. Analysis of results showed that children with SLI used correct verb agreement markers 77% of the time, while both age and language-matched TD groups were performing near ceiling (93% and 99.80% respectively). There was no significant difference in performance on past or present agreement inflections. However, there was a significant difference on both person and gender, with children with SLI performing better on first person and masculine forms. So, while in Hebrew, children with SLI did not show significant deficits in tense and agreement, Arabic speaking children with SLI in Abdalla's (2002) study presented with problems in both agreement and tense, compared to their controls.

This section presented an overview of some of the syntactic, morphological, and morphosyntactic deficits in children with SLI. These deficits are considered the main characteristics of SLI in many languages, such as English, where children were found to have difficulties with tense, agreement, and production of inflectional morphemes. Manifestations of these deficits differ across languages. For example, studies of verb morphology in Arabic and Hebrew SLI had different views on whether these children have significant deficits in inflectional morphology.

1.3.2 Phonological deficits in children with SLI

While more attention has been paid to morphosyntactic problems in children with SLI, it is widely acknowledged that phonological impairments are an essential part of SLI and can be used reliably to identify children with SLI (Tager-Flusberg, 2004). In his review of phonological deficits of children with SLI, Leonard (1998) explained that these children exhibit more errors compared to their typically developing age controls on different measures of phonological abilities, such as segment and feature accuracy and phonological processing. These children show, however, typical stages of phonological processes (e.g., consonant cluster deletion), especially in their early years of development (Lahey, Flax, & Schlisselberg, 1985). Children with expressive SLI show lower percentages of consonant correct (PCC), have a smaller phonetic inventory and their overall intelligibility is less than their controls, despite not showing a delay in their vocal development (Rescorla & Bernstein-Ratner, 1996; Roberts, Rescorla, Giroux, & Stevens, 1998). Shriberg et al. (Shriberg, Tomblin, & McSweeny, 1999) calculated the sample-wide comorbidity of language impairment and speech disorders based on the epidemiological study conducted by Tomblin et al. (1997) and estimated it to be 2%. They estimated that 5-8% of children with persistent SLI had speech delay (Shriberg et al., 1999). They attributed this relatively low level of comorbidity, when compared to previous studies they reviewed, to possible effects of contemporary early intervention programs that usually target phonological disorders, especially since most of the studies they reviewed were based on data collected a decade before the Tomblin et al. (1997) epidemiological study (Shriberg et al., 1999). Another possible explanation is differences in the classification systems used to diagnose children with speech disorders.

When compared to children of equivalent levels of grammatical development and phonetic inventory (language controls), the scores of English-speaking children with SLI were inferior on measures of syllable structure (final consonant and weak syllable deletion) and consonant accuracy (Bortolini & Leonard, 2000). Similar findings were reported by Bortolini and Leonard (2000) when they looked at structural constraints in Italian-speaking children with SLI. They found that these children had reduced phonological skills even when compared to MLU and phonetic-inventory matched typically developing children. They also showed that despite some possible influences of grammatical morphology on phonological errors, there was evidence of phonological deficits that were independent of grammatical influence.

Children with SLI have significant impairments in nonword repetition (NWR) tasks (Bishop, North, & Donlan, 1996; Dollaghan & Campbell, 1998; Ellis Weismer et al., 2000; Gathercole & Baddeley, 1990) which are used to tap phonological processing skills of children with SLI and their phonological short term memory. These deficits have been found to be resilient and seemed to persist even when performance on standardised language tests fell within average range (Conti-Ramsden, Botting, & Faragher, 2001; Montgomery, 1995a). See also Chapter 5.

There are no known studies of phonological skills of Gulf Arabic speaking children with SLI; however there are some studies investigating Hebrew speaking children with SLI. Owen and colleagues examined the phonological abilities of Hebrew speaking children with SLI (HSLI) and their interaction with morphological difficulties and revealed that these children had deficits in phonological structures that have no morphological functions and those that affect morphology (Owen, Dromi, & Leonard, 2001). Children with SLI aged 4-6 years were presented with storybooks to elicit production of some specific verb forms that required morphophonological transformations, e.g. "Yoav went to the pool because he wanted to swim (*lisxot).* He entered the pool and___ (*saxa* "swam") (Owen et al., 2001, p.330). The authors analysed responses that involved phonological changes that had no effects on morphological accuracy of the word and those that had some bearing on morphology. Children with SLI showed higher percentages of substitutions of marked phonological structures such as consonant clusters and glottal stops when they were compared to their age and MLU controls. Overall, Hebrew speaking children with SLI produced more phonological errors when they were compared to the two control groups. Moreover, children with SLI showed higher percentage of morphological errors involving more complex phonological structures e.g., simplifying phonologically complex Binyamin (morphological patterns) by substituting a pattern like *mexabek* (hug) for the more complex *mitxabek* (hug each other) which involves consonant clusters (Owen et al., 2001). It seems that the poor phonological skills of Hebrew speaking children with SLI could have deleterious effects on their vulnerable morphological abilities.

1.3.3 Lexical-semantic deficits in children with SLI

Most children with SLI are slow in their lexical development and show less lexical diversity compared to their TD peers (Leonard, 1998; Owen & Leonard, 2002; Thal, O'Hanlon, Clemmons, & Fralin, 1999). They acquire words and start to combine words later than their typically developing peers and most of these children meet the general description of being late talkers (Leonard, 1998; Leonard, 2003). Thal et al. (1999) have found that the lexical diversity of children with SLI at age 2;2 years was comparable to typically developing children at age 1;4 based on a measure of parental report. This lack of lexical diversity seems to persist consistently across different ages (see also Leonard, Miller, & Gerber, 1999 and Owen & Leonard, 2002).

This protracted lexical development in children with SLI was investigated in order to analyse the underlying deficits in word learning. While typically developing children acquire words at a very high rate, children with SLI have delayed lexical development compared to their typically developing peers (Leonard, 1998). Studies conducted by Leonard and colleagues (Leonard & Schwartz, 1982; Schwartz, Leonard, Messick, & Chapman, 1987) have shown that, when compared with children of the same size of lexical inventory, children with SLI are not different in terms of learning novel words. However, when compared to chronological age controls, children with SLI showed reduced ability in associating a novel word to unfamiliar objects in production, but performed similarly to age controls on comprehension tasks (Dollaghan, 1987). Dollaghan (1987), however, demonstrated that language impaired children do not differ from typically developing children in fast mapping. Dollaghan (1987) defines fast mapping as "the initial step in lexical acquisition, in which a listener rapidly constructs a representation for an unfamiliar word on the basis of a single exposure to it" (p.218). In her study, children with language impairment were able to associate a novel word to a referent and comprehend the novel word. However, they were less efficient in producing these novel words. The cause of these difficulties in lexical development is not well understood yet as there are disagreements on whether children with SLI show difficulties in learning new words. Other researchers have shown that children with SLI are less efficient at verb and noun learning. Riches and colleagues (2005) showed that children with SLI were sensitive to the frequency of novel verb presentations and the interval of these presentations (spacing) compared to language-matched children and they exhibited poor retention of these novel verbs. Gray (2003b) showed that children with SLI needed more presentation frequency to learn novel nouns compared with control children in both production and comprehension tasks. The finding that children with SLI do not benefit from syntactic cues in lexical acquisition (Rice, Cleave & Oetting, 2000) is not surprising considering the significant problems children with SLI find in acquiring their native morphosyntactic system. Conti-Ramsden (2003a) attributed poor lexical development in SLI to their poor processing skills. These children require more frequent exposure to novel items before they can acquire them.

Overall, children with SLI typically lag behind their peers in lexical development and across lifespan. Researchers, however, disagree about what causes this protracted lexical development and whether children with SLI have deficits in acquiring novel nouns and verbs.

1.3.4 Pragmatic skills in children with SLI

Pragmatics refers to the use of language in a social-communicative context. Therefore, pragmatic disorders refer to difficulties in “the recognition and application of the social rules for language and discourse” (van Balkom & Verhoeven, 2004, p. 283). Researchers have mixed views on whether children with SLI have primary deficits in their pragmatic skills. Van Balkom and Verhoeven (2004) suggested that there are two competing approaches to the nature of pragmatic language disorders, namely the modular approach and the functional approach.

The modular approach assumes that pragmatics is an independent component of the language system, alongside syntax, morphology, phonology, and semantics. Hence, children with primary pragmatic impairments are expected to have intact linguistic structures in other linguistic components. However, some studies have shown that pragmatic deficits in children with SLI are linked to lexical or grammatical difficulties (see van Balkom & Verhoeven, 2004 and references therein). The functionalist approach, on the other hand, views pragmatics as a system for "linking linguistic forms to discourse functions"(van Balkom & Verhoeven, 2004, p. 287). It seems that the functionalist approach is more capable of capturing the pragmatic deficits seen in children with SLI who seem to display a wide range of appropriate communicative functions but are hindered by their deficits in linguistic structures. This was demonstrated in van Balkoma and Verhoeven's (2004) study of pragmatic deficits in higher and lower functioning children with SLI. They found that children with SLI evinced appropriate social-communicative functions that were not different from their peers, but showed excessive use of some atypical linguistic forms (e.g. ellipsis, imitations, self-repetitions) where their pragmatic functioning was influenced by their limited linguistic knowledge. Van Balkom and Verhoeven (2004) found that, in comparison to typically developing children, children with SLI showed a highly significant number of communication breakdowns, decreased discourse coherence and an increased number of parental repairs. This hypothesis of the influence of linguistic forms on pragmatic skills was strengthened by the fact that children in the higher functioning SLI group scored significantly better than the lower functioning SLI group on measures of pragmatic functioning.

As for the types of pragmatic difficulties seen in children with SLI, it has been shown that although children with SLI evince some difficulties in certain aspects of pragmatics, their performance is comparable to those with the same level of language functioning (Fey & Leonard, 1983). Children with SLI are known to have problems in some areas of social communication, for example they tend to participate in fewer peer interactions compared to typically developing children, and their initiatives are more likely to be ignored than the initiatives of their typically developing peers. However, children with SLI show similar levels of communication acknowledgements to peers' initiatives (Hadley & Rice, 1991).

These difficulties have been shown to vary according to the modality of linguistic deficits present in children with SLI. Children with SLI show different pragmatic difficulties depending on whether they have receptive language impairments (Craig & Evans, 1993). Children with receptive and expressive SLI (R-E SLI) used less cohesive strategies, showed reduced ability to take turns in conversations and used more ambiguous references when they were compared to children with expressive SLI only (E-SLI). However, Craig and Evans (1993) showed that children with E-SLI were different from their chronologically matched group, but not from their language-matched controls, in their use of cohesive strategies. Therefore, they cautioned against combining children with E-SLI and those with R-E SLI when studying pragmatic skills in children with SLI.

Craig and Evans (1993) reviewed the peer interaction skills of children with SLI and noted that they have difficulty with the following areas of social communication:

- Initiation of social communication with peers and adults and joining of interactions. They also tend to prefer to initiate interaction with adults.
- Dispute management: Children with SLI use fewer verbal strategies to resolve conflicts than their typically developing peers.
- Responsiveness to communication partners: though their responsiveness is not significantly different from their peers, it has inadequate properties compared to their peers.
- Assertiveness: Children with SLI rarely use turn-interruptions to secure their speaking turns.
- Discourse adjustments: though children with SLI are able to modify their communication style to suit their communication partners, they are less efficient at monitoring the conversation and recognizing when some adjustments are needed.

Tomblin and colleagues (Tomblin, Zhang, Catts, Ellis Weismer, & Weiss, 2004) showed that 85% of children with SLI have better social communication skills than semantic or syntactic skills. Rice and Warren (2004) argue that the relatively mild pragmatic difficulties displayed by children with SLI could be secondary to their primary language impairments. They, however, explained that this is still a contentious issue and more research needs to be conducted.

Schaeffer (2003) takes a Chomskyan perspective to language and argues that there is a separate pragmatic module of the language system that, although it interacts with the lexicon and the computational system, exists as an independent system. She tested this hypothesis in Dutch children with SLI. By investigating the interaction of pragmatics and syntax in object scrambling in 20 Dutch speaking children with SLI (ages 4;2-8;2). Scrambling involves reordering of sentence constituents and object scrambling in Dutch involves both pragmatic and syntactic operations. Results showed that children with SLI performed well (96% accuracy) on referential objects (when speaker and hearer talk about a particular member of a class, e.g., a particular car they both saw vs. non-referential nouns, when they refer to cars in general) despite lagging behind in their grammatical development. While their overall grammar was comparable to typically developing children aged 2-3 years, they showed age-appropriate interface pragmatics. Younger typically developing children, on the other hand, do not develop interface pragmatics before the age of 3;4 years (Schaeffer, 2003). Therefore, Schaeffer (2003) concluded that "pragmatic principles….develop as a function of age, rather than as a function of grammar developmental stage" (p.147).

van der Lely (2003) argues that children with pragmatic SLI are characteristically different from the G-SLI children she has studied. Children with G-SLI do not have problems with conversational inference or referential pronouns. Therefore, the co-occurrence of pragmatic and grammatical deficits (syntactic, morphologic and phonological) should not preclude us from concluding that they are independent components that may interact with each other in certain situations.

Overall, researchers disagree as to whether pragmatic deficits in children with SLI exist as an independent deficit additional to the linguistic deficits observed in other linguistic components or they are secondary to the primary deficits in the language system.

1.4 Genetic nature of SLI

The study of the genetic basis of SLI is still in its infancy, due to the complex relationship between the genotypic and phenotypic characteristics of human language. Twin and familial studies have been two of the most instrumental methods used to explore the genetic basis of SLI. It has been noted that concordance rates of SLI were higher in monozygotic (MZ) or identical twins than those in dizygotic (DZ) or fraternal twins (Bishop, North, & Donlan, 1995). A meta-analysis of 10 twin studies found that concordance rate of language and reading disorders (i.e. both twins have language/reading disorders) was very high (almost two-third of the concordance); and in the case of twin studies of language impairment the concordance rate varied between 25% and 100% depending on what aspects of language were tested (Stromswold, 2001). Conti-Ramsden et al. (2006) used two different assessment procedures (direct assessment tests and interviews) and found that the rate of familial aggregation of language and literacy disorders was 35% on both procedures. This is identical to the number cited by Stromswold (1998) in her review of seven familial aggregation studies. She reported an incidence rate of 35% in families of children with language impairment, while the rate for control families was 11%.

The genetic basis of SLI is further supported by studies of parents of children with SLI, who mostly do not present with existing speech and language disorders, though their overall performance lags behind parents of children with no history of language deficits. In a study that looked at the performance of parents of children with SLI on a group of language and literacy measurements, these parents (n=34) performed significantly worse than parents of typically developing children (n=33) (Barry, Yasin, & Bishop, 2007). The prevalence of history of language impairment in parents of children with SLI was 33%, while it was 6% in parents of TD children. Among the different cognitive and language tests used with these parents, it was found that nonword repetition test provided the best discrimination and sensitivity for familial language impairment, though the authors cautioned against assuming that all children with SLI had deficits in nonword repetition as there is evidence against this (See for example, Catts, Adlof, Hogan, & Ellis Weismer, 2005)

With respect to which aspects of the linguistic profile of children with SLI are more affected by genetic inheritance, there is some evidence of inheritance of grammatical impairments in children with SLI. A study by Rice and colleagues found a higher incidence of language impairment in families of a group of children with SLI, who exhibited significant deficits in marking tense, when these families were compared to a control group of families of unaffected children (Rice, Haney, & Wexler, 1998). This was also reported in a previous study by van der Lely and Stollwerck (1996), who found an autosomal dominant pattern of inheritance in a group of 12 children with grammatical-SLI (G-SLI), aged between 9;0 and 12;0. They revealed that the incidence rate of positive familial history of language impairment in parents and siblings of children with G-SLI was 78%, while the control group had a rate of 29%. This familial aggregation has been documented in SLI probands and twin studies (Bishop, Adams, & Norbury, 2006; Conti-Ramsden, Simkin, & Pickles, 2006; Rice et al., 1998; Stromswold, 1998; Tallal et al., 2001). Bishop et al. (2006) found evidence of heritability of nonword repetition difficulties and verb inflection deficits; however, no evidence supported the heritability of auditory deficits in children with SLI (Bishop et al., 1999). When the impact of verb inflections and nonword repetition on the heritability of language impairment was examined, weak genetic overlapping was obtained between phonological short-term memory (PSTM) and verb inflections despite apparent strong phenotypic association that might had been caused by selection bias or some common environmental factors (Bishop et al., 2006). This poor genetic overlapping indicates that poor PSTM cannot be the underlying cause of tense marking deficits in children with SLI.

Interest in the genetic basis of abnormal language acquisition arose with the widely-reported discovery of the first gene to be implicated in speech and language disorders, known as FOXP2, which was located on chromosome 7 (7q31) (Fisher, Vargha-Khadem, Watkins, Monaco, & Pembrey, 1998; Lai, Fisher, Hurst, Vargha-Khadem, & Monaco, 2001) in a four-generational British family, known as the KE family. Earlier reports indicated that this family's main deficits were in linguistic domains (Gopnik, 1990), but later investigations have shown that there were other general cognitive and motor deficits in members of the KE family such as having low IQ and oral and verbal apraxia (Hurst, Baraitser, Auger, Graham & Norell, 1990; Shriberg et al., 2006; Vargha-Khadem, Gadian, Copp & Mishkin, 2005; Vargha-Khadem, Watkins, Alcock, Fletcher & Passingham, 1995). However, Marcus and Fisher (2003) maintain that the cognitive ability of family members is not affected by the FOXP2 gene, as there is an overlap with the IQ scores of unaffected members. It seems that verbal abilities of affected members are highly influenced by genes, unlike their non-verbal abilities. Shriberg et al. (2006) reported a case of a mother and her daughter who had mutations on FOXP2 and present with severe speech and oral motor disorders (spastic dysarthria and apraxia of speech) that resemble the motor speech disorders reported in the KE family. Other reports, however, have indicated that this region on 7q31 might be related to language impairment in children with SLI and other developmental disorders, such as autism

(O'Brien, Zhang, Nishimura, Tomblin, & Murray, 2003; for a different point of view, see Newbury et al., 2002)

In addition to the FOXP2 gene on chromosome 7, the other loci linked to language abilities have been found on chromosomes 2, 3, 16, and 19 (for an overview see Fisher, Lai & Monaco, 2003). The SLI consortium (2004) conducted a full genome scan of 840 children with SLI and their 184 families, and reported a significant linkage of six language and reading ability measures to chromosome 16 (16q) and chromosome 19 (19q). The language measures used were the receptive and expressive parts of the Clinical Evaluation of Language Fundamentals (CELF) test (Semel, Wiig & Secord, 1996) and a nonword repetition test. These same loci were also linked to some reading skills tasks. These links were not present across all tasks: chromosome 16 showed significant interaction only with the nonword repetition task, while the expressive parts of the CELF were correlated with chromosome 19 (The SLI Consortium (SLIC), 2004). Newbury et al (Newbury, Bishop & Monaco, 2005) have suggested that an area on chromosome 16 (known as SLI1) might be implicated in deficits in phonological short-term memory in children with SLI. However, there have been no clear links between the phenotypes and genotypes of SLI. It is believed that the link is very complex and may involve the interaction of many genes amongst each other, in addition to their interaction with the environment (The SLI Consortium (SLIC), 2004).

The debate in domain-general versus domain-specific explanations of the behavioural symptoms of SLI extends as well to the theories of genetic basis of SLI. Researchers like Plomin and Kovas (2005) argue that language impairments share with other learning disabilities a set of ‘general genes’ that are responsible for many of their symptoms. This explains, according to Plomin and Kovas (2005), the commonalities between many learning disorders. Proponents of domain-specificity, on the other hand, contend that there are ‘specific genes’ that are linked to linguistic abilities in both normal and abnormal language acquisition (Stromswold, 2001). Based on her analysis of the contributions of different environmental factors, Stromswold (2006) argues that the linguistic knowledge is innately programmed and controlled by specialised neural circuitry. She suggests that the difference found in the linguistic performance of MZ twins (who share 100% of the alleles) can be accounted for by pre- or perinatal/ biological factors (e.g., whether the MZ twins share one placenta and amniotic sac, their birth weight…etc), while postnatal factors (e.g., home environment) account for the cognitive/psychosocial aspects of development. Hence, she concludes that language is affected by perinatal/biological factors, while cognitive development is more influenced by postnatal factors; thus supporting the nativist view and arguing against the empiricist view of language acquisition (Stromswold, 2006).

1.5 Why study SLI

There are various reasons for the increased interest in studying SLI. The most obvious reasons are clinical ones; SLI with a prevalence rate of 7% is considered one of the most common childhood disorders, and is closely related to other reading and learning disorders, such as dyslexia. Therefore, this group of children and their families will benefit from any insights on the cause, characteristics and remedial approaches of this disorder. SLI prevents many individuals from reaching higher academic goals, and it negatively affects their communication skills, jeopardising their careers in modern societies where successful careers are increasingly becoming reliant on communication skills. SLI may be linked to subsequent behavioural problems caused by inefficient communication skills. While young preschool children with SLI do not evince behavioural problems more than their age peers do, the incidence of behavioural problems increases as children with SLI get older. One study found that 48% of children with SLI were reported by parents and/or teachers to have significant behavioural problems at 8-12 years old and the authors cautioned against not dealing with the impact of language disorders on the children's social-communicative well being (Goorhuis-Brouwer et al., 2004).

SLI is of great significance to those interested in language acquisition and the nature of human language and how it is represented in the mind/brain. SLI has fascinated researchers since first reported cases in 1800's (see Leonard, 1998) due to the discrepancy between language skills and IQ that are not explained by any neurological, sensorimotor, or social-emotional impairments. The existence of an impairment that is specific to language components lends great support to theories claiming there are domain-specific modules that handle linguistic representations (Fodor, 1983). While domain-general theories of cognition necessarily assume that

SLI should not exist because the deficits are caused by general processes that are shared with other cognitive processes; these theories face the challenge of explaining the gap between linguistic and non-linguistic abilities of children with SLI assuming that they all rely on the same cognitive operations (Leonard, 1998). Deciding whether the underlying causes of SLI are based in linguistic operations or general cognitive processes will undoubtedly reflect on the assessment and intervention approaches employed to help children with SLI.

Researchers who study the genetic basis of language are interested in SLI as it could shed light on understanding genetic influences on language acquisition in general and abnormal language acquisition in particular. However, any genetic investigation of SLI should be based on clear links between phenotypes and genotypes. Therefore, one of the important reasons of studying SLI is to reach a reliable phenotype for normal and abnormal language acquisition, without which proper investigations of the genetic basis of language cannot be conducted. Therefore, SLI researchers endeavour to describe specific phenotypic characteristics that are shared by all languages. Rice (1996, pp. xviii-xxiii) lists some criteria for a phenotype of SLI:

- It must be consistent with universal features of language.
- It must yield to reliable measurement.
- It should differentiate affected from non-affected individuals.
- It shows variation where none is expected.
- It is relatively resistant to environmental effects.
- This behavioural phenotype applies over the age span.
- It can be specified in terms of biological mechanisms and functioning.

Therefore, researchers of SLI have been striving to identify clinical markers of SLI. A clinical marker is an “aspect of the linguistic functioning that may uniquely define the phenotype of the disorder” (de Villiers, 2003, p. 247). This is a crucial step towards tracing genetic contributions to normal and abnormal acquisition of language. It is suggested that studying processing-based phenotypes rather than knowledge based measures can lead to better understanding of the phenotypic characteristic of SLI (Bishop, 2004; Campbell, Dollaghan, Needleman, & Janosky, 1997), especially since the former are less sensitive to previous knowledge and socio-economic status. A qualitative marker that has shown great promise for children with SLI is nonword repetition (NWR), which has a high rate of heritability and low correlation with environmental factors (Bishop et al., 1996; Newbury et al., 2005). Bishop et al. (1996) showed that NWR provides an excellent behavioural marker due to its high sensitivity, as has been desmonstrated in studies of parents of children with SLI (Barry et al., 2007). Nonword repetition has shown poor or nonsignificant correlation with nonverbal IQ (Bishop et al., 1996) or cultural background (Campbell, et al., 1997; Ellis Weismer et al., 2000; Rice, 2000). It has been found to dissociate from cognitive skills as shown in children with Williams syndrome (Karmiloff-Smith et al., 1997). Another possible candidate for a phenotype of SLI that is commonly implicated in English speaking children with SLI is the difficulty with tense marking or verb inflection (Rice, 2003; 2007; Rice & Wexler, 1996b; Rice et al., 1995; Rice, Wexler & Hershberger, 1998). Bishop and colleagues found that both NWR and verb inflection have strong hereditary basis, and despite the presence of overlapping, they are distinct factors (Bishop et al., 2006). However, these two clinical markers do not present a complete explanation for the deficits seen in children with SLI. For example, despite the high sensitivity of NWR, it has poor specificity, as NWR tests showed that some children with SLI performed within normal range on tests of NWR. Conti-Ramsden (2003b) found that despite having good specificity, both NWR and tense inflection had less than satisfactory sensitivity, with 59% for the former and 52% for the latter. Moreover, both deficits in NWR and tense inflections did not stand cross-linguistic examination, as both have been found to be relatively intact in other languages as discussed in the following section. Therefore, researchers continue to search for a good clinical marker (or markers) for SLI in English and other languages that can serve as a reliable indicator of the language deficits in this population.

1.5.1 Why study SLI cross-linguistically?

Studies of SLI across languages have enriched our understanding of the nature of SLI and how it manifests differently across languages. Until the 1990's most of the studies investigating SLI were conducted with English speaking children. When researchers started investigating SLI in other languages, insightful contributions have been added. Many languages have properties that are useful to investigate theories initially proposed for English speaking children with SLI. In some cases, new theories of SLI were conceived on the basis of data from other languages (e.g., the grammatical agreement deficit, Clahsen, 1989 and the sparse morphology account; Dromi et al., 1993; Leonard, 1998). For example, most studies of Germanic languages have supported the proposal for a deficit in marking finiteness (see Rice, 2007). However studies of Romance languages (Italian and Spanish) did not support the initial proposal by Rice et al. (1995) and this lead to a revision in the EOI theory (Wexler, Gavarro & Torrens, 2004) to account for these findings. Another example is Crago and Paradis' finding that children learning Inuktitut have significant problems in mastery the inflectional morphology of this morphologically-rich language (Crago & Paradis, 2003), a conclusion that does not support Leonard and colleagues' proposals (Bedore & Leonard, 1998; 2005; Bortolini, Casalini & Leonard, 1997; Dromi et al., 1993; Leonard, 1998) that were based on Italian, and Hebrew. While most studies of nonword repetition of English speaking children with SLI documented that children with SLI had deficits in phonological short term memory (Archibald & Gathercole, 2006a; 2006b; Bishop et al., 1996; Dollaghan & Campbell, 1998; Ellis Weismer et al., 2000; Gathercole & Baddeley, 1990), the study of Cantonese speaking children with SLI by Stokes et al. (2006) revealed no significant difference between children with SLI and age and language controls on nonword repetition.

In summary, investigations of SLI in other languages, not only helped in characterising this language impairment in other languages, but they increased our understanding of the nature of SLI across languages, including English. Cross-linguistic investigation of SLI has helped modify and improve some of the theories explaining SLI in English and other languages, and some theories were conceived and developed based on how SLI manifests in other languages.

1.5.2. Why study SLI in Arabic

Arabic can be a potentially valuable testing ground for theories of SLI. Arabic is a morphologically rich language, characteristically different from other morphologically rich European languages (e.g., Italian, French, Spanish) that have been used to investigate some accounts of SLI. Like other Semitic languages, Arabic is a nonconcatenative, root and pattern language with complex interaction between syntax, morphology, and phonology. Arabic has a relatively flexible word order where SVO and VSO are commonly used, with the former being the neutral order in modern dialects while VSO is considered the unmarked word order in classical Arabic (CL) and Modern Standard Arabic (MSA). Moreover, Arabic has a group of clitic pronouns that attach to different categories (nouns, verbs, and prepositions). All direct and indirect object pronouns are cliticised and there are no freestanding object pronouns like in European languages. Unlike clitics in Romance languages that can precede or follow verbs, Arabic clitics always affix to the end of the category they attach to and are commonly placed at the end of the utterance, and thus are in a more perceptually salient position.

Tense distinctions in Arabic are quite different from English and other European languages. Arabic has a distinction based on whether the action is completed or not, so the distinction is more aspectual than tense-based and more tense distinctions could be made by using optional phrases.

Another interesting characteristic of Arabic is the phenomenon of diglossia whereby spoken dialects of Arabic exist alongside the Modern Standard Arabic (MSA) (Ferguson, 1959). These spoken dialects have syntactic, morphological, phonological, and lexical properties that are distinct from MSA and they are used in most social situations, while MSA is generally used in written or formal situations (i.e., formal speeches, news, some TV and Radio programmes). Therefore, while children grow up speaking and listening to spoken dialects, they will be exposed to MSA mainly when they start going to school, although even before starting formal schooling most children watch some foreign cartoons translated into MSA.

Arabic has many characteristics that set it apart from most of the languages studied so far in the SLI literature. Some of these properties will be investigated in this thesis, such as variable word order and the effects of roots and patterns on NWR. Therefore, this investigation of SLI in Gulf Arabic might contribute to the theory and clinical practice in assessment of SLI.

1.6 Summary

This chapter began by defining what SLI is and discussing the criteria used to diagnose it. SLI is an expressive/receptive language impairment that is not explained by known cognitive, neurological or social deficits. An overview of the main morphosyntactic, phonological, lexical and pragmatic impairments sheds light on the heterogeneity of SLI that makes it difficult to classify into coherent, stable and well-defined sub-groups. The link between phenotypic manifestations of SLI and underlying genetic factors was discussed as well as their implications for the process of (ab)normal language acquisition. This chapter was concluded by assessing the relevance of investigating SLI to cognitive science in general and the interaction between cognition and language in particular. This chapter shows that cross-linguistic investigations of SLI have had an important role in developing theories of SLI and testing their propositions regarding the processes and linguistic operations implicated in children with SLI. Typological characteristics of Arabic will render it very useful in examining different accounts of SLI. It has been demonstrated that SLI has received great attention in the last few decades due to its role in investigating the two main accounts of language acquisition and language representation in the mind/brain, namely the domain-general and domain-specific accounts of SLI, which are examined in chapter 2.

2. Theories of Specific Language Impairment

This chapter reviews different theories of SLI by grouping them into two main camps: domain-general and domain-specific. These two different approaches to language impairment are differentiated by their stance on the domain-specificity of language and the relationship between language and cognition. Domain-specific theories are based on the premise that the language faculty exists as a module within the cognitive system (Chomsky, 1986; Fodor, 1983; Pinker, 1999). Fodor (1983) divides cognitive systems into modular and non-modular systems. Non-modular systems are those that handle central executive functions, while modular systems are specialised systems that have evolved to process specific type of data. He cites the language system as an example of a modular system. According to Fodor (1983), a module has the following defining properties: it is domain-specific, informationally encapsulated, and subserved by dedicated neural architecture that is genetically determined (i.e., innate), although this neural circuitry could be functionally defined and not necessarily anatomical in nature. Despite their acknowledgment that some aspects of language (such as the lexicon, and pragmatics) are not necessarily specific to the language faculty, proponents of domain-specificity emphasise that there are core aspects of language that are modular and dissociable from general cognitive abilities, such as computational operations or recursion (Hauser, Chomsky, & Fitch, 2002; Hauser, Fitch, & Chomsky 2002) or morphosyntatic rules (Pinker, 1999).

In contrast, domain-general theories claim that there is no single module or function of cognition that is innately specialised, rather this specialisation emerges as a result of development (emerging modularity) (Elman et al., 1996; Gathercole & Thomas, 2005; Karmiloff-Smith, 1992; Karmiloff-Smith & Thomas, 2003). Proponents of domain-general accounts criticise the proposition that there are selective deficits that inflict certain components of the system in isolation while the rest of the system functions properly. They claim that this ‘residual normality’ cannot be maintained in developmental disorders because it ignores the process of ontogenetic development (the process of the development of the individual from conception to adulthood) (Karmiloff-Smith & Thomas, 2003). Instead, their view is that developmental disorders proceed “developmentally under different neurocomputational constraints, not as demonstrations of static modularity" (Karmiloff-Smith & Thomas, 2003, p. 970). Modularity is a result of the process of development, and not a starting point in language development (Karmiloff-Smith & Thomas, 2003).

Examples of different domain-specific and domain-general theories that try to explain the deficits observed in children with SLI are reviewed in this chapter as well as other approaches that are not clearly domain-specific or domain-general , such as the mapping theories of development (Chiat, 2001) and the procedural deficit hypothesis (Ullman & Pierpont, 2005).

2.1 Domain-general theories of SLI

Research in the domain-general theories of SLI has identified several possible areas of deficits in the information-processing skills of children with SLI (Gillam & Hoffman, 2004). Among the information processing operations implicated in language disorders are: speed of processing, capacity of processing, perceptual processing, and verbal working memory (Montgomery, 2002b). According to these approaches, the deficits in SLI are due to processing delay or slowness. It is argued that there is generalised slowing across all types of processes (Edwards & Lahey, 1996; Kail, 1994; Miller et al., 2001), while others argue for language-specific slowing (Windsor, 2002). The temporal processing deficit account, (Tallal & Piercy 1973), on the other hand, does not assume limited capacity, as it is based on the premise that SLI (and other developmental language disorders) is caused by deficits in low level auditory processing. Other proponents of information processing deficits suggest that there are process-specific deficits, such as limitations in verbal working memory (Baddeley & Gathercole, 1990; Dollaghan, 1998). Bishop (1992), on the other hand, proposes that children with SLI can have limitations in both working memory capacity and speed of processing.

The following section reviews some of the accounts that try to explain SLI by referring to deficits in three main areas of information-processing: speed of processing, processing capacity and working memory.

2.1.1 Speed of processing accounts

2.1.1.1 The auditory (temporal) processing theory of SLI

The temporal processing theory of SLI proposes that children with SLI have difficulties in processing rapid auditory information that is not language-specific (Tallal et al., 1996; Tallal & Piercy, 1973; Tallal, Stark, Kallman & Mellits, 1981). These deficits lead to the children's inability to integrate auditory information that converges in the central nervous system. These temporal and perceptual deficits seem to affect both linguistic and non-linguistic processes and it is the short duration of the material that causes perceptual deficits, which will eventually lead to poor comprehension and processing of linguistic information. The linguistic problems described in children with SLI, according to Tallal and colleagues, are secondary to primary temporal processing deficits. Based on this assumption of temporal deficits in SLI and other developmental disorders, Tallal and colleagues claim that if rapid speech sounds are manipulated to make them more salient and children with language impairment are exposed to intensive training using synthesised speech through computer programmes, dramatic improvements in speech and language performance can be achieved (Merzenich et al., 1996; Tallal et al., 1996). Corriveau and colleagues found that children with SLI had auditory processing difficulties (Corriveau, Pasquini & Goswami, 2007). However, these were not specific to brief, rapidly successive acoustic cues. They showed that sensitivity to durational and amplitude envelope cues can better predict language and literacy outcomes (Corriveau et al., 2007).

Gillam and Hoffman (2004) examined the findings of Tallal and her colleagues (Tallal & Piercy, 1973; 1974; 1975) and reported that children with language impairment had difficulty remembering the order of synthesised CV syllables when they were presented rapidly, but not when the interstimulus duration was increased. They commented that these tasks were not assessing temporal processing; but they measured children's memory for sounds. They concluded that

> " it is not clear from these studies whether the problem with recalling rapidly presented tones and syllables involves perceiving the differences between sound correctly, rapidly creating well-specified mental representations of

sound…and /or retrieving representations accurately when it is time to produce a motor response" (Gillam & Hoffman, 2004, p. 139).

Moreover, these deficits in auditory processing have not always been replicated in other studies of children with language impairment. Hanson et al (2002) investigated the role of temporal processing in lexical processing skills of children with SLI by examining their reaction times (RTs) on a lexical recognition task that involved listening to sentences loaded with stop consonants and sentences with no stop consonants. According to the temporal processing deficits hypothesis, stop consonants are more difficult to process because they carry brief durations and are expected to pose more challenges to the temporal processing system. However, the predictions of the temporal processing deficits hypothesis were not borne out and there was no difference in RTs in the two types of sentences. Similar findings of lack of evidence for temporal or acoustic-phonetic deficits have been reported by other studies (Hanson & Montgomery, 2002; Montgomery, 2006).

Bishop and colleagues (1999) used three auditory measures with a group of language-impaired children and a control group and found no significant differences on any of these auditory measures (detection of a brief backward-masked tone (BM), detection of frequency modulation (FM) and pitch discrimination using temporal cues) (Bishop, Carlyon, Deeks & Bishop, 1999). More interesting was the finding that some control children had poor auditory processing skills (Bishop et al., 1999). Bishop et al. (1999) explained any association between language impairment and auditory deficits as a result of the latter being a moderating variable that is "neither necessary nor sufficient for causing LI [language impairment], but which exerts an effect on language development only in children who are already at genetic risk" (Bishop et al., 1999, p. 1308). Rosen (2003), in his review of auditory processing theory of SLI and dyslexia, comes to the conclusion that few of the children with SLI and dyslexia have deficits in auditory processing and there is no clear relationship between deficits in auditory processing and linguistic problems in these two populations.

Studies of the efficacy of computer-based intervention programmes based on temporal processing theory as developed by Tallal and her colleagues (Merzenich et al., 1996; Tallal et al., 1996) have not found sufficient evidence for improvement in language performance that is directly linked to acoustic modifications as proposed by Tallal and her colleagues (Borman & Benson, 2006; Cohen et al., 2005; Gillam, Crofford, Gale & Hoffman, 2001; Gillam, Frome Loeb, & Friel-Patti, 2001; Marler, Champlin, & Gillam, 2001; Pokorni, Worthington & Jamison, 2004; Rouse & Krueger, 2004).

Another domain-general account that has sprung from the auditory processing deficit hypothesis is the phonological-deficit hypothesis (Joanisse, 2004; Joanisse & Seidenberg, 1998). This hypothesis suggests that these perceptual deficits lead to subsequent phonological impairments that are considered the underlying cause of linguistic deficits in children with SLI. Joanisse (2004) used connectionist model to show that phonological deficits have an impact on the child’s ability to generalise from known to unknown (novel) phonological representations. This explains, according to Joanisse (2004), the oft-mentioned deficits in nonword repetition in children with SLI. These phonological impairments have a negative impact on all word forms and this can cause the grammatical (e.g., morphosyntactic) deficits in children with SLI. However, this account was not able to address the criticism directed towards the temporal processing account.

The auditory processing account of SLI claims there are primary perceptual deficits that cause the symptoms of SLI, however evidence for these deficits has not been found in many studies of children with SLI. Moreover, the connection between these limitations in processing of transient speech and nonspeech sounds and language deficits in SLI has not been adequately explained or justified.

2.1.1.2 The surface hypothesis

Another account that attributes linguistic deficits in children with SLI to poor information processing is the surface hypothesis account, proposed by Leonard and his colleagues (Le Normand, Leonard & McGregor, 1993; Leonard, 1989; Leonard, 1992; Leonard, 1998; Leonard, Eyer, Bedore & Grela, 1997). They argue that limitations in general processing capacity can account for the cross-linguistic differences in grammatical morphology in children with SLI. Leonard (1998) posits that children with SLI have difficulty with certain grammatical morphemes that are characterised by non-salient perceptual features. Therefore, it is this combination of poor auditory processing of these sounds and their grammatical functions that cause them to be of greater difficulty to children with SLI. Leonard (1998), however, does not propose that these children have perceptual deficits like the ones suggested by the auditory processing account (Tallal, 1996). Instead, he maintains that children with SLI are capable of perceiving weak non-final brief syllables, but their morphological functions increase processing demands and cause them not to be fully processed or placed into morphological paradigms (Leonard et al., 1997).

The surface hypothesis proposes that English-speaking children with SLI have difficulties with inflectional morphemes that are unstressed and non-syllabic, such as third person singular *–s*, and past *–ed* (Leonard & Bortolini, 1998). The shorter duration and non-salient acoustic properties of these morphemes make them more difficult to process and less likely to have adequate morphological representations, in contrast to more salient morphemes, such as "*-ing*". Leonard and colleagues reported less impairment in grammatical morphology in Hebrew and Italian because grammatical morphemes in these languages are mostly syllabic and occur at the end of words or phrases (Dromi et al., 1993; Leonard, 1992). They also noted that Italian grammatical morphemes with less salient properties, such as direct object clitics and articles, are more vulnerable as they mostly have short durations and occur in non-salient positions of the phrase (they mostly occur in medial positions) (Bortolini, Caselli, Deevy & Leonard, 2002). In a group of studies that looked into potential clinical markers of SLI in Italian, Leonard and colleagues found that Italian children with SLI performed significantly worse than age and language controls on production of direct object clitics, articles and third person plural inflections; all of them involve weak syllables in vulnerable positions (outside strong-weak syllables or in initial weak syllables) (Bortolini, Casalini, & Leonard, 1997; Bortolini & Leonard, 1996; Leonard & Bortolini, 1998). These children, however, performed as well as MLU controls on other grammatical morphemes, such as production of noun plural inflections, first person singular, plural verb inflections and third person singular verb inflections.

Montgomery and Leonard (1998) tested this hypothesis in an on-line task and found that children with SLI showed similar RTs for both grammatical and ungrammatical sentences that lack inflectional morphemes of the non-salient nature (e.g. third plural-s), while control children showed faster RTs for inflected sentences when these were compared to non-inflected sentences. Similar findings supporting the surface hypothesis have been reported in the use of articles in both French (Le Normand et al., 1993) and Spanish (Restrepo, Iacute & Gutierrez-Clellen, 2001).

Critics of the surface account contend that it fails to provide an adequate explanation of how morphemes with the same phonetic substance (e.g., 3PS –*s*, and plural –*s*) behave differently in the linguistic system of children with SLI, whereby the former, but not the latter, is frequently dropped in the language of children with SLI (Rice & Wexler, 1996a). Leonard (1998), however, explains that his original proposition (Leonard, 1989) was conceived within Pinker's (1984) Learnability theory, which states that grammatical morphemes with low phonetic substance are introduced later in the process of paradigm building. Leonard (1998) maintains that surface phonetic features are only one factor among other crucial factors such as semantic features of the morphemes. For example, verbs in Italian are inflected not only for number, but also for tense and gender which makes their verbal inflections more complicated than noun inflections. According to Leonard (1998), "two phonetically identical morphemes can be acquired at different rates because the child will hypothesize the grammatical function of one before the other" (p.253).

However, Leonard (1998) admitted that the surface hypothesis, with its assumption of intact grammatical knowledge, cannot explain the problems children with SLI face in auxiliary inversion (*what mommy is making?) and problems in case assignment (*Me take that). Moreover, findings from other languages, such as Dutch, do not support the surface hypothesis. De Jong (1999; 2003) reported that some of the substitutes for tense and agreement markers in Dutch were equally low in phonetic substance.

2.1.1.3 The generalised and process-specific slowing hypotheses

It has been reported that many children with language impairments perform slower than their age controls on many linguistic and non-linguistic tasks. For example, they are slower in bead threading, peg moving (Bishop, 1990) and picture naming tasks (see also Edwards & Lahey, 1996; Windsor, Milbrath, Carney & Rakowski, 2001 and references therein). Therefore, the generalised slowing hypothesis was proposed by Kail and colleagues as a parsimonious account to explain the differences in processing linguistic and non-linguistic tasks between typically developing children and children with language impairment (Kail, 1994; Leonard et al., 2007; Miller et al., 2001; Miller et al., 2006). Kail (1994) suggested that processing linguistic and non-linguistic tasks consists of several sub-operation or processes (e.g., acoustic-phonetic perception, encoding, decision making...etc) and each one of these processes is performed slower by children with SLI compared to their typically developing peers. Kail (1994) bases his theory on five previous studies that looked at 22 pairs of mean reaction times (RTs) of five different linguistic and non-linguistic tasks performed by children with SLI and TD children. He found that the RTs of children with SLI increased linearly as a function of the RTs of control children and regardless of the task involved. Therefore, Kail (1994) argues that children with SLI have slower general processing skills that are not related to specific tasks.

Miller et al. (2001) examined the performance of 77 children with SLI on 10 linguistic and non-linguistic tasks (such as grammatical truth-value judgements, picture matching, mental rotation, tapping) and compared their performance to a group of children with non-specific language impairment and a group of typically developing children. Both of the language-impaired groups showed RTs that were significantly slower than the control group, with the SLI group performing better than the non-specific language impairment group. The results support the theory of generalised slowing processes in children with SLI, though some children with SLI did not show slowed RTs. Five years later, Miller and her colleagues (2006) followed the same group of children at the age of 14 and found both groups of children with language impairment had slower RTs than the control children, with no significant differences between the language-impaired groups on both linguistic and non-linguistic tasks. Once again, some children with SLI showed RTs that were within the normal range. This finding of normal RTs in some children with SLI and the finding that children with SLI’s scores on linguistic and non-linguistic tasks were close to each other need to be addressed by proponents of the generalized slowing hypothesis. Proponents of the generalised slowing processing hypothesis will have to define a clearer relationship of the effects of slowed processing speed on language development and whether language impairment and slowed processing merely co-occur or there is a causal relationship between the two. A generalised slowing hypothesis faces the challenge of explaining why language processes seem to be more vulnerable than general cognitive processes in children with language impairment (Windsor, 2002).

Windsor and her colleagues (2001) analysed RT data from 25 previous studies that looked into 20 different linguistic and non-linguistic tasks. Using a different statistical method that is described as superior to the one used in the previous studies supporting the general slowing hypothesis (Kail, 1994; Miller & Leonard, 1998), Windsor et al. (2001) report great variability across these studies, indicating lack of significant general slowing across studies. Based on these methodological limitations and the unsystematic way of collecting RT data across previous studies, Windsor et al (2001) stated that there was a lack of sufficient support for the generalised slowing hypothesis.

While the generalised slowing hypothesis assumes invariant slowing regardless of the tasks involved, Montgomery's (2002) premise is based on the idea that slowness is process-dependent, i.e., inefficient processing of linguistic information is more critical than processing of other types of data such as acoustic-phonetic information. Montgomery (2002a; 2000b; 2002c; 2005; 2006) argues that children with SLI have limited capacity in processing high-order linguistic information. Montgomery (2002a; 2002b) looked at the interaction between linguistic information and processing using online word recognition tasks with children with SLI. He demonstrated that children with SLI had slower processing times not only compared to chronological age controls, but also compared to children with the same language level. He found that children with SLI took longer time to recognise words embedded in sentences compared to their age and language control children, though they had the same accuracy levels. Therefore, he argues that processing of linguistic material rather than deficits in linguistic knowledge or auditory processing is what underlies comprehension deficits in children with SLI (Montgomery, 2002a). Montgomery explains that children with SLI are slower to "recruit and complete the various linguistic operations involved with accessing and integrating the linguistic properties of incoming words into an evolving sentence meaning" (Montgomery, 2005, p. 172). Montgomery (2006) investigated the real-time language processing performance of children with SLI to assess the contributions of higher-order linguistic processes and phonetic-perceptual factors, using RT tasks. Children with SLI were compared to an age control group and a group matched on receptive language level on their performance on isolated lexical processing task and sentence-embedded lexical processing task. Children with SLI were not significantly different in processing isolated words from the control groups, but were slower on processing similar words that were embedded in sentences. Montgomery (2006) argues that these findings suggest that children's performance is influenced by linguistic operations. So, this approach differs from other processing accounts in its proposal that slow processing pertains to inefficient processing of linguistic components (e.g., poor lexical retrieval or lexical recognition) (Montgomery, 2002a; 2002b; 2005) and is not explained by perceptual deficits (Tallal 1993) or generalised slowing mechanisms (Kail, 1994).

2.1.2 Capacity accounts

2.1.2.1 The sparse morphology hypothesis

The sparse morphology account is another domain-general theory that posits that children with SLI have limited processing capacity. It proposes that the characteristics of language deficits in children with SLI are contingent upon typological properties of the language they acquire (Leonard, 1992). In English, there are few grammatical morphemes and hence children acquiring English dedicate most of their attention and processing capacity to other crucial cues such as word order. On the other hand, children acquiring languages with rich inflection (such as Italian and Hebrew) pay more attention and use more processing resources to handle grammatical morphology (Dromi et al., 1993; Leonard, 1998). According to the sparse morphology account, not only inflectional morphemes are affected in morphologically scarce languages like English, but also freestanding morphemes, such as articles, are vulnerable due to limited resources dedicated to morphology in general.

In their investigation of SLI in Hebrew speaking children (HSLI), Dromi and her colleagues (1993) found that these children performed better than their English speaking counterparts on measures of inflectional morphology. Leonard's (1998) analysis of the present-tense inflection in various languages reveals a mean use of inflection in English of 21%, followed by German with 53%, Italian with 94% and Hebrew with 93% and 88% for masculine and feminine singular, respectively. In fact, no differences were reported in the use of present tense between children with SLI and MLU controls for Italian and Hebrew (Leonard, 1998). Another cross-linguistic support for the sparse morphology account comes from studies of French children with SLI who were not found to be particularly poor at grammatical morphology (Thordardottir & Ellis Weismer, 2001; Thordardottir & Namazi, 2007). Though their performance was weaker than their age controls, their use of grammatical morphemes (e.g., tense morphemes, direct object clitics, verb inflections, and noun phrase morphology) was comparable to their MLU controls (Thordardottir & Namazi, 2007). Therefore, the evidence used by proponents of the sparse morphology is based on the relatively good performance of children with SLI in languages with rich morphological systems.

However, other studies have reported difficulties in tense marking, grammatical morphology, and object clitics in French-speaking children with SLI (Franck, Cronel-Ohayon, Chillier, Frauenfelder & Hamann, 2004; Hamann et al., 2003; Jakubowicz, 2003; Jakubowicz & Nash, 2001; Paradis & Crago, 2000,, 2001; Paradis, Crago, Genesee & Rice, 2003). In addition to results from French, there are other richly inflected languages such as Arabic and Inuktitut in which significant inflectional problems have been documented in children with SLI acquiring these languages. Abdalla (2002) reveals a pattern incongruent with the sparse morphology account as she noticed that children with SLI acquiring Hijazi-Arabic demonstrated marked difficulty with verb inflections. Moreover, Crago and Paradis (2003) reported on an Inuktitut child with SLI who had problems with inflectional morphology in this highly inflected language.

The sparse morphology account of SLI has attempted at providing an explanation of the significant difficulties children with SLI have in acquiring morphology in languages such as English, and their relative strengths in morphologically rich languages where more resources are presumably dedicated to morphology. However, this account does not explain the difficulties seen in children with SLI acquiring some morphologically rich languages. Furthermore, the explanation that children acquiring languages with sparse morphology dedicate less processing resources to morphology needs further explanation.

2.1.3 Working memory accounts

2.1.3.1 Phonological short term memory (PSTM) hypothesis

According to current models of working memory (WM), it consists of the following components: the phonological loop, which is responsible for storing verbal-acoustic information; the visual-spatial sketchpad, which retains visual information; the central executive system, which regulates attention in the working memory; and the episodic buffer, which has limited capacity for storage and relies heavily on the central executive (Baddeley, 2003). Baddeley (2003) claims that deficits in the phonological loop component, which includes a phonological store and a subvocalic rehearsal process, are the main cause for language deficits in children with SLI. The phonological loop is responsible for processing and storing novel sound combinations and it is thought to be implicated in children with SLI. Deficits in this part of the WM can cause problems in forming appropriate phonological representations and learning new words (Archibald & Gathercole, 2006a; Baddeley, Gathercole & Papagno, 1998; Gathercole & Baddeley, 1990; Gathercole, Service, Hitch, Adams, & Martin, 1999). Deficits in phonological short-term memory can be assessed using nonword repetition tasks such as the ones developed by Gathercole and Baddeley (1996) or Dollaghan and Campbell (1998). In their famous study, Gathercole and Baddeley (1990) showed that children with SLI have poor nonword repetition scores and they inferred that these children had limited phonological short term memory. Children with SLI demonstrated more difficulty in repeating longer nonwords than shorter ones, indicating, according to the authors, limited PSTM capacity. This limited or rapidly decayed representation of phonological elements, they argue, affects other linguistic representations and is the main factor behind language problems in children with SLI. According to Gathercole and Baddeley, SLI is essentially a disorder of phonological short-term memory (1990).

Later studies of nonword repetition in children with SLI found evidence supporting Gathercole and Baddeley's (1990) claims (Bishop et al., 1996; Dollaghan

& Campbell, 1998; Ellis Weismer et al., 2000; Montgomery, 1995b, 2004). Bishop et al (1996) propose the use of nonword repetition task to identify children with SLI as they argue that limitations of short-term memory can be a main clinical marker of SLI. Furthermore, deficits in children's ability to retain phonological representations over time could be the underlying cause of some syntactic deficits, such as difficulty assigning anaphora by children with SLI (Joanisse & Seidenberg, 2003).

Although Baddeley's (1986; 2003) model of working memory is perhaps the most influential model, it is by no means the only one. Another model of WM is the one suggested by Just and Carpenter (1992), which differs from Baddeley's model in how working memory capacity is specified. While Baddeley's model assumes that working memory is defined in terms of new phonological information to be stored at one time (phonological working memory hypothesis or PWM) , Just and Carpenter (1992) on the other hand define working memory capacity in terms of its ability to store *and process* verbal information (the functional working memory hypothesis or FWM). These processes are involved in the computations necessary to achieve language comprehension; therefore the FWM corresponds to the central executive of Baddeley (1986). Just and Carpenter (1992) do not conceptualise a role for a modality-specific storage system such as the phonological loop (Just & Carpenter, 1992). The capacity of working memory, according to Just and Carpenter (1992) is expressed in terms of the amount of activation available to support the storage and processing functions of the WM.

Montgomery examined the role of both PWM (1995a; 1995b) and FWM (Montgomery, 2000b) in sentence comprehension of children with SLI. He found a strong relationship between PWM (as measured by a nonword repetition task) and comprehension of short and long sentences (1995b) but little role for FWM in sentence comprehension in children with SLI. Montgomery's (1995b) was one of the first studies to link deficits in PSTM in children with SLI to their problems in comprehension. This was established by finding a strong correlation between the performance of children with SLI on a NWR task and their performance on comprehension of redundant (long) and nonredundant (short) sentences. Children with SLI showed significantly poorer skills on NWR and long sentences, but not on short ones. This correlation between NWR and sentence comprehension was attributed to the hypothesis that comprehension of sentences involves integrating new information to previous information that are already stored in the WM. Therefore, inefficient working memory could lead to breakdown in this integration process causing increasing difficulties with sentence comprehension (Montgomery, 1995a).

Despite the extensive evidence of deficits in nonword repetition skills in children with SLI, the findings of distinct genetic bases for grammatical deficits and the argument that NWR may not be a pure measurement of phonological short term memory have questioned the core premises of the PSTM account of SLI. Bishop et al. (2006) found little genetic overlap between phonological short term memory and verb inflections, thus posing a major challenge to the claim held by proponents of PSTM account that NWR deficits cause grammatical deficits in children with SLI. Moreover, Chiat and colleagues have criticized the view that NWR is a pure measurements of PSTM as they argue that NWR tests involve various components, and phonological storage is only one of them (Chiat; 2001; Snowling et al., 1991). Furthermore, studies of phonological complexity in children with SLI have shown that manipulations of this variable independent of phonological storage can explain some of the performance of children with SLI on NWR tests (Marshall & van der Lely, 2009; Gallon et al., 1997; see also chapter 5).

2.1.3.2 The Relationship between processing speed and WM

Leonard and colleagues (Leonard et al., 2007) examined the relationship between two groups of domain-general accounts, namely those that attribute language impairments to slow processing and those that argue that working memory limitations can explain the underlying deficits in language disorders. Leonard et al. (2007) explored the possible correlation between different speed and working memory factors and whether there are linguistic vs. non-linguistic distinctions among speed and working memory. They used confirmatory factor analysis to compare four models that assessed the impact of WM and processing speed on various measures of language in children with SLI. These models were: a speed=WM model, speed ≠ WM model, a model that subdivides speed and WM into linguistic/verbal vs. non-linguistic/non-verbal factors, and a general speed model.

Regression analysis showed that the models which treated WM and general speed factors as different factors accounted for 62% of the variance in language test scores with WM playing a larger role than general speed in predicting test scores (Leonard et al., 2007). Their analysis illustrates that non-linguistic tasks (e.g., visual search task) constitute a separable dimension and hence they argue that any account of language impairment should consider these factors. This study reveals that these two processing factors are related, but not interchangeable and each one of them adds unique contributions to language test scores.

Similarly, Montgomery and Windsor (2007) investigated the intercorrelatedness of processing speed and phonological short term memory in children with SLI by testing the influence of nonword repetition (as an indicator of PSTM) and auditory detection task (an indicator of processing speed) on receptive and expressive language performance. Results showed significant correlation between both speed of processing and PSTM in children with SLI and TD children. However, when the effect of age was accounted for this correlation was significant only for TD children. Montgomery and Windsor (2007) note that the contributions of these measures to off-line language performance are different, with PSTM accounting for significant proportion of the variance in off-line language performance in the SLI group, but not the TD group. Speed, on the other hand, does not have any significant contribution to off-line language tasks. As for on-line language tasks, processing speed has more important predictive value, as it correlated better with the performance of both groups on the online word recognition task. Therefore, it seems the relationship between processing speed and working memory is determined by the type of task being performed, with WM playing a more important role than speed in off-line tasks.

2.2 Domain-specific accounts of SLI

Proponents of domain-specific accounts of SLI suggest that deficits in certain aspects of the linguistic system are the main cause of SLI. According to them, SLI is an example of a deficit affecting some specific components of the language faculty and therefore provides a window into understanding the language faculty and how it is represented in the human brain. Unlike domain-general accounts, which suggest that linguistic impairments are caused by general deficits in the cognitive system;

domain-specific accounts explain the causes of SLI by referring to linguistic theories. For example, most of the accounts that try to explain the syntactic deficits in children with SLI are discussed in the framework of the Principles and Parameters theory (Chomsky, 1981) or the Minimalist Program (Chomsky, 1995).

However, these linguistic accounts differ as what part(s) of the linguistic system are affected and whether the linguistic system itself is missing some grammatical features (Gopnik, 1990; Gopnik & Crago, 1991) or some aspects of the system are optional, such as tense (Rice & Wexler, 1995). Others suggest that children with SLI have deficits in subject-verb agreement (Clahsen, 1998) or they have difficulties in syntactic, morphological, and phonological structures that involve complex grammatical computations (van der Lely, 2005).

The following section presents an overview of the main domain-specific theories of SLI, with special emphasis on accounts that focus on grammatical complexity.

2.2.1 The Agreement-deficit hypothesis

The agreement-deficit (AD) hypothesis is the only account of SLI that has been conceived in another language and then was tested in English and other languages. It was first proposed by Clahsen (1989; 1991) to account for linguistic deficits in German speaking children with SLI. Clahsen (1989;1991) explains that the strong features of Comp (complementiser), Tense, and the phi features (person, gender and number features) of subject DPs are present, but are noninterpretable. For verbs, these Agreement (phi) features are absent or underspecified in the linguistic system of children with SLI. Agreement features such as person and number are more vulnerable because they are controlled by the subject and not inherent in the verb. Though it was initially framed in the Generalized Phrase Structure Grammar (GPSG), Clahsen et al. (1997) re-presented the AD hypothesis in the Minimalist framework (Chomsky, 1995). According to the Minimalist Program, all verb features are non-interpretable features (i.e., they are not relevant for semantic interpretation) and so they must be checked off before the Logical Form (LF). Based on this, Clahsen et al. (Clahsen, Bartke & Goellner, 1997), argue that these non-interpretable features of the verb (i.e., person, number and agreement features) are affected. This deficit in verbal agreement, Clahsen argues, is the core deficit in the linguistic system in children with SLI. Clahsen et al. (1997) examined data of English speaking children with SLI and found that third-person singular affix –s is more affected than tense markers (e.g., past tense -ed). While these children use finite verbs with an accuracy level between 76%-89%, their use of agreement marker –s is always at (49%). According to the AD hypothesis, the use of third person singular –s, but not tense markers, is controlled by the phi-features of the verb and is expected to be more challenging. In German, the agreement system is more complex, as there are four agreement affixes that are used in all tenses as well as on modal verbs and auxiliaries. When examining the German data, Clahsen et al. (1997) compared children's performance on verbs marked for tense (preterite verb) and those marked for agreement. They found that German children scored significantly worse on subject-verb agreement verbs with an accuracy of 64% compared to an accuracy level of 99% for preterite verbs. Therefore, Clahsen et al. (1997) concluded that agreement errors are more prominent than tense errors in the linguistic system of children with SLI and that their account presents a better explanation of the German and English data when compared to the Extended Optional Infinitive (EOI) account of Rice et al. (1995). Moreover, Clahsen and Dalalakis (1999) found that deficits in marking subject-verb agreement were more prominent in Greek, thus arguing that while in English and some other languages, children have difficulty with both tense and agreement, they cited Greek and German as languages where agreement but not tense is affected.

These findings are contrasted with studies of SLI in Italian and Hebrew speaking children with SLI who were reported to produce few errors of subject-verb agreement. Leonard and colleagues have shown that Italian children with SLI did not have marked deficits in subject-verb agreements when their performance was compared to an MLU matched group (Bortolini, Leonard, & Casalini, 1998). Similar findings where children with SLI did not show significant deficits in subject-verb agreement were reported by Dromi and Colleagues (Dromi, Leonard, Adam, & Zadunaisky-Ehrlich, 1999; Dromi et al., 1993) for Hebrew speaking children with SLI. Therefore, the agreement deficit account of SLI does not provide adequate explanation for such data from these languages.

2.2.2 The missing features hypothesis

The missing feature hypothesis has been proposed by Gopnik and colleagues (Crago & Allen, 1996; Gopnik, 1990; Gopnik & Crago, 1991; Ullman & Gopnik, 1999) to account for the linguistic deficits in a four generational English family, the KE family, whose language disorder is characterised by hereditary deficits transmitted through autosomal dominant genetic factors (Gopnik; 1990). Gopnik (1990) shows that family members affected with "dysphasia", another term for SLI, evinced significant deficits in using rule-based structures, such as past tense *–ed* or regular plural inflections. She argues that some semantic-syntactic features like number, person, and tense are missing from the grammar of children with SLI. These errors are evident in all modalities (speaking, writing, grammaticality judgment and repetition), indicating that the fault lies deeply within the language system (grammar) and not in the "peripheral processing system" (Gopnik, 1990, p. 715). Ullman and Gopnik (1999) investigated inflectional morphology (e.g., past tense) in affected members of the KE family children to see if their linguistic system was in line with domain-specific accounts (Pinker, 1994) or domain-general accounts (Elman et al., 1996; A. Karmiloff-Smith, 1998). They found that these family members lacked suffixation rules and instead relied on associative memory. Ullman and Gopnik (1999) explain that affected members did not use regular past tense *–ed* morphemes when presented with novel verbs (plam-plammed) because these features are missing from their linguistic system. However, affected members had no apparent difficulty in generating novel overregularisations (e.g., crive-crove) as these are influenced by associative memory and not by the rule-based linguistic system as predicted by the dual-route mechanism of inflectional morphology (Pinker 1994).

Subsequent studies of the KE family have refuted the main argument of Gopnik and her colleagues, especially since these affected members are far from having pure linguistic deficits. Many of these family members show various non-linguistic, motor and cognitive deficits (Hurst et al., 1990; Shriberg et al., 2006; Vargha-Khadem et al., 2005; Vargha-Khadem et al., 1995). Many children with SLI do show some use of many linguistic rules (Rice et al., 1995) that is not predicted by this account of featureless grammar that assumes children with SLI are missing some syntactic rules (such as regular past tense *–ed* and plural *–s*). For example, Leonard et al. (1992) and Rice et al. (1995) found that when children with SLI do use correct forms, they tend to use them appropriately and in a manner not different from MLU matched children. This behaviour is observed in production of various grammatical morphemes, such as articles, regular plurals, copula and auxiliary be (Leonard et al., 1992). It is reported that children with SLI make overregularisation errors, a pattern not congruent with the missing feature hypothesis (see Leonard, 1998 chapter 3 and references therein). The missing feature hypothesis would predict severe deficits in learning inflectional rules in general, but cross-linguistic studies have revealed that children with SLI learning Hebrew and Italian do not have the same difficulties that English children with SLI have in learning the grammatical morphemes of their language (Bortolini et al., 2000; Dromi et al., 1993; Leonard, 1998) . This account, furthermore, cannot explain the common optionality shown in the grammar of children with SLI, who produce for example both *she likes* and *she like* (Bishop, 1994; Leonard, 1998; Leonard, Eyer, Bedore, & Grela, 1997). Overall, it seems that this hypothesis does not generalise well to cases other than the KE family, who show severe form of speech and language disorders. Moreover, it has many theoretical and empirical challenges (Leonard, 1998).

2.2.3 The extended optional infinitive (EOI) hypothesis

The extended optional infinitive (EOI) account of SLI was developed by Rice, Wexler and colleagues to account for the morphosyntactic deficits in English speaking children with SLI (Rice & Wexler, 1995; 1996b; Rice et al., 1995). This theory has been developed to account for the well-documented evidence for poor marking of tense in English speaking children with SLI. It has been shown that most English speaking children with SLI fail to mark tense on finite verbs and some never master tense inflections. Therefore, Rice and Wexler proposed the extended optional infinitive (EOI) as a clinical marker of SLI in English speaking children (Rice & Wexler, 1996b).

The EOI is based on Wexler's theory of optional infinitive (OI) of language acquisition in typically developing children. According to Wexler (1994; 1996; 1998) typically developing children go through a stage where they fail to mark finiteness on main verbs. This stage, termed the optional infinitive (OI), is estimated to finish by the age of three. The Optional Infinitive account is used to explain how children's productions are non-adult like not because of faulty learning, but because certain aspects of the grammatical knowledge mature or grow with age (See Borer & Wexler, 1987; 1992). Wexler maintains that children in the OI stage have no problems in setting the appropriate parameter; but they omit tense (TNS) or agreement (AGR) in their syntactic representation. This model known as AGR/TNS Omission Model (ATOM) was used to explain the OI stage in children's grammar (Schutze & Wexler, 1996). The driving force for ATOM is the Unique Checking Constraint (UCC) (Wexler, 1998). According to Wexler (1998), OI occurs because the child cannot perform two checking operations for the D-feature on the DP (determiner phrase), namely to check the D-feature of the DP against the D-features on AGR and TNS. In the OI stage, the D-feature of DP cannot check against more than one functional category. This checking theory is based on the Minimalist Program theory of Chomsky (1995). This constraint on typically developing children's grammar is considered "a case of pure growth" and maturation (Wexler, 1998, p. 63).

The EOI assumes that finiteness is genetically determined and not influenced by learning factors. Neither IQ nor mother education can predict finiteness (Rice et al., 1998). Mabel et al. (1998) found that the time factor was the most important factor in determining the growth of finiteness. This model argues that the use of finiteness matures as the child grows and it takes a genetically determined time course. Therefore, Wexler (2003) uses finiteness as a prime example for the growth of grammar and as a proof against the learning of it. Apart from this delay in marking finiteness, Wexler (2003) suggests that other aspects of the grammatical development are intact. While in typically developing children the OI stage ends by age 3;0 when their grammatical system matures, children with SLI extend this stage and may never grow out of it.

The EOI predicts that grammatical features marking tense, such as third person present singular *–s* (as in she walk*s*), past tense *–ed* (as in walk*ed*), auxiliaries (e.g., "*do*"), and auxiliary and copula *b*e as in ("he *is* walking", and "he *is* a student") are treated as optional grammatical features by children with SLI. This optional use of tense markers is not unlike the pattern seen by younger typically developing children in the OI stage (estimated between age 2;0 and 3;0) (Rice & Wexler, 1996b; Rice et al., 1995).

In their 1996 study, Rice and Wexler looked at the performance of 36 children with SLI, aged between 52-68 months and compared it with two control groups matched on chronological age and MLU. Language samples and linguistic probes were used to assess the children's mastery of different grammatical morphemes. Results showed that only TNS marking morphemes could reliably distinguish children with SLI from other control groups. There was no overlap between the SLI group and other control groups on this measurement, an argument for TNS as a clinical marker of SLI. Other morphemes (e.g., "*a/the*", "*-ing*") did not yield similar or even closer accuracy levels (Rice & Wexler, 1996b). Grammaticality judgements tasks show that English speaking children with SLI tend to accept sentences lacking tense marking (Rice et al., 1999) as correct; this was used by Rice et al (1999) to argue against any production constraints account, such as the one proposed by Bishop (1994).

This EOI phenomenon has been reported in other languages, such as: German, Swedish, Danish, Norwegian, Icelandic, Faeroese and French (Wexler, 2003). However, the OI stage is contingent upon the properties of each language. Wexler (1998) predicts lack of the OI stage in INFL-licensed null-subject languages. He argues that in these languages, AGR has +D feature and therefore the D-feature on the DP does not have to check against two functional categories. He explains this based on data from null-subject languages (like Italian). In such languages, AGR is pronominal and this property is what licenses null-subjects. Therefore, while children learning non pro-drop languages are expected to obey the UCC constraint and fail to produce finite main verbs in the OI stage, children learning null-subject languages do not pass through this stage.

Moreover, Wexler and his colleagues (Tsakali & Wexler, 2004; Wexler, 2000) propose that some children acquiring languages with direct object clitics can pass through a stage similar to the OI stage where they omit clitics (and he termed this stage the Clitic Omission Stage (ClO). They argue that the ClO could explain the variation seen in the acquisition of pronominal clitics across languages. He states that the interaction between UCC and the syntactic properties of certain languages can explain why some languages drop clitics while others do not. Wexler (2000) proposes that the omission of clitics is caused by the UCC in languages that show participle agreement (e.g., French and Italian) where there is double D-checking. In languages such as Greek and Spanish only single D-checking is required and so clitic omission is not expected. To test their UCC hypothesis which states that the D-feature of the DP can only check against one functional feature, Wexler et al. (2004) studied object clitics in Catalan and Spanish. Both have closely related syntax; however, the former has participle agreement, while the latter does not. They predicted that UCC would lead to more object clitics omission in Catalan because the D-feature of the DP has to check against two functional features, while in Spanish it has to check against one functional feature only. Their predictions were borne out. They revealed that in Catalan object omissions occur at the same stage in which OI occurs in non-null subject languages (like English) (Wexler et al., 2004).

A study of SLI in French-English bilingual children found that bilingual children with SLI showed EOI in both English and French (Paradis, Crago, Genesee & Rice, 2003). In their study of SLI in Quebec French, Paradis and Crago (2001) explain that SLI children's use of verbal inflections differed from both age matched and younger language matched children, indicating a tendency similar to that in EOI. However, some of the substitutions used were not always non-finite, but included finite forms too. Therefore, Paradis and Crago (2001) recommended the use of the term extended optional default instead of optional infinitive to describe children's difficulties with tense. Crago and Paradis (2003b) examined the different forms used by children with SLI and typically developing children to substitute for the finite form of verb in different languages (Swedish, German, Dutch, French, Arabic and Inuktitut). They concluded that these forms differ across languages. These non-finite substitutions cannot be described as an extended optional infinitive (some language do not have an infinitive, e.g. Arabic, Inuktitut). However, there was a common error pattern characterised by the use of a default form of the verb, which is described as "the most minimally inflected forms in the present tense paradigm" (Crago & Paradis, 2003b, p.104).

Findings from other languages have also pointed to some weaknesses in EOI. For example, both de Jong for Dutch (1999) and Hansson and Leonard (2003) for

Swedish, found that some children with SLI tended to substitute present tense for past tense; a substitution that cannot be explained in the framework of EOI. The fact that children with SLI have significant problems with regular-past tense but not with present tense is not accounted for by EOI (Hansson et al., 2000; Hansson and Leonard, 2003). This divergent performance on present vs. past tense has also been reported in Hebrew (Dromi, Leonard, Adam, & Zadunaisky-Ehrlich, 1999). For Greek-speaking children with SLI, Clahsen and Dalalakis (1999) and Tsimpli (2001) found intact use of tense with varying degrees of difficulty in subject-verb agreement.

In summary, though the EOI account of SLI provides a convincing explanation for the types of prevalent errors of tense in English that is supported by genetic studies of SLI (Bishop et al., 2006), findings from other studies question the existence of these marked deficits in tense in children with SLI acquiring others languages.

2.2.4 Grammatical complexity accounts

van der Lely and colleagues have investigated grammatical deficits in children with SLI and a sub-group of children identified as Grammatical-SLI (G-SLI) and revealed that most of these grammatical problems are attributed to a core deficit in manipulation of grammatical components that involve computational complexity (van der Lely, 1998; 2003; 2005; van der Lely & Stollwerck, 1996). Children with G-SLI constitute a sub-set of children with SLI. This subgroup is more homogeneous and characterised by severe deficits in receptive and expressive grammatical skills, relatively good non-grammatical language abilities (e.g., in pragmatics, linguistic reasoning and vocabulary). These children have a persistent language impairment and are typically diagnosed with G-SLI after the age of 9 (van der Lely, 1998; van der Lely & Battell, 2003; van der Lely & Stollwerck, 1996). G-SLI, is considered an impairment in components of grammar (syntax, morphology, phonology), especially with respect to hierarchical structural complexities that involve dependencies, such as wh-movement, tense and agreement marking, assigning thematic roles to NPs (as in passive voice sentences), and assigning of coreference (anaphora) based on syntactic cues only. Children with G-SLI were described as having Representational

Deficits in Dependent Relations (RDDR) (van der Lely, 1998; van der Lely & Battell, 2003), which mainly affected their syntactic system. Later adaptations of the RDDR lead to the expansion of the complexity area to include other grammatical components. So the core deficits in computations of complex grammatical structures are not limited to syntax, but also encompass phonology and morphology (van der Lely, 2005). This approach to grammatical complexity in children with SLI is known as the computational grammatical complexity (CGC) account of SLI. Unlike the EOI account that assumes core deficits in SLI to be in tense marking, CGC proposes that children with G-SLI have deficits in "the computations underlying hierarchical, structurally-complex forms in one or more component of grammar" (van der Lely, 2005, p. 55). The following section reviews some of the findings of CGC in areas of syntax, morphology, and phonology.

2.2.4.1 Syntactic deficits in grammatical complexity accounts

In syntax, computational complexity is manifest in structures that involve syntactic dependency, especially those that involve movement, such as in object wh-questions (for example, "who did Joe see__?"). Object questions involve two movements: an I- to-C movement (head-to-head movement) and a second movement that creates a filler-gap dependency between the wh-operator and its trace (van der Lely, 2005; van der Lely & Battell, 2003). Other examples of computational complexity in syntax, which were found vulnerable in children with SLI, are structures requiring pronominal reference, such as reflexives (van der Lely & Stollwerck, 1997) and comprehension of passive sentences that involve reversible actions (van der Lely & Harris, 1990). van der Lely and Battell (2003) investigated the production of subject and object wh-questions using wh-operators (who, what, and which) in 15 children with G-SLI and compared their performance to 24 age and language matched children. Children with G-SLI showed significant impairments relative to age and language control children in both subject and object wh-questions and the wh-movement in object wh-questions is longer than in subject wh-questions. However, their performance on object wh-questions was significantly worse than subject questions as there is an additional I-to-C movement in object questions. van der Lely and Battell (2003) concluded that children with G-SLI treat wh-movement as optional, as indicated in the presence of both correct and incorrect questions. This "optionality" of some syntactic operations, such as movement is considered one of the underlying deficits in children with G-SLI (van der Lely, 1998).

Another domain-specific account that is based on the premise that children with SLI have deficits in syntactic complexity is the one put forward by Jakubowicz and colleagues (Jakubowicz, 2003; Jakubowicz & Nash, 2001; Jakubowicz, Nash, Rigaut & Gérard 1998). They propose their computational complexity (CC) hypothesis to account for the different ways in which French-speaking children with SLI treat past and present tenses (Jakubowicz, 2003; Jakubowicz & Nash, 2001). They found children with SLI had problems with past tense but not with present tense. They attribute this divergence in tense to the hypothesis that present tense is an easy-to compute functional category as it does not introduce further semantic modifications, while past tense (passé composé) involves more complex computations. They define complexity with reference to the "properties of functional elements of the language the (ab)normal child is acquiring" (Jakubowicz and Nash 2001, p.324). They maintain that less complex computations involve a functional category that "must be present in EVERY sentence" (p, 324), i.e. it is syntactically necessary; while more complex computations involve a functional category that is seen only in some sentences. These complex computations add semantic information to the obligatory syntactic information and thus pose extra level of difficulty for children; an example of this is the passé compose in French:

> In French, a second functional head Past is projected only in the past. The failure to produce the past tense can result from the failure to project this second functional head, which renders the syntactic computation more complex. (Jakubowicz and Nash 2001, p .336)

They argue that this can account for patterns of language acquisition in TD children and patterns of breakdown in children with SLI. Jakubowicz and Nash (2001) illustrate that computational complexity effects are evident in both production and comprehension of past tense (passé composé) in French-speaking children. Another support for the CC hypothesis comes from comparing definite determiners and accusative clitics. In French these are homophonous, yet children with SLI have more difficulty with clitics. Computational complexity predicts that definite determiners involve less computations as "nominal projection must be headed by D in order to be fully licensed in the clause" (Jakubowicz and Nash 2001, p.337). They also noted that not all complex computations are equal, so computations of accusative clitics are more complex than past tense, because accusative clitic computations involve "noncanonical projection of the pronominal object argument" (Jakubowicz and Nash 2001, p.337), while INFL computations are generally less complex.

The CGC account of SLI offers an attempt to explain a variety of deficits in children with SLI and across different linguistic components. However, it is criticised for not providing a narrow definition of grammatical complexity and structural dependency. Moreover, like other linguistic accounts of SLI, some of its predictions may not hold in other languages. For example, while object relative clauses and movement structures were found challenging in languages such as Greek and Hebrew, agreement deficits are not seen in other languages (see references section 2.2.1 on agreement deficit hypothesis) even though they involve a structural dependency. Leonard (1998) argues that pronominal determiners in English do involve dependent relationships (as in *those cats*), yet children with SLI do not present with significant problems using these determiners. Moreover, direct object pronouns are assigned the accusative case by AGR_0, another form of structural dependency, yet they are usually used correctly by children with SLI.

van der Lely (1998, 2005) and Jakubowicz and Nash (2001) provide two examples of a computational account of SLI. They describe some specific syntactic structures that involve extra or more complex computations and they argue that these will be more challenging for children with SLI whose vulnerable grammatical system will make it difficult for them to carry out the necessary computations.

2.2.4.2 Morphological deficits in grammatical complexity accounts

A hallmark of SLI is the deficits in grammatical morphology (Bishop, 1997; 2000; Leonard, 1998) and children with G-SLI are not different from other groups of children with SLI in having problems with inflectional morphology. van der Lely (1996) found that children with G-SLI aged between 9-12 years still had problems with third person singular-s (as in using "he sing" instead of "sings"). They also evinced difficulties in using regular past –ed as in (Yesterday I walk) and showed some overregularisation errors (Yesterday I swimmed) (van der Lely, 1996; van der

Lely & Ullman, 1996). Similar patterns were seen in grammaticality judgement tasks where children with G-SLL tended to accept stem forms (e.g., walk) and overregularised forms (falled) as correct forms of past tense (van der Lely, 1996; van der Lely & Ullman, 1996).

The CGC proposes that children with G-SLI store all regular and irregular past tense forms and do not use the dual route mechanism proposed by Pinker's (1999) "words and rules theory". This explains the presence of frequency effects for both regular and irregular past tense forms in children with SLI, unlike typically developing children (van der Lely & Ullman, 1996). Marshall and van der Lely (2006) provide further evidence supporting the dual mechanism approach to past-tense inflection and the conclusion reached by van der Lely and Ullman's (2001) about the representation of past tense morphology in children with G-SLI. van der Lely and Ullman (2001) argue that while typically developing children use two mechanisms for past-tense inflection, i.e., inflectional rules for regular past and lexical storage for irregular past-tense, children with G-SLI retrieve all past-tense forms from the lexicon. Marshall and van der Lely (2006) have shown that children with G-SLI performed significantly better on 'monomorphemic legal clusters', i.e. clusters that can occur both at the end of inflected verbs as well in monomorphemic words (e.g., miss*ed*/mist) than they did on verbs with illegal clusters, i.e., clusters that occur only in inflected words (e.g., hugg*ed*). Legal clusters are more frequent than illegal ones and children with G-SLI performed significantly better on them showing that they are retrieved from storage using a single mechanism. Typically developing children, on the other hand, showed no difference on their performance on both legal and illegal clusters, as predicted by the dual route mechanism. (Marshall, 2004; Marshall & van der Lely, 2006). Moreover, Marshall and van der Lely (2006) re-analysed Thomas et al.'s (2001) data on the performance of children with Williams Syndrome on forming past tense according to the legal/illegal cluster distinctions and found that their performance was not affected by these phonotactic factors. Unlike children with SLI, children with Williams syndrome seemed to have intact representations of the phonotactics of past tense morphology, a finding that supports the words and rules approach to morphology (Pinker, 1999).

According to CGC, inflectional morphology is more affected than derivational morphology in the grammar of children with SLI, as it involves more morphosyntactic operations (e.g., past tense –ed), while derivational morphology depends more on lexical operations (e.g., changing a verb to a noun to create a new word) (Marshall, 2004; Marshall & van der Lely, 2007).

Marshall and van der Lely (2007) used two experiments to investigate whether derivational morphology (e.g., the use of comparative and superlative suffixes) is impaired in children with G-SLI. They also looked at the effects of phonological and morphological complexity on the performance of children with G-SLI. In the first experiment they found that the G-SLI group made few omissions errors, hence indicating relatively intact derivational morphology in comparison with their performance on inflectional morphemes, such as the regular past tense –ed. In the second experiment, children were expected to produce adjectives by adding the suffix '-y' to a noun (e.g., spot $\rightarrow$ spotty). Some of these nouns were already inflected with the plural '-s' (e.g., scales, spots) while the others were not. They found that most children with G-SLI had no problem with adding the derivational morpheme '-y'; however, some tended to keep the plural marker between the stem and the derivational morpheme (e.g., *scalesy), especially with less frequent adjectives. Therefore, it seems that phonological and inflectional complexity (exemplified in this case by the presence of regular plural –s on the adjectives) can cause problems in forming suffixations, despite the presence of generally intact derivational suffixations in the absence of inflectional or phonological complexity. Therefore, Marshall and van der Lely (2007) argue that children with G-SLI tend to store plurals in their lexicon and then add the suffixes to these plurals, leading to the production of these incorrect forms that are not seen in typically developing children. This finding supports previous reports that showed that children with G-SLI produce compound nouns that have regular plurals in them (e.g.,*rats-eater) (van der Lely & Christian, 2000) .

2.2.4.3 Phonological deficits in grammatical complexity accounts

Working in the framework of CGC, Marshall and colleagues found that children with SLI had significant problems repeating nonwords with complex prosodic characteristics, such as in nonwords with weak syllables and nonwords with consonant clusters (Marshall et al., 2002). Moreover, Gallon et al. (2007) used the Test of Phonological Structures (van der Lely & Harris, 1999) with adolescents with G-SLI and showed that these children had significant difficulties with nonwords containing marked stress patterns (e.g., unfooted syllables) and cluster consonants, when they were compared to TD age and language matched children. These children with G-SLI still presented with difficulties at these marked phonological structures even at single syllable length (see chapter 5 for more on nonwords repetition skills in children with SLI).

Marshall (2004) looked at the morpho-phonological interface in children with SLI and revealed that phonological factors affect past tense morphology. For example, she demonstrated that past-tense morphemes that create consonant clusters (e.g., *hugged*) or involve syllabic allomorph (*rested*) pose special difficulty for children with G-SLI. While the use of progressive *-ing* is generally intact in the grammatical system of children with G-SLI, Marshall (2004) found that phonological complexity could increase error rates of production of *–ing* in children with G-SLI. She used an elicitation task for production of progressive *–ing* and revealed that the *–ing* morpheme was omitted significantly more when it was affixed to a disyllabic verb with initial weak syllable (e.g., return).

The computational grammatical complexity (CGC) account of SLI presents a theory that tries to account for deficits spanning syntax, morphosyntax, phonology, and morphology in children with SLI. It proposes that complex grammatical structures that involve hierarchical dependency and require extra grammatical computations pose extra levels of difficulty for children with SLI in general, and for those with G-SLI in particular.

2.3 Alternative approaches

While the study of SLI and developmental cognitive science in general have been dominated by domain-general and domain-specific accounts of language, these are not the only theories proposed to account for developmental language disorders. This section examines two prominent alternative accounts that take different approaches to the question of domain-specificity and therefore add their unique contributions. The two accounts discussed here are the mapping theory of developmental language disorders (Chiat, 2001) and the procedural deficit hypothesis (PDH) (Ullman and Pierpont, 2005).

Chiat (2001) takes a rather different approach to deficits in SLI and proposes an alternative account known as the phonological theory account of SLI, based on a general theory that assumes there are mapping deficits in developmental language disorders. Chiat (2001) maintains that it is the mapping process connecting form to meaning that is disrupted in children with SLI. The mapping theory posits that phonological processes are crucial for lexical, morphological and syntactic acquisition and any breakdown in phonological processes will cause deficits in other levels of language (Chiat, 2001). During mapping processes, the child uses her phonological skills to help in the segmentation and storage of lexical items and the identification of their semantic and syntactic functions. Limitations in this mapping process can cause the linguistic deficits seen in children with SLI. Therefore, inefficient phonological processing, rather than deficits in linguistic knowledge or general processing limitations, is what characterises language deficits in children with SLI. These deficits in mapping processes are predicted to span different levels of phonological processing and affect different levels of language development depending on the impact of phonological processing on these levels. For example in semantic development, concrete nouns are less affected in children with SLI than verbs and prepositions, because phonological processes play a more important role in the mapping process of the latter categories (Chiat, 2001). Nonword repetition, on the other hand, is highly influenced by early phonological processes and representations that develop early, therefore those with less sensitivity to the phonology of their language will develop not only deficits in NWR, but also problems in lexical development and sentence structure. According to Chiat (2001) deficits in phonological and mapping processes are the source of impairments in children with SLI.

Ullman and Pierpont (2005) propose an alternative proposal to account for both linguistic and non-linguistic deficits in children with SLI, namely the procedural deficit hypothesis (PDH). According to PDH, most individuals with SLI have deficits in their “procedural memory system” that is subserved by certain brain structures, especially Broca’s area and the basal ganglia (Ullman & Pierpont, 2005).

Ullman (2001) proposes that linguistic representations and processing involve two distinctive brain memory systems that he termed the lexical/declarative system and the procedural memory system. The lexical/declarative system constitutes the neural basis of the lexicon and is situated in the temporal and parietal lobes, while the procedural memory system represents the neural basis of the grammar (i.e., computational rules) and is situated in the frontal/basal ganglia region (Ullman & Gopnik, 1999). According to Ullman (2001) "the learning, representation, and processing of aspects of grammar depend largely upon procedural memory…[which is] implicated in the learning and expression of motor and cognitive skills and habits" (p. 117). Moreover, the procedural memory system is closely related to other functions, such as lexical retrieval, mental imagery, temporal processing and working memory, even though that relationship has not been clearly elucidated (Ullman & Pierpont, 2005). Ullman and Pierpont (2005) cite anatomical studies, event related potential (ERP) studies, and behavioural studies to support their argument for brain abnormalities in the procedural system (mainly the frontal lobe and basal ganglia) in individuals with SLI. They suggest that these abnormalities in the procedural system can explain both linguistic and non-linguistic deficits in children with SLI in a way that has not been done by domain-general and domain-specific accounts of SLI. Their model explains the "mysterious" link between grammatical and non-linguistic deficits in children with SLI, which have never been accounted for in a clear manner. The PDH predicts that rule-governed processes (e.g. in syntax and morphology) will be especially vulnerable as well as learning novel phonological sequences (Ullman, 2004; Ullman & Pierpont, 2005). Thus, PDH expects deficits to manifest in different grammatical components. It also predicts that lexical learning will be affected due to the neural connection between the lexical and procedural systems.

In summary, the theory of mapping process (or phonological theory of SLI) of Chiat (2001) and the procedural deficit hypothesis (PDH) hypothesis (Ullman & Pierpont, 2005) endeavour to provide two alternative frameworks for domain-specific and domain-general theories of SLI. Both try to present an origin for the type of deficits seen in children with SLI. While, Chiat (2001) argues that SLI is caused by deficits in early phonological skills that cause lexical and syntactic deficits, the PDH posits that deficits in the procedural memory system, especially in the frontal lobe/basal ganglia circuitry can account for difficulties children with SLI have in syntax, morphology, phonology, short-term memory and temporal processing. It provides neuroanatomical explanations for both linguistic and non-linguistics deficits in children with SLI. Both accounts try to explain many of the characteristics of SLI by referring to these basic deficits and therefore provide interesting and testable predictions that are worthy of further investigation.

2.4 Summary

This chapter reviewed various domain-general and domain-specific accounts of SLI. Two alternative approaches that operate outside domain-general and domain-specific boundaries were also examined.

One of the challenges that any theory of SLI will encounter is the heterogeneity of the disorder itself. It is now generally accepted that SLI is unlikely to be one but many disorders. Heterogeneity of SLI makes it very difficult to present one comprehensive theory that can explain the heterogeneous nature of SLI (Bishop, 1997; Leonard, 1998; van der Lely, 2003). Any single-factor explanation cannot provide adequate coverage of the heterogeneity of the disorder. Therefore, it is possible that there are domain-general and domain-specific factors that cause different manifestations of SLI. Hence, some have cautioned against treating SLI as a single unified disorder. Others have emphasised dimensions of impairment in SLI rather than looking at subtypes of SLI (Bishop, 2007). Some children exhibit deficits throughout the language system (Bishop, 1997), whereas others exhibit primary deficits in vocabulary: Lexical (L)-SLI. There are those who present with word finding difficulties (Dockrell et al. 2001), or pragmatics (P)-SLI (Bishop, 2000; Botting and Conti-Ramsden 2003). Furthermore, some children's core deficits have been described as being restricted to grammar (syntax, morphology, phonology), e.g. G-SLI (van der Lely 1998, 2005), or even components within grammar, such as morphosyntax (Rice et al. 2000). Furthermore, some children with SLI show additional deficits in auditory, cognitive or even motor abilities (Bishop, 1997; Bishop & McArthur, 2005; Leonard, 1998) suggesting that some forms of the disorder, and by implication some mechanisms underlying some aspects of the language systems, are not so specific as once thought, although no direct link has been established between any of these non-linguistic deficits and language deficits in children with SLI (van der Lely, 2005; van der Lely et al. 2004).

Overall, it is agreed that both domain-general and domain-specific accounts have contributed to the advancement of the study of how language is represented in the brain in general and the study of SLI in particular. It is most likely that both linguistic and processing factors contribute to causes of language impairment in SLI. So, it is not unusual for domain-specific proponents to recognise the relationships between syntax and other cognitive abilities and shun the strong localist view of language representation in the brain (Newmeyer, 1997), nor for proponents of domain-general theories of SLI to acknowledge the distinct genetic basis of morphosyntactic (e.g., verb inflections) deficits in children with SLI (Bishop et al., 2006).

It is hoped that the investigation of SLI in Arabic can contribute to this debate about the causes of SLI. Studies of SLI across languages have enriched our understanding of the nature of SLI. Therefore, results of this investigation of SLI will be analysed with reference to some of the accounts of SLI in order to advance our understanding of the disorder, its nature, its manifestations across languages, and viability of some of the accounts of SLI.

3. Tests developed to assess language skills in Gulf Arabic speaking children

3.1 Introduction

This chapter describes the tests developed to assess the language skills of Gulf Arabic (GA) speaking children and the results of the administration of these tests with typically developing children and children with SLI. Gulf Arabic is the variety of Arabic spoken in the eastern parts of Arabia in the modern states of Bahrain, Kuwait, Qatar, United Arab Emirates, and the Eastern Province of Saudi Arabia. The chapter starts by introducing the challenges of working in a language where assessment tools are not available and it explains the steps taken to deal with these challenges. The tests developed were designed to assess different linguistic skills, such as sentence comprehension, production of morphosyntactic structures, sentence repetition, and receptive vocabulary. Other procedures used to identify children with SLI are also described, such as non-verbal IQ tests, and screenings of oral-motor and articulation skills. The results of testing GA speaking children are presented. Distributional characteristics, reliability and validity of these measurements are examined as well as developmental trends. Since these are general language tests that cover broad areas of receptive and expressive language skills, only general remarks about item difficulty for some linguistic structures in Arabic are examined in order to provide some initial information about characteristics of SLI in GA speaking children with SLI. Initial results show appropriate levels of reliability and validity and support the usefulness of these tools to diagnose children with SLI, whose performance on the tests was mostly consistent with findings in other languages.

3.1.1 Challenges of conducting research in Gulf-Arabic

Since the aim of this study is to investigate SLI in GA speaking children aged between 6 and 9 years old, the first challenge faced by any researcher studying this population is the lack of standardised tests, criterion-referenced measures or any other tools for diagnosing children with SLI. There are no published tests for any of the GA varieties and there are no systematic investigations of language acquisition in this population. The only published study on language acquisition was an investigation of the development of tense and agreement of three toddlers in Kuwaiti GA (Aljenaie, 2001). There are no studies of SLI in Gulf Arabic, though Abdalla (2002) looked at morphosyntactic deficits in preschoolers with SLI acquiring Hijazi-Arabic, a variety of Arabic that is different from Gulf Arabic. Abdalla (2002) based her diagnosis of children with SLI on MLU as well as adaptations of English tests and clinical judgments of speech-language therapists.

With regard to availability of language tests, there has been only one systematic attempt to create a comprehension test for typically developing children in Saudi Arabia (Al-Akeel, 1998), though the test has not been published yet. The test was developed to assess language comprehension skills of Saudi children aged 3;0-6;0 years old and was meant to be used with children using different regional dialects of Saudi Arabia. The test was designed to assess children's understanding of 24 morphosyntactic structures that were selected from three sources: spontaneous language samples of typically developing children interacting with their fathers; morphosyntactic structures that the author added himself based on his linguistic knowledge of Arabic and some morphosyntactic structures were modified from existing English language tests.

Al-Akeel (1998) reported some of the difficulties he faced when collecting data for his project in Saudi Arabia. These were related to difficulties having access to participants, especially young children. He reported that it was difficult to obtain permission from authorities to carry out his research. There were also some cultural practices that made data collection difficult for a male researcher, since most kindergartens and primary schools were staffed by female personnel only and no male visitors were allowed to enter these schools.

Some of these challenges were observed and encountered in this current project in Qatar. When the investigator approached the Ministry of Education to obtain necessary permissions to carry out research, the application process took more than four months, at the end of which no permission was provided, with no explanation given to justify this decision. There was a clear lack of cooperation and understanding of the nature and importance of this research project at various departments of the Ministry of Education, despite holding meetings to explain the nature of the project. When no permission was obtained, the investigator approached publicly funded independent schools, which are supervised by the Education Institute (EI) of the Supreme Education Council. The EI did not give permission though this message was not given in writing. Instead, the EI directed the investigator to approach each school individually. Luckily, some owners of Independent schools agreed to participate in the project and allowed the investigator to conduct his research.

Moreover, the investigator had some difficulties accessing the all-female schools and each visit was arranged carefully and on an individual basis. It was not possible to arrange a schedule for visits or assessment sessions and therefore, the assessment sessions started from December 2006 and continued until April 2009. During visits to schools, the investigator was situated in one quiet room in the all-female schools and children were brought by the special need coordinators or social workers, who were mostly very cooperative. Despite seeing children in more than six schools, most children participating in the experiments belonged to two schools whose staff members were very cooperative and understanding throughout the period of testing and conducting the experiments.

Some of the children participating in the experiments and data collection were acquaintances of friends and family and some were previous clients of the investigator, when he was working as a speech-language therapist in Qatar. In summary, any investigator conducting research in Qatar and other Gulf countries might consider the challenges of conducting research in Gulf Arabic and the importance of social networks in participants' recruitment.

3.1.2 General remarks about testing

The test battery used throughout this project consists of the following tests: Sentence Comprehension test, Expressive Language test, Sentence Repetition test, and the Arabic Picture Vocabulary Test. Two nonverbal IQ tests were conducted, as well as two screening tests for oral-motor functioning and articulation skills.

The time it took to finish all testing ranged between 45-60 minutes, depending on child's age, participation, and whether he or she asked for a break or not. Children aged between 4;0 and 5;0 were given frequent breaks and most of the time testing was done in two 30-minutes sessions. Most children enjoyed testing and were praised for their performance. At the end of each session, each child received some stickers as a reward.

Testing usually started with a short chat with the child to establish rapport. This was followed by less verbally demanding tasks, such as the nonverbal IQ test, which was followed by the sentence comprehension test. Then the Expressive Language test and the Sentence Repetition test were conducted. The Arabic receptive vocabulary test would follow these two expressive tasks and the session usually ended after running the articulation and oral-motor test. Children's responses were scored on individual record forms (see Appendices M-S). All individual test scores were transferred to the Arabic Language Test record form, which is shown in Appendix L.

Before describing the design of these tests and the results of administering these tests with GA children, the selection criteria for children with SLI are discussed in the following section.

3.1.3 Selection criteria for children with SLI

The criteria adopted and the cut-off scores for typical vs. atypical language performance for this study were largely based on Tomblin et al. (1997). These include having within normal range performance on one of two nonverbal IQ tests and the absence of any motor, neurological, or socio-emotional deficits. The criteria for inclusion in the group of children with SLI were having a score of – 2.0 standard deviations (SD) or more on one out of four language tests, or -1.5 SD or more on two tests. Due to lack of normative data in typical and atypical language acquisition in children acquiring Gulf Arabic, and lack of tests that could be used with typically and atypically developing children, the project had to start by collecting normative data from typically developing children before conducting the experiments with children with SLI and the age and language controls. The data collection proceeded as follows:

A meeting was held with coordinators/social workers at each school/nursery to explain the purpose of the project and how to recruit children. Consent forms were already distributed by then and only children whose parents agreed to participate were seen.

Coordinators were informed of the criteria for the study and the type of population being targeted. For example, teachers were informed that the project did not include children with stuttering, articulation, or autism spectrum disorders. Coordinators informed teachers that children with average academic performance and those at risk of language-learning difficulties would be good candidates for the study. When children came to the testing room, the examiner engaged them in a short conversation that was followed by a nonverbal IQ test. Children who scored lower than average on nonverbal IQ or showed evidence of social-emotional problems (e.g., ASD) were not included in the study. Those who passed these two initial criteria and who had uneventful developmental history, based on their history forms, were asked to complete the test battery. A few children were not included because of low nonverbal IQ or due to presence of other problems, such as stuttering.

The targeted age groups for children with SLI were ages 6;0 – 8;11 years old. Therefore, the project started by conducting the full battery of tests with 20-30 children in these age brackets, to identify 'norms' for these four language tests. Therefore, most of the children were recruited from year 1 to year 3. Following testing of at least 20 children in these age groups, means, standard deviations and z-scores were separately calculated for each age group and for each language test. Cut-off scores of -1.5 and -.2.0 standard deviations were established for each group. This was followed by adding children below the age 6;0 as these were needed to act as language controls.

Based on the criteria developed for the first three age brackets, children with SLI, who ranged in age between 6;0 and 8;11 years old at time of initial testing were diagnosed based on comparing their performance with the normative sample for their age brackets.

More children were added to all group bands, depending on availability and time constraints. Due to difficulties with scheduling and access to schools, the total number of typically developing children in each age bracket ranged between 19 and 24, falling below the initial target of 30 TD children in each age group.

Children with SLI were identified by comparing their performance with typically developing children on four language tests that were conducted with all typically and atypically developing children. The tests were the Sentence Comprehension test, the Expressive Language test, the Sentence Repetition test, and the Arabic Picture Vocabulary test. Other screening tools were used to rule out oral-motor dysfunction and abnormal intelligence. No other informal measures of spontaneous speech were conducted due to time constraints, however all children were engaged in a 5-minute conversation before administering the tests. Though two of the children with SLI had been previously diagnosed with developmental language disorder when they were aged 4;0, the rest of them were not diagnosed. However, most of them came with concerns about their academic performance. These concerns were expressed by class teachers and social workers. It is not uncommon that children with SLI are not identified, even in countries with better speech-language services. Tomblin et al. (1997) in their epidemiological study of SLI, reported that 71% of the children they diagnosed with SLI had not been previously diagnosed with language impairment.

In the rest of this chapter, the language tests used to diagnose children with SLI are explained and their reliability and validity are discussed. Furthermore, this chapter discusses other screening and testing procedures used to rule out any other deficits in nonverbal IQ, articulation/oral motor functioning, and hearing ability.

3.2 Test 1: The Sentence Comprehension (SC) test

3.2.1 Method

3.2.1.1 Participants

The Sentence Comprehension (SC) test was administered to 88 typically developing children and 26 children with SLI, whose characteristics are described in Table 1. Children with SLI met the selection criteria mentioned in the first section of this chapter as they all scored -1.5 SD or more on two out of four language tests or -2.0 SD on one test. They all had within normal scores on one of the two nonverbal

IQ tests used throughout the project, namely the Test of Nonverbal Intelligence (TONI-3) (Brown, Sherbenou, & Johnsen, 1997) for children aged 6 years and above or the Block Design and Picture Completion subtests of the Wechsler Preschool and Primary Scale of Intelligence (WPPSI-III) (Wechsler, 2002). All children with SLI passed a hearing screening at 20 dB for frequencies between 500-2000 Hz. Moreover, they had uneventful developmental history with no sensory, motor, or social-emotional problems. These children were recruited from two kindergartens and four primary schools in Doha, the capital of Qatar and some were recruited through personal acquaintances. Most participants came from what can be described as middle-class families and Qatari Arabic was the language spoken at home. However, most of these children had some exposure to English, which is taught at kindergarten level in Qatar and is widely spoken at the community due to the large number of expatriates in Qatar.

Table 1: Summary of the characteristics of participants in the Sentence Comprehension test.

Age Groups	Typically Developing Children	SLI
Age Band 1: 4;6 - 5;11 years		
Number of participants (Male:Female)	24 (13:11)	5 (2:3)
Mean age in months (years)	64.0 (5;3)	62.6 (5;2)
Range in months (years)	54-71 (4;6-5;11)	58-70 (4;10-5;10)
Age Band 2: 6;0 - 6;11 years		
Number of participants	23 (15:8)	8 (7:1)
Mean age in months (years)	77.6 (6;6)	78.9 (6;7)
Range in months (years)	72-83 (6;0-6;11)	73-83 (6;1-6;11)
Age Band 3: 7;0 - 7;11 years		
Number of participants	22 (14:8)	5 (4:1)
Mean age in months (years)	90.6 (7;6)	88.8 (7;5)
Range in months (years)	84-99 (7;0-7;11)	85-94 (7;1-7;10)
Age Band 4: 8;0 - 9;4 years		
Number of participants	19 (13:6)	8 (5:3)
Mean age in months (years)	103.1 (8;7)	103.0 (8;7)
Range in months (years)	96-112 (8;0-9;4)	99-107 (8;3-8;11)
Total Number of participants	88 (54:33)	26 (18:8)
Mean age in months (years)	82.6 (6;10)	85.1 (7;1)
Range in months (years)	54-112 (4;6-9;4)	58-107 (4;10-8;11)

3.2.1.2 Materials and procedure

The Sentence Comprehension test examines the comprehension of different syntactic, morphological, and morphosyntactic structures in Gulf Arabic. Table 2 lists all the different linguistic structures used in the SC test.

Since this project examines the language performance of typically developing and GA speaking children and children with SLI aged between 4-9 years, all language tests were organised into two sections: A and B (each in a different booklet as shown in Appendices: M and N) where section A targeted structures expected to be mastered between 3-5 years old and B contains more advanced items. This division served an organisational purpose only (i.e., it helped divide the battery of tests into smaller coherent units where breaks could be taken preferably between sections of tests and not within sections). Section A and B in the Sentence comprehension test consisted of 22 and 18 items respectively, for a total of 40 items. The distractor items used in the experiment were not systematically controlled. The incorrect pictures generally displayed items that were semantically related to the correct picture. For example, for item 14 (‘the girl is painting’), the distractors show girls performing different actions (e.g., writing, playing).

Table 2: Distribution of items used in the Sentence Comprehension (SC) test, n= 40.

Category		Item Number	Total
Negative		14, 23	2
Modification		12,13, 24	3
Prepositional Phrase		2, 3,29,39	4
Indirect Object		8,21,31	3
Verb Phrase	present	1,5,18,26	4
	past	6,4,	2
	future	16,40,34	3
Relative Clause		10,22, 25,28	4
Subordinate Clause		7,17,30,35,36,37	6
interrogative		11,38	2
Passive		20,33	2
Indirect Request		32	1
Coordinated sentence		9,27	2
Imperative		15	1
Topicalisation		19	1

All children were required to attempt all test items and no basal or ceiling items were set, due to lack of normative data for typical and atypical language development in GA speaking children. In section A, the child was required to listen to a sentence produced by the examiner and point to the right answer from three different pictures on each sheet, while for section B the child selected the correct picture from an array of four pictures. An artist from the Gulf region drew some of the pictures, while others were taken from some English tests, such as CELF-3

(Semel, Wiig, & Secord, 1996) and were carefully examined to ensure they were culturally appropriate for this population.

During testing, children were presented with two trial items and were given instructions in GA equivalent to the following in English: “We are going to look at this book and I will show you some pictures. I want you to point to the picture I am talking about. For example: “show me ‘the girl is sleeping’”. Instructions were repeated if necessary, and there were two trial items to familiarise children with the procedures. All children understood instructions and answered all questions. Self-corrections were accepted and the second answer was considered the final one. Children were given 0 for incorrect scores and 1 for correct answers. The score was written on the record form for the SC test. The highest possible raw score was 40/40. Children were praised for their compliance and not for the accuracy of their answers.

3.2.2 Results and discussion

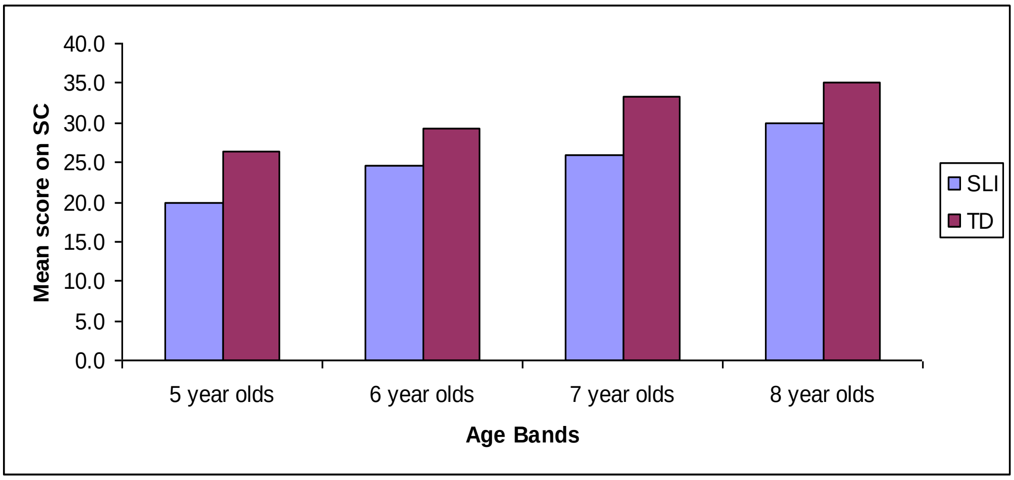
Table 3 summarises the performance of all children on the Sentence Comprehension (SC) test. It shows that children with SLI consistently lagged behind their typically developing (TD) peers in their scores on the SC test, as shown in Figure 1.

One way ANOVA of the scores of the four TD age bands (Age Band 1: 4;6-5;11 years; Age Band 2: 6;0-6;11 ; Age Band 3: 7;0-7;11 ; Age Band 4: 8;0-9;4.) showed there was a significant difference among their performances, $F(3,84)=31.8$, $p<.001$. Post-hoc tests with Bonferroni correction showed that all age groups were significantly different from each other, except the 7 and 8 year old groups. The 5 year old group was significantly different from the 6 year old group $t(45)=-2.89$, $p=.02$ and the 6 year old group had significantly lower score when compared to the 7 year old group $t(43)= -4.0$, $p<.001$. However, there was no significant difference between the 7 and 8 year old groups $t(39)= -1.73$, $p=.54$. This shows that the test was sensitive to age factors in typically developing children, especially for younger children from 4:6 to 7 years old. These differences cease to be significant in children aged between 7 and 8 years old, because the test becomes less challenging at this age.

Table 3: Means (and standard deviations) for performance on the Sentence Comprehension test.

Age Groups	Typically Developing Children	SLI
4;6-5;11 years		
Number of participants (Male:Female)	24 (13:11)	5 (2:3)
Mean Raw Score of SC Test (SD)	26.4 (3.65)	19.80 (4.65)
Range of SC scores	20-33	15-26
6;0-6;11 years		
Number of participants	23 (15:8)	8 (7:1)
Mean Raw Score of SC Test (SD)	29.3 (3.38)	24.63(4.56)
Range of SC scores	24-37	18-31
7;0-7;11 years		
Number of participants	22 (14:8)	5 (4:1)
Mean Raw Score of SC Test (SD)	33.3 (3.41)	26.00(4.52)
Range of SC scores	27-38	20-31
8;0-9;4 years		
Number of participants	19 (13:6)	8 (5:3)
Mean Score of SC Test (SD)	35.1 (4.05)	30.00 (5.19)
Range of SC scores	32-39	21-35
Total Number of children	88 (54:33)	26 (18:8)
Mean Raw Score of SC Test (SD)	30.8 (4.64)	25.62 (5.78)
Range of SC scores	20-39	15-35

Figure 1: Comparison between children with SLI and their typically developing (TD) peers on the Sentence Comprehension (SC) test.



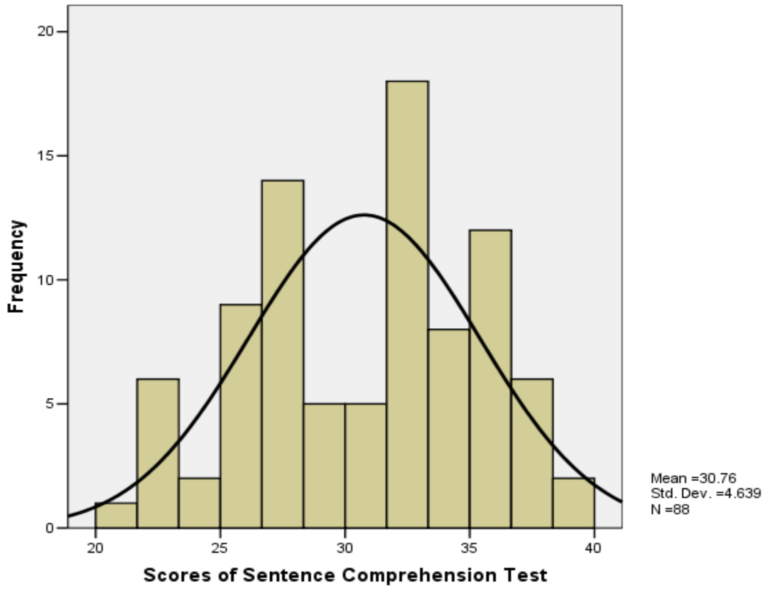
A T-test was performed to compare the overall means of the two groups. It showed that the TD group was significantly better than the SLI group on the SC test $t(112)=4.6$, $p<.001$. Children with SLI obtained a mean score equivalent approximately to their TD peers who were 2 years younger, as depicted in Figure 1.

ANOVA of the SC scores of the four groups of children with SLI showed a significant effect of group $F(3,22)=4.8$, $p=.01$. Multiple comparisons with Bonferroni correction showed that only the 5 and 8 year old groups were significantly different from each other, $t(11)= - 10.2$, $p<.01$. No comparisons among the other age groups were significantly different from each other. These null findings may be explained by a combination of small sample sizes and lack of developmental effects, whereby severity level might have influenced performance more than chronological age.

3.2.2.1 Distribution of Test Scores

One of the important psychometric properties of a test is the distribution of test scores. The following figure shows the distribution of the scores of all typically developing children on the Sentence Comprehension test. It depicts a broadly normal distribution of these scores across the whole sample of TD children.

Figure 2: Distribution of typically developing children on scores of the SC test.



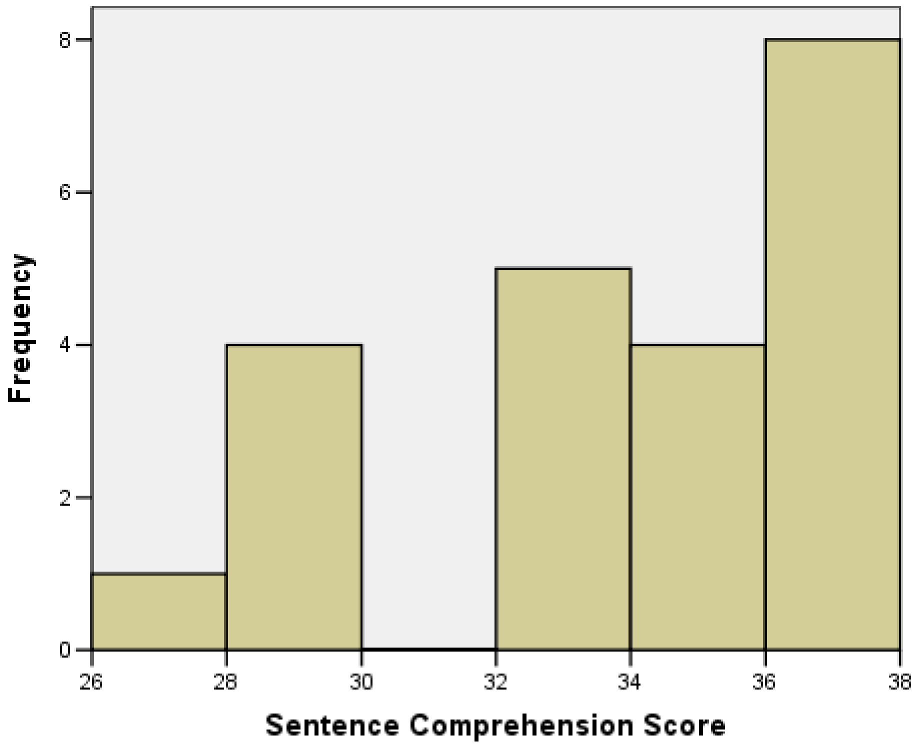
The Shapiro-Wilk test of normality was conducted on data for each age group. Only for the 7 year old group was there a significant departure from normality.

Table 4: Results of the Shapiro-Wilk normality test for the Sentence Comprehension test.

Age Group	Statistic	Significance
5 year olds (n=24)	.96	.48
6 year olds (n=23)	.96	.47
7 year olds (n=22)	.90	.044
8 year olds (n=19)	.93	.17

The following histogram depicts the distribution of the scores of the TD 7 year old children on the SC test. It shows a positively skewed distribution with a high proportion of children reaching ceiling level scores on the test. It is worth noting that although these 7 year old children showed positively skewed distribution, their older counterparts (the 8 year old group) had a normal distribution.

Figure 3: A histogram showing the distribution of the Sentence Comprehension scores for the 7 year old group of TD children.



3.2.2.2 Reliability

Reliability refers to the ability of a test to yield consistent measures when used under identical conditions. It is usually divided into three types of measures: split-half analysis, Cronbach's alpha, and test-retest reliability.

Split-half Analysis. A split-half analysis was conducted to examine the correlation between the scores obtained from odd-numbered items with scores from even-numbered items. This was more appropriate than measuring the correlation between the first and second half of the test as items were arranged in terms of difficulty. Correlation coefficients of .70-.80 are considered acceptable (Field, 2005). The results of split half-analysis with Spearman-Brown coefficient showed a correlation of .89 between odd and even items, indicating a significant level of internal consistency.

Cronbach's alpha. While the Split-half analysis groups the items into one way only (e.g., odd vs. even), Cronbach's α splits the data into two in every possible way and then computes the correlation between these items. Therefore, it is considered a better measurement of internal reliability. Cronbach's alpha is considered acceptable if it falls between .70 and .80 (Field, 2005). The Cronbach's α for the sentence comprehension test was .79, indicating the presence of a good level of internal reliability. Cronbach's alpha provided alpha value for each item on the test and how it correlated with the overall score. Table A- 1 in Appendix A shows the alpha values for each test item and it shows no item needs to be deleted to improve the Cronbach's alpha.

Tests re-test reliability. Test-retest reliability is used to measure the stability of the test when used with the same individual over time. To examine test-retest reliability of the sentence comprehension test, 6 children were retested one week after they took the test for the first time. This group of children consisted of five male students and one female student, aged 75 to 107 months (6;3-8;11 years old). Five of them were typically developing and one was diagnosed with SLI. Results of test-retest reliability showed a Pearson correlation coefficient of .95, p=.003, indicating that the test is stable over time. It is noteworthy that the highest variability was noticed in a child with SLI.

Table 5: Test and re-test raw scores for the Sentence Comprehension test.

Subject	Sex	Age in months	Test	Re-test	Language Ability
625	Male	75	39.00	45.00	Typically Developing
626	Male	77	64.00	65.00	Typically Developing
628	Male	79	56.00	59.00	Typically Developing
624	Female	82	55.00	56.00	Typically Developing
11627	Male	83	44.00	53.00	SLI
823	Male	107	54.00	55.00	Typically Developing

3.2.2.3 Validity

Validity refers to the extent to which a test measures what it intends to measure. Two types of validity are usually assessed: content validity and concurrent validity.

Content Validity. Content validity refers to what extent the test items are relevant and representative of the targeted constructs being assessed (Haynes, Richard, & Kubany, 1995). To ensure that the Sentence Comprehension test possesses an appropriate level of content validity, all structures used in the test were chosen based on the same criteria adopted by Al-Akeel (1998) in his test of comprehension of morphosyntactic structures in Saudi Arabic. Therefore, the structures were selected based on these criteria: they appeared in the language samples of TD GA speaking children; they were chosen by the investigator based on his native knowledge of the language and his clinical experience as a speech-language therapist. Thirdly, some structures were carefully chosen from English language tests (such as CELF-3 (Semel, Wiig, & Secord, 1996) or PLS-3 (Zimmerman, Steiner, & Pond, 1992), provided they also appear in Gulf Arabic and are culturally appropriate.

Concurrent validity. Concurrent validity measures the correlation of the novel test with other tests taken by the same group of children at the same time (Anastasi & Urbina, 1997). Ideally, these tests should tap into the same skill, e.g. various vocabulary tests are expected to correlate with each other. However, due to lack of any standardised tests in Gulf Arabic, the Sentence Comprehension test had to be compared to tests developed in this project that measure various aspects of language abilities, such as the Expressive Language test, the Sentence Repetition test, and the Arabic Picture Vocabulary Test. Results of the Pearson Correlation revealed significant correlations ($p<.001$) between the Sentence Comprehension Test and all of these measures, as illustrated in Table 6.

Table 6: Correlation between the standard score of the Sentence Comprehension (SC) Test (*n*=114) and other tests.

	SR (*n*=111)	EL (*n*=111)	APVT (*n*=105)
The SC Test	.43**	.54**	.63**

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

Note. SR=Sentence Repetition test, EL=Expressive language test, APVT=the Arabic Picture Vocabulary Test.

3.2.2.4 Item Analysis

Item analysis was conducted to assess the overall item difficulty. Table A-2 in Appendix A shows the proportion of TD children who answered each item correctly. These proportions ranged between 0.0 (no one answered a particular item correctly) and 1.0 (all children had correct answers).

Table A-2 was examined to identify items that were not consistent with the general pattern of decreasing item difficulty in TD older groups in comparison with younger groups. Based on this, those items with a substantial difference in favour of the younger group were identified. A younger group is said to have a significant difference when its proportion of correct responses on a particular item is .05 or higher in comparison to an age band that is two years older than it is. Employing this criterion, item 8 was identified as an item that may warrant revision in future versions of the test. The 5 year old group (Age band A) had a proportion of .79 correct responses, while the children in the 8 year old age group had .74. These differences were accepted if they were restricted to consecutive age groups (e.g., 5 vs. 6 year olds or 7 vs. 8 year olds), as this was not considered as a substantial deviation. Apart from item 8, item 40 was found very difficult by all groups and therefore may warrant further revision.

Table A-3 in Appendix A shows the proportion of correct responses for children with SLI. Generally, children with SLI did relatively well on the first 10 items, however as the test proceeded, the items became more challenging for children with SLI. Due to the nature of the test, which is a general test of sentence comprehension, it was difficult to compare the TD and SLI groups on items due to limited exemplars from each linguistic structure. However, some items of interest for further follow up could be relative clauses (items 22 and 28), negation (14 and 23), and passive (20). These are linguistic structures that are known to be challenging for children with SLI in other languages, however, the small number of exemplars in the test does not allow for in-depth analysis. Table 7 lists these items as they appeared in the SC test and the proportion of correct responses for TD children and children with SLI.

Table 7: Proportion of correct responses of the TD and SLI participants on some linguistic structures of the Sentence Comprehension test.

Item	Linguistic structure	The item as it appeared in the SC test	SLI (n=26)	TD (n=88)
22	Subject relative clause	Il-walad shaaf ɪl-bɪnt ɪlli qaʕda tʃi:l matˤrɪga The-boy saw the-girl who is carrying hammer The boy saw the girl who is carrying a hammer.	.32	.76
28	Subject relative clause	Il-mara ɪlli sha:yla lɪ-bnayya tˤayyaħat ʃantˤat-ha The-woman who carrying the-girl dropped bag-her The woman who is carrying the girl dropped her bag.	.34	.58
14	negation	Il-bɪnt mub gaʕda tarsɪm The-girl not doing drawing The girl is not drawing	.63	.88
23	negation	Il-walad mub gaʕid yɪ-lʕab The-boy not doing playing The boy is not playing	.68	.90
20	(truncated) passive	Il-bɪnt gaʕda tɪndaz The-girl doing being pushed The girl is being pushed.	.50	.74

Children with SLI seem to have difficulties with all these structures as their proportion of correct responses is at least .20 less than TD children. This discrepancy is even higher in subject relative clauses. However, the limited number of items per structure does not allow for more than noting these general trends, that could be examined in depth in future studies.

3.3 Test 2: The Expressive Language (EL) Test

3.3.1 Method

3.3.1.1 Participants

The Expressive Language (EL) test was conducted with 112 Qatari speaking children aged between 4;6 and 9;4. Eighty six were typically developing and 26 were diagnosed with SLI. Children with SLI were selected based on the criteria used throughout this project, i.e., they passed hearing screening, had no history of sensory, motor, or social impairments and scored two or more standard deviations below the mean on one test or -1.5 SD or more on at least two of the four language tests developed for this project. All children scored within the normal range on one of the two nonverbal IQ tests employed, namely; the Test of Nonverbal Intelligence (TONI-3) (Brown et al., 1997) for children aged 6 years and above or the Block Design and Picture Completion subtests of the Wechsler Preschool and Primary Scale of Intelligence (WPPSI-III) (Wechsler, 2002). These children were mostly the same children who completed the Sentence Comprehension (SC) test. See *Table 8* for a summary of participants' characteristics.

Table 8: Participants' characteristics for the Expressive language (EL) test.

Age Groups	Typically Developing Children	Children with SLI
Age Band 1: 4;6 - 5;11 years		
Number of participants (Male: Female)	24 (13:11)	5 (2:3)
Mean age in months (years)	64.0 (5;3)	62.6 (5;2)
Range in months (years)	54-71 (4;6-5;11)	58-70 (4;10-5;10
Age Band 2: 6;0 - 6;11 years		
Number of participants	23 (15:8)	8 (7:1)
Mean age in months (years)	77.6 (6;6)	78.9 (6;7)
Range in months (years)	72-83 (6;0-6;11)	73-83 (6;1-6;11)
Age Band 3: 7;0 - 7;11 years		
Number of participants	21 (13:8)	5 (4:1)
Mean age in months (years)	90.4 (7;6)	88.8 (7;5)
Range	84-95 (7;0-7;11)	85-94 (7;1-7;10)
Age Band 4: 8;0 - 9;4 years		
Number of participants	18 (12:6)	8 (5:3)
Mean age in months (years)	103.4 (8;7)	103.0 (8;7)
Range	96-112 (8;0-9;4)	99-107 (8;3-8;11)
Total No. of participants	86 (53:33)	26 (18:8)
Mean age in months (years)	82.2 (6;10)	85.1 (7;1)
Range in months (years)	54-112 (4;6-9;4)	58-107 (4;10-8;11)

3.3.1.2 Materials and procedure

The Expressive Language (EL) test measured the production of various morphosyntactic structures commonly used by Gulf Arabic speaking children. It consisted of 68 items divided into two sections: EL (A) examined early developing morphosyntactic structures and comprised 24 items, while EL (B) targeted more advanced language learners and consisted of 44 items. Each section was in a separate booklet (see Appendices O and P). However, this division was used for organisational purposes, as all children were required to answer all items. The distributions of all EL items are listed in Table 9. These linguistic structures were chosen based on structures seen in language samples of TD children, the investigator's native knowledge of Gulf Arabic, his experience as a speech language pathologist and previous research on Gulf Arabic (e.g., Aljenaie, 2001) or varieties that are close to Gulf Arabic, such as those spoken in Saudi Arabia (e.g., Abdalla, 2002; Al-Akeel, 1998). Some English language tests, such as the Clinical Evaluation of Language Fundamentals-CELF3 (Semel, Wiig, & Secord, 1996) and Preschool Language Scale-PLS4 (Zimmerman, Steiner, & Pond, 1992) were consulted and some structures that appear in Arabic were used while ensuring their ecological validity (e.g., superlatives). Other clinicians working with Gulf Arabic children in Qatar were consulted about appropriate structures to be used with this population and their input was incorporated in the choice of items used in the test.

All testing was performed in a quiet room at school or home and was performed along with other tests in the battery. Usually, the EL test followed the IQ test and the sentence comprehension test when a good rapport had already been established between the examiner and the child. The testing started with two practice items and the instructions were as follows (in Arabic): "Together, we will look at some pictures. I will show you some pictures, I will say something and I want you to complete what I say. For example (showing the child a picture of one strawberry): "here we have a strawberry, and here (pointing to the picture of three strawberries in the second page) we have three… (Child is expected to say 'strawberries')". Example 3-1 illustrates elicitation procedure for another item. Children would get a score of 1 for a correct answer or 0 for an incorrect one. In this test, single repetition was allowed and a specific prompting procedure was permitted. When a child did not reply, her/his score was considered as 'no response' (NR) and she/he would get a score of 0.

Example: 3-1

Examiner (pointing to a picture showing a girl sleeping)
'Hni: ɪl-bɪnt targɪd' ('Here: the girl is sleeping')
Turning to another page with a picture of a girl playing
'Hni: ɪl-bɪnt................. ('tɪlʔab') ('Here: the girl is....... 'playing'')

Table 9 : Distribution of Items in the Expressive Language (EL) test.

Linguistic Structure				Item Number
Possessive pronouns+Cl		3rd Person Feminine singular		25
		3rd person plural		26
		3rd Person Masculine singular		27
		2nd person masculine Singular		28
Subject Pronouns		Plural		15
Demonstrative Pronouns				61, 62
Reflexive Pronouns		Plural		10, 67
		Masculine Singular		68
Prepositions				1, 4
Possessive particle				9,29, 30
Plurals		Regular	Feminine	3,31, 33
			Masculine	32
		Irregular		34, 35, 36
Dual			Masculine	37, 38
			Feminine	39
Verb Markers	Present	3rd Person Masculine Plural		7, 40
		2nd person Feminine Singular		16, 41
		2nd Person Plural		42
		3rd Person Feminine Singular		2, 8
	Past	3rd Person masculine singular		11
		3rd person Plural		50
		3rd person Feminine Singular		51
		3rd person masculine singular		52
	Future			13
Construct State				43, 44
Derivation of Nouns				43, 44
Derivation of Adjectives				47, 48,49
Adjective	Plural			18,20
	Feminine			22,23
	Dual			17,19,
Clitic Pronouns				
	Dative Clitic 3rd Person MS			53, 54
	Object pronoun clitic 3rd FS			55,56
	Genitive (Possessive) clitic 2nd MS			57
	Genitive (Possessive) clitic 2nd MP			58
	Genitive (Possessive) clitic 3rd FS			5,6,59
	Genitive (Possessive) 3rd MS			60
	Object clitic 3rd FPl			12
	Object clitic 3rd MS			14
Comparative and Superlative				
Comparative				63, 64
superlative				65, 66
Negation				21, 4

3.3.2 Results and discussion

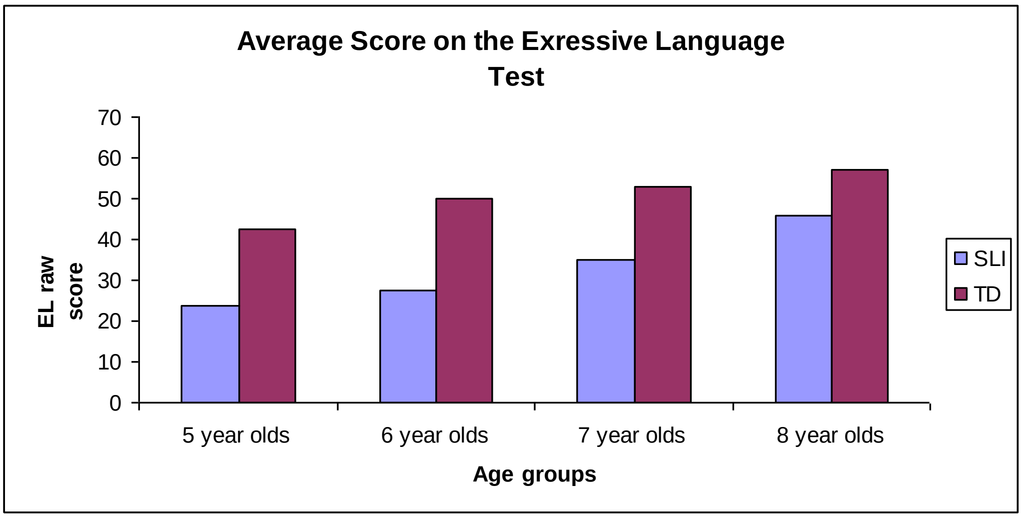
Table 10 summarises the results of all children on the Expressive Language test. It shows that children with SLI were consistently lagging behind their TD peers, and that performance of the TD groups improved consistently with age.

Table 10 and *Figure 4* show that the oldest group of children with SLI (8 years old) had a score that was close to the score achieved by the youngest TD group (4;6-5;11 years old), indicating that production of various syntactic and morphological structures constitute a major area of deficits in GA children with SLI.

Table 10: Results of all participants on the Expressive Language (EL) test.

Age Groups	Typically Developing Children	SLI
Age Band 1: 4;6 - 5;11 years		
Number of participants (Male: Female)	24 (13:11)	5 (2:3)
Mean Raw Score of EL Test (SD)	42.6 (7.8)	23.6 (3.0)
Range of EL scores	30-55	20-27
Age Band 2: 6;0 - 6;11 years		
Number of participants	23 (15:8)	8 (7:1)
Mean Raw Score of EL Test (SD)	50.0 (6.6)	27.4 (13.6)
Range of EL scores	37-61	10-43
Age Band 3: 7;0 - 7;11 years		
Number of participants	21 (13:8)	5 (4:1)
Mean Raw Score of EL Test (SD)	52.8 (5.4)	35.2 (4.3)
Range of EL scores	44-62	29-39
Age Band 4: 8;0 - 9;4 years		
Number of participants	18 (12:6)	8 (5:3)
Mean Raw Score of EL Test (SD)	57.2 (4.0)	45.9 (9.1)
Range of EL scores	51-66	32-59
Total Number of children	86 (53:33)	26 (18:8)
Mean Raw Score of EL Test (SD)	50.1 (8.1)	33.8 (12.7)
Range of EL scores	30-66	10-59

Figure 4 : Comparison of the overall Expressive Language raw scores by children with SLI and typically developing (TD) children across different age groups.



T-test showed that the group of TD children was significantly better than the SLI group on the EL test $t(31.4)=6.6$, $p<.001$. One-way ANOVA of the scores of the four TD age groups, showed a significant group effect, $F(3,82)=20.4$, $p<.001$, indicating the presence of a developmental factor. Multiple comparisons with Bonferroni effect showed that the 5 year old group scored significantly lower than all three older groups. The 5 year old group was significantly different from the 6 year old group $t(45)=-7.41$, $p=.001$. However, the 6 year olds' scores were not significantly different from the 7 year old group, but they had significantly lower scores when they were compared with the 8 year old group $t(39)= -7.12$, $p=.003$. There was no significant difference between the 7 and 8 year old TD groups.

3.3.2.1 Distribution of test scores

The following figure shows the distribution of the scores of all typically developing children on the Expressive Language (EL) test. It depicts a normal distribution of these scores, though with some positive skewness. Moreover, The Shapiro-Wilk test of normality was conducted and it was not significant for any of the groups, indicating that they had normal distributions, except for the 5 year old group (see Table 11).

Figure 5: Distribution of typically developing children on scores of the Expressive Language test.

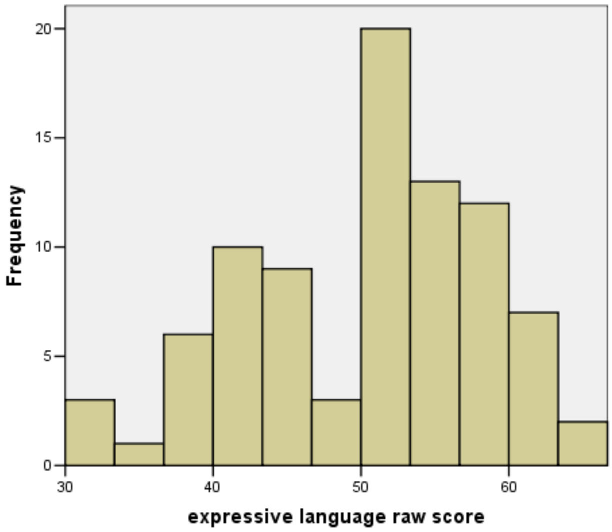
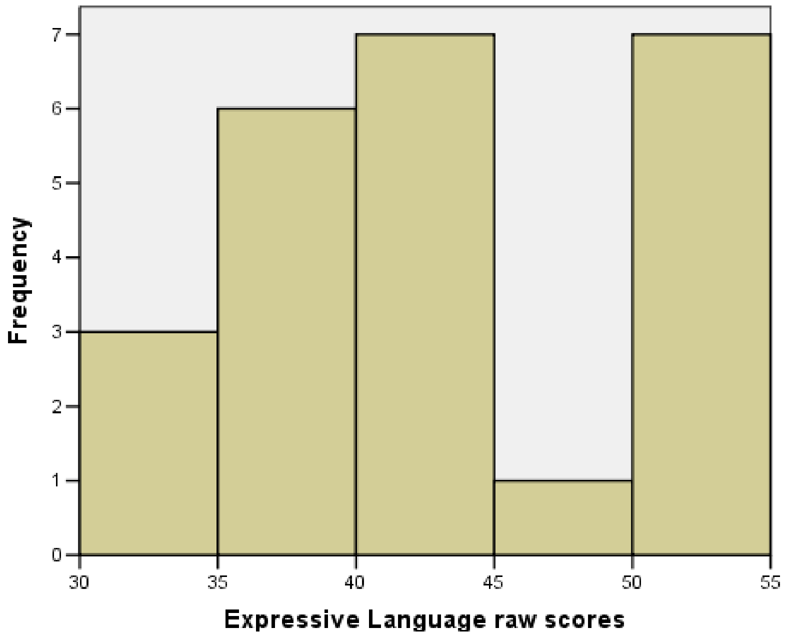


Table 11: Results of the Shapiro-Wilk test of normality for the four age groups.

Age Group	Statistic	Degrees of freedom	Significance
5 year olds (n=24)	.90	24	.03
6 year olds (n=23)	.97	23	.90
7 year olds (n=21)	.95	21	.42
8 year olds (n=18)	.95	18	.50

The following histogram (Figure 6) shows the distribution of the 5 year old TD children. It shows a relatively negative skewed distribution of the scores.

Figure 6: Distribution of the raw scores of the 5 year old TD group on the Expressive Language test, n=24.



3.3.2.2 Reliabilityy

Split-half analysis. A split-half analysis was conducted to examine the correlation between the scores obtained from odd-numbered items with scores from even-numbered items. Results of split-half analysis with Spearman-Brown coefficient showed a correlation of .94 between odd and even items, indicating a significant level of internal consistency.

Cronbach's alpha. The Cronbach's α for the Expressive Language test was .93, indicating the presence of a significant level of internal reliability. Examination of the total items correlation as shown in Table A- 4 in Appendix A showed no one item could have increased alpha if deleted.

Test-re test reliability. Six children were re-tested on the Expressive Language test a week after they took the test for the first time. This group of children consisted of five male students and one female student, aged 75 -107 months (6;3-8;11). Results of test-retest reliability showed a Pearson correlation coefficient of .95, p=.003, indicating that the test is stable over time. Individual data for each child are shown in the table below.

Table 12: Raw scores for the Expressive Language test and re-test by six children

ID	Sex	Age in months	Test	Re-test	Language Ability
625	Male	75	28	31	Typically Developing
626	Male	77	28	31	Typically Developing
628	Male	79	28	33	Typically Developing
624	Female	82	22	25	Typically Developing
11627	Male	83	22	23	SLI
823	Male	107	36	37	Typically Developing

3.3.2.3 Validity

Content validity. Most of the linguistic structures were included in the Expressive Language test based on language samples taken from more than 35 Gulf Arabic speaking children, whose age ranged between 2;11 and 4;11 years old (Khater & Shaalan, 2007; Shaalan & Khater, 2006). Some linguistic structures were based on the investigator's knowledge of Gulf Arabic, as a native speaker, and on his experience as a speech-language therapist working with Gulf Arabic speaking children with and without language impairment. Additionally, a group of Gulf Arabic speaking clinicians and linguists were asked to examine the structures in the Expressive Language test before conducting the tests. Their overall responses were positive and they gave some suggestions about elicitation that were incorporated in the test. Furthermore, some standard English language tests were consulted in order to examine linguistic structures that were relevant for Gulf Arabic. For example, the addition of comparative and superlatives was based on their existence in some English tests, such as the CELF-3 (Wiig et al., 1996). All structures were included after making sure they were linguistically and culturally appropriate. For example, while English has regular and irregular plurals, Arabic's plural system is characterised by the presence of a minority of regular plural nouns and majority of irregular plurals. Moreover, the regular plurals in Arabic are divided into feminine regular plurals and masculine regular plurals and Arabic has a dual structure. Therefore, when examining the performance of GA speaking children on plurals, items representing this complex plural system had been included.

Concurrent validity. Results of the correlation study of the Expressive Language (EL) test showed that this test correlated significantly with the rest of the language tests used with Gulf Arabic speaking children. Table 13 showed that the EL test had a correlation coefficient of .69 (p<.001) with the Sentence Repetition test, .54 (p<.001) with the Sentence Comprehension Test and .50 (p<.001) with the Arabic Picture Vocabulary test.

Table 13: Correlation coefficients between the Expressive Language (EL) test standard score and standard scores of other language tests.

	SC (*n*=114)	SR (*n*=112)	APVT (*n*=105)
EL test (*n*=112)	.54**	.69**	.50**

**Correlation is significant at the 0.001. level (2-tailed).

Note. SC= Sentence Comprehension test, SR=Sentence Repetition test, , APVT=the Arabic Picture Vocabulary Test.

3.3.2.4 Item Analysis

Table A- 5 and Table A- 6 in Appendix A show the proportion of examinees who answered each item correctly. Some of the items were very difficult for all children, including TD children, so they warrant revision or deletion in future versions of this test. These were items: 18, 19, 47 and 64.

Though this is considered a general expressive language test that was designed to assess various linguistic structures that exist in Gulf Arabic, some initial conclusions can be drawn about structures that were well represented in the test, such as verb morphological markers, plurals, and clitic pronouns. Table 14 lists the numbers of items representing each of these linguistic structures and how children with SLI fared in comparison with TD children.

Table 14: Proportion of correct responses of the TD and SLI participants on some linguistic structures of the Sentence Comprehension test.

Linguistic Structure			Item Number	SLI n=26	TD n=86
Verb Markers					
	Present	3rd Person Masculine Plural	7, 40*	.95	1.0
		2nd person Feminine Singular	16, 41*	.73	.93
		2nd Person Plural	42	.73	.95
		3rd Person Feminine Singular	2, 8*	.65	.90
	Past	3rd Person masculine singular	11	.38	.71
		3rd person Plural	50	.40	.63
		3rd person Feminine Singular	51	.21	.63
		3rd person masculine singular	52	.58	.76
	Future		13	.26	.49
Plurals					
	Regular	Feminine	3,31, 33*	.83	.98
		Masculine	32	.21	.50
	Irregular		34, 35, 36*	.36	.73
Clitic Pronouns					
	Prepositional Clitic 3rd Person MS		53, 54*	.24	.67
	Object pronoun clitic 3rd FS		55,56*	.36	.82
	Genitive pronoun clitic: 2nd MS		57	.27	.49
	Genitive clitic 2nd MP		58	.32	.72
	Genitive clitic 3rd FS		59	.11	.36
	Genitive clitic 3rd MS		60, 5,6*	.51	.80
	Obj. 3rd Feminine Plural		12	.11	.42
	Obj. 3rd Masculine Singular		14	.27	. 73
Possessive particle			9,29, 30*	.83	.99
Possessive particle+Clitics					
		3rd Person Feminine singular	25	.52	.81
		3rd person plural	26	.56	.91
		3rd Person Masculine singular	27	.48	.81
		2nd person masculine Singular	28	.38	.71

*Where there is more than one item, the proportion represents the means of these items.

Based on the results shown in Table 14, the following initial remarks about expressive language abilities of children with SLI are suggested:

Verb inflections: The following table shows the distribution of various verb inflectional markers in Gulf Arabic for present (imperfective) and past (perfective). The future is represented by adding an auxiliary 'ra:ħ' (will) with the imperfective inflections. While the imperfective takes mostly prefixes to mark person, number, and gender, the perfective takes suffixes.

Table 15: Perfective (past) and Imperfective inflections for the Arabic verb 'yɪlʕab' (play), which has the consonantal root (-l-ʕ-b).

Person	Number	Gender	Imperfective	Perfective
1st	S	M+F	ʔ-alʕab	laʕab-t
	Pl	M+F	nilʕab	laʕab -na
2nd	S	M	t-ilʕab	laʕab-t
	S	F	t-alʕɪbiːn	laʕab-ti
	Pl	M+F	t-alʕobou:n	laʕab -tau
3rd	S	M	y-ilʕab	laʕab (no suffix)
	S	F	t-ɪlʕab	lʕaba-t
	Pl	M+F	ya-lʕib-oun	lʕab-au

Note. S=singular, Pl=plural, M=masculine, F=feminine.

Examination of the performance of the groups of children with SLI and TD children on verb markers in Table 14 shows that generally children with SLI seem not to have significant difficulties with verb inflections, especially for non-finite (present) verb inflections. They tended to perform less well on finite verbs (past). This trend, however, is based on a few verb inflections and therefore further examination of verb inflections is needed. When both tense and agreement markers used in the test were combined, children with SLI had a mean of 58% percent while TD children had a mean of 80%. Most of the errors in finite tense were based on substitution of the non-finite tense. Abdalla (2002) examined spontaneous language samples of Saudi Arabic speaking children with SLI, who were between 4;0-5;3

years and found that they had difficulties with both tense and agreement markers. These children had percent correct production of 68% and 77% on tense and agreement markers respectively, while their language and age controls score 93% and 100% respectively with no difference between tense and agreement. However, Dromi et al. (2003) found that Hebrew speaking children with SLI generally did well on verb morphology, though they had some difficulties in verb forms that involve complex morphophonological and semantic structures, a pattern seen in the verb markers in the present study, where verbs with present markers seem to be easier than past tense verbs that involve more complex morphophonological manipulation in Arabic.

Plurals: In Arabic, plural nouns refer to more than two countable entities, as the dual is used for two entities. The plural system in Arabic is characterised by the existence of a majority of irregular plural patterns whereby the regular plural is a minority. In regular plural, the default marker is the feminine regular plural, which involves adding the suffix *–a:t*, as in *sayyara* (car S), which takes the plural *sayyara:t* (cars Pl). The other type of regular plural is regular masculine plural which involves adding the suffix –i:n, as in *mudarrɪis* (male teacher S), whose plural form is *mudarrɪsi:n* (male teachers Pl). The irregular plural involves inserting various forms of consonantal/ vowel patterns to the roots of the singular noun, e.g., *kitaab* (book S) takes the plural *kutub* (books Pl) (Holes, 2004). Based on the results of the test, regular masculine plural seems to be acquired much later than the default regular feminine plural. TD children as old as 6 years old showed some difficulties using this form of plural and they resorted to using the regular feminine plural *mudarɪsa:t* (female teachers Pl), when referring to a picture of three male teachers.

Table 14 shows that children with SLI seemed to do relatively well on the early developing regular feminine plural nouns (0.83 vs. 0.98 for TD children), though they had more difficulties with irregular plurals that required morpho-phonological manipulation of the singular noun. There was only one exemplar of masculine plurals, and both groups had difficulties with this type of plural noun. Therefore, this initial examination of plurals in SLI showed that these children may not have difficulties with regular feminine plurals at this age, but may have more problems using irregular plurals that involve inserting various consonantal/vowel patterns that require good morpho-phonological skills. However, these observations need more systematic investigation of the acquisition of regular and irregular plurals in Gulf Arabic.

Clitics: Clitics in Arabic can appear on all lexical categories, and sometimes on some functional categories (e.g., prepositions) and they always appear to the right of their hosts (Shlonsky, 1997). The Expressive Language test contained many examples of various clitic pronouns. An example of a direct object clitic and how it was elicited is shown in 3.2.

3.2

(Child is presented with a picture of boys eating pizza), then the examiner says:

‘Haðeila Il-awlad yaklu:n pizza’

‘These children are eating pizza’,

(then the examiner shows another picture of the eaten pizza and says:

‘Yaʔni haðI ɪl-pizza ɪlli al-awlaad…(kalu:-*ha*)’

This means this is the pizza that the children … (ate- it- CL FS).

The most common error was dropping the clitic (–*ha* in the example above), this was followed by errors of substitution, where children used another type of clitic, mostly third person masculine singular clitic. Table 14 shows that generally children with SLI did much worse than TD on all types of clitics, and it seemed that they did best on 3rd masculine singular clitic, which could be a default clitic. These difficulties in clitics, especially object clitics in children with SLI have been reported in other languages (Bedore & Leonard, 1998; Jakubowicz, Nash, Rigaut & Gérard, 1998; Leonard & Bortolini, 1998). Therefore, based on these preliminary results, clitic pronouns in Arabic might prove very difficult for children with SLI and these deficits might be one of the major characteristics of SLI in Arabic

Possessive particles and pronouns: The possessive particle *maal* is used in Gulf Arabic, as in: ‘ɪl kɪtaab *maal* Khalɪd’ (Khalid’s book). The same particle can be used to indicate possession with a clitic attached to it, as in 3-3:

3-3

Mohamed ɪʃtara ɪl-kɪtaab... yaʕni haðа ɪl-kɪtaab…..(*maal-**ah**)*
Mohamed bought the book, this means this book (is)…*his).*

An examination of the proportion of correct responses showed that children with SLI seemed to have difficulty with possessive particles mainly when a clitic is attached to them. This pattern is consistent with their performance on all kinds of clitics, which seem to be very challenging for children with SLI.

In summary, analysis of some of the linguistic structures in the Expressive Language test can only give initial and general impressions about both typical and atypical acquisition of some structures that were well represented in the test. Children with SLI seemed to have significant difficulties with clitic pronouns and irregular plurals. They presented with less difficulties with present tense markers and regular feminine plurals. However, these were not systematically investigated due to the nature of the EL test and therefore the results are far from being conclusive.

3.4 Test 3: The Sentence Repetition (SR) Test

3.4.1 Method

3.4.1.1 Participants

The Sentence Repetition (SR) test was conducted with 112 Qatari speaking school children aged between 4;6 and 9;4. Eighty six of these children were typically developing (TD) and 26 were diagnosed with SLI based on the same criteria used throughout this project. They all passed hearing screening, articulation and apraxia screening and scored within normal range on nonverbal IQ tests (the Test of Nonverbal Intelligence (TONI-3) (Brown et al., 1997) for children aged 6 years and above; or the Block Design and Picture Completion subtests of the Wechsler Preschool and Primary Scale of Intelligence (WPPSI-III) (Wechsler, 2002). Children with SLI were diagnosed based on a Z-score of -2.0 or more on one test or -1.5 or more on at least two of the four language tests developed for this project. See Table 16 for a summary of participants' characteristics.

Table 16: Participants' characteristics for the Expressive language (EL) test.

Age Groups	Typically Developing Children	SLI
Age Band 1: 4;6 - 5;11 years		
Number of participants (Male: Female)	24 (13:11)	5 (2:3)
Mean age in months (years)	64.0 (5;3)	62.6 (5;2)
Range in months (years)	54-71 (4;6-5;11)	58-70 (4;10-5;10
Age Band 2: 6;0-6;11 years		
Number of participants	23 (15:8)	8 (7:1)
Mean age in months (years)	77.6 (6;6)	78.9 (6;7)
Range in months (years)	72-83 (6;0-6;11)	73-83 (6;1-6;11)
Age Band 3: 7;0-7;11 years		
Number of participants	21 (12:9)	5 (4:1)
Mean age in months (years)	90.4 (7;6)	88.8 (7;5)
Range	84-99 (7;0-7;11)	85-94 (7;1-7;10)
Age Band 4: 8;0-9;4 years		
Number of participants	18 (12:6)	8 (5:3)
Mean age in months (years)	102.8 (8;7)	103.0 (8;7)
Range	96-112 (8;0-9;4)	99-107 (8;3-8;11)
Total No. of participants	86 (54:34)	26 (18:8)
Mean age in months (years)	82.2 (6;10)	85.1 (7;1)
Range in months (years)	54-112 (4;6-9;4)	58-107 (4;10-8;11)

3.4.1.2 Materials and procedure

The Sentence Repetition (SR) test consisted of 41 sentences divided into two sections: A (items 1-18) and B (items 19-41), see Appendices Q and R. The sentences were arranged in a least-to-most difficult order. Sentences increased in length and grammatical complexity as the child progressed through the test. *Table 17* shows the distribution of the SR test items. Some of the linguistic structures used in the test were similar to the ones used in the Sentence Comprehension test. However, their length increased (e.g., by using relativisations and passive structures). The following table shows the distribution of these items.

Table 17: Distribution of the items used in the Sentence Repetition test (*n*=41).

Category			Item number
Simple	Active		2,4, 10
	With noun modification		9, 16, 23, 40
	With negation		12, 15
	With coordination		6, 14, 25, 35
	Conjunction deletion		39
	Negation		21
Imperative			1, 8
Interrogative	What/where		3, 5 , 17
	With noun modification		7, 19
	With coordination		11
Complex	With relativisation		17, 24, 27, 28, 30, 33, 34, 41
	With subordination		13, 18 26, 36, 37
	Passive		
		negative	22
		with subordinate clause	38, 31,32
	Topicalisation		20, 29

All testing was done at a quiet room at school or home and the test was part of the battery of tests used in this project. The instruction was the equivalent of the following in Arabic: “You will hear some sentences and I will say each one once only. I want you to repeat them exactly the way I say them”. This was followed with two practice items. Most children did not have problems understanding the instructions; in a few cases, a third example was needed. The scoring method used was adapted from the one used in the Clinical Evaluation of Language Fundamental-3 (CELF-3). Therefore, children would get 3 points if they repeated the whole sentence with no errors, 2 points when there was one error, 1 point when there were two errors and 0 if they produced three or more errors or when they provided no response. Error was defined as any change in the sentence that is not of articulatory nature. No repetition of any sentence was allowed. There was no basal or ceiling and children were required to attempt repeating all sentences.

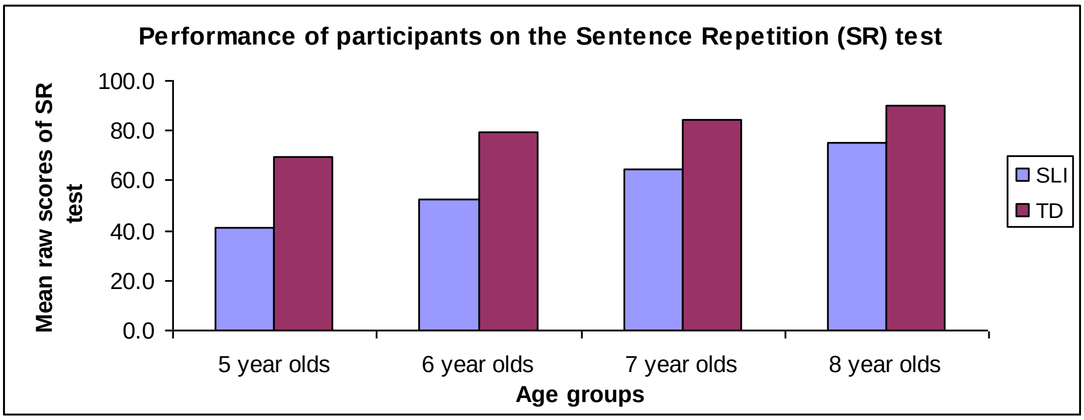
3.4.2 Results and discussion

Results of all children on the Sentence Repetition test are summarised in Table 18. It shows that children with SLI were consistently worse than their TD peers across different age groups. The mean score of all children with SLI was less than the mean score of the youngest typically developing group. Table 18 shows that children with SLI had a performance comparable to TD peers who were two years younger than they were. Figure 7 compares the performance of all groups of children on the Sentence Repetition test.

Table 18: Results of all participants on the Sentence Repetition (SR) test.

Age Groups	Typically Developing Children	SLI
Age Band 1: 4;6-5;11 years		
Number of participants (Male: Female)	24 (13:11)	5 (2:3)
Mean Raw Score of SR Test (SD)	69.8 (13.9)	40.8 (7.8)
Range of SR scores	49-94	29-49
Age Band 2: 6;0-6;11 years		
Number of participants	23 (15:8)	8 (7:1)
Mean Raw Score of SR Test (SD)	79.3 (10.6)	52.5 (17.6)
Range of SR scores	62-101	28-76
Age Band 3: 7;0-7;11 years		
Number of participants	21 (13:8)	5 (4:1)
Mean Raw Score of SR Test (SD)	84.3 (8.4)	64.6 (9.6)
Range of SR scores	68-99	50-76
Age Band 4: 8;0-9;4 years		
Number of participants	18 (12:6)	8 (5:3)
Mean Raw Score of SR Test (SD)	90.4 (9.9)	75.0 (12.4)
Range of SR scores	71-111	32-59
Total Number of children	86 (53:33)	26 (18:8)
Mean Raw Score of SR Test (SD)	80.5 (13.0)	59.5 (17.9)
Range of SR scores	49-111	29-59

Figure 7: comparison of the overall Sentence Repetition (SR) raw scores of children with SLI and typically developing (TD) children across different age groups.



T-test showed that the TD group was significantly better than the SLI group on the Sentence Repetition test t(33.3)=5.5, p<.001. One way ANOVA of the scores of the four age groups of TD children showed a significant group effect, F (3,82)=13.9, p<.001. Multiple comparisons with Bonferroni correction showed that the 5 year old group scored significantly lower the 6 year old group t(45)= -5.56, p=.02 and the other older groups. The 6 year old group had a significantly lower score than the 8 year old group t(39)=3.39, p=.01. However, there were no significant differences between the 6 and 7 year old groups on one hand and between the 7 and 8 year old groups on the other hand. Overall, these results are consistent with developmental trends, where groups of TD older children perform better than younger groups.

3.4.2.1 Distribution of Test Scores:

The following figure shows the distribution of the scores of all typically developing children on the Sentence Repetition (SR) test. It depicts a normal distribution of these scores, though with some positive skewness, as shown in *Figure 8*. Moreover, The Shapiro-Wilk test of normality was conducted and it was not significant for any of the groups, indicating the presence of normal distribution, as shown in Table 19.

Figure 8: Distribution of typically developing children on scores of the Sentence Repetition (SR) test, *n*=86.

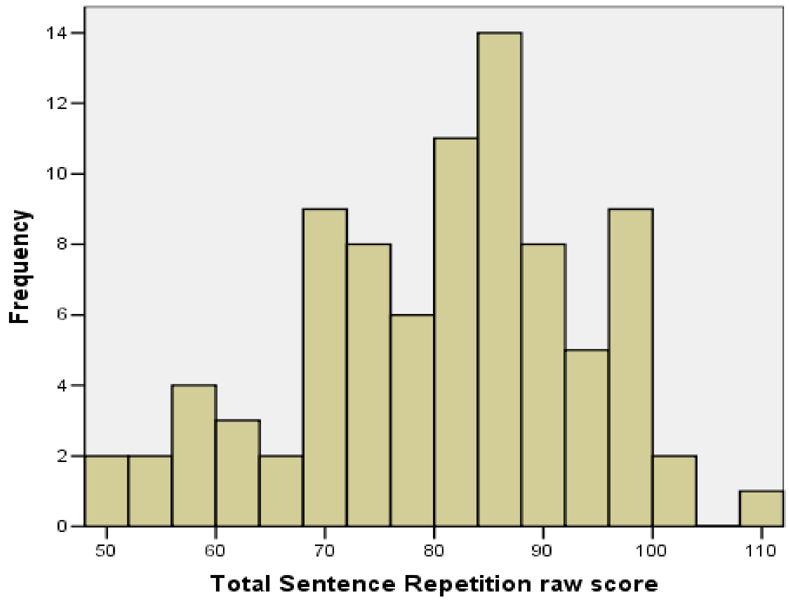


Table 19: Results of the Shapiro-Wilk test of normality for the four age groups.

Age Group	Statistic	Degrees of freedom	Significance
5 year olds (n=24)	.95	24	.31
6 year olds (n=23)	.97	23	.72
7 year olds (n=21)	.97	21	.72
8 year olds (n=18)	.97	18	.76

3.4.2.2 Reliability

Split-half analysis. The results of split half-analysis with Spearman-Brown coefficient for the Sentence Repetition (SR) test showed a correlation of .96 between odd and even items, indicating a significant level of internal consistency.

Cronbach's alpha. The Cronbach's α for the Sentence Repetition (SR) test was .89, indicating the presence of significant level of internal reliability. Examination of the total items correlation showed no one item could have increased alpha if deleted (see Table A-7 in Appendix A).

Test-re test reliability. Six children were re-tested on the Sentence Repetition test a week from the first time they took the test. This group of children consisted of 5 male students and 1 female student, aged 75 -107 months (6;3-8;11). Results of test-retest reliability showed a Pearson correlation coefficient of .97, p=.002, indicating that the test is stable over time. Individual data for each child are shown in the table below.

Table 20: Raw scores for the Sentence Repetition test and re-test.

ID	Sex	Age in months	Test	Re-test	Language Ability
625	Male	75	76.00	89.00	Typically Developing
626	Male	77	99.00	105.00	Typically Developing
628	Male	79	92.00	101.00	Typically Developing
624	Female	82	87.00	93.00	Typically Developing
11627	Male	83	75.00	89.00	SLI
823	Male	107	86.00	95.00	Typically Developing

3.4.2.3 Validity

Content validity. The linguistic structures included in this test were chosen because they appeared in the language samples taken from more than 30 Gulf Arabic speaking children and they include a variety of Gulf Arabic sentence types, such as statements, questions, negations, sentences with coordination, relative clauses, passive voice…etc.

Concurrent validity. In order to assess the concurrent validity of the Sentence Repetition (SR) test, it was compared to tests developed in this project to measure various aspects of language abilities, such as the Sentence Comprehension test, the Expressive Language test, and the Arabic Picture Vocabulary Test. Results of the Pearson Correlation revealed significant correlations (p<.001) between the SR test and all of these measures as illustrated in Table 21.

Table 21: Correlation between the Sentence Repetition (SR) test standard score and standard scores of other language tests.

	SC (*n*=114)	EL (*n*=112)	APVT (*n*=105)
The SR test (*n*=112)	.43**	.69**	.34**

**Correlation is significant at the 0.001. level (2-tailed).

Note. SC= Sentence Comprehension test, EL= Expressive Language test, APVT=the Arabic Picture Vocabulary Test.

The results above show that SR and EL tests were highly correlated with each other as they both tapped into the expressive abilities of all children.

3.4.2.4 Item Analysis

Item analysis was conducted to assess the overall Item difficulty index, i.e. proportion of examinees who answered each item correctly. This ranged between 0.0 and 1.0. Tables A-8 and A-9 in *Appendix A* show the item difficulties for all the SR items for the TD children and children with SLI. On examining these two tables, it is clear that the last 6 items (items 36-41) were too difficult for all children and therefore they might be omitted in future revisions of the test. Generally, most of the item difficulty rates were consistent with age groups and no item was identified where younger children were significantly outperforming older children. Due to the nature of the test, being a general language test, linguistic complexity was manipulated simultaneously with sentence length, which makes it difficult to pinpoint specific relations between linguistic complexity and proportion of correct responses, however, it might be useful to show how all children performed on a linguistic structure commonly cited as being very challenging in children with SLI, namely relative clauses (Adams, 1990; Friedmann & Novogrodsky, 2004; Stavrakaki, 2001). Relative clauses were represented in the test with nine exemplars, however item 41 was not included as it was among the most difficult items for all children. The proportions of correct responses for all children on relative clauses are shown in Table 22.

Table 22: Proportion of correct responses of the TD and SLI participants on relative clause items in the Sentence Repetition test.

Item No	The relative clause sentence	SLI *n*=26	TD *n*=68
17	Weɪ:n ɪʃ-ʃɪrtˤɪ ɪlli mɪsak ɪl-ħara:mɪ Where the-policeman who caught the-thief Where is the policeman who caught the thief?	.79	.98
24	Il-waləd ɪlli rfɪs-a lɪ-ħˤsa:n tˤa:ħ da:xɪl ɪl-ħufrə The-boy who kicked-him the horse fell into the ditch The boy whom the horse kicked fell into the ditch	.35	.77
27	ɪl-waləd ɪʃtarə kɪta:b ħag sˤadi:qə illi yħɪb ɪl-qɪsass ɪl-bu:li:ciyyə The-boy bought book for friend-his who likes stories detectives The boy bought a book for his friend who likes detective stories	.02	.24
28	ɪl- waled ma: kalləm ɪl-mudarrɪs ɪlli ysˤaħɪħ ʕala:ma:t isˤsˤaf ɪl-xa:mis The boy did not talk the-teacher who marks grades the class the fifth They boy did not talk to the teacher who marks the fifth class's grades	.07	.34
30	ɪl-bɪnt ɪlli taskɪn wara bei:tnə maʕa:y b-nafs ɪl-madrɪsə The-girl who lives behind house-our with me at the same school The girl who lives behind our house is with me at the same school	.29	.59
33	ɪtˤ-tˤɪla:b ktɪbau rɪsa:lə ħag sˤaħɪbhum ɪlli sa:fər bɪl-fasˤl ɪl-ʔawwəl The-students wrote letter for friend-their who travelled in the term the-first The students wrote a letter to their friend who travelled in the first term	.01	.2
34	ɪl-waləd ɪlli y-sou:g ɪs-sayyarə iz-zargə ħa:tˤ naðˤa:rə sou:də The-boy who 3MS-drives the-car the blue put glasses black The boy who drives the blue car wears black glasses.	.16	.59

All the relative clauses in *Table 22* are subject relative clauses (except item 24, which is an object relative clause), i.e., the relative operator occupies the subject position in the relative clause, and no attempt was made to systematically manipulate relative clauses. Therefore, these observations are based on these examples and may not necessarily apply to other types of relative clauses. As with other sentences in the Sentence Repetition test, grammatical complexity was confounded with sentence length due to the nature of the test. This analysis of item difficulty might provide some information about items that may pose more difficulty for children with SLI, however this information needs to be thoroughly and systematically investigated in future research.

3.5 Test 4: The Arabic Picture Vocabulary Test (APVT)

3.5.1 Method

3.5.1.1 Participants

The Arabic Picture Vocabulary Test (APVT) was administered to 107 Gulf Arabic speaking school children in Qatar, who ranged in ages between 4;6 and 9;4 years old. Eighty one of these children were typically developing (TD) and 26 were diagnosed with SLI based on the same criteria used throughout this project. All of these children participated in the previous tests and they all passed nonverbal IQ tests, articulation, and developmental verbal dyspraxia screenings. Children with SLI were diagnosed based on a Z-score of -2.0 or more on one test or -1.5 SD or more on at least two of the language tests developed for this project. Table 23 shows a summary of participants' characteristics.

Table 23: Characteristics of participants in the Arabic Picture Vocabulary Test (APVT).

Age Groups	Typically Developing Children	SLI
Age Band 1: 4;6 - 5;11 years		
Number of participants	22 (11:11)	5 (2:3)
Mean age in months (years)	64.1 (5;4)	62.6 (5;2)
Range in months (years)	54-71 (4;6-5;11)	58-70 (4;10-5;10
Age Band 2: 6;0 - 6;11 years		
Number of participants	22 (14:8)	8 (7:1)
Mean age in months (years)	77.7 (6;5)	78.9 (6;7)
Range in months (years)	72-83 (6;0-6;11)	73-83 (6;1-6;11)
Age Band 3: 7;0 - 7;11 years		
Number of participants	19 (10:9)	5 (4:1)
Mean age in months (years)	90.1 (7;6)	88.8 (7;5)
Range	84-95 (7;0-7;11)	85-94 (7;1-7;10)
Age Band 4: 8;0 - 9;4 years		
Number of participants	18 (12:6)	8 (5:3)
Mean age in months (years)	103.3 (8;7)	103.0 (8;7)
Range	96-112 (8;0-9;4)	99-107 (8;3-8;11)
Total No. of participants	81 (47:34)	26 (18:8)
Mean age in months (years)	82.6 (6;10)	85.1 (7;1)
Range in months (years)	54-112 (4;6-9;4)	58-107 (4;10-8;11)

3.5.1.2 Materials

The test consists of 132 items, which were arranged in terms of difficulty into 10 groups with 12 items per group. A booklet was made that has 134 pages (2 pages for practice items and 132 for test items). Each page depicted 4 pictures that were mostly taken from either the BPVT (Dunn et al., 1997) or from non-copyrighted material (e.g., free clip arts). However, all four pictures on one page were taken from the same source to ensure no single picture stood out pictorially. All answers were transferred to a record form (see Appendix S).

3.5.1.3 Procedure

All testing was done in a quiet room at school or home. Children were given the following instructions (in Arabic). “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’). 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. Because of the large number of stimuli, a ceiling criterion was employed in order to reduce fatigue, especially in younger children. The ceiling criterion used was a minimum number of eight errors in one group before stopping the test. Children were encouraged to continue if they seemed to like the test even when ceiling was established, however testing stopped at the ceiling item with many children. Fourteen TD children continued until the last item despite reaching a ceiling at a previous item. Only 3 children with SLI reached the last item, with two of them reaching a ceiling at a previous one; so only one child with SLI did not have a ceiling. All responses were recorded on a score sheet and children received (1) for a correct answer and (0) for incorrect answers. 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 60 and had total errors of 14 would have a raw score of 46.

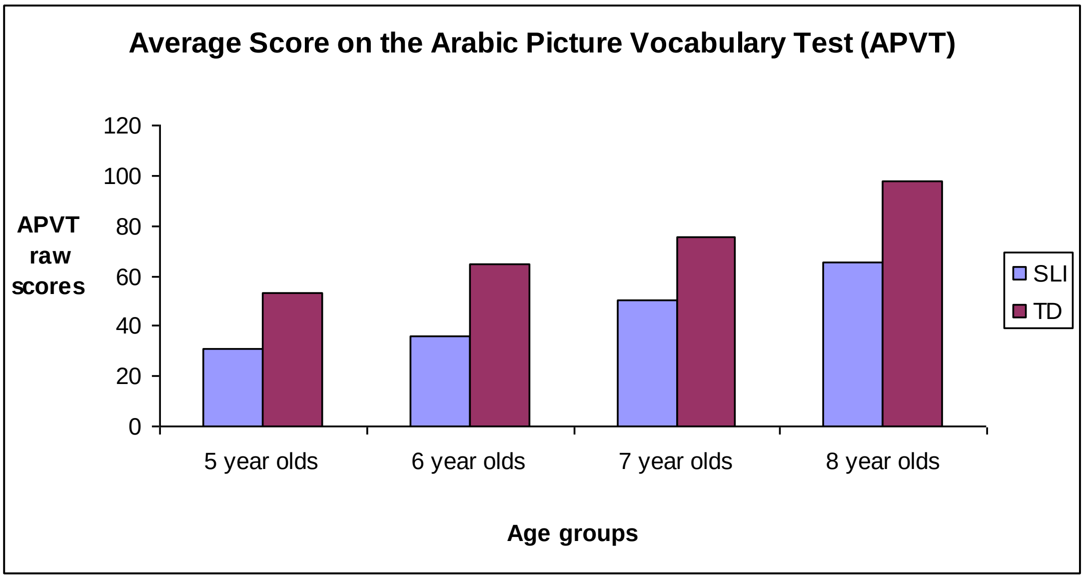
3.5.2 Results and discussion

Table 24 summarises the results of all children on the Arabic Picture Vocabulary Test (APVT). It shows that typically developing children (TD) scored significantly higher than children with SLI and across different age groups, indicating that children with SLI have limited receptive vocabulary compared to their TD peers. Table 24 and Figure 9 show that children with SLI have generally scores similar to those of TD peers who were 2 years younger than they were, a pattern that has been noticed in the Sentence Repetition test and the Sentence Comprehension test.

Table 24: Summary of results of all participants on the Arabic Picture Vocabulary Test (APVT)

Age Groups	Typically Developing Children	SLI
Age Band 1: 4;6-5;11 years		
Number of participants (Male: Female)	22 (11:11)	5 (2:3)
Mean Raw Score of APVT Test (SD)	52.9 (13.7)	31.2 (7.8)
Range of APVT scores	37-89	21-42
Age Band 2: 6;0-6;11 years		
Number of participants	22 (14:8)	8 (7:1)
Mean Raw Score of APVT Test (SD)	65.0 (16.7)	36.0 (15.2)
Range of APVT scores	43-92	13-54
Age Band 3: 7;0-7;11 years		
Number of participants	19 (11:8)	5 (4:1)
Mean Raw Score of APVT Test (SD)	75.4 (17.7)	50.6 (14.1)
Range of APVT scores	48-101	37-72
Age Band 4: 8;0-9;4 years		
Number of participants	18 (12:6)	8 (5:3)
Mean Raw Score of APVT Test (SD)	98.0 (7.5)	65.3 (13.5)
Range of APVT scores	86-110	43-83
Total Number of children	81 (47:34)	26 (18:8)
Mean Raw Score of APVT Test (SD)	71.5 (21.8)	46.9 (18.8)
Range of APVT scores	37-110	13-83

Figure 9: comparison of the scores of the Arabic Picture Vocabulary Test (APVT) by children with SLI and typically developing (TD) children across different age groups.

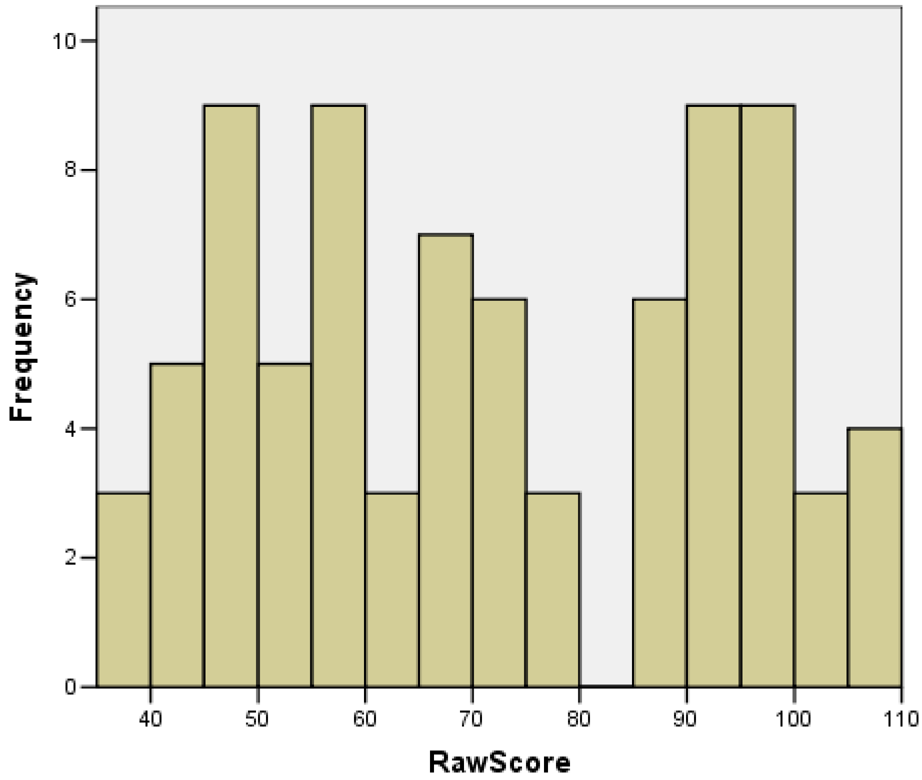


The TD group scored significantly better than the SLI group on the APVT, t(105)=5.2, p<.001. One-way ANOVA of the scores of the four age groups of TD children showed a significant group effect, F (3,77)=33.6, p<.001. Multiple comparisons with Bonferroni correction showed that the 5 year old group scored significantly lower than all three older groups. The difference between 5 and 6 year olds was significantly in favour of the latter t(42)=12.10, p=.045; the 7 year olds had higher scores than those obtained by the 5 year olds t(39)=22.56, p>.001 and so did the 8 year olds t(38)=45.13, p<.001. The difference between 6 and 7 year olds was not significant, but the difference between the 6 year olds and the 8 year old group was significant (38)=33.04, p=.001. Finally the 8 year olds were significantly better than the 7 year olds, t(35)=22.48, p<.001. Overall, there was a clear developmental pattern with each age group obtaining higher scores on the APVT than the age group preceding it, except for the difference between 6 and 7 year olds, which was not significant.

3.5.2.1 Distribution of Test Scores

Figure 10 shows the distribution of the scores of all typically developing children on the Arabic Picture Vocabulary Test. It depicts a broadly normal distribution of the test scores.

Figure 10 : Distribution of typically developing children on scores of the Arabic Picture Vocabulary Test (APVT), *n*=81.



The Shapiro-Wilk test of normality was conducted and it was significant for two groups (5 and 6 year olds). However, it was not significant for the 7 and 8 year old groups, as shown in the table below.

Table 25: Results of the Shapiro-Wilk test of normality for the four age groups.

Age Group	Statistic	Degrees of freedom	Significance
5 year olds (*n*=22)	.90	22	.040
6 year olds (*n*=22)	.90	22	.045
7 year olds (*n*=19)	.92	19	.134
8 year olds (*n*=18)	.92	18	.130

Figures 11 and 12 depict the distribution of the scores of the 5 and 6 year old groups on the APVT. They both showed some mild negative skewness.

Figure 11: Distribution of the scores of the TD 5 year olds on the APVT.

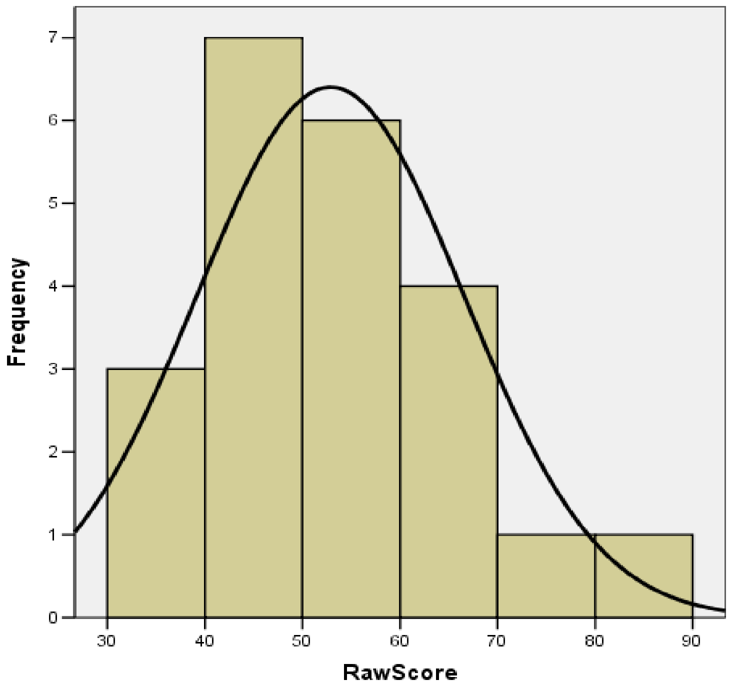
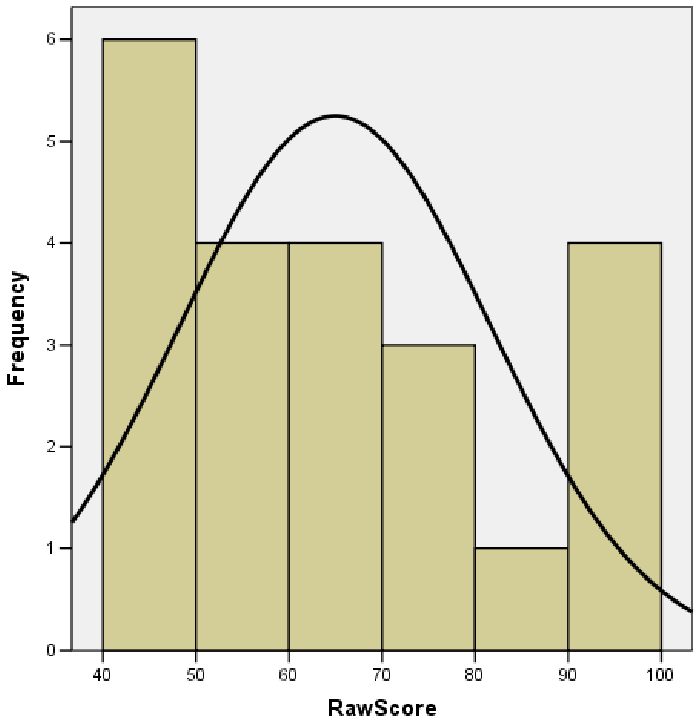


Figure 12: Distribution of the scores of the TD 6 year old children on APVT.



3.5.2.2 Reliability

Split-half analysis. The results of split half-analysis with Spearman-Brown coefficient showed a correlation of .71 between odd and even items, indicating a significant level of internal consistency.

Cronbach's alpha. The Cronbach's α for the Arabic Picture Vocabulary Test (APVT) was .82, indicating the presence of significant level of internal reliability.

Test-re test reliability. Six children were re-tested on the Arabic Picture Vocabulary Test (APVT) a week from the first time they took the test. This group of children consisted of five male students and one female student, aged 75 -107 months (6;3-8;11). Results of test-retest reliability showed a Pearson correlation coefficient of .97, p=.002, indicating that the test is stable over time. Individual data for each child are shown in Table 26.

Table 26: Raw scores for the Arabic Picture Vocabulary Test (APVT) test and re-test.

ID	Sex	Age in months	Test	Re-test	Language Ability
625	Male	75	52	58	Typically Developing
626	Male	77	93	96	Typically Developing
628	Male	79	63	66	Typically Developing
624	Female	82	92	95	Typically Developing
11627	Male	83	45	48	SLI
823	Male	107	91	99	Typically Developing

3.5.2.3 Validity

Content validity. While previous tests were based on structures from language samples, clinician's feedback, investigator knowledge of his native language, and some adaptations of English material, the Arabic Picture Vocabulary test (APVT) had to be developed de novo. The process started by asking 24 adult speakers of Qatari Gulf Arabic to rate 600 words in terms of familiarity. Each word received a rating from 1-5 (1= rarely heard or used, 5=very familiar and used very frequently). These words belonged to 20 different semantic categories (e.g., verbs, animals, occupations, adjectives….etc) following the same practice used in the development of the British Picture Vocabulary Test (BPVT) (Dunn, Dunn, Whetton & Burley, 1997). Out of these 600 words, 132 words were chosen and organised into 11 groups of 12 words per group ranked according to their difficulty, which was determined based on the familiarity rating of each item. The criteria for choosing these words were similar to those used in the BPVT. Hence, all the words included were functional, easy to depict pictorially and common in everyday life, except for the advanced vocabulary where some were taken from Classic and Modern Standard Arabic.

In order to examine the extent to which the familiarity ratings were able to predict the proportion of correct identification of each item, linear regression was performed. Results showed that adults' ratings of these words accounted for 29% of the variance in the performance of children on the test item. R^2=.29. ANOVA showed that the goodness of fit of this model was significant $F(1,130)=54.0$, $p<.001$. Regression analysis identified six items whose scores on the actual test items were poorly predicted by familiarity rating. These were items no: 28, 58, 75, 90, 92, 104, 111, and 128. There could be a number of explanations for discrepant performance on these items. For example, an adult may rate a vocabulary item as high in familiarity, whilst children may not use such word frequently. There are also other test factors, such as clarity of pictures, the role of foils (distracters) in terms of giving cues and the ability of these children to eliminate incorrect responses in order to reach a correct identification of vocabulary items. For example, most children did not correctly identify Item 28 (post stamp), which was frequently confused with the picture of a 'rubber stamp", because while adults use the word *tˤabiʔ* to describe a post stamp and *xatim* to describe a rubber stamp, children know the verb *ytˤabiʔ* (to rubber stamp) and therefore most of them pointed to the picture of a rubber stamp. Furthermore, many of them may not be familiar with post stamps as generally people in the Gulf do not use mail services in the same manner and frequency they are used in European countries. Therefore, any possible revisions of the test should consider such factors.

Concurrent validity. In order to assess the concurrent validity of the Arabic Picture Vocabulary Test, it was compared to tests developed in this project to measure various aspects of language abilities, such as the Sentence Comprehension test, the Expressive Language test, and the Sentence Repetition test. Results of the Pearson Correlation revealed significant correlations ($p<.001$) between the APVT and all of these measures as illustrated in Table 27 below.

Table 27: Correlation between the Arabic Picture Vocabulary Test (APVT) standard scores and standard scores of other language tests.

	SC	EL	SR
APVT(*n*=107)	.62**	.50**	.34**

**Correlation is significant at the 0.001. level (2-tailed).
Note. SC= the Sentence Comprehension test, EL= the Expressive Language test, SR= the Sentence Repetition test

3.5.2.4 Item Analysis

Item analysis was conducted to assess the overall Item difficulty and identify items that may need to be revised or omitted in future revisions of the tests. Due to the length of the test, which consists of 132 items, a ceiling was set, which was based on scoring eights errors in a group of 12 items. However, children were encouraged to continue the test if they did not show signs of fatigue or impatience. Most of the younger children did not complete all the test items. The following tables show the number of TD children and children with SLI who completed each group of items in the APVT test.

Table 28: Number of TD children who completed each group of items in the APVT, *n*=81.

Items	Number of children	%
1-12	81	100
13-24	81	100
25-36	81	100
37-48	81	100
49-60	81	100
61-72	72	88.9
73-84	60	74.1
85-96	54	66.7
97-108	49	60.5
109-120	40	49.4
121-132	39	48.1

Table 29: Number of children with SLI who completed each group of items in the APVT, *n*=26.

Items	Number of children	%
1-12	26	100.0
13-24	26	100.0
25-36	25	96.2
37-48	22	84.6
49-60	21	80.8
61-72	16	61.5
73-84	14	53.8
85-96	9	34.6
97-108	6	23.1
109-120	4	15.4
121-132	3	11.5

Only the older group of TD children (8;0-9;4 years old) did complete all the test items, therefore, the proportion of correct responses for the 132 test items were calculated for this group. The results are depicted in Table A-10 in Appendix A. Moreover, Tables A-11 and A-12 in Appendix A show the proportion of correct responses for the first 60 items for TD children and children with SLI respectively. These items were chosen because approximately 80% of children with SLI attempted identifying these vocabulary items and all TD attempted them. These tables show a general trend of increasing difficulty as the test proceeded from the first two groups onward, with all groups showing good scores of the first two groups (items 1-24) as they consist mostly of easier items. Starting from group 3 (items 25-36), the gap between TD children and children with SLI widens. This is a general trend with some exceptions as the vocabulary items were chosen and arranged in increasing difficulty order based on the ratings they received from adult native speakers.

The regression analysis showed six test items that may warrant revision or elimination in future revision of the test due to their poor correlation with the familiarity ratings. These were items: 28, 58, 75, 90, 92, 104, 111, and 128. Furthermore, a newer version may re-arrange the order of the items of the test based on the proportion of correct responses shown by the group of children aged 8;0-9;4.

This will be more appropriate than the original ordering, which was based on familiarity ratings of adult native speakers of Gulf Arabic.

3.6 Other tests

At the beginning of each testing session, all children received a non-verbal IQ test. The non-verbal IQ test conducted with children six years and older was the Test of Non-verbal Intelligence (TONI-3) (Brown et al., 1997). TONI-3 has very good psychometric properties and was found to be culture-friendly during piloting stage. The Standard Scores obtained from Gulf Arabic speaking children were not far off those reported in the normative data of US children (see Table 30), despite the much smaller sample size in the case of Qatari children. Therefore, the determination of normal performance IQ was based on US norms, despite the fact that the scores of TD Gulf Arabic speaking children were negatively skewed compared to US norms, especially for children aged seven and above. This performance IQ test may prove a very useful and culture-free way to assess the nonverbal IQ of Gulf Arabic speaking children; however, local norms need to be collected from a larger sample.

Table 30: Mean and standard deviations of the scores of typically developing children on the Test of Nonverbal Intelligence-3 (TONI-3).

Age Groups	TONI-3 Standard Scores
5;8-6;11 years (Age_M=76.4 months) (*N*=27)	
Mean Standard score (SD)	99.7 (10.1)
Range	85-119
7;0-7;11 years (Age_M=90.3 months) (*N*=22)	
Mean Standard score (SD)	94.7 (9.4)
Range	81-123
8;0-9;4 years (Age_M=103.0 months) (*N*=19)	
Mean Standard score (SD)	94.6 (9.6)
Range	83-118
Total Number of children	68
Mean Standard score (SD)	96.7 (9.9)
Range	81-123

As part of the battery of tests, all children were required to participate in an articulation screening test (see Appendix J) and a screening for oral-motor functioning and developmental verbal dyspraxia (see Appendix K). In the articulation test, children were asked to repeat 30 words that started with as many Gulf Arabic consonants. These 30 words were carefully selected to include most of the late developing sounds in medial and final word positions. The screening test for developmental apraxia of speech contained some oral motor tasks that were adapted from the Kauffman Speech-Praxis Test (Kauffman, 1995). All typically developing children and those with SLI passed these two screening tests.

3.7 General discussion

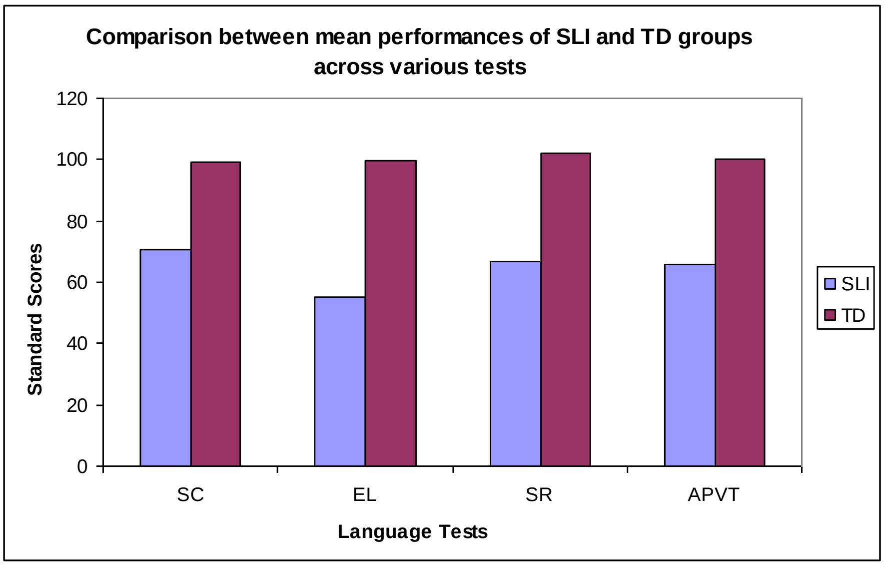
The results of the battery of language tests used with Gulf Arabic speaking children are encouraging. Analysis of the performance of typically developing children shows normal distribution on most of these tests. Moreover, all tests showed good levels of reliability and validity, as shown by their good levels of internal consistency and content and concurrent validity. These results validate the use of these tests to identify children with specific language impairment for this project, though these tests require some revisions to reorganise, omit, or modify some items as discussed in the item analysis section of each test.

While previous sections analysed the performance of TD and SLI children on individual language tests by examining their raw scores, it is important to examine and compare the standard scores of those children on the language tests developed for this project. Table 31 and Figure 13 compare the standard scores of TD and SLI children on the four language tests: the Sentence Comprehension (SC) test, the Expressive Language (EL) test, the Sentence Repetition (SR) test, and the Arabic Picture Vocabulary test (APVT).

Table 31: Descriptive statistics of the performance of typically developing children and those with SLI on various language tests.

Tests	Typically Developing Children (*n*=86)	SLI (*n*=26)
1- The Sentence Comprehension test		
Mean Standard Score (SD)	99.4 (14.5)	70.5 (26.2)
Range	74-132	0-106
2- The Expressive Language test		
Mean Standard Score (SD)	99.7 (14.3)	55.3 (24.8)
Range	76-133	6-107
3- The Sentence Repetition test		
Mean Standard Score (SD)	102 (16.8)	66.9 (20.3)
Range	69-177	30-111
4- The Arabic Picture Vocabulary Test		
Mean Standard Score (SD)	100.2 (12)	65.7 (25.6)
Range	77-128	33-102

Figure 13: Means of Standard scores of TD and SLI children on various language tests.



Note. SC=the Sentence Comprehension test; EL=the Expressive Language test; SR= the Sentence Repetition test; APVT=the Arabic Picture Vocabulary test.

The results of comparisons of the language tests reveal that Gulf Arabic speaking children with SLI have relative weakness in their expressive language test, as shown by their low scores on the Expressive Language test. This has been consistently reported in studies of SLI studies in other languages (Bishop, 1997; Leonard, 1998), where the expressive language abilities of children with SLI typically lag behind other language abilities, especially their receptive language skills. These children with SLI's performance on the receptive vocabulary test might seem incongruent with many studies that found better performance on vocabulary than on sentence repetition, but there is evidence that some children with SLI were reported to have comparable results on these two tests in English (e.g., Leonard, 1998). A possible explanation for the relatively poor performance on the receptive vocabulary test might be attributed to the root-and-pattern nature of Arabic. In Arabic, a semantic root undergoes many morphological and phonological transformations to generate various vocabulary items derived from that root. For example, the root 'd-r-s' (study) is used to derive finite and infinite verbs 'daras' (he studied) and 'yadris' (he studies), the nouns 'madrasa' (school), 'dira:sa' (studies), mudarris' (teacher)…etc. Therefore, children with SLI, who have been shown to have syntactic, morphological, morphosyntactic, and phonological deficits, might be less competent at using these roots to derive more vocabulary items out of them. However, it is difficult to have conclusive remarks based on these tests only, as they need further revision and should be used in projects with larger number of participants.

It is commonly acknowledged that the SLI population is very heterogeneous as these children come with varying individual linguistic abilities (Bishop, 1997; Bishop, 2004; Conti-Ramsden & Botting, 1999; Leonard, 1998). Therefore, Tables 32 and and 33 display the individual performance of all TD and children with SLI, respectively, on the four language tests used in this project. These two tables depict the number of TD children who showed some below average performance on some tests, without meeting the SLI criteria. They also show the number of children with SLI who scored significantly below the mean of TD children on more than two tests. It is noteworthy that none of the TD children scored more than -1.5 standard deviation below the mean on the APVT test. Moreover, only three children with SLI had been included because they scored -2.0 SD on one test only, while the rest failed at least two tests.

Table 32: Standard Scores of TD participants on various language tests

ID	SC	EL	SR	APVT	ID	SC	EL	SR	APVT
401	114	117	116	146	628	92	109	110	127
403	106	96	104	91	629	83	100	88	94
404	90	101	95	104	632	79	91	114	126
405	82	92	104	85	702	121	117	111	100
501	102	92	87	missing	704	99	81	91	117
508	110	96	112	108	705	112	119	102	
503	106	94	94	99	706	82	117	120	84
504	94	94	112	94	707	112	92	71	118
506	94	92	85	113	708	78	103	84	92
507	102	120	111	116	709	108	106	98	96
509	110	115	121	100	710	78	92	115	92
511	125	122	101	100	713	116	125	117	117
502	86	80	79	missing	714	104	87	113	83
505	98	78	86	100	718	112	92	115	116
514	94	96	89	116	719	95	122	117	109
515	82	80	85	82	720	74	81	116	99
516	86	96	88	85	721	104	95	114	96
518	118	117	120	132	722	99	76	97	81
519	122	120	127	99	727	121	98	93	124
520	86	101	99	83	725	95	111	126	84
521	122	120	106	99	715	78	81	91	79
522	75	87	79	84	717	99	106	104	97
523	82	106	112	119	726	112	98	95	118
524	106	96	111	93	711	112	missing	113	Missing
602	83	102	78	98	729	92	98	missing	Missing
603	118	122	118	Missing	801	92	missing	78	Missing
604	92	106	82	104	901	92	114	92	97
605	106	87	106	111	902	115	88	93	100
606	79	100	97	110	803	85	99	108	94
608	101	80	94	87	804	85	84	71	109
609	97	124	96	126	806	92	77	missing	missing
611	114	91	101	85	807	85	84	96	88
612	106	106	111	114	811	107	107	108	110
613	97	83	77	84	812	100	96	77	94
614	114	104	89	99	813	115	103	104	99
616	110	115	130	95	814	85	96	111	97
617	79	96	86	81	815	129	133	131	110
621	74	85	108	85	816	77	80	96	88
620	92	72	93	103	903	92	107	116	98
622	110	96	89	90	819	107	107	110	109
623	110	122	121	121	820	107	234	180	100
624	132	111	111	128	822	100	99	95	86
625	92	87	94	83	823	107	126	113	93
626	92	113	118	95	826	129	99	101	99

Standard Score of 79 and above
70-78
Less than 70

SC=Sentence Comprehension test; EL=Expressive Language test; SR=Sentence Repetition test,

APVT=Arabic Picture Vocabulary Test.

Table 33: Standard Scores of children with SLI on various language tests.

ID	SC	EL	SR	APVT
11510	71	20	42	88
11512	59	27	49	79
11515	98	21	29	72
11402	55	26	46	62
11513	86	24	38	72
11601	48	26	41	92
11607	74	46	56	76
11618	83	18	46	67
11619	57	13	30	52
11610	101	61	85	89
11631	106	85	67	63
11615	83	83	88	67
11627	66	76	96	90
11701	91	46	67	97
11703	44	35	36	70
11712	61	62	69	77
11723	65	62	84	67
11716	87	54	58	84
11805	100	47	48	50
11810	18	51	81	43
11809	55	92	81	66
11802	100	55	75	73
11818	62	81	74	78
11821	100	107	111	62
11824	77	25	59	83
11825	92	6	84	67

Standard Score of 79 and above
70-78
Less than 70

SC=Sentence Comprehension test; EL=Expressive Language test; SR=Sentence Repetition test, APVT=Arabic Picture Vocabulary Test.

3.8 Conclusion and summary

Gulf Arabic speaking children with SLI were identified based on a battery of language tests that were developed specifically for this project, due to lack of formal and informal language assessment tools or language development norms. This chapter described the tests used to assess the language abilities of Gulf Arabic speaking children. These tests included a Sentence Comprehension test, an Expressive Language test, a Sentence Repetition test and an Arabic Picture Vocabulary Test. Since all of these tests were new, measures of validity and reliability were examined to validate the use of these tests. The results show good levels of reliability and validity; therefore, supporting the use of these tests in the identification of children with SLI for this project. Moreover, the pattern of performance of typically and atypically developing Gulf Arabic speaking children on various language tests is consistent with findings reported in other languages. However, while children with SLI acquiring European languages tend to have relative strength in receptive vocabulary, Arabic speaking children with SLI showed poor performance on the receptive vocabulary test. This is probably due to the root-and-pattern nature of the language. Overall, Gulf-Arabic speaking children with SLI showed variable abilities on the four language tests used in the project, hence confirming the heterogeneous characteristics of SLI seen in other languages. However, since this is the first attempt at developing such tests, all these assessment tools warrant further revisions and should be administered with a larger number of participants. Moreover, these general tests of language abilities tap into general and broad groups of receptive and expressive language skills. In the following chapters, experimental tasks are employed to examine how children with SLI perform on some more specific linguistic tasks.

4. Comprehension of Complex Sentences: Comprehension of Sentences with Fronted Noun Phrases (NPs) in Gulf Arabic Speaking Children with SLI

4.1 Introduction

This chapter investigates the comprehension of simple SVO sentences and complex sentences that involve fronting of the direct object noun phrase (NP) in Gulf Arabic speaking children with SLI, whose performance is compared to both age and language-matched groups. Children with SLI have been shown to have deficits in understanding some complex syntactic structures such as wh-questions (Deevy & Leonard, 2004; van der Lely, 1998; van der Lely & Battell, 2003), relative clauses (Adams, 1990; Friedmann & Novogrodsky, 2004; Stavrakaki, 2001), pronominal references (Bishop, Bright, James & van der Lely, 2000; Montgomery & Evans, 2009; van der Lely & Stollwerck, 1997) and passives (Bishop, 1979; 1997; Montgomery & Evans, 2009; van der Lely, 1996; van der Lely & Harris, 1990). These studies of sentence comprehension skills in children with SLI were conducted in the frameworks of domain-general (processing) or domain-specific (linguistic) accounts of SLI. Domain-general accounts of SLI explain that sentence comprehension deficits occur because of limitations in working memory, processing speed, or general cognitive load, while domain-specific theories argue for the presence of language-specific deficits in children with SLI. Both domain-general and domain-specific accounts acknowledge that children with SLI and typically developing children have access to various cues in sentence comprehension, such as animacy cues, lexical cues, world knowledge, and subject-first strategy. Therefore any systematic investigation of sentence comprehension in children with SLI should control these cues or manipulate them systematically.

4.1.1 Comprehension of complex sentences: processing-based perspectives

Processing accounts of SLI attribute difficulties in sentence comprehension to limitations in some processing capacities, such as working memory (Montgomery, 1995b, 2000a), speed of processing (Kail, 1994), or general cognitive load (Deevy & Leonard 2004). According to processing accounts, children with SLI cannot recall all of the input they hear or they have delays in processing these sentences (Montgomery; 2002).

Montgomery (1995b, 2000a) argued that limitations in verbal working memory could explain most of the deficits seen in sentence comprehension skills in children with SLI. There are two main models of verbal working memory; the Baddeley (1986) model of working memory, where WM is defined in terms of the amount of phonological material stored in the phonological short term memory (PSTM), and Just and Carpenter's (1992) model, which defines WM in terms of its ability to simultaneously store and process phonological information. They called the latter type of working memory the functional working memory (FWM). Functional working memory is measured with tasks such as the Competing Language Processing Task (CLPT), where subjects are required to listen to sentences while trying to recall the final word in each sentence. Montgomery showed that the sentence comprehension skills of children with SLI correlated with their scores on PSTM and FWM skills (Montgomery, 1995b; Montgomery, 2000b). While TD children showed similar performance on longer (redundant) and shorter (nonredundant) sentences, children with SLI had more difficulties understanding longer sentences than shorter ones matched for syntactic complexity (Montgomery, 1995b). Therefore, Montgomery (1995b) concluded that part of the difficulties seen in sentence comprehension in SLI could be associated with limited phonological short-term memory. When Montgomery (2000b) examined the relationship between sentence comprehension and FWM using CLPT tasks, results showed that children with SLI had lower scores on CLPT tasks in comparison to age-matched groups, but their performance was not significantly different from younger typically developing children (Montgomery, 2000a; 2000b).

Montgomery and Evans (2009) examined the relationship between the comprehension of simple and complex sentences and two WM mechanisms, the phonological short term memory (PSTM) and the attentional resources capacity/allocation in 24 children with SLI, 18 age controls (AC) and 16 language and memory controls (LM). These children completed a NWR task (as a measure of PSTM) and a CLPT task (as a measure of attention capacity/allocation) and a simple and complex sentence comprehension task. The complex sentences consisted of reversible passive sentences (e.g., 'the boy was kissed by the girl') and sentences containing a pronoun or an anaphor (e.g., 'Daffy Duck says Bugs Bunny is tickling himself') that were modelled after the sentences used by van der Lely and colleagues (Bishop et al., 2000; van der Lely, 1996; 1998; van der Lely & Stollwerck, 1997). Simple sentences, on the other hand, did not include any nonlocal dependency and conformed to the canonical word order (e.g., 'the clown is hugging the tiny white elephant'). Results showed comparable performance by all groups on simple sentence comprehension, but on the complex sentence comprehension task, the SLI and LM groups performed worse than the AC group. Results showed that for the SLI group, there was a correlation between the NWR task and simple sentence comprehension, while the CLPT task correlated with complex sentence comprehension. In the AC group none of these two tasks correlated with simple or complex sentence comprehension, while for the LM group only the CLPT correlated with complex sentences. Therefore, Montgomery & Evans (2009) concluded that WM limitations, especially in the attention capacity/allocation played an important role in complex sentence comprehension in children with SLI, while PSTM capacity did not. The correlation found between PSTM and simple sentence comprehension in children with SLI, but not in TD children, suggests according to the authors that even this simple task requires great cognitive effort in this population. Evans and MacWhinney (1999) investigated the comprehension strategies used by 14 children with SLI, who were divided into a group of children with severe expressive and receptive impairments (ER-SLI) and a group with expressive impairments only (E-SLI). Children were presented with three types of word orders: Noun Verb Noun (NVN), NNV, and VNN and they were asked to determine the agent in each sentence. Evans and MacWhinney (1999) found that those with ER-SLI performed significantly less well than children with E-SLI. They reported that these atypically developing children employed different comprehension strategies. Children with ER-SLI were not able to use word order cues and relied on animacy cues, while children with E-SLI relied exclusively on first-noun-phrase-as subject strategy regardless of the type of word order they encountered. They also reported that performance on word order cues was correlated with these children's receptive language skills. Evans and MacWhinney (1999) claimed that these difficulties in using word order cues in children with ER-SLI were possibly caused by limitations in their working memory, as they had difficulties maintaining the sequential order of words while trying to use word order cues during comprehension. Overall, these limitations in working memory are used by some proponents of domain-general accounts of SLI to explain the difficulties these children have in sentence comprehension.

The generalised slowing processing theory (Kail, 1994) argues that linguistic deficits in children with SLI are caused by their slowed processing abilities. Many studies have shown that children with SLI have performed slower than their typically developing peers on a variety of linguistic and non-linguistic tasks (Edwards & Lahey, 1996; Johnston & Ellis Weismer, 1983; Lahey & Edwards, 1996). Kail (1994) reported that children with SLI had slower response time (RT) on a variety of linguistic and non-linguistic tasks; hence, he argued that these differences in processing speed between typically developing children and children with SLI reflect a general delay in the cognitive processing abilities of children with SLI. Kail (1994) therefore, predicts that there is a linear relationship between the RT of children with SLI and their TD counterparts, which reflects a nonspecific slower response. This slower processing of input could influence the sentence comprehension of children with SLI, whose slower processing means they will be less efficient in extracting necessary information from the input, leading to losing part of the information that follows. Lahey and colleagues (2001) argued that such an account would predict a linear relationship between processing speed and severity of language impairment, as the slower the processing speed is, the higher the impact it would have on language abilities of children with SLI. However, Lahey and colleagues did not find evidence supporting this linear relationship when they examined the correlation between RT of various tasks with the scores of some standardised language tests in 66 children with SLI (Lahey et al., 2001).

Deevy and Leonard (2004) argue that limitations in both working memory and processing capacity might explain deficits in complex sentence comprehension in children with SLI. Deevy and Leonard (2004) examined the comprehension of subject and object wh-questions in 16 children with SLI (mean age= 5;1) and 28 age matched and language matched typically developing children. They compared the predictions of a domain-specific account of SLI (van der Lely and Battel, 2003) with a processing account that argues that children with SLI have reduced capacity to process object wh-questions because in these questions, the distance between the wh-word in subject questions and its trace is much shorter than those in object questions, thus the latter consume larger processing capacity and working memory. The difference between wh-subject and wh-object question, according to them, is illustrated in the following examples from Deevy and Leonard (2004):

4-1

a. Wh-subject question: Who [x is washing the dog]?

b. Wh-object question: Who is [the dog washing x]?

Deevy and Leonard (2004) argue that the wh-phrase in these questions could not be interpreted with the main verb or its arguments until the gap has been identified. This long dependency creates more burden on the already vulnerable working memory in these children.

In order to test these two hypotheses, Deevy and Leonard (2004) showed participating children pictures of three animal characters engaging in reversible actions. Deevy and Leonard (2004) manipulated sentences by adding extra information (adjectives) that increased processing load without affecting syntactic complexity. Therefore, children were asked short and long questions either about the subject or the object of the sentence, as in ‘Who is washing the dog?’ (short subject question)’ and ‘who is the happy brown dog washing?’ (long object questions). In order to control for effects of syntax, the SLI group included only those who showed above chance performance on short subject and object questions (Deevy & Leonard, 2004). Linguistic accounts of SLI would predict syntactic complexity effects will influence performance, but increasing sentence length will have a lesser impact. Deevy and Leonard (2004) found that children with SLI performed significantly worse on long object questions because of the increase in the distance between the moved wh-word and the position it moved from, attributing this poor performance to limited processing capacities. According to Deevy and Leonard (2004), children with SLI cannot process wh-questions efficiently due to the distance between the moved wh-word and the position it moved from, this relationship has to be held and kept in the short-term memory of children with SLI, which constitutes a difficult task for these children.

All these domain-general accounts of SLI agree that there are no language-specific deficits in children with SLI; instead the linguistic profiles of children with SLI are manifestations of more general, cognitive deficits. However, proponents of domain-general hypotheses of SLI differ on which cognitive processes underlie the linguistic deficits of children with SLI. Some argue that children with SLI process linguistic and non-linguistic information slower than normal, while others argue that these children have limited capacity working memory. Other researchers posit that slower processing and limitations in working memory can explain the linguistic deficits in children with SLI. These domain-general accounts have not been able to identify the relationship between linguistic and non-linguistic tasks in children with SLI and cannot explain the presence of grammatical deficits that have distinct genetic basis that is independent of processing factors (see chapter 3 for a critique of the different domain-general accounts).

4.1.2 Comprehension of complex sentences: domain-specific (linguistic) accounts of SLI

Most of the studies that investigated the abilities of children with SLI to comprehend sentences with complex syntactic structures were carried out by van der Lely and colleagues (Marinis & van der Lely, 2007; Marshall & van der Lely, 2006; van der Lely, 1996; 1998; 2005; van der Lely & Harris, 1990; van der Lely & Stollwerck, 1997). van der Lely and colleagues argue that deficits in understanding syntactic structures are due to limitations in the ability of the grammatical system to compute structural dependencies between constituents, a theory known as Computational Grammatical Complexity (CGC) (Marshall & van der Lely, 2006; van der Lely, 2005), which is a descendant of a previous account of SLI known as the representation deficit for dependent relationships (RDDR) (van der Lely, 1998). van der Lely and Harris (1990) used an acting- out task and picture pointing to examine reversible active and passive sentences in children with G-SLI aged between 4;10 and 7;10. There were 14 participants in the acting-out experiment and 16 in the picture-pointing task. These children's performance was compared to age and language control groups. In active sentences, thematic roles (e.g., agent, theme, goals…etc.) corresponded to the canonical (typical) word order, e.g., '(the girl carries the baby). However, passive sentences contained a noncanonical word order where the agent was assigned to the object position and the theme to the subject position (e.g., the baby is carried by the girl). Children with SLI performed significantly worse than both control groups only on sentences with noncanonical word order. Therefore, van der Lely and Harris (1990) argue that children with SLI have great difficulty assigning the roles of agent or theme based on syntactic structures alone.

This difficulty with establishing dependencies among syntactic constituents in children with G-SLI was replicated in an on-line experiment using the cross-modal picture-priming paradigm (Marinis & van der Lely, 2007). In this study, Marinis and van der Lely (2007) looked at the comprehension of wh-questions in 14 G-SLI children and compared their performance to 14 age control and 17 language control children. The aim was to test if their performance on comprehension of filler-gap dependency in wh-questions was better explained by a generalised slowing account (Kail, 1994) or the Computational Grammatical Complexity account (CGC) (van der Lely, 2005). In the cross-modal picture priming method, children listened to experimental and filler sentences while looking at a picture and deciding if it showed a living or a non-living object (Marinis & van der Lely, 2007). Both the domain-general and domain-specific accounts predicted elevated reaction times (RT), though they had different underlying reasons. The generalised slowing theory (Kail, 1994) perceived late RT as a limitation of processing capacity, while CGC assumed that filler-gap dependency was the cause of the delayed response. Children listened to sentences like those in 4-2:

4-2

Baloo gives a long carrot to the rabbit$_i$.

Who$_i$ did Baloo give the long carrot to t$_i$ at the farm? (Marinis & van der Lely, 2007)

The performance of children with G-SLI was compared to age control (AC) and language control (LC) children in terms of accuracy, reaction time (speed of processing), and qualitative differences. Results showed that the comprehension skills of G-SLI were significantly inferior to that of the two control groups. Their speed of processing was significantly lower than the AC group, but not the LC group that matched the SLI group on memory span. However, the sentence comprehension skills in G-SLI children were qualitatively different from the two typical groups.

While both the AC and LC controls showed priming effects on the position of the trace, the G-SLI children did not show any priming on the trace position, as predicted by the CGC. Instead, a priming effect was seen on the verb, which indicated that the G-SLI children used lexical/thematic cues instead of syntactic cues (Marinis & van der Lely, 2007). These findings in G-SLI were reported in other SLI children. van der Lely and colleagues looked at another example of complex structures that involved structural dependency, namely pronominal references, such as pronouns (e.g. "*him*") and anaphors (reflexives, such as "*herself*") (Bishop et al., 2000; van der Lely & Stollwerck, 1997). van der Lely and Stollwerck (1997) examined Binding Principles A and B of the Government and Binding theory (Chomsky, 1986). Principle A states that an anaphor (a reflexive pronoun) must be bound in its governing category, while Principle B states that a pronoun must be free in its governing category (Chomsky, 1986). In 4-3 *himself* can refer to Peter Pan, which is in the same local binding domain (governing category), but it cannot refer to Captain Hook. An example of Principle B is shown in 4-4 where "him" cannot be bound by Peter Pan.

4-3

Captain Hook $_{i}$ says [Peter Pan $_{j}$ can hurt himself $_{i*/j}$] (van der Lely & Stollwerck, 1997)

4-4

Captain Hook $_{i}$ says [Peter Pan $_{j}$ can hurt him $_{i/*j}$] (van der Lely & Stollwerck, 1997)

van der Lely and Stollwerck (1997) examined how a group of G-SLI children performed on comprehension of binding principles A and B in comparison to three control groups matched with the experimental group on different language abilities. The experimental task involved a yes-no sentence-picture judgement task where children listened to sentences like: "This is Mowgli. This is Baloo Bear. Is Mowgli tickling himself?" Results of the study show that unlike typically developing children, children with G-SLI incorrectly accepted coreference between an anaphor and a non-local antecedent. This illustrates that while typically developing children use both lexical-semantic knowledge and syntactic knowledge to assign reference to anaphors and pronouns, children with G-SLI rely mainly on their lexical-semantic knowledge.

These findings by van der Lely and colleagues have been replicated by studies in other languages, such as Hebrew, which is a Semitic language, like Arabic. Friedmann and colleagues (Friedmann, Gvion & Novogrodsky, 2006; Friedmann & Novogrodsky, 2004; 2007) studied grammatical impairments in a group of Hebrew speaking children with syntactic-SLI (S-SLI), whose linguistic deficits are similar to those seen in G-SLI children. Friedmann and colleagues found that these children's difficulties in understanding complex structures were caused by deficits in syntactic movement (Friedmann et al., 2006; Friedmann & Novogrodsky, 2004; 2007). Friedmann and Novogrodsky (2007) examined linguistic deficits in a group of 15 Hebrew speaking children with syntactic-SLI (S-SLI), aged between 9;3 to 14;6 years and compared their performance to 50 typically developing children. Children were presented with subject relative clauses and object relative clauses, such as the ones below:

4-5

a. Subject relative:

Zot ha-safta she-menasheket et ha-yalda

This the-grandmother that-kisses ACC the-girl

(This is the girl$_i$ that t_i is kissing the grandmother)

b. Object relative:

Zot ha-safta she-ha-yalda menasheket.

This the-grandmother that the girl kisses

(This is the girl$_i$ that the grandmother is kissing t_i.)

(Friedmann & Novogrodsky, 2004)

Comprehension of these types of sentences requires activation of the trace position of the moved element and establishing a relationship (a chain) between the trace and the new position of the moved element in order to assign the appropriate thematic role for the moved element. In the case of object relatives, it involves both movement and non-canonical word order, while subject relatives are less challenging due to their canonical word order (subject first strategy correctly assigns thematic role to the agent). Results showed that children with SLI presented with significant deficits in understanding of object relative clauses compared to age and language controls, although they showed good performance on subject relative clauses when they were compared to their language controls. As for the nature of these deficits in comprehension of object relative clauses, Friedmann and colleagues showed that these difficulties were due to deficits in assignment of thematic roles to moved elements (Friedmann et al., 2006; Friedmann & Novogrodsky, 2007) and not in the creation or activation of the trace.

This chapter will shed light on how domain-general and domain-specific accounts of SLI will explain the performance of Gulf Arabic speaking children with SLI on simple and complex sentence comprehension, a task that has not been undertaken yet in this language. Therefore, a definition of what constitutes a grammatically complex structure is warranted.

4.1.3 Defining grammatical complexity

Most researchers who investigated complex grammatical structures have not provided a definition of grammatical complexity that is restrictive and specific; instead they usually give examples of types of complex structures (van der Lely, 2005, Montgomery & Evans, 2009). These include reversible passive sentences, pronominal sentences, wh-questions, and object relative clauses. What is common among these structures is the presence of grammatical movement and anaphoric binding. Other linguistic accounts of SLI propose that the core linguistic deficits in SLI are in tense (Rice & Wexler, 1996a) or subject-verb agreement (Clahsen, 1989).

The definition of grammatical complexity adopted in this thesis is mainly based on van der Lely (1998, 2005), which provides one of the few attempts to define grammatical complexity. According to the Computational Grammatical Complexity (CGC) account of SLI (Marshall & van der Lely, 2006; van der Lely, 1998; 2005), a grammatically complex structure is one that requires the computation of nonlocal grammatical dependency between different constituents. According to this definition, the following structures are examples of grammatical complexity in syntax:

- Reversible 'be' passive sentences (e.g., 'the baby was kissed by the grandfather').
- Long distance fronting, such as in object relative clauses (e.g., 'this is the baby that the grandfather kissed') and object wh-questions (e.g., 'who is

the boy tickling?') Sentences with subject movement (e.g., 'this is the girl that is hugging the teacher') may not cause the same level of complexity as in these sentences the agent still precedes the theme and therefore children may comprehend them based on their canonical order, even if syntactic movement is impaired (Friedmann & Novogrodsky, 2004).

- Pronominal sentences with anaphoric binding (e.g.," Mowgli says Baloo Bear is tickling himself" (van der Lely and Stollwerck, 1997).
- Topicalised sentences that involve fronting of an object (e.g., 'the baby, the grandfather kissed').

This definition of grammatical complexity is contrasted with other definitions that characterise complexity in SLI as any utterance that consists of more than one clause, through coordination or subordination (Schuele & Dykes, 2005). The notion of grammatical complexity can be defined in relation to unmarked or canonical structures. For example, a word order is canonical when it emerges in early stages of language acquisition, used in neutral contexts, and requires fewer derivations (Fassi-Fehri, 1993). Therefore, most noncanonical structures are considered grammatically complex. Likewise, in phonology a CV syllable is considered simple and canonical, while a CVCC is considered complex and noncanonical.

This chapter uses the term grammatical complexity to refer to word order types that involve movement and computations of grammatical dependency between syntactic constituents. It does so by evaluating how typically developing children and children with SLI perform on a comprehension task that involves a canonical SVO word order and two marked word orders that involve fronting of direct object NP's.

4.1.4 Word order in Gulf Arabic

Arabic is a classical example of a diglossic language, where there are two varieties of the language, one that is literary and spoken in formal situations, while the other is the spoken variety that that child gets exposed to since birth (Ferguson, 1959). The formal variety is called Modern Standard Arabic (MSA) and is typically taught when children enter school, while there are regional spoken varieties, such as

Gulf-Arabic. One of the main differences between varieties of Arabic spoken in different Arab countries and MSA is the lack of case marking in spoken varieties and the status of Subject Verb Object (SVO) structure as the unmarked word order, while it is VSO in MSA. The issue of word order in MSA and Classic Arabic has received attention by researchers because of the apparent asymmetry in subject-verb agreement in SVO and VSO. In SVO, there is full agreement between the subject and verb (4-6a), but not when the subject follows the verb (VSO) as in (4-6b)

4-6

a. Al-rijal-u namuu?
the-men-NOM slept 3MPl
"The men slept".

b- nama al-rijal-u
slept 3MS the-men-NOM
"The men slept"

In MSA and various spoken varieties of Arabic, these two word orders are commonly encountered. While it is believed that VSO is the unmarked word order in MSA, SVO is considered the basic word order in most modern spoken varieties of Arabic Such as Palestinian Arabic (Ouhalla & Shlonsky, 2003; Shlonsky, 1997), Tunisian Arabic (Mahfoudhi, 2002), Moroccan and Lebanese Arabic (Aoun, Benmamoun & Sportiche, 1994), and Gulf Arabic (Aljenaie & Farghal, 2009; Holes, 1989). Others have demonstrated that SVO is the unmarked word order in MSA and spoken varieties of Arabic (Shlonsky, 1997).

This study examines the comprehension of three types of word order in Gulf Arabic: SVO word order, which is considered the basic word order in GA and two derived word orders: OSV and OVS. In OSV, the object moves to the beginning of the sentence, while in OVS, both the object and the verb move to the left of the subject. In both OSV and OVS, a clitic is placed in the position of the object, and this clitic is coindexed with the fronted object. This type of syntactic construction seen in OSV and OVS is referred to as clitic-left dislocation (CLLD) (Cinque, 1990), which expresses a topic-comment function. In Arabic, CLLD involves a resumptive clitic that is coindexed with the topicalised CLLD-element. In GA, like in Italian (Rizzi, 1997), this resumptive clitic is obligatory when the topicalised constituent is a direct object.

The analysis of CLLD construction adopted in this study is based on the work of Aoun and colleagues in their study of Lebanese Arabic (LA) (Aoun & Benmamoun, 1998; Aoun, Choueiri & Hornstein, 2001; Li & Aoun, 2003). It is believed that the generalisations and examples cited in Aoun and Benmamoun (1998) for LA can hold for GA, with some minor lexical or phonetic differences. For example, in (1) the only difference between this CLLD construction in 4-7 in LA and its equivalent in GA is in the pronunciation of the verb and the clitic. In LA they are pronounced (sheef-a) while in GA it is (shaaf-ha).

4-7 An Example of CLLD construction (Aoun & Benmamoun, 1998)

Naadya ʃeef-a Kariim mbeeriħ.
Nadia saw- her CL Karim yesterday
'Nadia, Karim saw her yesterday.'

All the clitics in this experiment are direct object clitics, though CLLD constructions can involve dative clitics, genitive clitics, or adnominal clitics. In a series of studies Aoun and colleagues (Aoun & Benmamoun, 1998; Aoun et al., 2001; Li & Aoun, 2003) have demonstrated that the type of CLLD construction used in the current experiment does involve movement of the direct object, which is coindexed with the clitic. In LA and GA, CLLDed elements are interpreted as sentence topics or aboutness topics.

4.1.5 Sentence comprehension in Arabic

Very few studies have investigated sentence comprehension and comprehension of various word orders in Arabic. In Al-Akeel's (1998) comprehension test, some tasks examined the comprehension of some reversible sentences that have various cues, such as lexical cues and agent-action-object cues. Al-Akeel (1998) found that 50% of these children (60 children) passed the reversible active comprehension task showing that this structure was mastered only by the age group 4;6-4;11 years. Results of the performance of these children on these tasks showed that they used world knowledge and lexical cues more than they employed agent-action-cues. Less than a quarter of the subjects used lexical cues, but failed to rely on assigning the agent to the first NP. Therefore, Al-Akeel (1998) concluded that in Saudi Arabic word order was not a strong cue since 19% of the children did not assign the role of agent to the first noun although there was no conflict with world knowledge. However, Al-Akeel used very few exemplars and due to the nature of the sentence comprehension test, no systematic control of various cues was established.

Aljenaie and Farghal’s (2009) study is the only published investigation of the comprehension of different word orders in typically developing Gulf Arabic speaking children. They examined the comprehension of three types of word orders in Kuwaiti GA speaking children aged between 4 and 8 years old. These were: a canonical SVO, a marked VSO and a more marked type known as Topic-Comment structure (T-C). The T-C structure involves fronting an object to the initial (topic) position with the presence of a resumptive pronoun that is corefrential with the fronted NP and agrees with it in person, gender, and number. Moreover, the verb moves to the left of the subject. This structure corresponds to the $OV_{CL}S$ word order subsequently used in the present study. Aljenaie and Farghal’s (2009) aims were to examine which word orders were acquired earlier and how children used different strategies to help in comprehension, i.e., whether they relied on word order or gender agreement cues.

Aljenaie and Farghal (2009) used an acting out procedure where children listened to 27 reversible and non-reversible sentences and acted out the three types of sentences. An example of these three types of sentences is mentioned below. The sentences below all correspond to the English sentence ‘the dog is chasing the cat’ (Aljenaie & Farghal, 2009, p.499).

4-8

a. SVO	ɪl-tʃalb	yɪ-lħag	ɪl-gatˁwa
	The-dog-MS	3 MS-chases	the-cat- F
b. VSO	yɪ-lħag	ɪl-tʃalb	ɪl-gatˁwa
	3MS-chases	the-dog-MS	the-cat-F
c. T-C	ɪl-gatˁwa	yɪ-lħag-ha	ɪl-tʃalb
	the-cat-F	3 MS-chases-her CL	the-dog-M

Sentence (4-8a) represents canonical SVO order, while sentence (4-8b) shows a VSO order, commonly found in Standard Arabic. In both (4-8a) and (4-8b), the subject agrees with the verb in person, number, and gender and this agreement is marked as a prefix on the verb. In the Topic-Comment structure in sentence (4-8c), there are gender cues (both the subject and the prefix on the verb are masculine) since subject agrees with verb in Arabic, while the other cue comes from the resumptive pronoun that agrees with the feminine object (the cat) in gender and number. The authors explained that both SVO and T-C sentences showed a Noun Verb Noun (NVN) configuration, while in VSO, the configuration was VNN (Aljenaie & Farghal, 2009).

Results of the experiment showed that these TD children did relatively well on both SVO and VSO word orders, where subjecthood is assigned to the first NP and objecthood to the second NP. These typically developing GA speaking children had a score of 91% on SVO sentences and 85% on VSO sentences. These findings supported the acquisitional order of SVO then VSO. The authors explained that the higher score on SVO was due to unmarkedness of this structure since it is encountered early in development, VSO, on other hand, is less frequent in Gulf Arabic, and T-C is very infrequent. They also found significant differences between younger children (4-6 years) and older ones (6-8 years) in favour of the latter group on the comprehension of the three types of word orders. Younger children had a correct score of 85%, 71%, and 54% on SVO, VSO, and T-C sentences respectively, while older children scored 96%, 96%, and 83% on SVO, VSO, and T-C orders respectively. They argued that correct interpretation of SVO and VSO sentences could be reached by relying on word order strategy whereby subjecthood is assigned to the first NP, while objecthood is assigned to the second NP, though morphological cues in the form of subject verb agreement can play a role. Only in T-C sentences was the syntactic cue of word order neutralised. The T-C sentences posed some difficulty for these TD children because relying on word order alone will not derive the correct interpretation; children had to have efficient utilisation of morphological cues (gender and number agreement) in order to understand these sentences. They also found that older children benefited from agreement cues more than younger children, as older children had lower scores on reversible sentences where there were no gender cues (both subject and object were masculine nouns) in comparison to those with mixed genders. Apart from word order cues and morphological cues, the authors found strong evidence for animacy cues and little evidence for pragmatic cues in younger children, as many children experienced difficulties understanding sentences whose reverse interpretation was deemed pragmatically inappropriate in Gulf Arabic (e.g., the cat chases the dog).

Both Al-Akeel's (1998) and Aljenaie and Farghal (2009) examined sentence comprehension in typically developing children. While Al-Akeel's test was a general sentence comprehension test that included simple and some complex sentences, Aljenaie and Farghal's (2009) task examined three types of word order to establish the developmental pattern of comprehension of marked vs. unmarked word orders. Both studies, however, did not systematically control animacy, agreement, and world knowledge cues. Therefore, an examination of sentence comprehension of children with SLI in Gulf Arabic can increase our understanding of how typically developing GA children understand complex and reversible sentence in Arabic and how children with SLI will perform when they are compared to age and language controls.

4.1.6 Comprehension of sentences with fronted NPs in Gulf Arabic speaking children with SLI

This chapter examines the comprehension of three types of word order in Gulf Arabic: the first one is the canonical and most frequent word order, i.e., SVO, and the other two are less frequent and more marked, namely OSV_{CL} and $OV_{CL}S$. In OSV_{CL}, the object moves to the initial position of the sentence, while in $OV_{CL}S$ both the object and the verb move to the front. In both sentences a clitic is coindexed with the fronted NP and agrees with it in person, gender and number.

The following derivations are suggested for these three types of word order in GA. The following sentences are taken from the stimuli used in the present experiment and they all convey the meaning: "the elephant is kicking the dog', produced in different word orders.

4-9

a- Canonical SVO:

ɪl-fi:l	ya-rfɪs	ɪl-tʃalb
the-elephant M	3MS-is kicking	the-dog M

b- O S V_{CL}_ (one movement)

ɪl-tʃalb ɪl-fi:l ya-rfɪs-ah

the-dog M the-elephant M 3MS-is kicking-him (Object CL)

c- O V_{CL} S __ __ (two movements) : 1- OSV_{CL}_,2- $OV_{CL}S$_ _

ɪl-tʃalb ya-rfɪs-ah ɪl-fɪ:l

the-dog (M) 3MS-is kicking-him (Obj CL) the-elephant M

As shown in the three examples above, subject-verb agreement is always indicated by a prefix on the verb that agrees with the subject in gender, person, and number. Examples (4-9 b) and (4-9 c) differ from (4-9 a) in that the direct object is fronted to the beginning of the sentences, leading to the presence of a clitic that agrees with the object in gender, person, and number.

These structures are best described as Clitic Left Dislocation (CLLD) structures, which have been found to involve movement (Aoun & Benmamoun, 1998; Aoun, 2001; Li & Aoun, 2003). Evidence supporting the above mentioned movements of verb and object can be found in sentences where some type of adverbs, such as manner adverbs, are inserted. Manner adverbs can be inserted before the verb and the subject (4-10a) and can precede or follow the subject (4-10c) and (4-11b). However, grammaticality is significantly degraded if manner adverbs are inserted between the verb and the object in SVO sentences (4-10b). Example 4-11 shows fronting of both the object and verb leading to $OV_{CL}S$.

4-10 SVO

a. Ali bsɪrʕa daz ɪs-sayyara

Ali quickly-ADV pushed the-car F

Ali quickly pushed the car.

b.* Ali daz bsɪrʕa ɪs-sayyara

Ali pushed quickly ADV the-car F

Ali pushed quickly the car.

c. bsɪrʕa Ali daz ɪs-sayyara

quickly-ADV Ali pushed the-car F

Quickly Ali pushed the car.

4- 11 $OV_{CL}S$

a. Is-sayyara daz-ha bsɪrʕ Ali

The-car F pushed-it F quickly ADV Ali

The car Ali pushed it quickly.

b. Is-sayyara daz-ha Ali bsɪrʕ

The-car F pushed-it F Ali quickly ADV

The car Ali pushed it quickly.

It is important to note that adverbs in Arabic behave differently depending on their types. Therefore, while some types may appear between the verb and direct object in SVO sentences (e.g., Shlonsky, 1997), others like manner adverbs in Gulf Arabic are significantly degraded if they appear in such position.

Although there is another word order which is encountered in MSA and most modern spoken varieties of Arabic, namely the VSO order, it is not included in this experiment because VSO order is mostly used in narratives in spoken dialects of Arabic (Brustad, 2000) and therefore may not be appropriate for this type of experiment where isolated sentences are used without previous context. However, in Aljenaie and Farghal's (2009) study, they found that VSO order was easier than $OV_{CL}S$. They concluded that VSO sentences were more marked than SVO sentences, but less marked than the $OV_{CL}S$ structures (Aljenaie & Farghal, 2009). As for the difference in meaning between the two marked structures $OV_{CL}S$ and OSV_{CL} , some argue that these changes in configuration do not involve changes in meaning, as fronting of the object in the two structures has the function of topic in Arabic (Bulk, 2008; Ford, 2009). However, there is a contrastive focus on the subject in the OSV_{CL} sentences, as in the following sentence, where the subject 'the elephant' receives a contrastive focus.

4- 12 ɪl-tʃalb ɪl-fi:l ya-rfɪs-ah

The-dog M the-elephant M 3MS-is kicking-him (Obj CL)

The dog, the elephant kicked him (the elephant and not the tiger).

This study assumes that children with SLI do not have problems interpreting the discourse functions of topic-comment sentences, especially in spoken Arabic where SVO is frequently interpreted as topic-comment (Brustad, 2000). Botting (2004) found that children with SLI, who had an average age of 11 years old, scored within the normal range on a standardised test of pragmatics: the Children's Communication Checklist (Bishop, 1998) and concluded that children with SLI generally do not have a primary deficit in pragmatics. Typically developing children as young as 1;11 year old correctly interpreted dislocated elements as topics (De Cat, (2003). Moreover, Schaeffer (2003) reviewed pragmatics skills in Dutch children with SLI and showed that Dutch speaking children with SLI had no problems with scrambling of referential objects such as pronouns, a syntactic operation that is driven by pragmatic considerations. She, therefore, concluded that children with SLI were similar to their age matched children in terms of interface pragmatics (the concept of non-shared knowledge) and they performed better than younger TD children who were at similar stage of grammatical development (Schaeffer, 2003). Overall, based on what is known about the pragmatic skills of children with SLI in general, and the variability of word order in Gulf Arabic, where even SVO sentences are sometimes interpreted as topic structures (Brustad, 2000), it is very unlikely that pragmatic factors will influence the results of this comprehension task.

4.1.7 Aim and predictions of the present study

Firstly, the investigator is not aware of any study that examined the sentence comprehension skills of children with SLI in Gulf Arabic or any other variety of Arabic and in general there are very few studies of sentence comprehension skills of typically developing Gulf Arabic speaking children (e.g., Aljenaie & Farghal, 2009). The results will be examined to see if they confirm the pattern seen in other languages, where children with SLI show greater deficits in comprehension of complex sentences vs. simple sentences when they are compared to age control children, and sometimes in comparison to their language controls. If children with SLI score significantly worse than the language control children on the comprehension of complex sentences, this type of task could be potentially employed in identifying children with SLI.

Secondly, the comprehension of complex sentences in Gulf Arabic may serve as a testing ground for domain-general and domain-specific accounts of SLI, as these two theories of SLI have different predictions for the performance of the typically and atypically developing groups.

Predictions of processing-based accounts. Most domain-general (processing accounts) will predict that SVO sentences will be the easiest as they are canonical, and frequent in parental input. Moreover, the SVO sentences carry less processing load than the other two word orders. The three types of sentences are of relatively short length (mainly three constituents: subject, verb, and object, with a clitic in the fronted sentences), however, OSV_{CL} and $OV_{CL}S$ sentences could pose more difficulty due to the extra processing involved in waiting to process the clitic and its referent before reaching a full interpretation, thus creating extra load on the working memory. However, these accounts do not specify which of the two noncanonical word orders is more challenging, though based on Deevy and Leonard (2004), one may suggest that OSV_{CL} should be more challenging due to the increased distance between the clitic and its referent (the direct object). Deevy and Leonard (2004) argue that any increase in the distance between the wh-word and its trace/gap (dependency) would prove more challenging for SLI children than it is for typically developing children. Therefore, such an account may predict a slight advantage for $OV_{CL}S$ sentences over OSV_{CL}, or at least comparable performance on the two.

Overall, the hierarchy of difficulty (from easiest to most difficult) according to processing accounts is as follows: $SVO>OV_{CL}S>OSV_{CL}$

Predictions of linguistic accounts. Domain-specific (linguistic) accounts of SLI would predict, like domain-general accounts, that SVO sentences will be the easiest, because they follow canonical word order in Gulf Arabic and the object is not moved. Moreover, there are less syntactic dependencies when compared with $OV_{CL}S$ and OSV_{CL}. However, linguistic accounts diverge from processing accounts in predicting that $OV_{CL}S$ sentences will be more challenging than OSV_{CL} sentences, especially for children with SLI. The $OV_{CL}S$ sentences involve movements of both the object and the verb, while there is only object movement in the $OSV_{CL.}$ Though this increased syntactic complexity can affect all children, it is predicted to have more detrimental effects on the vulnerable linguistic system of children with SLI. If children with SLI have less efficient syntactic system, it is predicted that they will perform less well on the types of sentences that rely on good syntactic ability. Linguistic accounts of SLI predict that children with SLI will perform worse on these two complex word orders (OSV_{CL} and $OV_{CL}S$).

Between these two complex word orders, linguistic accounts predict that children with SLI will find the $OV_{CL}S$ word order significantly more difficult than the OSV_{CL} order due to the increased number of movements in $OV_{CL}S$. One of the factors contributing to the complexity of the OSV_{CL} and $OV_{CL}S$ sentences is the number of movements involved in these two sentences. In both OSV_{CL} and $OV_{CL}S$ sentences, the object moves to the front creating a more complex structure than the unmarked SVO, however the $OV_{CL}S$ involves an extra number of movement, i.e., verb movement, therefore $OV_{CL}S$ according to linguistic accounts of SLI should be more challenging than the other two types, with the group of children with SLI being more susceptible to this increased grammatical complexity. Linguistic accounts predict that children with SLI, when faced with more complex sentences, will resort to using the common Subject- first strategy, where the first NP is assigned the agent role, while the second NP is assigned the theme role. This is contrasted with typically developing children, who will rely on their syntactic knowledge and interpret sentences based on syntactic structures in a more efficient way when compared to atypically developing children.

Linguistic accounts of SLI predict the following hierarchy of difficulty (from easiest to most difficult) in children with SLI:

$SVO > OSV_{CL} > OV_{CL}S$.

The hierarchy SVO, $OS_{CL}V$, $OV_{CL}S$ will be strictly observed in children with SLI, while typically developing children are expected to perform well on both SVO, where the word order is canonical, and OSV_{CL}, where they can use their efficient syntactic abilities to reinterpret these sentences, where an NNV configuration is encountered. They will realise that a strict word order will not yield the right interpretation and therefore they will rely on syntactic relations among constituents and agreement. The order $OV_{CL}S$ will be challenging for TD children, especially the younger ones who may not use agreement cues efficiently at this stage of their language development.

Thirdly, this experiment endeavours to investigate how word order cues and agreement cues are employed by younger and older typically and atypically developing GA speaking children. Apart from manipulating word order, gender agreement is manipulated in this experiment to examine its influence on the results of all groups on the sentence comprehension task and compare it with the role of word order cues. In Arabic, all nouns are essentially either feminine or masculine and subject and verb agree in gender, number and person. Therefore, the results will be analysed to see whether agreement cues can help in facilitating the comprehension of complex sentences, where subjects and objects have all possible combinations of gender. The design of the experiment will aim to neutralise other cues, such as animacy, reversibility and frequency. Therefore, children will rely on mainly two cues: word order and gender agreement cues. There is disagreement about which cues are preferred when interpreting reversible sentences in Arabic. The two studies examined so far point in different directions, with Al-Akeel (1998) arguing that word order is not a strong cue, while Aljenaie and Farghal (2009) argue that word order cues are stronger than agreement and pragmatic cues. Moreover, Aljenaie and Farghal (2009) found that very young children (less than 6 years old) had difficulties utilising agreement cues, as it is argued that in pro-drop languages, like Arabic, gender agreement is more difficult as the subject is commonly deleted (Sokolov, 1989). However, older children can efficiently use gender agreement cues when interpreting reversible sentences as has been found in Aljenaie & Farghal's (2009). None of these studies, however, systematically manipulated agreement cues as they included other types of cues, such as animacy and pragmatic cues. Therefore, it is predicted that this experiment, where animacy, world knowledge, and prosodic cues are arguably controlled, would provide useful information about the role of word order and agreement cues in interpreting reversible sentences in Arabic.

4.2 Method

4.2.1 Participants

Thirty nine Qatari Gulf Arabic speaking children participated in this experiment. Thirteen of them were diagnosed with SLI (Age_M= 8;0 [years;months]), 13 were typically developing age control (AC) children (Age_M =7;11), and 13 were typically developing language control (LC) children (Age_M =5;8) matched with the SLI group based on their score on the Sentence Comprehension test. All children received a battery of tests that included the following: the Sentence Comprehension test, the Expressive Language test, the Sentence Repetition test, and an Arabic Picture Vocabulary test (see chapter 3 for descriptions of these tests). All children scored within normal range on either the Test of Nonverbal Intelligence (TONI-3) (Brown et al., 1997) for children aged 6;0 and above or two subtests of the Wechsler Preschool and Primary Scale of Intelligence (WPPSI) (Wechsler, 2002) for children aged less than six years old. These subtests were the Block Design and the Picture Completion subtests, which were recommended as an appropriate short form of nonverbal IQ (LoBello, 1991; Tomblin et al., 1997). All participants younger than 6 years old had a total score of 16 and above on the Block Design and the Picture Completion test of the WPPSI, as 16 was the cut-off score recommended by LoBello (1991), which corresponds to a standard score of 87.

All children with SLI met the SLI criteria adopted for this study, which were based on Tomblin et al. (1997). Therefore, these children's performance was characterized by the following:

- The presence of significant receptive/expressive impairment defined as having -1.5 standard deviations (SD) and above on at least two tests or -2.0 (SD) or more on one out of the four language tests developed for this project, namely; the Sentence Comprehension test, the Expressive Language test, the Sentence Repetition test, and the receptive Arabic Picture Vocabulary test.
- Non-verbal IQ of 85 and above for children aged 6;0 years old and above or a combined score of 16 on the shorter form of WPPSI (Wechsler, 2002)
- Normal Hearing. All children passed a hearing screening at 20dB at frequencies 500-2000 Hz, performed by the investigator, who is a certified speech-language therapist.
- Uneventful developmental history (e.g., no diagnosis of attention deficit hyperactivity disorder (ADHD), autism spectrum disorder). Moreover, all Children passed the developmental verbal dyspraxia and oral-motor screening tests.

While Tomblin et al. (1997) set the criterion scores for children with SLI at -1.25 SD below the mean on two tests or -2.0 SD on one test, a more stringent criterion of -1.5 SD was used in this experiment due to novelty of the tests used to diagnose children with SLI in Gulf Arabic.

The children with SLI were group-wise matched with the LC based on their raw score on the Sentence Comprehension test (both groups had a mean raw score of 25.8). This test comprised 40 sentences that ranged in difficulty from simple to most difficult and the test had a good level of reliability and validity (see Chapter 3). Leonard (1998) recommends that SLI and language control groups be matched based on the dependent variable being tested. Therefore, matching on a general test of language comprehension might be more appropriate when the measure of interest is the comprehension of a specific grammatical structure (Leonard, 1998), as is the case in this experiment.

There was no significant difference in nonverbal IQ score on the TONI-3 test between children with SLI and their age-matched group $t(1,24)=.51$, $p=.61$. The younger typically developing children were presented with a different performance IQ test and therefore it was not possible to compare them with the older participants. However, they all scored within normal range on the Block Design and the Picture completion subtests of the Wechsler Preschool and Primary Scale of Intelligence (Wechsler, 2002).

The SLI group consisted of 10 boys and 3 girls, while the AC group had 9 boys and 4 girls, and the LC group comprised 7 boys and 6 girls. All participants were monolingual native speakers of Qatari Gulf Arabic. Most children were recruited from two kindergartens and four primary schools in Doha, the capital of Qatar, except for three (two children with SLI and one AC child) who were recruited through personal acquaintances. Only one of the children with SLI had previously received speech-language therapy and no child was receiving speech-therapy at the time of testing. None of the schools visited provided speech-therapy services for its students. See Table 34 for descriptive summary data of the three groups.

Table 34 : Descriptive summary data for the children with SLI (SLI; *n*=13), age-control group (AC; *n*=13) and language control group (LC; *n*=13).

Group	Age (months)	SC	EL	SR	APVT	TONI-3
SLI						
M	95.5	25.8	33.2	58.2	48.2	94.03
SD	11.4	3.9	11.1	16.3	12.0	7.4
Range	75-112	20-31	16-55	29-78	29-72	85-109
AC						
M	95.4	32.8	54.8	84.8	79.2	95.8
SD	11.8	2.3	14.0	11.4	20.0	7.5
Range	74-112	29-38	38-94	61-101	47-110	86-111
LC						
M	68.3	25.8	43.2	70.1	55.2	a
SD	7.4	3.3	8.1	13.3	18.5	a
Range	58-78	22-33	31-55	49-89	37-90	a

Note. SC=Sentence Comprehension test raw score; EL=Expressive Language test raw score; SR=Sentence Repetition test raw score; APVT=Arabic Picture Vocabulary test raw score; DS=digit span task from the Wechsler Preschool and Primary Scale of Intelligence (Wechsler, 2002**)** raw score; TONI-3=Test of Nonverbal Intelligence-3 standard score.

[b] All the language control children scored at or above the cut-off score of 16 on the shorter version of Wechsler Performance IQ (see LoBello, 1991 and Tomblin et al., 1997).

4.2.2 Materials and Procedure

A picture-pointing task was used to test children's comprehension of reversible sentences with fronted NP's. The stimuli consisted of 54 drawings with an array of four pictures per drawing. Each picture depicted two animal characters involved in a reversible action, with a third picture depicting an adjectival interpretation, and a fourth distracter picture. An example of the sentences and the corresponding stimuli used is shown in 4-13 and Figure 14.

4-13

Il-fi:l y-maʃɪt lɪ-bgarə

The-elephant M 3M- comb the-cow F

The elephant combs the cow

Figure 14: An example of the sentences used: 'The elephant combs the cow' (SVO).

The stimuli used were designed to measure the influence of two variables on the comprehension skills of all participants, namely the type of word order used and gender agreement cues. The design was a 3 (group: SLI, LC and AC) X 3 (word order types: SVO, OSV_{CL} and $OV_{CL}S$) X3 (agreement cues: no cues, masculine cue and feminine cue). The dependent variable was the percentage of correct answers, while the independent variables were word order type and agreement cues, with three levels in each of them. The 54 sentences were divided into three word order types: SVO, OSV_{CL} and $OV_{CL}S$ with 18 sentences in each type. The SVO order is considered the basic word order in spoken Arabic, while the OSV_{CL} and $OV_{CL}S$ orders involve fronting of the direct object to the initial position leading to the creation of a clitic pronoun that is prosodically attached to the verb. Gender agreement was manipulated in this experiment. All nouns used were singular and all the animals were either masculine or feminine. In Arabic, all nouns are either feminine or masculine and the subject agrees with the verb. This agreement is marked by a prefix on the verb, which is either 'y-' for masculine subjects or 't-' for feminine subjects. Furthermore, the object in fronted sentences (OSV_{CL}, and

$OV_{CL}S$) agrees with the clitic in number and gender. The clitic marking masculine objects is '-a', while the '-ha' clitic is used for feminine objects. Therefore, there were 3 types of word order sentences, with three types of gender agreement cues in each word order type. The notation AGR1, AGR2, and AGR3 are used to indicate agreement: (AGR1) stands for no gender agreement cue: where subject and object have the same gender as they were equally distributed as both feminine or masculine. Agreement type 2 (AGR2) represents masculine agreement cue, where the subject is masculine and the object is feminine and therefore the verb has the masculine prefix 'y-'). Gender agreement type 3 (AGR3) stands for feminine agreement cue, where the subject is feminine and the object is masculine and the verb in these sentences starts with the feminine marker 't-'. There were six items for each of the nine conditions, bringing the total number of sentences to 54. The gender of each animal was checked with two native Qatari speakers and they unanimously agreed on the distribution of masculine and feminine animals. Most of the female animals ended with the feminine marker traditionally known in Arabic as *taa al-taneeth* (the feminine /t/), which is phonetically realized as /a:t/ or /ə/ in spoken dialects of Arabic, such as the underlined 'ə' in "ħam:am-<u>ə</u>": a pigeon-FS). Table 35 shows examples of the nine types of sentences used in this experiment. For a complete list of all the sentences used in the experiment, see Appendix C.

Table 35: Examples of the nine types of sentences used in the experiment.

Word Order	Gender agreement cue Subj	 Obj	Examples
SVO1	M	M	ɪl-fi:l ya-rfɪs ɪl-tʃalb the elephant M 3M- kicks the dog M The elephant kicks the dog
SVO2	M	F	ɪs-sɪndʒa:b y-ħɪk ɪs-sɪlħɪfa:t the-squirrel M 3M-scratches the-turtle F The squirrel scratches the turtle
SVO3	F	M	ɪn- namlə t-arsɪm ɪl-fa :r the-ants F 3F- draws the-mouse M The ant draws the mouse
$OSV_{CL}1$	M	M	Id-dub ɪl-?asad y-dʒɪr-a the-bear M the-lion M 3M-pulls-him CL The elephant, the dog kicks him
$OSV_{CL}2$	M	F	ɪs- sɪlħɪfa:t ɪs- sɪndʒa:b y-ħɪk-ha the-turtle F the squirrel M scratches-her CL The turtle, the squirrel scratches her
$OSV_{CL}3$	F	M	lɪ-ħma:r ɪl-ħamamə t-alħɪq-a the donkey M the-pigeon F 3F-chases-him CL The donkey, the pigeon chases him
$OV_{CL}S1$	F	F	ɪl- fara:ʃə t-aʁsɪl-ha ɪn-naħla The butterfly F 3F-washes-her-CL the-bee F The butterfly, the bee washes her.
$OV_{CL}S2$	M	F	ɪd-dou:də y-bou:s-ha iðˤ-ðˤɪfdaʕ The-worm F 3M-kisses-her-CL the-frog M The worm, the frog kisses her
$OV_{CL}S3$	F	M	ɪl-fa:r t-arsɪm-a ɪn-namlə The-mouse M F-draws-him CL the-ant F The mouse, the ant draws him (draws a picture of him)

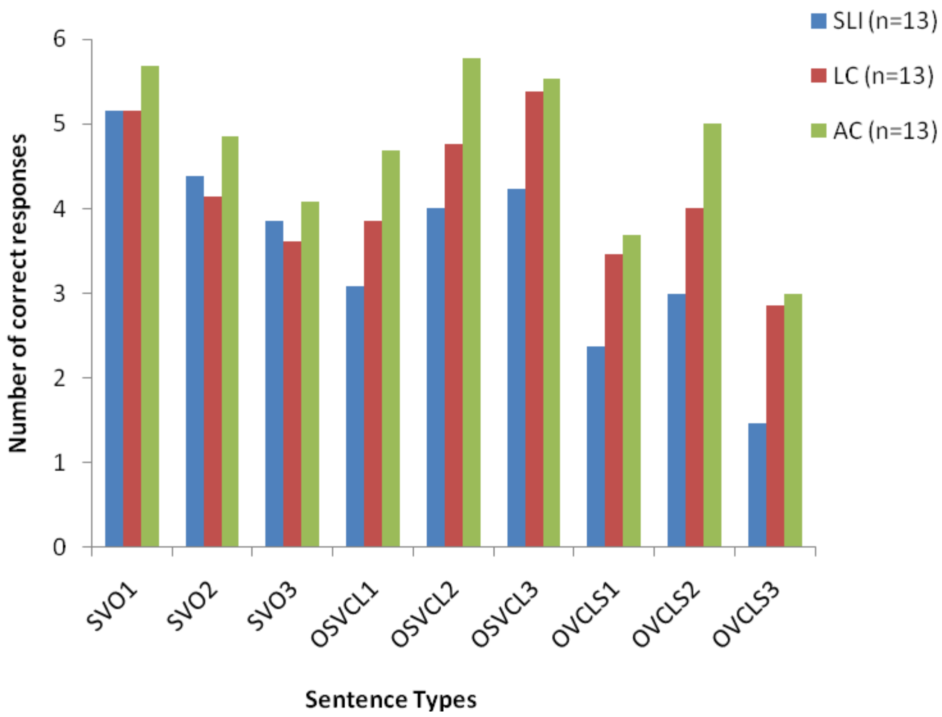
The sentences consisted of actions being performed by 24 different animals (12 feminine and 12 masculine) that were counterbalanced in terms of number of appearances in each condition. In each sentence, the agent and theme were of similar level of familiarity. These animals were shown performing 12 familiar verbs. No frequency data were available of Gulf Arabic and therefore a familiarity-rating questionnaire was conducted to assess the familiarity of all nouns and verbs used in the experiment. Twenty two native speakers of Qatari Gulf Arabic were given 650 words that belonged to 20 semantic categories (the same procedure was used in designing the Arabic Picture Vocabulary Test) and were asked to rate these words in terms of familiarity. All the animals chosen for this experiment were of high familiarity, i.e., they received an average rating of 3 and above on a rating scale of 1-5 where 1 referred to words that were rarely used and 5 used to described words that were used all the time and in different situations. Most of the verbs used received high familiarity ratings (See Appendix B for familiarity ratings of the verbs and animals used in the experiment). The actions depicted by the pictures showed animate characters to avoid any animacy bias. The animals used and the actions illustrated were carefully chosen to avoid any semantic, pragmatic, or world knowledge influences. All sentences and the position of the correct picture were then randomised and counterbalanced to avoid any consistent pointing at certain positions of the pictures. Two different lists of the stimuli were created and children were randomly assigned to one of the two lists. The sentences and the instructions were recorded by a female native speaker of Qatari Gulf Arabic in Doha, Qatar. The stimuli were presented to children through a computer program where they were requested to listen to sentences delivered through a headset and press on the correct answer on a touch screen installed on a laptop. The instruction was as follows (in Gulf Arabic): "You will listen to a sentence and I want you to look at the four pictures on the screen and press on the one corresponding to the sentence you hear".

In the beginning of the experiment, children were asked to identify the animals from a small booklet to establish that they were familiar with these animals. All children identified all animals. Children started the experiment by listening to instructions and they were shown how to use the touch screen to choose the correct picture. There were three practice items to help children understand the testing routines, which involved listening to the sentence, pressing on one of the four pictures on the touch screen, and pressing a "Go" button (a green check mark) to proceed to the next screen. All children seemed to understand the testing procedures and enjoy the experiment. Half-way through the experiment (after item 27), the testing was paused and children were given visual reinforcements (pictures of balloons and cheering sound) and were given the option of taking a break or continuing the experiment. All of them wanted to continue the experiment. On average, it took 15 minutes to finish the experiment.

4.3 Results and Analysis

Scoring was based on the number of correct sentences identified and each item received a score of 1 (correct) or 0 (incorrect). There were 9 types of sentences with 6 exemplars in each type, with an overall total of 54 sentences. The results for the scores of all groups on each word order type are shown in Figure 15 and Table 36.

Figure 15: Overall results of all the participants on the nine types of sentences



Note: SVO: Subject-Verb-Object, OSV_{CL} (Object-Subject-Verb + Clitic), $OV_{CL}S$ (Object-Verb + Clitic + Subject). 1=no agreement cue, 2=masculine subject agreement cue, 3=feminine subject agreement cue.

Table 36 : Descriptive statistics of the comprehension of nine types of sentences in Gulf Arabic speaking children.

	SLI (*n*=13)			LC (*n*=13)			AC (*n*=13)		
Word order type	M	SD	Range	M	SD	Range	M	SD	Range
SVO 1	5.15	1.34	2.0-6.0	5.15	0.99	3.0-6.0	5.69	0.63	4.0-6.0
SVO 2	4.38	0.96	2.0-5.0	4.15	0.99	2.0-5.0	4.85	0.55	3.0-5.0
SVO 3	3.85	0.55	3.0-5.0	3.62	0.77	2.0-5.0	4.08	0.49	3.0-5.0
Total SVO /18	13.38	2.20	8.0-16	12.92	1.89	10.0-15.0	14.62	0.96	13.0-16.0
$OSV_{CL}1$	3.08	1.50	0.0-5.0	3.85	1.14	1.0-5.0	4.69	1.11	3.0-6.0
OSV_{CL} 2	4.00	1.63	1.0-6.0	4.77	1.17	3.0-6.0	5.77	0.44	5.0-6.0
OSV_{CL} 3	4.23	1.69	1.0-6.0	5.38	0.77	4.0-6.0	5.54	0.78	4.0-6.0
Total OSV_{CL}/ 18	11.31	4.42	3.0-17.0	14.00	2.27	10.0-17.0	16.00	2.12	12.0-18.0
$OV_{CL}S$ 1	2.38	1.12	1.0-5.0	3.46	1.56	1.0-6.0	3.69	1.44	1.0-6.0
$OV_{CL}S$ 2	3.00	1.15	1.0-4.0	4.00	1.47	2.0-6.0	5.00	0.82	4.0-6.0
$OV_{CL}S$ 3	1.46	0.88	0.0-3.0	2.85	1.52	1.0-5.0	3.00	1.08	1.0-5.0
Total $OV_{CL}S$/18	6.84	2.11	3.0-11.0	10.31	3.52	6.0-16.0	11.69	2.50	8.0-17.0
Overall Total/54	31.53	6.64	19.0-40.0	37.23	6.18	31.0-47.0	42.3	3.35	37.0-47.0
Total %	58.40	12.30	35.20-74.07	68.95	11.44	57.41-87.04	78.34	6.20	68.2-87.04

Note: SVO= Subject-Verb-Object, OSV_{CL}: Object-Subject-Verb + Clitic, $OV_{CL}S$= Object-Verb+Clitic+Subject.1=no agreement cue, 2=masculine subject agreement cue, 3=feminine subject agreement cue.SLI=children with SLI, AC=age controls, LC=language controls.

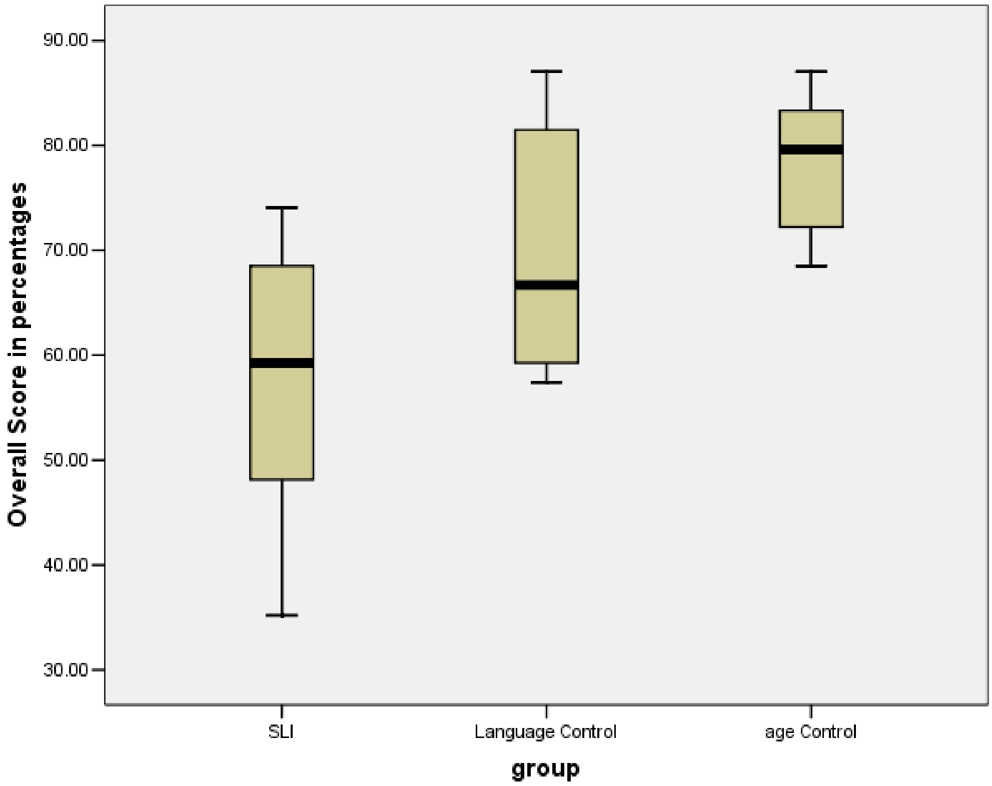
A 3 (group: SLI, AC and LC) X 3 (word order type: SVO, OSV_{CL} and $OV_{CL}S$) X 3 (Agreement cue based on subject's gender: no agreement cue, masculine subject cue, and feminine subject cue) ANOVA was conducted. Results showed that there was a significant main effect of word order type F(2,72) =42.23, p<0.001, η^2=.54, agreement F(2,72)=15.91, p<.0.001, η^2=.30 and group F(2,36)=12.10, p<001, η^2=.40. This showed that all independent variables (word order type, agreement cue, and the group to which each child belonged) had an effect on the performance of children. Moreover, there was a significant word order type*group interaction F(4,72)=3.62, p=.01, η^2=.16, and word order type*agreement interaction, F(4,144)=35.17, p<.001, η^2=.49. There was not group*agreement interaction.

In the following, main effects of the different variables are discussed, followed by an examination of interaction effects.

4.3.1 Main effect of group

Pairwise comparison with Bonferroni correction showed there was a significant difference between AC children and children with SLI (p<.001). In general the AC group had an overall score that was significantly higher than the score obtained by the children with SLI (p<.001). Moreover, the group of younger typically developing children (LC) scored significantly higher than the SLI group (p=.04), even though they matched the SLI group on Sentence Comprehension test scores. The difference between the two typically developing groups did not reach significance. Figure 16 shows that there was a little overlap between the scores of the AC group and those of the SLI group and that the LC group had a significantly higher score than the SLI group. It also shows that the SLI group had a very heterogeneous profile compared to the other two typically developing groups, with the children with SLI having a wider range of scores (between 35% and 72%).

Figure 16: A Boxplot summarising the overall score (in percentages) of each group of children on the comprehension of sentences with fronted NP's.



4.3.2 Main effect of word order

Table 37 and Figure 17 show the means of the three types of word orders across all groups.

Table 37: Percentages of correct sentences based on the types of word order used.

Word order type	Mean	SD	N
SVO	75.78	10.55	39
OSV_{CL}	76.50	20.05	39
$OV_{CL}S$	53.41	18.89	39

Figure 17: Percentage of correct responses of all groups on the three types of word order.

Pairwise comparisons with Bonferroni correction for the three types of word order showed that scores on SVO sentences were not different from OSV_{CL}. However, children's performance on $OV_{CL}S$ was significantly different from both SVO and OSV_{CL} word orders (p<.001 for both types). This showed that the noncanonical word orders were not treated equally, as one of them (OSV_{CL}) was as easy as the canonical (SVO), while the other one, ($OV_{CL}S$) was most difficult to understand.

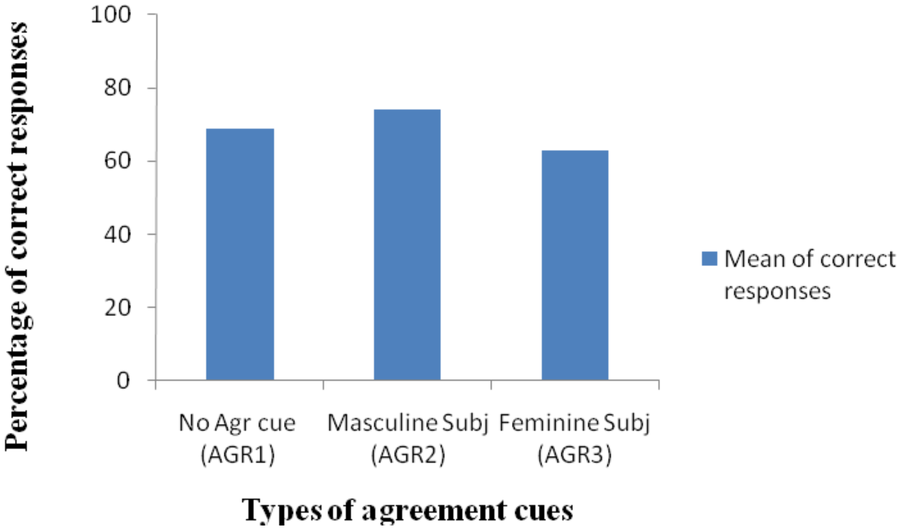
4.3.3 Main effect of agreement

The performance of all children on the three types of agreement cues is summarised in Table 38 and Figure 18.

Table 38: Means of the three types of agreement cues based on the gender of the subject (in percentages).

Subject-verb agreement cue	Mean	SD	N
No agreement cue (AGR1)	68.80	15.11	39
Masculine subject (AGR2)	73.93	16.18	39
Feminine subject (AGR3)	62.96	12.96	39

Figure 18: Means of correct responses of all participants on sentences with different agreement cues (AGR1 *n*=39, AGR2 *n*=39, AGR3 *n*=39).



Pairwise comparisons with Bonferroni showed that all sentences with agreement cues were significantly different from each other. Sentences that had no agreement cues (i.e. half of the time both subjects and objects were masculine, and they were both feminine the other half) were significantly more difficult than sentences with masculine subject agreement cue (i.e., the subject was masculine and the object feminine) ($p<.05$). These non agreement cue sentences were significantly easier than sentences with feminine subject agreement cues ($p<.05$). Sentences with masculine agreement cue where the subject was masculine and object were feminine were significantly easier than sentences with feminine agreement cues ($p<.001$). Based on these results, the following order of difficulty is suggested (from easiest to most difficult) based on gender agreement cues:

Masculine subject cues > No gender cues> Feminine subject cues.

4.3.4 Group * word order type interaction

There was a significant group by word order type interaction $F(4,72)=3.62$, $p=01$, $\eta^2=.16$. The performance on types of word order differed depending on the group to which the child belonged. The following table and bar chart show the means of groups' performance on the three types of word order: SVO, OSV_{CL}, and $OV_{CL}S$.

Table 39: Means and standard deviations of percentage of correct sentence comprehension of each group on each type of word order.

Sentence Type	SLI Mean (SD)	LC Mean (SD)	AC Mean (SD)	Total Mean (SD)
SVO	74.36 (12.73)	71.80 (10.50)	81.20 (5.33)	75.78 (10.55)
OSV_{CL}	62.82 (24.57)	77.78 (12.63)	88.89 (11.79)	76.50 (20.55)
$OV_{CL}S$	38.03 (11.75)	57.26 (19.56)	64.96 (13.89)	53.41 (18.89)
Overall mean	58.40 (12.30)	68.95 (11.44)	78.34 (6.20)	

Figure 19: The percentages of correct responses by type of word order for all three groups.

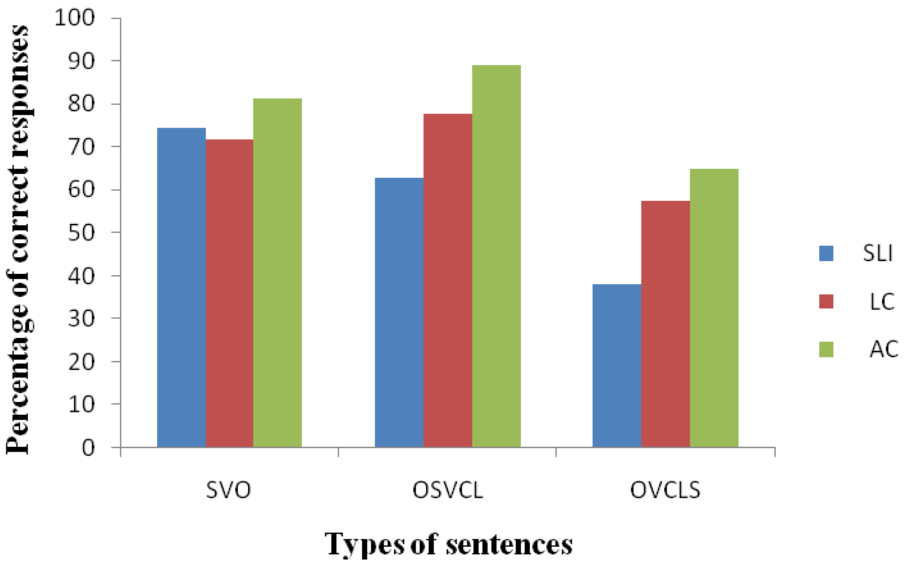
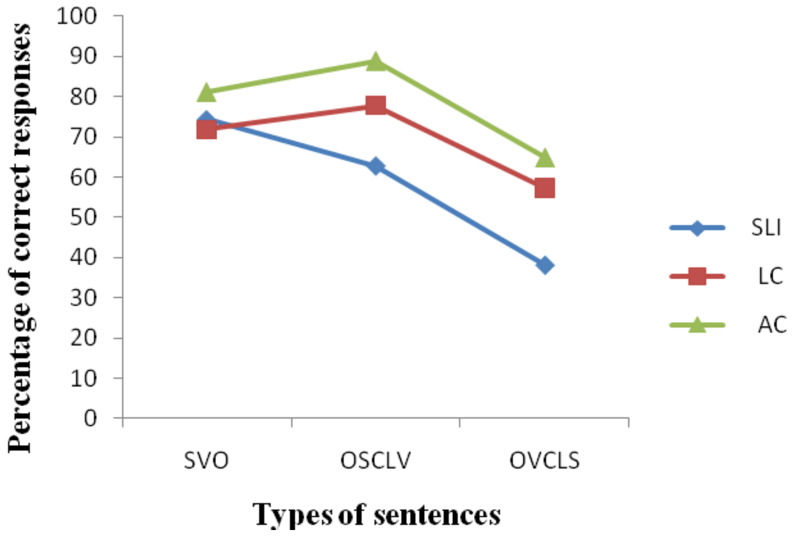


Figure 20 illustrates the interaction between word order types and groups. It shows that the performance of the group of children with SLI followed a pattern that is characteristically different from those of the two typically developing groups. While the three groups of children seemed to have relatively comparable performance on the canonical SVO sentences, the group of children with SLI diverged significantly from the two TD groups on OSV_{CL} and $OV_{CL}S$ sentences, with their performance degrading dramatically on the $OV_{CL}S$ word order.

Figure 20: Group by word order interaction



A one way ANOVA was performed to examine the effect of groups on each type of word order used. Table 40 summarises the results of these ANOVAs.

Table 40: Results of One way ANOVAs for types of sentences

Word order type	ANOVA results	Significance	Size effect
SVO	$F(2,36)=3.60$	$p=.06$	$\eta^2=.15$
OSV_{CL}	$F(2,36)=7.40$	$p<.01$	$\eta^2=.30$
$OV_{CL}S$	$F(2,30)=10.52$	$p<.001$	$\eta^2=.37$

This shows that all participating groups had comparable performance on canonical reversible SVO sentences. All children with SLI, who had an average age of 8 years, their age controls and their language controls who were on average 28 months younger, did well on reversible sentences of the order SVO. There was an advantage for the AC group over the other two groups, and slightly higher correct responses by the group of children with SLI over the LC group, but in both cases the difference was not significant.

On OSV_{CL} sentences, the SLI group had a score that was not significantly different from the LC group, though the LC had a higher score this time. However, the SLI group's performance on this type of sentences was significantly worse than the AC group ($p=.001$). The two typically developing groups were not significantly different from each other.

It was the third type of sentences, the $OV_{CL}S$ type that posed most difficulty to children with SLI, as their performance was significantly different from both typically developing groups, with a p value of .009 when compared to LC group, and $p<.001$ when the SLI group was compared to AC group. The two typically developing groups were not different from each other.

To sum up the performance of groups on sentence types, it was seen that all three groups performed well on canonical reversible SVO sentences. On the noncanonical word orders, the SLI group consistently performed worse than the AC group on both noncanonical word orders, while they were significantly worse than the LC on the $OV_{CL}S$ type of sentences only. When examining the profiles of each group the following generalizations hold:

SLI: $SVO=OSV_{CL}>OV_{CL}S$

AC: $SVO=OSV_{CL} >OV_{CL}S$

LC: $SVO=OSV_{CL} >OV_{CL}S$

To examine the significance of these differences paired sample t-tests with Bonferonni correction were conducted for each group. The results were as follows:

SLI:

SVO and OSV_{CL} the difference was NOT significant $t(12)=1.90$ ($p=.08$).

SVO and $OV_{CL}S$: the difference was significant $t(12)=8.0$ ($p<.001$).

OSV_{CL} and $OV_{CL}S$ the difference was significant $t(12)= 3.63$($p<.01$).

AC:

SVO and OSV_{CL} the difference was NOT significant $t(12)=-1.66$ ($p=.12$).

SVO and $OV_{CL}S$: the difference was significant $t(12) =3.15$($p<.01$).

OSV_{CL} and $OV_{CL}S$ the difference was significant $t(12)=4.03$ ($p=0.01$).

LC:

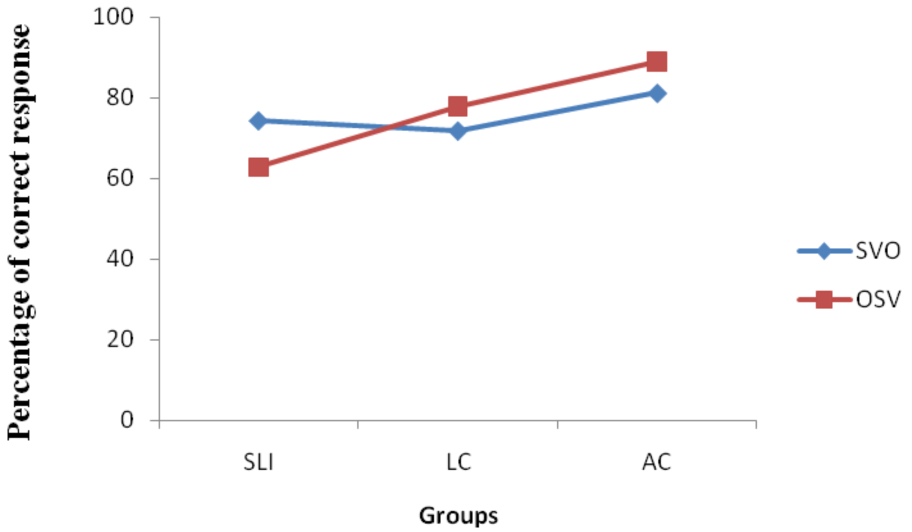
SVO and OSV_{CL} the difference was NOT significant $t(12)=-1.81$($p=.10$).

SVO and $OV_{CL}S$: the difference was significant $t(12) =3.63$ ($p<.01$).

OSV_{CL} and $OV_{CL}S$ the difference was significant $t(12)=5.59$ ($p<.001$).

Despite these apparent quantitative only differences, some qualitative differences are masked by the substantial variations in the SLI group. While these differences suggest some common pattern of performance among these groups, visual inspection of Figures 20 and 21 suggest some contrast may be operating as seen in the cross-over pattern of SVO and OSV sentences, with the group of children with SLI scoring higher than the LC group on SVO sentences, while the LC group scored higher on OSV sentences.

Figure 21: The crossover pattern seen in the performance of children with SLI on SVO and OSV sentences.



To further examine this cross-over, a repeated measure ANOVA of the SLI and LC groups was conducted. Results showed there was a significant word order type*group interaction $F(1,24)=6.16$, $p<.05$, $\eta^2=.20$, despite lack of word order type effect which is caused by the wide variations in the SLI group (almost double that of the typically developing groups). Therefore, this analysis shows that there are qualitative differences between the groups, with the SLI showing the following trend: SVO>OSV_{CL}>$OV_{CL}S$, while both TD groups showing the pattern: OSV_{CL}>SVO>$OV_{CL}S$.

The results of each group on different word order types show that all groups found SVO and OSV_{CL} sentences easier than the $OV_{CL}S$ sentences. The only difference in their profile is that the SLI group scored slightly higher on the SVO than OSV_{CL} sentences, unlike the TD groups who both scored higher on the OSV_{CL} sentences.

4.3.5 Word order type*agreement interaction

Results showed there was a significant word order by agreement interaction, $F(4,144)=35.17$, $p<.001$, $\eta^2=.49$. Children's performance on the sentence comprehension task changed as a function of not only the type of order employed in sentences (SVO, OSV_{CL}, and $OV_{CL}S$), but also of the gender agreement cues. There were three types of agreement cues: AGR1, where subjects and objects had the same gender (either both masculine or both feminine); AGR2, where the subject was masculine and the object feminine, and AGR3, where the subject was feminine and the object masculine. All nouns in the experiment were singular and therefore, no number cues were involved.

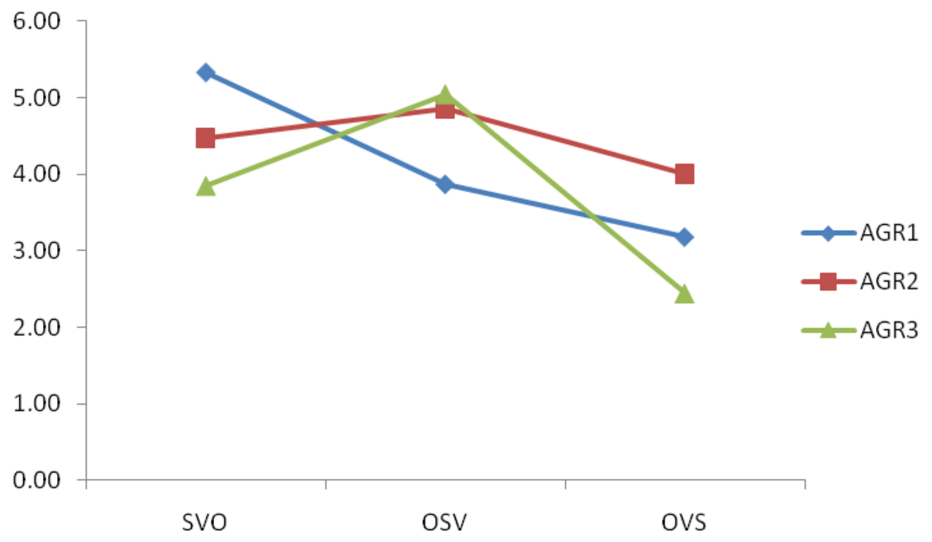
The following table and figure show the participants' overall performance on different word orders and agreement cues.

Table 41: Means and standard deviations of number of correct responses as a function of word order type (SVO, OSV_{CL} , and $OV_{CL}S$ and gender agreement cues (AGR1, AGR2, and AGR3).

Word order type	SVO Mean (SD)	OSV_{CL} Mean (SD)	$OV_{CL}S$ Mean (SD)
AGR1	5.33 (1.03)	3.87 (1.40)	3.18 (1.47)
AGR2	4.46 (0.88)	4.85 (1.37)	4.00 (1.41)
AGR3	3.85 (0.63)	5.05 (1.28)	2.44 (1.35)

Note. AGR1=both subject and object share the same gender. AGR2=masculine subject cues (feminine object), AGR3=feminine subject cue (masculine object).

Figure 22: word order type by agreement interaction



A repeated measure ANOVA was performed to examine the effect of agreement on each type of word order. Table 42 summarises the results of these ANOVAs.

Table 42: Repeated measure ANOVAs for word order type by agreement

Word order type	ANOVA results	Significance	
SVO	$F(2,76)=41.61$	$p<.001$	$\eta^2=.52$
OSV_{CL}	$F(2,76)=27.79$	$p<.001$	$\eta^2=.42$
$OV_{CL}S$	$F(2,76)=22.42$	$p<.001$	$\eta^2=.37$

This table shows that agreement played a role in sentence comprehension in all types of word order. To investigate whether the type of agreement cue had an effect on different word orders, pairwise comparisons with Bonferroni corrections were conducted. They examined the effects of the three different agreement cues: AGR1, AGR2 and AGR3.

In SVO word order: AGR1 was significantly easier than AGR2 (p<.001) and AGR3 (p<.001), while AGR2 was significantly easier than AGR3 (p=.001), which was the most difficult one. Therefore, the order of difficulty from easiest to most difficult gender agreement cues was the following: AGR1>AGR2> AGR3.

In OSV_{CL}: AGR1 was significantly more difficult than AGR2 ($p<.001$) and AGR3 ($p<001$), but there was no significant difference between AGR2 and AGR3. The order of difficulty from easiest to most difficult was: AGR2/AGR3>AGR1.

In $OV_{CL}S$: AGR1 was more difficult than AGR2 ($p=.005$) and AGR2 was easier than AGR3 ($p<.001$), with $OV_{CL}S$ sentences with AGR3 being the most difficult of all types of sentences. The order of difficulty from easiest to most difficult was AGR2>AGR1> AGR3.

Therefore, it seems that the effects of agreement cues differed as a function of the type of word order used. In sentences with N V N configuration (canonical SVO and the most marked $OV_{CL}S$, AGR3 sentences (with feminine agreement) were the most difficult type of agreement, while in OSV_{CL}, masculine and feminine agreement sentences (AGR2 and AGR3) were significantly easier than sentences where both subject and object shared the same gender (AGR1). In both sentences with fronted NP's, the sentences with masculine subject agreement (AGR2) were the easiest ones. Though AGR1 sentences in SVO word order were easier to understand than the rest, children in these sentences did not need to use agreement cues, as correct interpretations could have been reached by linear word order only.

4.3.6 Error Analysis

Children were presented with three distractors: a reversible distractor, an adjectival interpretation, and a random distractor. Table 43 shows the distribution of errors made by all children. It shows that all three groups of children tended to choose the reversible distractor more than the other two.

Table 43: Frequency of types of errors for the three groups.

	Reversible	Adjectival	Distractor
SLI			
No. of errors	237	31	24
%	80.4	11.2	8.5
LC			
No. of errors	173	31	14
%	76.0	16.6	7.4
AC			
No. of errors	101	28	16
%	67.2	21.4	11.4

In order to test whether the distribution of errors is influenced by the group to which children belonged, a Chi square analysis was performed by comparing the frequency of each type errors across groups. Result of the initial Chi-square test show a trend approaching significance $X^2(4)=9.35$, $p=.053$. To further investigate this trend, taking account of observed and expected frequencies, a Chi-square was performed on the reversible and adjectival interpretation interpretations only. Results showed that there was a significant association between group and error type ($X^2(2)=7.04$, $p=.03$. Results showed that children with SLI produced more reversible interpretation errors than expected, while the AC produced less of these errors.
The language control group produced error rates as expected, given the overall distribution of errors. This shows that the SLI group was more likely to resort to use First-subject strategy than the other TD groups.

4.4 Discussion

This study shows that Gulf Arabic speaking children with SLI have difficulty understanding complex sentences that involve fronting of the direct object. While they have a score that is not significantly different from their TD peers on canonical reversible SVO sentences, they face greater difficulties as the complexity of sentences increases in OSV and OVS sentences. These findings are better explained by theories that argue for domain-specific deficits in children with SLI. The analysis of results has shown that grammatical complexity, as defined in terms of the number of movements involved in the noncanonical word orders better explains the pattern of difficulties seen in comprehension of sentences with fronted NP’s in Gulf Arabic. Moreover, this study illustrates that both word order and gender agreement play a role in the comprehension of reversible sentences in Gulf Arabic, though word order might have a more important contribution. These significant deficits in the comprehension of fronted sentences might help in identifying children with SLI in Gulf Arabic and should be considered as a possible clinical marker of SLI in this population. The results demonstrate that noncanonical sentences with feminine agreement cues constitute the most challenging type of reversible sentences for typically and atypically developing children.

4.4.1 Implications about the role of syntactic complexity in SLI

The results of the performance of GA children with SLI on comprehension of reversible sentences with fronted (moved) objects are consistent with most studies in other languages that found that this population has significant difficulties understanding sentences that involve complex linguistic structures, especially those that involve movement leading to noncanonical word orders (Deevy & Leonard, 2004; Marinis & van der Lely, 2007; van der Lely & Battell, 2003; van der Lely & Harris, 1990). Since this is the first study to examine sentence comprehension skills in GA speaking children with SLI, comparison will be made to studies of this task in English, the most widely studied language, and Hebrew, which is a Semitic language like Arabic and both share many syntactic characteristics (Shlonsky, 1997).

The results of this experiment are consistent with findings of studies of complex structures that involve movement of arguments leading to noncanonical words orders as reported in studies of SLI in English speaking children. The results of this study, which examines comprehension of reversible sentences with fronted objects in children aged between 4;10 and 9;4 years, are consistent with van der Lely and Harris’s (1990) findings of comprehension of reversible sentences in children aged between 4;10 and 7;10. van der Lely and Harris (1990) reported that children with SLI scored significantly less than typically developing children on comprehension of passive sentences, while all SLI, age and language matched children had comparable scores on canonical, reversible SVO sentences. Both the current experiment and that of van der Lely and Harris (1990) controlled lexical difficulties of the items used in the test. However, while van der Lely and Harris (1990) included transitive, locative, and dative sentences, the current one included transitive sentences only. The differences in word order typology do not allow for comparison of types of sentences, as Arabic has variable word order that includes SVO, VSO, $OV_{CL}S$ and OSV_{CL}, while English is a strict SVO language with little variation in word order, but it seems that movement associated with noncanonical word order can be challenging for children with SLI in both languages. Bishop (1979; 1982) showed that children with SLI had significant deficits in understanding reversible passive sentences; and children with SLI used the canonical word order (SVO) to interpret these passive sentences. However, she argued that these severe deficits in sentence comprehension were seen only in a minority of children with SLI, who presented with severe comprehension deficits. Overall, despite typological differences, children with SLI in both English and Arabic exhibited limitations in sentence comprehension when presented with structures that involve movement and noncanonical word orders.

The findings of this study are consistent with results of studies of SLI that investigated the comprehension of subject and object relative clauses in Hebrew (Friedmann & Novogrodsky, 2004; 2007). These relative clause structures in Hebrew have varying degrees of difficulty and the pattern observed in the present study is in line with the performance of Hebrew speaking children with SLI on relative clauses. Subject and object relative clauses in Hebrew present an ascending degree of difficulties. While subject relativisation involves a movement operation, using subject first strategy will yield the correct interpretation. Therefore, Hebrew-speaking children with SLI aged eight to 14 years old had good scores when they were compared to language controls, but not in comparison to age controls. This has been the case with SLI children in the present study who have good scores on OSV_{CL} sentences, when compared to language controls, as there are fewer movement operations in this structure. Moreover, this structure has the configuration NNV that is distinct from the canonical SVO (NVN) in Arabic. Similarly, both object relative clauses in Hebrew and $OV_{CL}S$ sentences in Arabic present a higher level of complexity. In $OV_{CL}S$ sentences in Arabic, complexity is caused by the fronting of object and verb and its NVN configuration that is not easily distinguished from the canonical NVN in SVO order, especially in the grammatically vulnerable system of children with SLI. Object relative clauses in Hebrew are complex due to the longer movement of the object and the noncanonical order in object relative clauses. On both object relative clauses in Hebrew and OVS sentences in Gulf Arabic, children with SLI have significantly low scores when their performance is compared to both age and language control groups. Therefore, in both Semitic languages, children with SLI present with an increasing level of challenges as the complexity of sentences increases.

4.4.2 Theoretical Implications

The results of this experiment are more consistent with linguistic accounts of SLI and less compatible with domain-general accounts. In the following, the predictions of each theoretical accounts of SLI are compared to the results obtained in this experiment.

Domain-general accounts of SLI, such as the one formulated in Deevy and Leonard (2004) and Montgomery (1995a; 2002a) propose that limitations in working memory and processing capacity and speed can explain most of the linguistic deficits seen in this population, including the well-documented difficulties in structures that involve syntactic movement. Theories like the one formulated in Deevy and Leonard (2004) predict the following order: SVO>$OV_{CL}S$>OSV_{CL} based on the assumption that both $OV_{CL}S$ and the OSV_{CL} sentences pose similar amount of processing load, though the latter could be more challenging due to the increase in distance between the object the clitic coindexed with it . However, this prediction is not borne out by the results, as these two sentences with fronted NP's were treated differently by the SLI group and the TD groups. There are significant differences between these two word orders as the sentences with the $OV_{CL}S$ word order are significantly more difficult than OSV_{CL} sentences, especially for children with SLI. Children with SLI have an overall mean of 62.8% on OSV_{CL} sentences that involve single movement, indicating they have good performance when they are compared to language controls, though they are significantly poorer than age controls. However, children with SLI scored 38.0% on $OV_{CL}S$ sentences, which shows that they have particular difficulties with these sentences that involve two movements. These differences in the performance of children with SLI on these two types of word orders cannot be explained by processing accounts of SLI, as both are short and consist of two NP's , a verb and a clitic only, hence they are less demanding in terms of working memory. Moreover, both of them are noncanonical and less frequent than canonical SVO word order in Gulf Arabic. Nevertheless the differences in the test results are consistent and substantial.

Domain-specific accounts of SLI predict that both OSV_{CL} and $OV_{CL}S$ sentences would be more challenging than the canonical word order of SVO, especially for children with SLI. They predict, moreover, that OSV_{CL} would be less challenging than the $OV_{CL}S$ because there is less movement in OSV_{CL}.

According to grammatical accounts of SLI (e.g. van der Lely, 2005, Friedmann and Novogrodsky, 2005), in OSV_{CL} sentences, the object is moved to the initial position to put more emphasis on it (topicalisation) and a chain is formed between the object and its trace. The verb in these sentences assigns the thematic role of the theme to the trace of the moved NP, this is followed by another process where the thematic role is transferred via a chain to the moved NP. Therefore, to arrive at the correct interpretation of sentences with moved NP's, the thematic role of the trace must be linked with the moved constituent. In $OV_{CL}S$ sentences, there is an additional movement of the verb to a position higher than the subject and lower than the moved object. This additional movement is expected to increase the level of complexity in the sentence, especially for children with SLI whose grammatical system might be vulnerable to such structural complexity (van der Lely, 2005).

It is noteworthy, that there might be some extra processing demands involved in these complex sentences as they might require some working memory resources, but the argument taken by linguistic accounts of SLI is that grammatical complexity better explains the difficulties seen in this population. In any structure involving hierarchical dependencies, some representations should be held in memory while other information in the sentence is being processed. However, domain-specific accounts and domain-general accounts differ as to where the primary deficits lie. While the former attributes these difficulties to deficits in the grammatical system, the latter argues that general processing mechanisms, such as limited working memory or slowed processing cause these deficits in children with SLI. The results of this experiment show that children with SLI responded in a qualitatively different way when presented with sentences with moved elements. While typically developing children performed as well on canonical SVO as they did on fronted OSV_{CL} sentences (that have one movement), children with SLI did worse on OSV_{CL} than they did on SVO and their performance was significantly lower than their age control group. When the level of complexity increases in $OV_{CL}S$, where there are two elements being fronted, children with SLI show a dramatic decrease in their performance, which was significantly worse than their age and language controls. The findings of this experiment demonstrate that grammatical complexity is a crucial factor in the distinction between children with SLI and their typically developing peers, as children with SLI seem to perform worse as the sentences increase in complexity, a prediction put forward by domain-specific accounts of SLI, such as that of van der Lely (2005).

When children with SLI are faced with this movement complexity and due to limitations in their grammatical system, they resort to the use of the Subject-first (NP-first) strategy, which is one of the common sentence comprehension strategies available for all children. Children with SLI are known to assign the subject role to the first NP they encounter in the sentence even in passive sentences (Bishop, 1992; van der Lely & Harris, 1990). In Gulf Arabic, this strategy can be employed with SVO sentences leading to correct interpretation. When used with OSV and OVS sentences, this strategy can lead to erroneous interpretation by children with SLI. Typically developing children, on the other hand, will also have access to movement strategy, which helps them reach the correct thematic role assignment. The poor performance of children with SLI on $OV_{CL}S$ (NVN) sentences supports the notion that children with SLI over rely on the use of the Subject-first strategy due to deficits in their grammatical system. On OSV (NNV) sentences, these children may rely also on Subject-first strategy as their default strategy in the absence of animacy and world knowledge cues. Their better performance on OSV could be due to use of agreement cues, which is triggered by the presence of NNV configuration in these sentences.

An alternative non-structural (processing) hypothesis that assumes there is no syntactic movement in TD children and those with SLI might argue that the performance of children with SLI is better explained by a non-movement strategy that assumes children with SLI have more difficulties with fronted sentences because these children have access to word order cues only (Subject-first strategy) in these sentences. According to this, SVO sentences are easier because parsing these sentences using Subject first strategy will yield the right interpretation. While both fronted sentences start with the object, therefore they are more difficult to interpret, especially as only word order and agreement cues are available. This hypothesis assumes that these children rely on this ‘default’ strategy because they cannot use other cues, such as agreement cues. However, the results of the experiment show that children with SLI used agreement cues as their performance was not distinguished from typically developing children on agreement cues, as indicated by lack of group*agreement interaction. Moreover, the fact that children with SLI scored significantly better on OSV sentences than they did on OVS sentences shows that they must have used agreement cues in these sentences. Overall, results show that children with SLI, like their TD peers, benefited from agreement cues. Children with SLI use word order cues not because these are the only cues available to them, rather because they had limitations in their grammatical system caused by movement complexity, which lead to them relying on this Subject-first strategy.

In summary, the findings of this study demonstrate that the comprehension deficits seen in children with SLI are caused by movement complexity. They are not caused by limitations in working memory or the use of non-syntactic strategies.

4.4.3 Implications for typical and atypical sentence comprehension in Arabic

This study is probably the first to examine the sentence comprehension abilities of Gulf Arabic speaking children with SLI, and among a very few studies that looked at comprehension skills of typically developing GA children (e.g., Aljenaie & Farghal, 2009).

The results of the present study, where typically developing children perform better on SVO and OSV_{CL} than they do on $OV_{CL}S$ sentences , replicate those obtained in Aljenaie and Farghal (2009), who reported that that SVO word order was the first order to be mastered by Gulf Arabic speaking children aged between four and eight. They showed that $OV_{CL}S$ order (or what they describe as Topic-comment structure) was the most challenging for these children. Similarly, this study reveals that TD age and language controls perform as well as children with SLI on SVO sentences, as all three groups have scores of 70% and above. On OSV_{CL} sentences, both typically developing children score above 70% (77% for LC and 88% for AC), while the SLI group have a score of 62%. The three groups performed significantly worse on $OV_{CL}S$ sentences, though the difference between the TD groups and the SLI group is significant (38% for the SLI group, 57% for the LC, and 64% for AC). Therefore, this study shows that TD children do much better on canonical (SVO) and noncanonical word orders that involve fronting of the object only, namely OSV_{CL} sentences. However, they find the order that involves fronting of both the object and the verb ($OV_{CL}S$) more difficult.

Moreover, the present findings of performance on SVO reversible sentences are consistent with those in Al-Akeel (1998) in Saudi Arabic, a variety that is very close to Gulf Arabic. Al-Akeel (1998) reported that active reversible sentences were mastered by the age of five, where mastery is defined as having a passing criterion score of 60% of more. The present study shows that the three groups of children (SLI, AC, LC) seem to have mastered the comprehension of reversible sentences, as all of them have a score of 70% and above on the SVO sentences. The children in the three groups are aged between 5;10 and 9;4 years old. However, the results of this experiment do not support the conclusion drawn by Al-Akeel about which cues children use when interpreting reversible sentences, which will be discussed in section 4.4.4.

The findings of this study have clinical implications for assessment and treatment of SLI in Gulf Arabic speaking children. The significant difference between children with SLI and TD children on OSV_{CL} and the significant difference between children with SLI and both TD groups on $OV_{CL}S$ have implications for assessment of children with SLI. The children with SLI who participated in this study were matched on a score of a sentence comprehension test with their language controls, yet they exhibited substantial deficits in comprehension of $OV_{CL}S$ sentences when they were compared to these language controls. This illustrates that a general test of comprehension may not be able to differentiate between children with SLI and their TD peers, if not paired with another task that examines in depth some complex linguistic structures, such as comprehension of sentences with fronted NP's in Gulf Arabic. Moreover, the sizeable difference in performance on this complex task, especially the lack of overlap between the group of children with SLI and their TD peers on $OV_{CL}S$ sentences, might indicate that this task could prove a very useful clinical marker of SLI in Gulf Arabic. However, further investigation and replication of these results is highly needed. Moreover, this linguistic structure could be targeted during treatment due to the importance of word order variation in Arabic, which has a highly variable word order.

Overall, this study adds to the scarce literature about the comprehension of Arabic reversible sentences in typically and atypically developing children. The results show that typically developing children (age range 5;10 - 9;4 years old) and children with SLI (age range 6;3 - 9;4 years old) have a good mastery (defined as having a passing criterion of 70%) of SVO word orders, with all groups scoring significantly worse on $OV_{CL}S$ word order, though the group of children with SLI scored significantly worse than the two TD groups on $OV_{CL}S$ sentences and scored worse than age controls on OSV_{CL} sentences. These results suggest that comprehension of sentences with fronted NP's is potentially more useful in differentiating between children with SLI and TD children than a general sentence comprehension test. However, more research is needed to examine the role of complex sentence comprehension in assessment and treatment of children with SLI in Gulf Arabic.

4.4.4 Role of word order and agreement in comprehension

This experiment shows that Gulf Arabic speaking children relied on word order strategy to understand reversible sentences of varying word orders and agreement cues. While there is a clear distinction between the TD groups and the group of children with SLI on sentences with varying word orders, all groups utilised agreement with no significant distinction among them. This indicates that word order plays the most significant role in sentence comprehension in reversible sentences in GA. This is consistent with the findings of Aljenaie and Farghal (2009), who argue that TD children aged between four and eight years old rely on word order when acting out reversible sentences, and only older children (aged 6-8 years old) benefit more from agreement cues. However, it is not possible to compare the results of this study with their findings, as they included seven reversible sentences only and they did not systematically control for pragmatics (world knowledge) and subject and object gender. Aljenaie and Farghal (2009) showed that generally sentences with masculine subjects were easier to understand than sentences with no gender agreement cues. However, it is difficult to draw a conclusive remark based on their findings as the authors did not systematically manipulate pragmatic and gender cues, and they did not have reversible sentences with feminine agreement cues. Similarly, Al-Akeel (1998) argues that TD children in his study relied on lexical knowledge instead of word order when presented with reversible sentences, however this was based on very limited stimuli as he used three sentences only. This study arguably presents more conclusive results about the role of both word order and gender due to the systematic manipulation of the these two variables and the control of other semantic and pragmatic cues.

The word order type by agreement interaction showed that $OV_{CL}S3$ (feminine agreement cue) sentences and $OV_{CL}S1$ (no gender agreement cue) sentences were the most difficult to understand across groups. Hence, it seems that word order strategy may be the first strategy to use in such sentences; especially due to lack of case markers in spoken varieties of Arabic. The increased size effect of word order by agreement interaction (n=.49) indicates that the type of agreement along with word order types account for a large percentage of the performance of typically developing children and those with SLI.

4.5 Summary

This study sheds some light on the performance of typically developing Gulf Arabic speaking children and those with SLI on a comprehension task that involves three different word orders: a canonical SVO and two noncanonical word orders: OSV_{CL} and $OV_{CL}S$. The results show that children with SLI showed quantitative and qualitative differences on this task when their performance is compared to TD age and language matched groups. Their performance was differentially affected and dramatically reduced when they were confronted with increasing levels of grammatical complexity, such as an increase in the number of object and verb movements and a change in canonical word order. They were, however, less affected by agreement, though a combination of marked word order and marked agreement may contribute to the poor performance of children with SLI on this task. These results seem to concur more with a grammatical account of SLI that maintains that these children have a primary deficit in their grammatical system that makes them more vulnerable when presented with complex syntactic structures that involve movement. These results are not congruent with domain-general accounts that attribute these difficulties to problems in working memory or general processing factors. Finally, this study provides some information about the hierarchy of difficulty for some word orders in Gulf Arabic and the task seems to differentiate between SLI and their TD peers better than a general test of sentence comprehension. However, more examination of the role of complex sentence comprehension in assessment and treatment in Gulf Arabic is suggested.

5. Investigating Phonology: Nonword repetition skills in Gulf Arabic speaking children with SLI

5.1 Introduction

The ability to repeat nonwords is considered a very potent predictor of language learning, especially during the early stages of language development (see Gathercole, 2006 for an overview). A vast number of studies have shown that children with SLI have significant problems in nonword repetition (NWR) (Archibald & Gathercole, 2006a; Bishop et al., 1996; Conti-Ramsden, 2003b; Dollaghan & Campbell, 1998; Edwards & Lahey, 1998; Ellis Weismer et al., 2000; Gathercole, 2006; Gathercole & Baddeley, 1990; Gray, 2003a; Montgomery, 1995a, , 2002c; Munson, Edwards, & Beckman, 2005; Oetting & Cleveland, 2006; Roy & Chiat, 2004; Snowling et al., 1991).

Findings of consistent deficits in NWR in children with SLI have led many investigators to consider it as a clinical marker of SLI (Bishop et al., 1996; Botting & Conti-Ramsden, 2001; Dollaghan & Campbell, 1998). Nonword repetition has been found to function as a phenotypic marker of SLI and was linked to genetic factors (Bishop et al. 1996). It is unaffected by dialectal, socio-economic differences (Burt, Holm, & Dodd, 1999; Dollaghan & Campbell, 1998; Engel, Santos, & Gathercole, 2008; Oetting & Cleveland, 2006) or differences in IQ (Bishop et al., 1996; Conti-Ramsden et al., 2001; Ellis Weismer et al., 2000). Therefore, it is considered a good tool for screening and identifying children with language impairment (Ellis Weismer et al., 2000). A good clinical marker is defined by the presence of significant impairment in a certain structure not only when compared to age matched children, but also in comparison to children matched on the same language level (see section 1.5 for more details on criteria to define a clinical marker). A clinical marker with known heritability estimate is preferred since it can shed light on the interaction between genes and environment.(Tager-Flusberg & Cooper, 1999). Together with nonword repetition, tense marking has also been proposed as a good clinical marker of SLI in English speaking children (Bedore & Leonard, 1998; Rice & Wexler, 1996a; Rice et al., 1995) with proven links to heritable language impairments (Bishop et al., 2006).

These consistent impairments in nonword repetition lead researchers to investigate processes that underlie these significant deficits in children with language impairments in general and those with SLI in particular. Some attribute difficulties in NWR to a 'central' deficit in phonological short term memory (PSTM) (Gathercole, 2006; Gathercole & Baddeley, 1990); while others have argued that along with PSTM deficits, there are other contributing factors to NWR deficits, such as deficits in phonological processing skills (Snowling et al., 1991), or phonological complexity (van der Lely, 2005). These two accounts that challenge the PSTM account have demonstrated that NWR tests not only tap into phonological short term memory, but are influenced by multiple factors, such as prosodic factors (syllabic and metrical complexities), wordlikeness, and phonotactic probabilities. Therefore, they question the argument that deficits in PSTM are the central cause of language and nonword deficits in children with SLI.

5.2 How is nonword repetition tested?

One of the contentious issues in the study of NWR is how the design of a particular NWR test influences the results, and therefore the conclusion drawn about the relationship between NWR and PSTM. In the following, four commonly used English NWR tests are reviewed.

The Children's Test of Nonword Repetition (CNRep; Gathercole & Baddeley, 1996*)*. The CNRep was one of the first tests developed to assess nonword repetition and is widely used in UK. It consists of 40 nonwords that are equally divided into 2-5 syllable nonwords. The test uses typical English stress patterns and half of the nonwords contain cluster consonants, while the other half does not include any consonant clusters. However, the CNRep includes many wordlike nonwords and syllables within nonwords that correspond to real words (e.g., <u>under</u>brantuand) as there was no attempt to control for real morphemes, because syllable length was the main variable along which nonwords were created. Therefore, many words have real morphemes in them (e.g., "*defer*mica*tion*") and many words have consonant clusters in various positions (e.g., "blonter*st*aping", "lodder*n*apish", and "ta*fl*est". This test reveals there is a significant effect of syllable length on NWR, a finding that has been replicated in most of the subsequent tests of NWR and thus forms the foundation for the phonological short-term memory (PSTM) account of SLI (Baddeley et al., 1998; Gathercole, 2006).

The Nonword Repetition Test (NRT; Dollaghan & Campbell, 1998). The NRT was developed in the United States where the authors tried to address some of the limitations of the CNRep, such as the presence of lexical and sublexical elements (e.g., words ending with "-ing") in some of the nonwords in the CNRep and the fact that consonant clusters were prevalent in CNRep. Therefore, the NRT consists of 16 nonwords equally divided into 1-4 syllable nonwords. These nonwords contain early developing consonants and vowels (i.e., all English consonants except the late ones: /s, z, l, r,ʃ,ʒ,ð, θ/) and do not contain any clusters (i.e., they have a CV syllable shape) and none of the syllables is a lexical element in English. All the vowels used are tense vowels and the stress pattern used is not typical of the English stress because, unlike English words, the nonwords used in NRT do not contain any weak syllables.

Both the NWT and CNRep are widely used in NWR studies, therefore Archibald and Gathercole (2006) presented a comparison between these two tests. Like CNRep, the NRT was found to distinguish between children with SLI and their age controls with a high level of accuracy. However, only on the CNRep, where there are many nonwords with clusters or sublexical unit, did children with SLI perform significantly less well than their language controls (Archibald & Gathercole, 2006).

The Preschool Repetition Test (PSRep, Roy and Chiat, 2004). While the above mentioned tests are commonly used with children aged 4 years and above, the Preschool Repetition (PSRep) Test developed by Roy and Chiat (2004) is used to measure phonological skills of children between 2 and 4 years. Another factor that is considered in the design of the PSRep, but not the CNRep and the NWR, is prosodic structure. The PSRep consists of 18 real words and 18 nonwords equally divided into 1-3 syllable words/nonwords, with systematic manipulation of stress, so that half of the words have strong/weak stress (SW), while the other half have WS stress. The words and nonwords are phonologically matched, and nonwords are created by alternating the vowel in single syllable words ('lamb' becomes 'lomm' /lɒm/) and reversing two consonants in each word to create a corresponding nonword (e.g., 'dinosaur' becomes 'sinodaur' /sɒinədɔ/). Chiat and Roy (2007) found that this test reliably differentiated between typically developing children and children at risk of language impairment and it had good psychometric properties. They demonstrated that this test was independent of gender and socioeconomic status, though it showed effects of age, prosody, and syllable length (Chiat & Roy, 2007)

The Test of Phonological Structure (TOPhS; van der Lely & Harris, 1999). The TOPhS is designed to account for prosodic variables (such as syllabic and metrical complexity). Therefore, it consists of 96 items organised into four sets of 24 nonwords. In each set, the nonwords vary in complexity from the simplest to most complex in terms of prosodic structure. Syllabic complexity is defined in terms of whether the onset starts with a simplex consonant or consonant cluster, the rhyme ends with a vowel (open) or a consonant (closed), and whether the word ends with a vowel or a consonant. Metrical complexity is defined in terms of the match between the edge of a word and the edge of the foot. In nonwords with unmarked metrical structures, the edge of a foot corresponds to the edge of a word (as in 'city'), while in marked metrical structures the two edges are different with an extra unstressed syllable being at the left edge (e.g., the first syllable /ba/ in '(banana') or right edge (e.g., the last syllable /si/ in 'fantasy'). The properties of the nonwords used in TOPhS are represented in Table 44.

Table 44: Syllabic and metrical parameters used in the TOPhS. Adapted from Gallon et al., (2007), p.440.

Parameter		Description	Real word example	Nonword example
Syllabic				
Onset	Unmarked	No consonant cluster	**c**i.ty	**p**i.fi
	Marked	Consonant Cluster	**pr**e.tty	**pr**i.fi
Rhyme	Unmarked	Open syllable	c**i**.ty	p**i**.fi
	Marked	Closed syllable	f**il**.ter	p**il**.fi
Word-end	Unmarked	Vowel-final	ci.t**y**	pi.f**i**
	Marked	Consonant-Final	si**t**	pi**f**
Metrical				
Left Adjunction	Unmarked	No initial unfooted syllable	**ci**.ty	**kɛ**.tə
	Marked	Initial unfooted syllable	**ba**.na.na	**fə**.kɛ.tə
Right Adjunction	Unmarked	No final unfooted syllable	ci.**ty**	kɛ.**tə**
	Marked	Final unfooted syllable	Ca.na.**da**	kɛ.tə.**lə**

Note. Full stops indicate syllable boundary. Strings in bold indicate the relevant parameter.

5.3 Theoretical issues in NWR

This study will investigate some of the various processes implicated so far in nonword repetition deficits in children with SLI. The influential accounts that will be investigated here are the phonological short-term memory account (Baddeley & Gathercole 1990, Gathercole, 2006) , the Computational Grammatical Complexity account (Gallon et al., Harris, & van der Lely, 2007; van der Lely, 2005) and the phonological processing account (Snowling et al. 1991, Chiat, 2001).

5.3.1 The Phonological Short Term Memory (PSTM) hypothesis

According to the model of working memory described by Baddeley (2003), working memory consists of the following components: the phonological loop, which is responsible for storing verbal-acoustic information; the visual-spatial sketchpad, which retains visual and spatial information; the central executive system, which regulates attention in the working memory and the episodic buffer, which acts as a temporary, limited-capacity storage system that relies heavily on the central executive and combines various information into episodes or chunks (Baddeley, 2003). Baddeley (2003) argues that deficits in the phonological loop component, and especially in the phonological store, are the main cause of language deficits in children with SLI. The phonological loop is responsible for processing and storing novel sound combinations. Deficits in this part of the working memory can cause problems in forming appropriate phonological representations and learning new words (Archibald & Gathercole, 2006a; Baddeley et al., 1998; Gathercole & Baddeley, 1990; Gathercole et al., 1999). Deficits in phonological short-term memory can be reliably assessed using nonword repetitions tasks such as the Children's Test of Nonword Repetition (CNRep) (Gathercole & Baddeley, 1996) or the Nonword Repetition Test (NRT) (Dollaghan & Campbell, 1998). In their seminal study, Gathercole and Baddeley (1990) showed that children with SLI had deficits in their nonword repetition skills, even when compared to younger children matched for reading age, which they attributed to limitations in their phonological short term memory. Children with SLI demonstrated proportionally more difficulty 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, 1990).

Later studies of nonword repetition in children with SLI found evidence supporting Gathercole and Baddeley's (1990) claims that children with SLI have significant deficits in nonword repetition (Bishop et al., 1996; Dollaghan & Campbell, 1998; Ellis Weismer et al., 2000; Montgomery, 1995b; 2004). Bishop et al. (1996) propose the use of nonword repetition tasks to identify children with SLI as they argue that limitations of phonological short-term memory can be a primary phenotypic marker of SLI. It is also argued that deficits in children's ability to retain phonological representations over time could be the underlying cause of some syntactic deficits in children with SLI, such as difficulty assigning anaphora (e.g., reflexives) (Joanisse & Seidenberg, 2003). According to proponents of the phonological short term memory account, these limitations in the phonological store of the phonological loop will constrain long term memory representations and affect other areas of the language learning process, such as syntax (Baddeley et al., 1998). Word learning difficulties caused by limited phonological short-term memory can lead to delay in syntactic development, as words are the building blocks for multiword utterances, which in turn constitute models for abstractions of syntactic rules. Baddeley et al. (1998) reviewed a few studies that found correlation between typically developing children's phonological short-term memory and their grammatical abilities. Adams and Gathercole (1996) found that younger children aged 3-5 years with better phonological working memory abilities produced more detailed and longer utterances when compared to low-phonological short-term memory group. However, no further evidence for a direct relationship between NWR and syntax has been presented. On the other hand, most studies of SLI have demonstrated that there were more significant deficits in syntax and morphology than in vocabulary (Leonard, 1998; Norbury, Tomblin, & Bishop, 2008)), a finding not to be expected if phonological short term memory can equally account for both vocabulary and syntactic deficits in children with SLI. Furthermore, when Bishop and colleagues (2006) examined the behavioural markers of SLI, they found that past tense and NWR had independent genetic causation, therefore dismissing the claim that NWR can act as a single factor explaining the language deficits in children with SLI (Bishop et al., 2006; Norbury, Bishop & Briscoe, 2002). In addition, Norbury et al. (2002) found poor correlation between nonword repetition and verb inflections in 7 to 10 year old children with SLI. Finally, in a longitudinal study Gathercole and colleagues reported that some children with poor nonword repetition skills had within normal range scores on language tests, which showed that having deficits in NWR was not sufficient to cause language impairment in some children (Gathercole, Tiffany, Briscoe, Thorn & The ALSPAC Team, 2005). Therefore, although children with SLI as a group present with significant impairments in their NWR skills, the lack of a strong correlation between NWR and other language components, such as syntax and morphology, and the absence of clear causal relationship between NWR and language impairment led many researchers to conclude that NWR is one among many factors involved in language impairment and an impairment in NWR alone would not necessarily cause disruption in language development (Bishop, 2006).

This conclusion should lead to a revision of understanding not only of the role of NWR in language function and dysfunction but also the nature of this task in general.

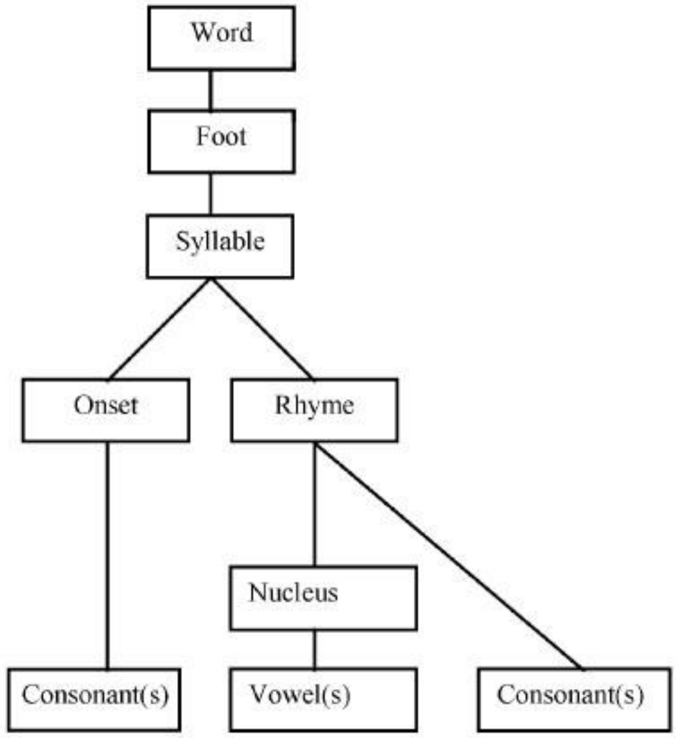
5.3.2 Challenges to the PSTM account of SLI

Though the findings of significant deficits in nonword repetition in children with language impairment, and especially children with SLI, is not controversial, the argument that nonword repetition is a “pure” measurement of phonological short-term memory is debatable. Many studies have shown that there are various processes involved in nonword repetition, such as speech perception, phonological awareness, and output processes (Chiat, 2001, Snowling et al., 1991, Bowey, 2006). Deficits in one or more of these processes, it is argued, might affect children’s performance on nonword repetition tasks. Another component that could be influential is the phonological complexity of the nonwords, as has been suggested in the Computational Grammatical Complexity account of SLI (Gallon et al., 2007; van der Lely, 2005).

5.3.2.1 The phonological complexity account of nonword repetition

Proponents of the Computational Grammatical Complexity account of SLI (CGC) argue that impairments in SLI are not restricted to the well-documented syntactic and morphological domains of the grammatical system, but that the phonological system in children with SLI is prone to disruption when confronted with hierarchical and complex structures (Gallon et al., 2007; Marshall et al., 2002; van der Lely, 2005). In phonology, complexity is defined in terms of complex syllabic and metrical structures, which when combined are called prosodic hierarchy (McCarthy & Prince, 1995; Selkirk, 1980, 1982). According to these theories of prosodic hierarchy, a core component of a phonological word is the foot, which consists of at least one syllable, which in turn consists of an onset, and a rhyme; onset and rhymes are then linked to individual phonemes. Figure 23 illustrates these different components of prosodic hierarchy.

Figure 23: Components of prosodic hierarchy (Gallon et al. p. 437, based on Selkirk (1980; 1982) and McCarthy & Prince (1995)



Marshall and colleagues (Gallon et al., 2007; Marshall, Ebbels, Harris & van der Lely, 2002) used the Test of Phonological Structures (TOPhS) (van der Lely & Harris, 1999) to measure the NWR skills in children with SLI and G-SLI because it was designed to account for prosodic factors (syllabic and metrical parameters), which were manipulated in a systematic way. Marshall et al. (2002) used the TOPhS to study the nonword repetition skills of four children with SLI, aged between 14-18 years old. The segmental parameters were syllable onset, rhyme and word end, while the metrical parameters included stress patterns that involved regular stress and unfooted syllables on the left or right edge of the nonword (see Table 44). They found that segmental and syllable structure errors, such as consonantal substitution and reduction, increased in the presence of marked foot structures. Results showed there were significant effects of foot markedness (both marked left and right adjunction) and syllable number on the accuracy of NWR. Marked structures on the syllabic level (e.g., onset with a consonant cluster or closed rhyme) did not affect repetition accuracy, which could be attributed to mastery of output phonological skills at this advanced age (14-18 years old) or the small size of the group. When examining the individual data of the four children, only one of them showed effects of syllable length (as a measure of PSTM). Moreover, Marshall et al. (2002) and Marshall & van der Lely (2009) showed that children with SLI not only dropped or simplified consonant clusters, but also created clusters in incorrect position, another argument according to Marshall et al. (2002) for the presence of deficits in phonological representations. Moreover, Marshall & van der Lely (2009) showed that children with SLI and dyslexia were more sensitive to stress and cluster position when they were compared to TD children. Therefore, Marshall and colleagues (Marshall et al., 2002 and Marshall & van der Lely, 2009) concluded that these deficits in phonological complexity were not accounted for by phonological short-term memory accounts, such as that of Gathercole and Baddeley (1990). The hypothesis that working memory limitations cause the deficits in nonword repetitions (Bishop et al., 1996; Gathercole & Baddeley, 1990) could not explain the fact that children with SLI had difficulty even with monosyllabic and bisyllabic words when these words had marked prosodic structures (Marshall, Harris, & van der Lely, 2003).

Gallon et al. (2007) studied nonword repetition skills in 13 participants with G-SLI aged between 12-20 years using the TOPhS (van der Lely & Harris, 1999). Each nonword was given a markedness value between 0 (containing no complex structure) and 4 (with four marked parameters). Results showed that children with G-SLI were significantly less accurate on nonwords with 1 or more marked structures, and they were not different from control groups on 0 marked structures. Therefore, Gallon et al. (2007) explained that prosodic complexity represented as marked metrical (stress) patterns (e.g., footed vs. unfooted syllables) or syllabic patterns (e.g., simple vs. complex onsets) posed significant difficulties to children with G-SLI when their performance was compared to age and language matched groups. Analysis of data showed that the number of phonological errors children with G-SLI made correlated with the increasing complexity of the nonwords (Gallon et al. 2007). The reported difficulties with consonant clusters in the children with SLI investigated by Marshall et al. (2002) and Gallon et al. (2007), who were mostly adolescents and young adults, contrasted with Gathercole and Baddeley’s (1990) results where the children with SLI did not have particular difficulty with consonant clusters. Gathercole and Baddeley explained that since the children involved in the study, whose mean chronological age was 8.0 years, had a language age of six years, they were expected not to have difficulty with consonant clusters, which are typically mastered at the age of five (Gathercole & Baddeley, 1989). Therefore, it seems that the age explanation as proposed by Gathercole and Baddeley (1990) does not explain the findings of Marshall et al. (2002) and Gallon et al. (2007).

5.3.2.2 The phonological processing account of NWR deficits

While nonword repetition was initially recommended as a pure measure of phonological short-term memory (Gathercole & Baddeley, 1990), Snowling et al. (1991) were among the first to point to the complexity of nonword repetition. They argued there are various phonological factors involved in nonword repetition, such as phonological awareness skills (e.g., segmentation skills), prosodic structure, articulatory instructions, and perceptual processing. Chiat (2001) explains that impaired phonological processing causes subsequent disruption of the mapping process, which is responsible for establishing word and sentence structures. Therefore, these basic phonological limitations affect lexical and syntactic development. According to this mapping theory, SLI can be conceived of as a deficit in these mapping processes, which constitute the building blocks of language development. This provides a better explanation for the various deficits seen in children with SLI than those offered by the phonological short term memory account or linguistic accounts of SLI that attribute language deficits in children with SLI to a specific grammatical structure (Chiat, 2001). Subsequent studies of some phonological factors found that nonword repetition performance was affected by some of these factors, such as wordlikeness, prosodic (syllabic and metrical) factors, such as stress and the presence of consonant clusters as discussed in the section below.

Wordlikeness. Results from the Preschool Repetition Test (Roy & Chiat, 2004) with 66 typically developing children aged 2-4 years showed that these children were sensitive to lexical familiarity and scored better on words than they did on nonwords. These results were replicated in a larger sample of 315 children (Chiat and Roy, 2007). Furthermore, Gathercole, Willis, Emslie, and Baddeley (1991) found that nonwords rated as being closer to real words were recalled more easily than those rated as less “word-like” on the CNRep test. Along with syllable number, a strong wordlikeness effect was reported by Dollaghan, Biber and Campbell (1993) in their studies of NWR in children with SLI.

Prosodic Factors. Roy and Chiat (2004) and Chiat and Roy (2007) showed that typically developing children aged 2-4 years old were sensitive to prosodic factors, such as stress, when they repeated words and nonwords. They showed that unstressed syllables were more likely to be omitted than stressed syllables, which were rarely dropped and that post-stress syllables were less likely to be omitted than pre-stress syllables. Moreover, Sahlen et al. (1999) found that unstressed syllables in weak-strong syllable combinations were omitted six times more in this position than they were in post-stress positions (strong-weak syllable combinations). Bortolini and Leonard (2000) found that prosodic factors affected children's production of consonants as they found that English-speaking children with SLI dropped more final consonants than TD children and both English and Italian speaking children with SLI omitted word initial weak syllables significantly more than their TD peers.

Another group of prosodic factors include the presence of consonant clusters. Gathercole and Baddeley (1989) studied the nonword repetition skills in 104 typically developing children between the ages of 4-5 years old and found that children at the age of four years were sensitive to the presence of consonant clusters; however by the age of five they were less affected by consonant clusters. Bortolini and Leonard (2000) found that consonant cluster effects and segmental inaccuracies were greater in English and Italian speaking children with SLI than they were in typically developing children matched on consonant inventory and mean length of utterance. They claimed that these limitations in phonological skills were caused by weak phonological representations and not by articulatory deficits (Bortolini & Leonard, 2000).

In addition, proponents of the Computational Grammatical Complexity (CGC) account reported significant difficulties with complex prosodic structures in the performance of children with SLI on TOPhS as discussed in the previous section.

5.4 Cross-linguistic studies of NWR and the nature of NWR deficits

In addition to shedding a light on the role of both phonological complexity and syllable length, investigating nonword repetition skills in children with SLI in Gulf Arabic will add a very important cross-linguistic perspective related to the various processes involved in nonword repetition. Most of the studies that found significant limitations in nonword repetition skills of children with SLI were conducted in English and other European languages (for Italian see Bortolini et al., 2006 and Casalini et al., 2007; Spanish: Gibrau & Schwartz, 2007; Dutch: de Bree, Rispens & Gerrits, 2007; Portuguese: Engel et al., 2008; and Swedish:, Hansson, Forsberg, Löfqvist, Mäki-Torkko, & Sahlén, 2004; Sahlen, Reuterskiold-Wagner, Nettelbladt, & Radeborg, 1999. To the best of my knowledge, the only non- European language that has been investigated to see if children with SLI have deficits in NWR was Cantonese (Stokes, Wong, Fletcher & Leonard , 2006), which came with very interesting findings. Cantonese is a tonal language that is characterised by a very simple syllabic structure (CV only) and limited possible syllabic combinations, with no irregular stress, difficult sounds, or consonant clusters

Stokes and colleagues (Stokes et al., 2006) studied the nonword repetition skills of Cantonese speaking children with SLI, aged between 4;7-5;7 years old and compared them to age controls (AC) and language controls (LC). Both the SLI and AC groups scored significantly better than younger LC group on 1-4 syllable length nonwords (Stokes et al., 2006). The finding of good performance of the SLI group on nonword repetition was used to infer a lack of phonological short-term memory limitations in Cantonese-speaking children with SLI. Their investigation of the repetition of IN syllables (CV combinations that are attested in the language) versus OUT syllables (CV combinations that do not appear in Cantonese) showed that children with SLI's scores on IN syllables were not different from their age controls. There was no significant difference between the two groups on OUT syllables (possibly due to lack of power). Both groups scored better than language controls on both IN and OUT syllables. This good performance on NWR task in Cantonese was explained by the simple structure of the phonological system in Cantonese, where words consist of a small set of CV combinations. Stokes et al. (2006), therefore, suggest that the weaker performance of English speaking children with SLI on nonword repetition tasks could be attributed to their less efficient use of redintegration strategy, which is the use of high probability CV combinations from long-term language knowledge when reconstructing nonwords. Especially, since most English NWR tests include highly predictable sound combinations, giving an advantage for TD vs. SLI children (Stokes et al., 2006).

Studies of NWR skills in other languages can be used to inform researchers of factors that are language specific and others that operate cross-linguistically. For example, findings of good NWR skills in Cantonese speaking children with SLI has shown that the typological properties of a language could have a decisive effect on the performance of children with SLI on NWR skills, and deficits in phonological short term memory or NWR are not necessarily present in children with SLI across all languages.

5.5 The relationship between NWR and other language abilities

Nonword repetition has been found to correlate with various aspects of language learning in both typically developing children and those with language impairments (Botting & Conti-Ramsden, 2001; Dollaghan & Campbell, 1998; Ellis Weismer et al., 2000; Gathercole, 1995; Gathercole & Baddeley, 1989; Montgomery, 1995a; 2000a). Among the various language and learning measures that correlate with NWR are: digit span, receptive and expressive vocabulary, sentence repetition, sentence comprehension and expressive language skills.

Digit span involves repeating single digit numbers of increasing length, while NWR involves repeating nonwords of varying length. However, they differ in that NWR does not use prior lexical knowledge. Both digit span and NWR involve retention of phonological information in short term memory and hence it is expected that they show strong correlation. In cross-sectional and longitudinal studies, Gathercole and colleagues found strong correlation between nonword repetition and digit span in children between the age of 4 and 9 years old (Gathercole & Baddeley, 1994; Gathercole, Willis, Baddeley & Emslie, 1994; Gathercole, Willis, & Baddeley, 1991; Gathercole et al., 1992).

Most studies of NWR in English found a strong correlation between NWR and vocabulary learning (Gathercole & Baddeley, 1989; Gathercole et al., 1992, see also Gathercole, 2006 for a review). It is argued that NWR resembles the process of novel word learning in early childhood. When children encounters a new word, they hold the phonological form of this new word in their phonological short term memory and later on this novel word is committed to stable long term memory representations. However, Snowling et al. (1991) explained that this correlation can be turned the other way around, i.e. vocabulary knowledge facilitates the learning of novel sound combinations. This may explain why words are easier to recall than nonwords (Gathercole, Pickering, Hall & Peaker, 2001). While most of these studies found a correlation with receptive vocabulary, Edwards and Lahey (1998) found that nonword repetition correlated with expressive vocabulary, but not with receptive vocabulary in a group of children with SLI.

There are few studies that examined the correlation between nonword repetition and sentence repetition and most of them reported strong association between the two tasks (Bishop et al., 1996; Bishop, 1999; Conti-Ramsden et al., 2001; Kamhi & Catts, 1986). This lead some to suggest that limitations in short term memory can explain the poor performance of children with SLI on the two tasks (Conti-Ramsden et al., 2001). However, Stokes et al. (2006), in their study of nonword repetition and sentence repetition in Cantonese speaking children with SLI, reported no significant correlation between nonword repetition and sentence repetition. In Cantonese, sentence repetition, but not nonword repetition, was a clinical marker of SLI as Cantonese speaking children with SLI performed significantly worse than their typically developing peers on the sentence repetition task, but not on NWR. This may indicate that sentence repetition seems to tap linguistic abilities as well as short-term memory; therefore, it frequently appears as a sensitive clinical marker of SLI (Bishop, 1996; 1999).

Another language measure that commonly correlates with NWR is sentence comprehension. Bishop et al. (1999) reported a strong level of correlation ($r=.37$, $p<.001$) between nonword repetition as measured by CNRep and the Test of Receptive Grammar (TROG, Bishop, 1989) in the performance of 280 twin children between the age of 7 and 13 years. Montgomery (1995) studied the relationship between sentence comprehension and nonword repetition in 14 children with SLI aged between 6 and 11 years old. Results showed there was a strong correlation (r=.62) between the two and he concluded that "a capacity limitation in phonological working memory compromised the comprehension efforts of the children with SLI" (p.194), i.e. their limited capacity to store speech material impacted upon their ability to form sentence representations (Montgomery, 1995; 2003). Similarly, significant correlation was reported between nonword repetition score and the receptive subtest of CELF (Semel et al, 1987) (r=.39, P<.01) (Montgomery, 2007) and between nonword repetition and simple sentence comprehension (r=.71, p<.01) (Montgomery & Evans, 2009).

In terms of correlation between NWR and expressive language skills, Botting and Conti-Ramsden (2001) found a strong association between nonword repetition and measures of productive grammatical abilities (such as third person singular -s and past tense). They, therefore, concluded that nonword repetition was closely related to expressive language skills, not only through phonological output, but also through more complex language skills. Montgomery (2007) found a significant correlation (r=.39, p<.01) between nonword repetition and the Expressive Language score of the Clinical Evaluation of Language Fundamental test (CELF-R) (Semel et al, 1987). Montgomery (2007) explains this correlation in that both NWR and language measures (receptive and expressive subtests of CELF) rely on linguistic knowledge as well as on phonological capacity.

This section shows that NWR skills in children with SLI correlate with various measures of linguistic abilities, such as receptive and expressive vocabulary, sentence comprehension, sentence repetition, and expressive language abilities.

5.6 Error patterns in NWR tasks

The few studies that investigated the error patterns of NWR in children with SLI found no qualitative differences in error patterns between children with SLI and their typically developing peers (Edwards & Lahey, 1998; Montgomery, 1995a). These studies found that children with SLI and TD children produced more substitution errors than omission errors, i.e., most of the errors were in the segmental level of the nonwords. Errors in the metrical level have not been prevalent in both

SLI and TD children as both seemed to preserve the number of syllables in each nonword, despite that fact that children with SLI exhibit more segmental errors than their TD peers (Dollaghan, Biber & Campbell 1995; Edwards & Lahey 1998; Roy & Chiat, 2004; Sahlén et al., 1999). Therefore, Marton (2006) argued that there are two separate levels of processing in nonword repetition: a level where the individual phonemes are processed and another level where the metrical frame is processed. Marton (2006), therefore, argued that children with SLI face great challenges processing these two levels simultaneously and this might explain the differential pattern of performance on segmental vs. metrical levels, where the former poses more difficulties than the latter. Marton (2006) suggested that these deficits in simultaneous processes could be caused by deficits in switching attention between these two tasks. Further evidence for the argument that children with SLI have deficits in simultaneous processing of segmental and metrical information was reported in Marton and Schwartz (2003). They explained that the although the overall error patterns in SLI and TD groups were similar, with substitution errors being more frequent than other types of errors and with errors increasing as the number of syllables increased, there was one qualitative difference between children with SLI and TD groups. This was related to the proportion of single errors vs. multiple errors within the same nonword. Children with SLI tended to produce more multiple errors than their TD peers without changes in the pattern of segmental vs. metrical errors. They also showed that while in TD children the proportion of multiple errors did not change relative to the increase in syllable length, children with SLI exhibited more multiple errors as the number of syllables increased (Marton & Schwartz, 2003). This, they argued, showed that children with SLI were able to preserve the metrical frame of nonwords, but could not add the phonemes to this level because of difficulties in the simultaneous processing of these two levels (Marton, 2006). However, this differential error pattern has not been reported by other studies.

In summary, there are large number of studies that investigated the nonword repetition skills of children with SLI in English and other languages. The vast majority of them showed that this population has persistent problems on this task. Though these results were used to indicate the presence of phonological short-term memory deficits in children with SLI, there are many confounding variables involved in the tests designed to measure NWR skills, such as wordlikeness, prosodic factors, the presence of consonant clusters, and CV combinations. These variables should be systematically examined and controlled for before suggesting a direct link between NWR and phonological short term memory skills in children with SLI. Moreover, the findings of within normal performance of Cantonese children with SLI on NWR tasks indicate that there also language specific factors that influence NWR, which is not expected if only deficits in phonological working memory operate in these children.

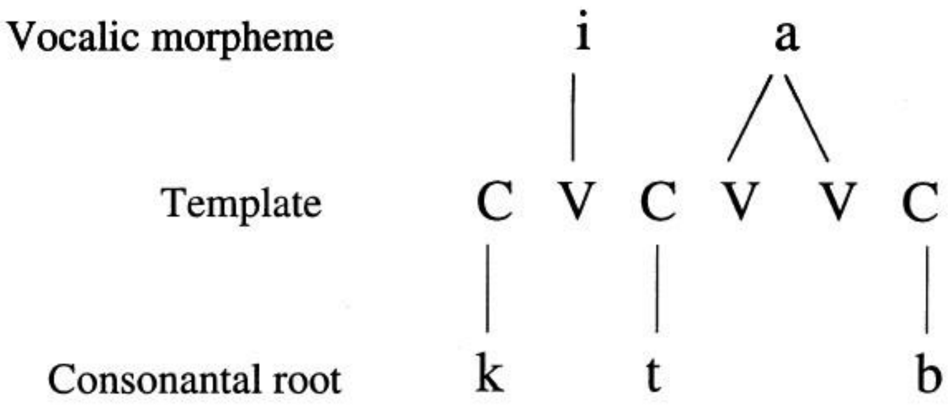
5.7 Developing a Gulf Arabic nonword repetition test

The aim of this chapter is to investigate the performance of Gulf Arabic speaking children with SLI on a nonword repetition test, the first time that such a task has been carried out with Gulf Arabic-speaking children. Therefore, a nonword repetition test was specifically designed to serve this purpose. Investigating this skill in this particular population will reveal if nonword repetition can act as a clinical marker of SLI as has been the case in many other languages.

5.7.1 The phonology of Gulf Arabic

To the best of my knowledge, there is no existing nonword repetition test for Arabic-speaking children, regardless of which dialect or variety of Arabic is used. Arabic poses some challenging issues to those who want to create a nonword repetition test. Arabic, like other Semitic languages, is a nonconcatenative language where words consist of intertwined roots and templates or patterns (McCarthy, 1982). Roots consist of three or four consonants that carry the basic semantic meaning. For example, the root ‘K T B’ (write) is used to derive words, such as ‘kitaab’ (book), ‘maktabe’ (library), ‘kaatɪb’ (writer)…etc. Vowels are inserted between these roots to form various words. Figure 24 depicts the representation of the word ‘kitaab’ (book) (see Béland & Mimouni, 2001).

Figure 24: Vocalic melody, template, and consonantal root of the word 'kitaab' (book), Béland and Mimouni, 2001, p.84.



In their study of a case of deep dyslexia in a Lebanese–Arabic/French bilingual patient with aphasia, Béland and Mimouni (2001) created a list of nonwords by replacing one consonant in existing words with a minimal pair. However, this will not be appropriate for use with children with language impairment as an effort was made to minimise wordlikeness effects.

5.7.1.1 The sound system of Qatari Gulf Arabic

Qatari Gulf Arabic has 30 consonants and 8 vowels as shown in Table 45 and Table 46.

Table 45: Qatari Consonants (adapted from Bukshaisha, 1985).

	Bi-labial	Labio-dental	Inter-dental	Alveo-lar	Palato-alveolar	Palatal	Velar	Uvular	Pharyn-geal	Glottal
Stop	b			t d tˤ			k g	q		ʔ
Affricate					tʃ dʒ					
Fricative		f	θ ð ðˤ	s z sˤ	ʃ			χ ʁ	ħ	h
Nasal	m			n						
Lateral				l lˤ						
Trill				r						
Approxi-mate						j	w		ʕ	

Table 46: The vowel system of Qatari Gulf Arabic (Mustafawi, 2006)

i:			u:
i			u
	e:		o:
		a	ɑ:

Bukshaisha (1985) listed 12 types of syllables in Qatari Gulf Arabic; ten of them are common, while the other two are not. The 10 common types are /cv/, /cv:/, /ccv/, /ccv:/, /cvc/, /cv:c/, /cvcc/, /ccvc/, /cv:cc/ and /ccv:c/.

Stress in Gulf Arabic. Like many other varieties of Arabic, stress in Gulf Arabic is regular and depends on syllable weight. The final syllable is stressed if it has a long vowel /cv:/ or consonant cluster (cvcc), including geminate consonants, otherwise stress falls on the penultimate syllable (Hole, 1989). This means that syllables with consonant clusters will carry the main stress and therefore, it was not possible to manipulate stress independently from consonant clusters in the current study. While tests like TOPhS can assess the influence of syllabic and metrical complexity, only syllabic complexity is manipulated in the NWR designed for this experiment.

5.7.2 Variables considered in the design of Arabic nonword repetition test

The design of nonword repetition tests and the stimuli chosen can affect children's performance on the test, as Archibald and Gathercole (2006) showed when they compared the performance of three groups of children on two of the most common nonword repetition tests in English, namely the Children Nonword Repetition test (CNRep) (Gathercole & Baddeley, 1990) and the Nonword Repetition Test (NRT) (Campbell & Dollaghan, 1998). Since the main objective for creating the Arabic nonword repetition test is to compare the effects of phonological storage (as measured by syllable length) and the effects of phonological complexity (as measured by consonant clusters), careful consideration was taken to control other variables that have been found to influence NWR. These include: articulatory complexity (output processes), lexicality effects, respecting phonotactic rules of Arabic, morphological information, syllable number, and wordlikeness.

Articulatory complexity. In order to control for articulatory complexity, all consonants chosen to form the stimuli were early developing consonants, except /s/ which was included as it is commonly found in many clusters across languages. Since there are no normative studies of phonological acquisition in GA, the consonants were chosen based on their emergence across languages and based on clinical experience of the investigator as a GA speaking speech-language therapist. Therefore, while Qatari Gulf Arabic has 30 consonants, only nine consonants were selected to form the nonwords. These sounds are /b/, /d/, /t/, /k/, /f/, /s/, /m/, /n/, and /l/. According to Bukshaisha (1985), all these consonants can occur in any position in Gulf Arabic words. Moreover, following the recommendation of Dollaghan and Campbell (1998), only tense vowels were chosen. Therefore, short tense vowels (a,u,ɪ) which are common in Standard Arabic and most spoken varieties of Arabic were employed to form the nonwords. No diphthongs or long vowels were included.

Lexicality effects. In order to reduce lexicality effects and neutralise the influence of previous vocabulary knowledge, an effort was made to minimise the number of syllables that are actual words inside the nonwords. Due to the design of the test and the limited number of consonants, it was difficult to eliminate all syllables that can be actual words. Therefore, out of the 140 syllables, 18 were possible words (i.e., 12.9% of the total number of nonwords). However, many of these syllables are words that may not be in the lexicon of these children at this time (e.g., /kɜd/ (“worked hard”), /mɜs/ (“touched”). As for sublexical effects, which are related to phonotactic probability of phoneme sequences (see Gathercole, 1999; 2006; Stokes et al. 2006), no database is available that lists consonant probabilities in Gulf Arabic and therefore it was difficult to determine the influence of phonotactic probability on the performance of children with SLI and typically developing children in this experiment.

Language-specific phonotactic rules. Alongside attempts to control articulatory complexity and lexicality effects, nonwords obeyed the phonotactic rules of Arabic. Therefore, no words with initial clusters were included, because most of these initial clusters are formed by shortening a vowel and then deleting it to form a consonant cluster, e.g. /ħɪsˤa:n/ 'horse' to /ħsˤa:n/ and sometimes a short vowel is introduced in front of initial consonant cluster /ɪħsˤa:n/ (Bukshaisha, 1985). To avoid this controversy of whether there is an initial cluster or not, none of the nonwords composed started with an initial cluster. Language-specific phonotactic rules were respected when forming the trilateral nonroots; therefore both the Obligatory Contour Principle on place of articulation (OCP-Place) and sonority principle were respected. The OCP-Place states that roots with homorganic consonants are extremely disfavoured or rare in Arabic (Frisch, Pierrehumbert & Broe, 2004; Frisch & Zawaydeh, 2001; Greenberg, 1950; McCarthy, 1986). Therefore, consonants that are produced at the same place of articulation are not found in proximity to each other. Pierrehumbert (1993) listed the following categories as major cooccurrence classes: labials, coronal obstruents, velars, gutturals, and coronal sonorants. For example, the cooccurrence of labial consonants (b,f,m) in the same root is infrequent compared to other consonants. Based on these phonotactic constraints and the consonants chosen, the following seven roots were selected: /S T L/, /K D F/, / D L S/, /S B N/, /D N F/, /K M S/, /D F L/, all these roots are nonexistent in Gulf Arabic. These roots were checked in Gulf Arabic dictionaries (Holes, 2000; Qafisheh, 1997) and their nonexistence was confirmed. When consulting the biggest dictionary of Classical Arabic compiled in thirteenth century (Ibn Manzur, 1290[1981]) , two of these roots were found, namely /D L S/ and /D N F/. Ten college-educated teachers were given these seven roots in the common a-a vocalic pattern (e.g., /dalas/, /kadaf/…etc). Two out of these ten teachers identified /D L S/ and knew its meaning, while the rest did not identify the meaning of any of the roots, though they recognised that they could be possible Arabic words. Therefore, it is very unlikely that any of the children in the study had encountered any of these two roots from Classical Arabic.

Morphological information. Another language-specific factor that was controlled was accessing morphological information. In Arabic, grammatical morphemes are affixed initially, medially, or finally to the root and hence the nonwords were carefully selected to avoid including such morphemes. Therefore, none of the nonwords started with /b/ (a preposition in Arabic as in ‘bi’ (in)); /f/ (a conjunction as in ‘fa’ (and) or preposition as in /fi/ ‘in’; /l/ (as in ‘li’ / (for), /n/ (a pronoun as in ‘naʔkul’ (we eat), /t/ (a feminine third person pronoun (she), and /m/, which is commonly used to derive nouns, places…etc. Possible suffixes, such as /m/ (used in plural third person pronouns), /k/ (2nd person pronoun), and /t/ (used in feminine pronouns and to indicate past tense) were avoided. Hence, all nonwords included in the test ended with the following consonants only: /b/, /f/, /l/, /n/, /s/, and /d/. Moreover, since Arabic has some infixes, the following consonants were not used in infix positions: /t/, /s/, and /n/.

Syllable number. Gulf Arabic can have up to seven-syllabic words (Bukshaisha, 1985), however most of the words longer than 3 syllables are formed by adding inflectional morphemes, therefore all the nonwords included in the task were either two or three syllables. Due to the root-and-pattern nature of the language, it was not possible to create monosyllabic nonwords that are phonotactically possible. Many studies found that differentiation in performance of children with SLI starts on three syllable words and upwards (Dollaghan & Campbell, 1998, Montgomery, 2004; Gathercole & Baddeley, 1990). Therefore syllable number effects can be examined by comparing the performance of children on two versus three syllable nonwords.

Wordlikeness. Since Arabic is a root-and-pattern language, triliteral roots cannot exist by themselves and need a pattern of vocalic sounds. However, patterns should respect phonotactic rules of Arabic and therefore using a non-existing pattern will violate these rules. Therefore, I opted to use patterns that are infrequent in Gulf Arabic. The following vocalic patterns were employed to generate the experimental nonwords in combination with the consonantal roots: a-u, a-u-a, u-ɪ, and u-ɪ-a. The vocalic pattern /a-u/ exists in some nouns (e.g., /tamur/ (dates), while the pattern /u-ɪ/ is used in Standard Arabic to form passive voices (e.g., /kusir/ ‘was broken’). Both these patterns are less common in Gulf Arabic. The test includes eight control nonwords, which have one of the most frequent vocalic patterns in Arabic, namely /a-a/ and /a-a-a/ (Holes, 2000). Therefore, the experimental nonwords are characterised by their low wordlikeness effect, while the control nonwords have higher wordlikeness effects as they have one of the most common vocalic patterns in Arabic.

5.8 Aims and predictions of the Arabic NWR task

The performance of children with SLI will be analysed and compared with typically developing children to see whether GA speaking children with SLI have significant impairments in nonword repetition skills, and whether NWR can serve as a clinical marker of SLI in Gulf Arabic. It is hoped that the results of this experiment will help explain how various processes contribute to nonword repetition, with special emphasis on comparing and contrasting the role of two components in NWR, namely syllable length, as a measure of phonological short-term memory, and consonant clusters, as a representative of phonological complexity. The results of the Arabic NWR test will be compared with results obtained from other languages in order to elucidate how various processes interact with language specific properties leading to deficits in NWR. Furthermore, this study will examine the correlation of NWR scores with various language measures, such as receptive and expressive vocabulary, sentence comprehension, sentence repetition and expressive language skills.

5.8.1 Clinical viability of NWR in Arabic

This experiment examines the performance of Gulf Arabic speaking children with SLI on a novel Arabic nonword repetition test in order to see if NWR is impaired in children with SLI in comparison to typically developing children. If this is the case, its viability as a clinical marker of SLI will be examined, especially if children with SLI exhibit significant difficulties on the test, not only when compared to age controls, but also in comparison with language controls. Nonword repetition has been found free of gender and socioeconomic factors and the stimuli chosen were selected to be appropriate with various dialects of Arabic, making it suitable for use in different varieties of Arabic and with children from different socioeconomic backgrounds. NWR is also characterised by easy administration and it is less time consuming than traditional language tests, which are already scarce in Arabic, making it an ideal tool for clinicians to assess and identify children at risk of language impairment.

5.8.2 Examining the predictions of competing theories of the nature of NWR deficits in SLI

The phonological short term memory (PSTM) account of SLI (Gathercole & Baddeley, 1990; Gathercole, 2006) predicts that children with SLI will have a major deficit in nonword syllables that are three syllables and longer and that they will not be significantly different from TD groups on nonwords with two syllables, as many studies found that divergence in performance starts at 3 syllables and above. Moreover, the PSTM account predicts that there will be a significant group by syllable length interaction, with the SLI group differentially affected as the number of syllables increases from two to three syllables.

Accounts such as the computational grammatical complexity (CGC) (van der Lely, 2005) and other accounts that argue for deficits in phonological skills of children with SLI (e.g., Snowling et al., 1991) predict that apart from the presence of syllable number effects, syllabic complexity, represented in this experiment in terms of consonant clusters, will differentiate between the performance of these groups. Therefore, they expect to see group by cluster interaction with the SLI group performing significantly less accurately on nonwords with consonant clusters. The CGC account predicts that as the number of consonant clusters increases, children with SLI will present with more difficulties in repeating these nonwords. The phonological processing account (Snowling et al., 1991; Chiat, 2001) predicts that children with SLI will be affected by articulatory processes and wordlikeness, since both of these processes are affected by basic phonological skills.

5.8.3 Comparing the results of the Arabic NWR with cross-linguistic findings

Comparing and contrasting the results of the Arabic NWR test with those reported in other languages can be useful in explaining the role of various psycholinguistic factors in nonword repetition test and the interaction among these factors. While NWR was initially thought to be free of language-specific effects, the fact that some children with SLI have exhibited within normal performance on NWR (e.g., Stokes et al., 2006) indicates that there are language specific properties that should be considered when examining NWR. Far from being a 'pure' measure of phonological short-term memory, NWR can be considered as a multidimensional process that involves a myriad of underlying processes and skills, such as phonological memory, wordlikeness, syllabic and metrical complexity, lexical and sublexical processes, phonotactic probabilities, and articulatory and output processes. It is conceivable that these various processes have different influences according to the properties of each language. Therefore, this study aims to assess the role of phonological short term memory, syllabic complexity, wordlikeness, and articulatory processes in the performance of Gulf Arabic speaking children with SLI and compare findings with other languages (e.g., English and Cantonese) in order to evaluate the roles of different processes across languages.

5.8.4 How NWR in Arabic correlates with other psycholinguistic abilities

This experiment will examine the correlation between nonword repetition and various measures, such as digit span, nonverbal IQ, sentence comprehension, expressive language skills, lexical development, and sentence repetition. Since this will be the first investigation of nonword repetition in both typically developing children and children with SLI in GA, the correlation between performance on nonword repetition and the various tests used in this project will be investigated. These tests measure various aspects of language competence, such as, sentence comprehension, expressive language skills, sentence repetition, and receptive vocabulary. Moreover, correlation between digit span and nonword repetition will be reported.

5.9 Method

5.9.1 Participants

Thirty three Gulf Arabic speaking children participated in this experiment; 11 diagnosed with SLI (Age_M =7;8 [years;months]), 11 typically developing age control (AC) children (Age_M =7;8), and 11 TD language control (AC) children (Age_M =5;8), who were matched with the SLI group based on their scores on the Gulf Arabic sentence comprehension test. The sentence comprehension test involves asking the child to listen to a sentence and point to the right answer in an array of four pictures. The test consists of 40 sentences and was found to have satisfactory levels of reliability and validity (for more details of the properties of this test, see chapter 3). Children were recruited from two kindergartens, four primary schools, and three children were recruited through personal acquaintance. All the participating children live in Doha, the capital of Qatar, and come from Qatari Gulf Arabic speaking households. They all received general language tests and children with SLI were selected based on the same criteria used in the syntactic experiment (see previous chapter), i.e., all of them scored -1.5 standard deviation (SD) or below on at least two out of the four language tests or -2.0 SD on one of these tests. The tests included the Sentence Comprehension test, the Expressive Language test, the Sentence Repetition test, and the Arabic Picture Vocabulary Test. All children scored within normal range on either the Test of Nonverbal Intelligene-3 (TONI-3; Brown et al., 1997) for children aged 6;0 and above or for children less than 6 years old, two performance subtests from Wechsler Preschool and Primary Scale of Intelligence (Wechsler, 2002) were conducted. The Wechsler Performance IQ subtests were the Block Design and the Picture Completion subtests. These were recommended as an appropriate short form of nonverbal IQ (see LoBello, 1991; Tomblin et al., 1997). Both the SLI and AC groups consisted of 8 boys and 3 girls, while the LC group had 6 boys and 5 girls. Table 47 summarises participants' characteristics.

Table 47: Descriptive summary data for the children with SLI (SLI; *n*=11), age control group (AC; *n*=11), and language control group (LC; *n*=11).

Group	Age (months)	SC	EL	SR	APVT	DS	TONI-3
SLI							
M	93.9	25.0	30.2	55.9	48.5	8.1	93.0
SD	10.5	4.5	8.7	15.8	15.6	1.2	7.4
Range	75-109	18-31	16-42	29-76	28-76	6-11	85-109
LC							
M	72.3	26.2	44.9	74.7	59.2	9.1	a
SD	8.3	2.9	6.4	12.5	16.6	2.9	a
Range	60-83	22-31	38-54	55-89	38-89	5-14	a
AC							
M	93.8	33.1	51.5	85.4	71.7	11.3	99.5
SD	10.7	2.3	6.8	9.3	19.9	1.3	7.8
Range	75-108	30-38	41-60	71-101	47-110	9-14	88-111

Note. SC=Sentence Comprehension Test raw score; EL=Expressive Language Test raw score; SR=Sentence Repetition Test raw score; APVT=Arabic Picture Vocabulary Test raw score; DS=Digit Span task from the Wechsler Preschool and Primary Scale of Intelligence (Wechsler, 2002) raw score; TONI-3=Test of Nonverbal Intelligence-3 standard score.
[a] All the language control children scored above the cut-off score of 16 on the shorter version of Wechsler Performance IQ (see LoBello, 1991 and Tomblin et al., 1997).

There was a good matching between the SLI group and the language controls that were selected based on their score on the Sentence Comprehension test. The SLI group was not significantly different from the LC group on the digit span subtest score (which included forward and backward digit span) of the Wechsler Preschool and Primary Scale of Intelligence (Wechsler, 2002). These two groups were not significantly different on the Arabic Picture Vocabulary Test (APVT), a test of receptive vocabulary. A one-way ANOVA for the scores of the groups on APVT revealed a significant difference: $F(2,30)=4.8$, $p,<.05$, post hoc test with Bonferroni correction showed that SLI scored significantly less than the AC group, but not the LC group. The LC group was not significantly different from the AC group. This was the case in the digit span task, $F(2,30)=6.8$, $p<.01$, with post hoc test with Bonferroni correction showing that the SLI group scored significantly less than the AC group ($p<.05$), but not the LC group. The LC group was not significantly different from the AC group. The difference on the nonverbal IQ scores between the SLI and AC groups was not significant $t(18)= -1.9$, $p=.07$.

5.9.2 Materials and Procedure

The nonword repetition test consisted of 56 nonwords: 48 experimental nonwords and 8 control nonwords. The experimental stimuli contained 6 nonexistent triconsonantal roots that do not appear in the Gulf Arabic lexicon and they were used to construct two and three syllable nonwords with four types of cluster conditions (No cluster, medial cluster, final cluster, and medial and final (M+F) clusters), so each root was used to construct 8 nonwords. The vocalic patterns used with these roots were existing but infrequent patterns in Gulf Arabic. See Table 48 for an example for one of the 6 triconsonantal roots.

Table 48: An example of a root and vocalic patterns used to create a list of two and three syllable nonwords. For a full list of all nonwords, see Appendix D.

Root	No. of syllables	Pattern	Syllable Type			
			No Cluster	Medial Cluster	Final Cluster	M+F Cluster
STL	2 syllables	a-u	Sa.tul	Das.tul	Sa.tulb	Das.tulb
	3 syllables	a-u-a	Da.su.tal	Das.bu.tal	Da.su.talb	Da.sum.talb

Note. Full stops indicate syllable boundary.

The eight control nonwords were created by taking a nonexistent root and using the same types of clusters, however the vocalic patterns used were (‘a-a’ and ‘a-a-a’), which are considered two of the most frequent vocalic patterns in Arabic.

These 56 nonwords were recorded by a female native speaker of Qatari Gulf Arabic. Recording of stimuli was conducted in a soundproof room using the Computerised Speech Lab (CSL 4300, Kay Elemetrics). The stimuli were then randomised and put into two lists and children were assigned randomly to one of these nonword lists (see Appendix E for list A & List B).

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 four trial items. Stimuli were presented from a laptop through a pair of external speakers. Children's productions were audiotaped through a Sony microphone attached to the laptop and using PRAAT software (Boersma & Weenink, 2004). Children's responses were transcribed online by the examiner. Each repetition was scored either correct (1) or incorrect (0). Minor misarticulations (especially distortion of /s/ or substituting /θ/ for /s/) were counted as correct. One typically developing child, aged 5 years old, was not included as his volume was very low and he did not respond to requests to increase his volume; his data were not subsequently included or analysed. There were no similar incidents with all other children as most of them enjoyed the experiment and found it amusing. No repetition of the stimuli was allowed.

5.10 Results and analysis

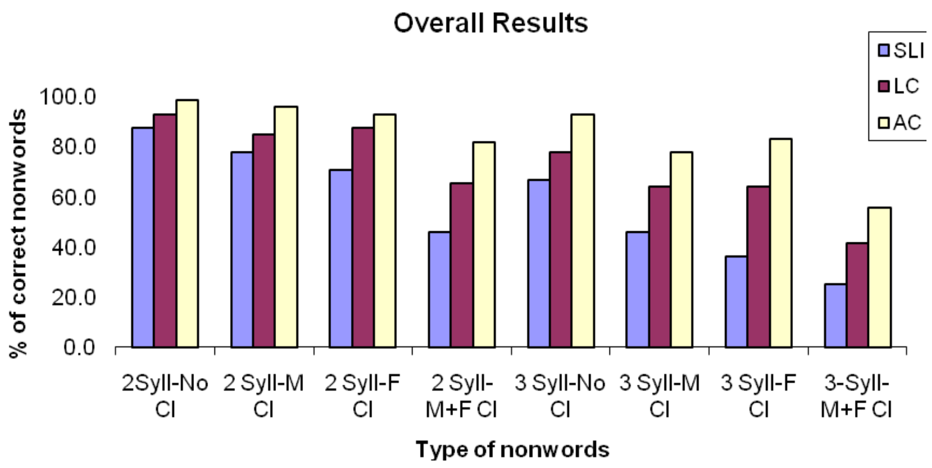
Nonword repetition accuracy was scored at word level, so each word received a score of 1 (correct) or 0 (incorrect). Raw scores were then converted to percentages. Table 49 shows the percentage of correctly recalled words for all groups. It is evident from Table 49 and Figure 25 that the SLI group found the nonword repetition harder than both control groups, especially as the number of marked structures increases.

Table 49: Group descriptive statistics (in percentages of correct repetitions) for the children with SLI (*n*=11), AC children (*n*=11), and LC children (*n*=11) in the nonword repetition (NWR) task.

	2 Syllables					3 syllables				
Group	Cluster Condition					Cluster Condition				
	0 Cl	M Cl	F Cl	M+F Cl	All 2 syll.	0 Cl	M Cl	F Cl	M+F Cl	All 3 syll.
SLI										
M	86.3	77.2	68.2	42.5	68.6	63.6	43.9	30.3	21.3	39.8
SD	16.4	23.9	21.7	25.1	26.9	20.9	27.2	25.6	22.5	28.4
Range	50-100	33-100	33-100	17-100	17-100	33-100	0-100	0-83	0-67	0-100
LC										
M	92.4	87.8	89.4	71.2	85.2	78.8	69.7	68.2	45.6	65.5
SD	11.5	10.8	17.1	31.7	20.7	15.0	31.5	22.9	22.4	25.9
Range	67-100	67-100	60-100	0-100	0-100	50-100	0-100	17-100	17-83	0-100
AC										
M	100	97.0	93.9	80.4	92.8	92.4	77.2	83.3	53.1	76.6
SD	0.0	6.9	11.2	22.1	14.5	11.5	17.1	16.7	26.7	23.3
Range	100-100	83-100	67-100	33-100	33-100	67-100	50-100	50-100	0-83	0-100

Note. 0 cluster=no cluster; M Cl=medial cluster; F Cl=final cluster; M+F Cl=medial+final cluster.

Figure 25: Overall performance of the three groups on the nonword repetition task.



Note. 2 Syll-No Cl= Bisyllabic nonwords with no cluster; 2 Syll-M Cl=Bisyllabic nonwords with medial clusters; 2 Syll-F Cl=Bisyllabic nonwords with final cluster; 2 Syll-M+F=Bisyllabic nonwords with medial and final clusters; 3 Syll-No Cl=3 syllable nonwords with no clusters; 3 Syll-M Cl=three syllable nonwords with medial clusters; 3 Syll-F Cl= three syllable nonwords with final cluster; 3 Syll-M+F=3 syllable nonwords with medial and final clusters.

The following histograms show the distribution of the scores for the three groups. They show an approximately normal distribution, with some limited positive skew in the TD groups.

Figure 26: Distribution of the nonword repetition scores for the SLI group, *n*=11.

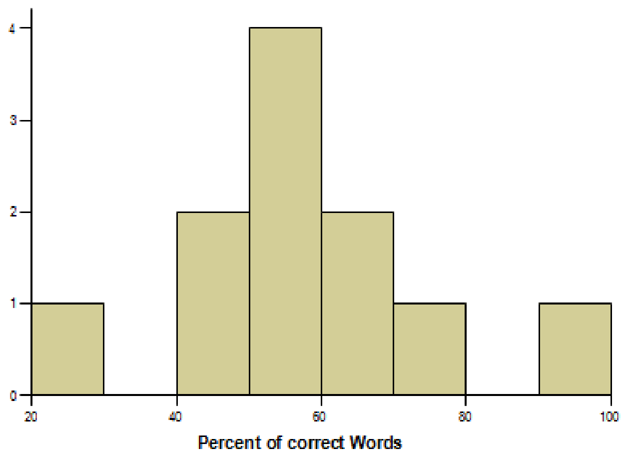


*Figure 27***:** Distribution of the nonword repetition scores for the LC group, *n*=11.

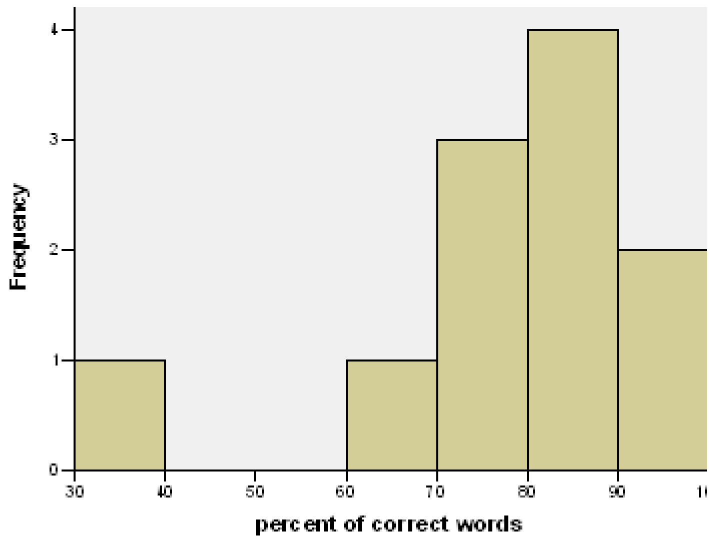
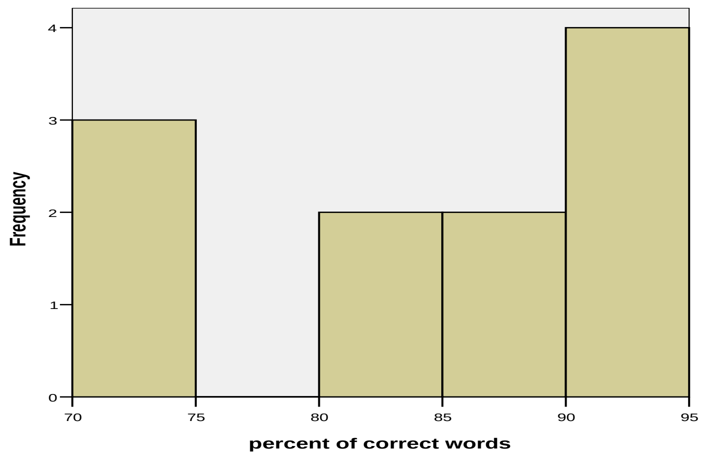


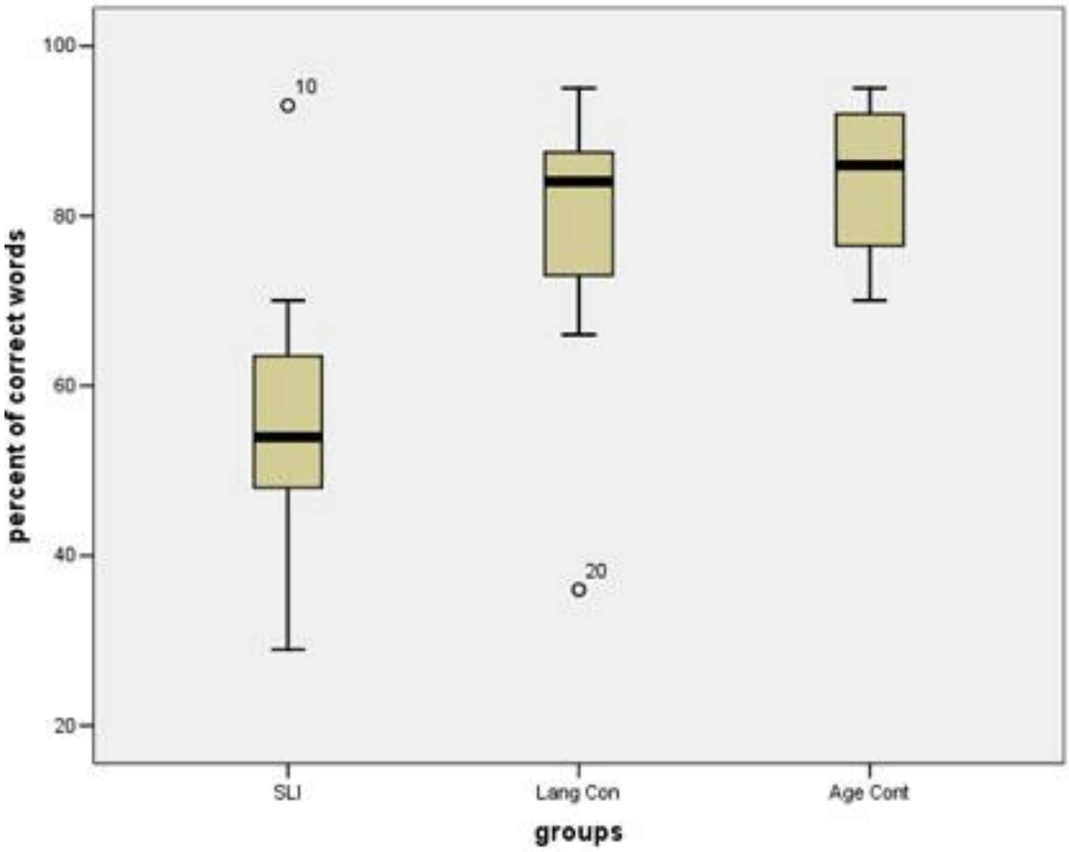
Figure 28: Distribution of the nonword repetition scores for the AC group, *n*=11.



The distribution of the scores of all children on the nonword repetition test is displayed in the boxplot in Figure 29. It clearly shows the significant difference between the group of children with SLI and the two typically developing groups, especially the clear lack of any overlap between AC children and those with SLI. This figure shows that there is one outlier score in the group of children with SLI and another one in the LC group. When examining the scores of the outlier in the SLI group, it was found that she had a Z-score of -2.6 on the Expressive Language test and her scores on other tests were on the lower range of typically developing children. Her score on the digit span task was near the mean of the SLI group (i.e., she had a score of 8 and the group had a mean score of 8.1). Therefore, there was nothing in this child's profile that explains her within normal performance on the NWR task. As for the outlier in the LC group, who had a score that was significantly below the range of typically developing children (i.e., 36% correct repetitions), his scores on all four language tests were within the normal range and nothing in his language abilities profile could explain this lower performance on the NWR task. However, his score on the digit span test was 5, which is considerably lower than the average score for the LC group, which is 9.1. Therefore, this child may have significant limitations in his working memory; however, this does not seem to have repercussions on his general language abilities, apart from this difficulty in nonword repetition. These two outliers are not inconsistent with findings of some studies of

NWR in SLI, where variable abilities have been reported across groups (Bishop, 2004). For example, Gathercole et al. (2005) reported that some children with normal scores on language tests were found to have poor NWR skills.

Figure 29: A boxplot showing the distribution of scores of all three groups.



A 3X2X4 ANOVA Group (SLI, LC, AC) X length (2 syllable, 3 syllable) X cluster type (no cluster, medial cluster, final cluster and medial+final cluster) was conducted. It revealed a significant main effect of group $F(2,30)=12.4$, $p=<.001$, $\eta^2=.45$, indicating there was a significant difference among groups on their overall accuracy of the nonword repetition test. Moreover, there was a significant main effect of syllable length $F(1,30)=71.7$, $p<.001$, $\eta^2=.70$ and cluster type $F(3,90)=60.9$, $p<.001$, $\eta^2=.67$, showing that both independent variables (syllable length and consonant clusters) had significant effects on the performance of the groups on the Arabic nonword repetition test. Cluster by group interaction was also significant $F(6,90)=2.5$, $p=.021$, $\eta^2=.15$ showing that the type of clusters had an effect that depended on which group the child belonged. Length by group interaction was not significant $F(3,90)=2.7$, $p=.08$, $\eta^2=.15$, indicating that across the three groups, length had a relatively comparable effect. Length by cluster interaction was not significant either $F(3,90)=1.5$, $p=.22$, $\eta^2=.049$. Nor was the group by syllable by cluster interaction significant $F(6,90)=1.25$, $p=.29$, $\eta^2 =.07$. In the following, the main effects of independent variables are discussed. This is followed by examining significant interaction effects.

5.10.1 Analysis of Main Effects

5.10.1.1 Main Effects of group

Post hoc test with Bonferroni correction revealed that the SLI group performed significantly worse than both the AC ($p<.001$) and the LC group ($p<.01$) on the overall accuracy of the nonword repetition test. There was no significant difference between the AC and LC groups. Not only did the NWR task differentiate the SLI group from their age control peers, but it also differentiated them from younger typically developing children who had matching scores on sentence comprehension, receptive vocabulary and digit span tasks.

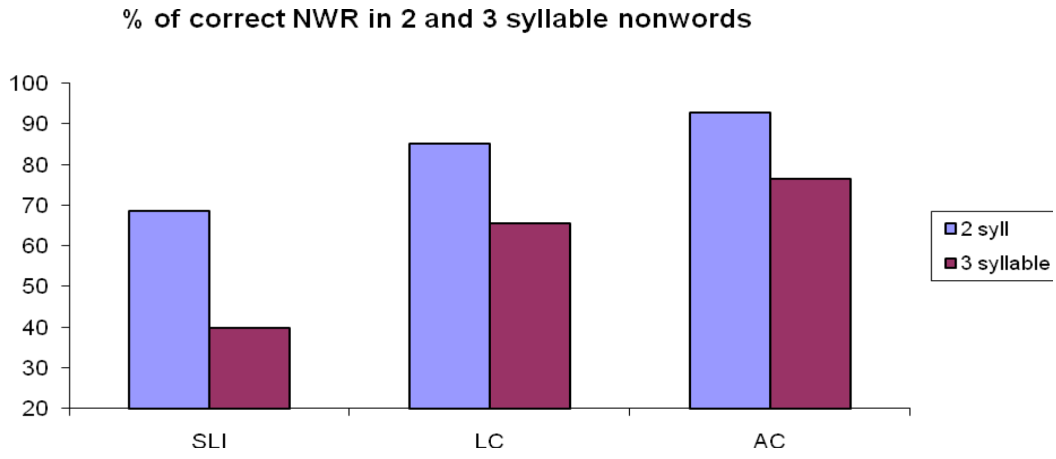
5.10.1.2 Main effects of syllable length

The nonword repetition task used bi and trisyllablic nonwords only. The general ANOVA showed a significant main effect of syllable length, $F(1,30)=71.7$, $p<.001$, $\eta^2=.70$. The following table and figure show the significant difference between two and three syllable nonwords.

Table 50: Percentage of correct nonwords by syllable length (2 vs. 3 syllables) for all participants (*n*=33).

Syllable No.	% NWR
2 syllable nonwords	
M	82.3
SD	16.7
3 syllable nonwords	
M	60.7
SD	23.6

*Figure 30***:** Percentage of correct nonwords based on the number of syllables in each nonword for the three groups.



5.10.1.3 Main effects of cluster types

The general ANOVA showed a significant main effect of cluster type F(3,90)=60.9, p<.001, η^2=.67. There were four types of nonwords used in the nonword repetition task: nonwords with no clusters (No Cl), nonwords with a medial cluster only (M Cl), nonwords with a final cluster only (F Cl), and those with medial and final clusters (M+F Cl). Table 51 summarises the overall performance of groups on the four types of clusters.

Table 51: Overall scores on cluster for all participants.

Cluster type	Mean % of NWR(SD)
No Cluster	85.6 (14.8)
M cluster	75.5 (20.4)
Final Cluster	72.3(23.4)
M+F cluster	52.3 (25.4)

Further analysis using multiple comparisons with Bonferroni correction revealed that there was a significant difference on the performance of all groups on nonwords with no clusters vs. all other types of clusters (see Appendix F). There was a significant difference between medial and finial cluster on one hand and M+F clusters on the other hand, t(30)=23.3 p<.001 and t(30)=20, p>.001 respectively. However, there was no significant difference between medial only and final only clusters. Therefore, the following generalisation about hierarchy of cluster difficulty holds:

0 cluster> 1 cluster (M or F)> 2 clusters

The more clusters a word has, the more challenging it becomes to recall and there was no significant difference between performance on medial vs. final cluster words showing that in Gulf Arabic the number of clusters matters more than the position of these clusters.

5.10.2 Analysis of Interactions

5.10.2.1 The group by cluster types interaction

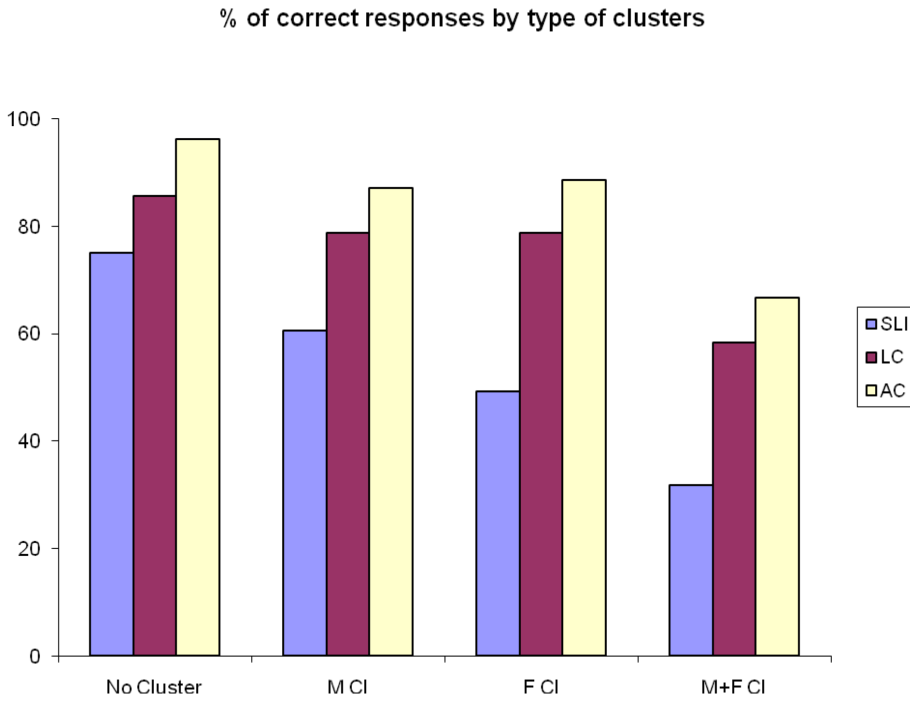
There was a significant group by cluster type interaction F(6,90)=2.5, p=.021 η^2=.15, which means the groups differed in their performance as a function of the type of cluster in nonwords. The three groups were compared with reference to the four types of nonwords: those with no clusters, those with medial or final clusters, and those with both medial and final clusters (M+F clusters). Table 52 and Figure 31 summarise the performance of each group on each type of cluster.

Table 52: Means and standard deviations (in percentage of) correct nonword repetitions for each type of cluster.

Cluster type	SLI Mean (SD)	LC Mean (SD)	AC Mean (SD)	Total Mean (SD)
No cluster	75.0 (16.2)	85.6(12.6)	96.2 (5.7)	85.6 (14.8)
Medial cluster	60.6 (22.8)	78.8 (18.0)	87.1 (8.8)	75.5 (20.4)
Final Cluster	49.3 (19.7)	78.8 (18.1)	88.6 (10.7)	72.2 (23.4)
M+F clusters	31.8 (21.0)	58. (22.5)	66.7 (19.7)	52.3 (25.4)
Overall mean	54.2 (23.0)	75.4 (20.4)	84.7 (14.0)	

Note. M+F Cl=nonwords with medial and final clusters.

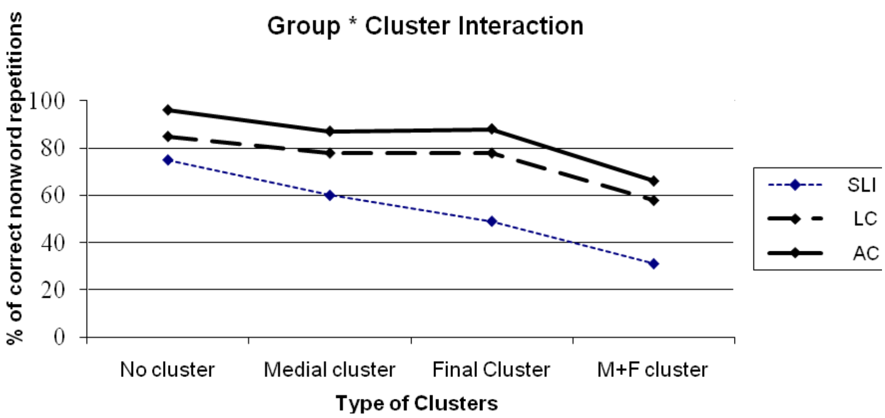
Figure 31: The performance of groups on different types of clusters.



Note. M Cl= nonwords with medial clusters, F Cl=nonwords with final clusters, M+F Cl=nonwords with medial and final clusters.

Figure 32 depicts the interaction between cluster type and groups. It clearly shows a pattern of increasing difficulty with nonwords as they increase in phonological complexity; this effect had a greater impact on the SLI group than it did on the other typically developing groups.

Figure 32: Group by cluster interaction.



One way ANOVAs were performed to investigate the effect of groups at each cluster level. Results showed that cluster effect was significant at each level as shown in the table below.

Table 53: Results of One way ANOVAs for types of clusters in the NWR task.

Cluster Type	ANOVA results	Significance
No Cluster	$F(2,30)=8.27$	$p=.001$
Medial Cluster	$F(2,30)=6.7$	$p=.004$
Final Cluster	$F(2,30)=16.8$	$p<.001$
M+F Cluster	$F(2,30)=8.2$	$p=.001$

Post hoc tests with Bonferroni correction revealed that the AC group consistently performed better than the SLI group on all types of nonwords (see the multiple comparison table in Appendix G). The AC and LC groups were not significantly different on any of the cluster levels. Therefore, the following section will focus on the difference between the SLI group and the LC group on each cluster condition and the interaction of syllable number and cluster types.

On nonwords with no clusters, the SLI group was not significantly different from the LC group $t(30)=-10.6$, $p=.15$. The difference between these two groups on medial cluster nonwords using Bonferroni correction was close to significance $t(30)= -18.3$, $p=.06$. However, on final cluster nonwords, the difference was significant, $t(30)=30.0$, $p=.001$. In addition, on nonwords with medial and final clusters (M+F cluster), the difference between the SLI and LC groups was significant, $t(30)=-9.0$, $p=.019$.

A repeated measure ANOVA was conducted to compare children with SLI with the LC group. The AC group was not included due to ceiling effects, i.e., they had scores of 80% and above on all types of clusters, except the M+F clusters. Results of the ANOVA showed a significant effect of group $F(1,20)=8.82$, $p=.008$, $\eta^2=.30$; cluster $F(3,60)=49.45$, $p<.001$, $\eta^2=.71$; syllable number $F(1,20)=45.40$, $p<.001$, $\eta^2=.69$; cluster*group $F(3,2)=3.42$, $p=.023$, $\eta^2=.14$. However, there was no group*syllable number interaction or syllable number*cluster interaction. The effects of both syllable number and cluster types seen in the overall ANOVA were also seen in this comparison between children with SLI and LC, with two syllable nonwords being significantly easier than three syllable nonwords, and increasing difficulties as number of clusters increases from zero (no cluster) to two (M+F cluster), with medial only and final only not being significantly different from each other.

Comparing the SLI and LC groups on syllable number by cluster interaction. Comparisons between SLI and LC group showed that the former group scored significantly less well than the younger typically developing group on both two syllable nonwords $t(20)= -2.4$, $p=.024$ and three syllable nonwords $t(20)=-3.0$, $p=.007$. The overall ANOVA did not show a cluster by syllable number interaction. Nevertheless, when cluster effects were removed and SLI and LC groups were compared based on performance on nonwords with no clusters only, some interesting results appeared. Results of T-tests showed that when cluster effects were removed, there was no significant difference between the SLI and LC groups on bisyllabic nonwords ($t(20)=-1.01$, $p=.32$) and the difference on three syllabic nonwords with no clusters, did not reach significance either ($t(20)=-1.94$, $p=.066$. However, when these two groups were compared on the average of all three types of nonwords with clusters, the difference was significant on bisyllabic nonwords, $t(20)=-2.59$, $p=.017$ and trisyllabic nonwords $t(20)= -3.16$, $p=.005$.

It is important to note that some of the nonsignificant results, for example the nonsignificant length*group interaction ($p=.08$) may suggest an effect that could be operating, but has not been seen due to the small sample size in this study ($n=11$) in each group). Therefore, such trends cannot be altogether dismissed and studies of larger size of population are needed to examine such possible trends.

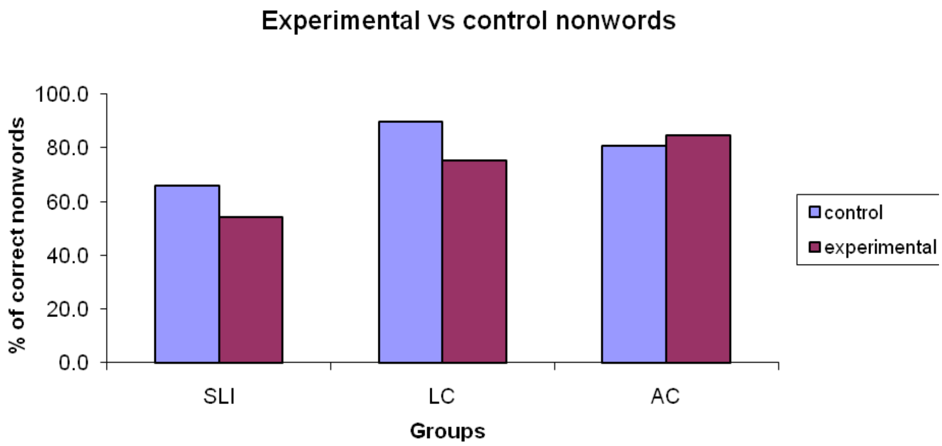
5.10.3 Wordlikeness effects

To compare the effects of wordlikeness on the performance of the three groups, their scores on the experimental (n=48) vs. control nonwords (n=8) were examined. While the experimental nonwords contained low frequency patterns, i.e., patterns that are rarely used, the control nonwords consist of highly frequent control patterns that are very common in Gulf Arabic and other varieties of Arabic. Indeed the control pattern (CaCaC) corresponds to the most frequent verbal pattern in Arabic (i.e., the so-called type 1 verbal form, which indicates past tense (perfective) in trilateral verbs (e.g., /ʔakal/ 'he ate'). It also corresponds to nouns that have the same vocalic pattern (e.g., /samak/ 'fish'). Table 54 and Figure 33 summarise the results of the three groups on both types of the nonwords patterns.

Table 54: Means and standard deviations (SD) (in percentages) of the scores of all groups on experimental nonwords that have non-frequent patterns (*n*=48) vs. control nonwords that have very frequent patterns (n=8).

Type of Patterns		SLI (*n*=11)	Language Control (*n*=11)	Age Control (*n*=11)
Non-frequent (Experimental nonwords)	Mean	54.2	75.4	84.7
	SD	17.2	17.2	9.4
Frequent Patterns (Control nonwords),	Mean	66.0	89.9	80.9
	SD	26.2	19.0	18.0

Figure 33: The performance of all groups on experimental nonwords (nonwords with non-frequent patterns) vs. control nonwords (nonwords with very frequent patterns).



A repeated measure ANOVA showed there was a main effect of wordlikeness F(1,30)=6.6, p=.015, η^2=.18. Overall, children found the high frequency patterns, which were more wordlike, easier to recall than the other less frequent patterns, t(32)=2.37, p=.02. Results showed that there was a significant effect of group, F(2,30)=7.0, p<.01, η^2=.31. Moreover, there was a significant wordlikeness X group interaction, F(2,30)=3.9, p=.03 η^2=.20. Subsequent analysis using Bonferroni correction showed that the SLI group did not benefit significantly from wordlikeness effects when compared to the LC group. T-test with Bonferroni correction for multiple comparisons showed that the difference between the SLI group's performance on high vs. low frequent patterns was not significant, t(10)=1.74, p=.11. The LC group on the other hand benefitted significantly from wordlikeness effect with its performance increasing significantly on the high frequency pattern, t(10)=5.0, p <.001. As for the AC group, there was no significant difference on their performance on low vs. high frequency patterns, t(10)=-.8, p=.4. No significant difference was found between the two typically developing groups on low vs. high frequency patterns.

5.10.4 Articulatory (output processes) effects

Though it was difficult to tease apart articulatory vs. phonological components in the nonword repetition task, an attempt was made to do so by comparing two different measures of calculating the overall score of all children. It is assumed here that since all the phonemes chosen to make up the nonwords were early developing phonemes and all the children with SLI passed the articulation and verbal dyspraxia screening tests, a measure that compares children on total phoneme accuracy, such as percentage of consonant correct (PCC) may be useful in providing information about the role of articulatory effects in the nonword repetition task. This type of measure may tap into articulatory competence more accurately than the measure used so far, i.e., the percentage of correct words recalled, which is closely linked to the design of the test where phonological complexity (consonant clusters) was manipulated. Therefore, it is hoped that by comparing the performance of the three groups on these two measures of articulatory and phonological competence, some insights can be drawn about the role of these two processes. The following table shows the performance of all groups using the percentage of consonants correct calculation method.

Table 55: Mean and standard deviations of all groups based on the percentage of consonant correct (PCC) as a measure of phoneme accuracy.

Group		PCC
SLI	M	89.4
	SD	4.4
LC	M	93.4
	SD	5.5
AC	M	96.3
	SD	2.0

When comparing the three groups on PCC using a one-way ANOVA, results revealed a significant difference among the groups F(2,3)=7.4, p=.002. The post-hoc test with Bonferroni correction showed there was no significant difference in the performance of the SLI and LC groups t(30)=-4.0, p=.18. The SLI group was however significantly less than the AC on PCC, t(30)=-7.0, p=.002. The LC and AC groups were not significantly different from each other t(30)=-6.1, p=.81

5.10.5 Correlations between nonword repetition and other language measures

The Pearson product-moment correlation coefficient was calculated to measure the correlation between nonword repetition as measured by the percentage of correct nonwords and other developmental, cognitive, and linguistic measures. Results showed that nonword repetition significantly correlated with the Expressive Language subtest (r=.50, p=.003), which measures the child's ability to produce various morphosyntactic structures in Gulf Arabic. Moreover it correlated with Sentence Repetition (r=.40, p=.02) and the Sentence Comprehension test (r=.37, p=.03). Nonword repetition correlated with the digit span score (r=.41, p=.01), a measure of working memory which combines both forward and backward span recall. However, NWR did not correlate with receptive vocabulary as measured by the Arabic Picture Vocabulary Test (r=.27, p=12) or age (r=-.16, p=35). Nor did NWR correlate with nonverbal IQ as measured by the Test of Non-verbal Intelligence (TONI-3) ((Brown et al., 1997). Table 56 shows the results of the correlation of NWR with various psycholinguistic measures.

Table 56: Correlations between nonword repetition scores as measured by percent of correct words (PCW) and other psychlinguistic measures in all participants (*n*=33).

	NWR (PCW) (*n*=33)	Age in Months (*n*=33)	Digit Span Total (*n*=33)	Arabic Picture Vocabulary Test (*n*=33)	Sentence Comprehension test (*n*=33)	Expressive Language test (*n*=33)	Sentence Repetition test (*n*=33)	Test of Non-Verbal Intelligence (*n*=20)
NWR (PCW)								
Pearson Correlation	1	-.16	.41*	.27	.37*	.50**	.40*	.26
Sig. (2-tailed)		.35	.01	.12	.03	.003	.02	.25
Age in Months								
Pearson Correlation	-.16	1	.23	.22	.36	-.02	.06	-.27
Sig. (2-tailed)	.35		.18	.21	.04	.87	.71	.24
Digit Span Total								
Pearson Correlation	.41*	.23	1	.37*	.56**	.55**	.54**	.11
Sig. (2-tailed)	.01	.18		.03	.001	.001	.001	.62
Arabic Picture Vocabulary Test, Pearson Correlation	.27	.22	.37*	1	.56**	.51**	.50**	.24
Sig. (2-tailed)	.12	.21	.03		.001	.002	.003	.29
Sentence Comprehension Test								
Pearson Correlation	.37*	.36*	.56**	.56**	1	.74**	.74**	.37
Sig. (2-tailed)	.03	.04	.001	.001		.000	.000	.14
Expressive Language Test								
Pearson Correlation	.50**	-.02	.55**	.51**	.74**	1	.91**	.26
Sig. (2-tailed)	.003	.87	.001	.002	.000		.000	.26
Sentence Repetition Test								
Pearson Correlation	.40*	.06	.54**	.50**	.70**	.91**	1	.36
Sig. (2-tailed)	.02	.71	.001	.003	.000	.000		.11
Test of Non-Verbal Intelligence								
Pearson Correlation	.26	-.27	.11	.24	.37	.26	.36	1
Sig. (2-tailed)	.25	.24	.62	.29	.14	.26	.11	

* Correlation is significant at the 0.05 level.

** Correlation is significant at the 0.01 level.

5.10.6 Error analysis

The errors produced by children were classified into eight types of errors and were grouped into two main categories: segmental and syllabic. Segmental errors are characterised by preserving the shape and order of the syllable with changes occurring at the segmental level only. There were two types of segmental errors: consonant and vowel substitutions. Syllabic errors, on the other hand, are characterised by changes in the shape of the syllable or the order of its components. They include final cluster reduction, medial cluster reduction, cluster creation, syllable omission, metathesis, and final consonant deletion. Examples of these types of errors are listed in Table 57.

Table 57: Examples of the various types of errors in NWR.

Type of error		Target Nonword	Child's Response
Segmental Errors	Consonant substitution	kus.mi.**b**an	kus.mi.**t**an
	Vowel Substitution	sad.l**u**.naf	sand.l**a.**faf
Syllabic Errors	Final cluster reduction	su.ki.da**fs**	su.ki.daf
	Medial cluster reduction	sa.du**n.n**afd	sa.du.nafs
	Cluster creation	du.ki.mas	ku.di.ni**fs**
	Syllable omission	ku.sib.**banf**	ku.sib.
	Metathesis	kad.lu**sb**	kad.lu**bs**
	Final consonant deletion	Suk.bi.da**f**	su.bi.ka

Table 58 shows the distribution of these types of error among the three groups. It shows that the three groups followed the same pattern of errors, with consonant substitution and final consonant cluster reductions accounting for almost 80% of the total number of errors across groups. Some notable exceptions where one group performed differently were cluster creation, where the SLI group created more clusters than the other TD groups and final consonant deletion, where the language controls omitted more final consonants than the other two groups. However, no concrete conclusions can be drawn due to small number of these three types of errors.

Table 58: Distribution of NWR errors for all participants in numbers and percentages

Type of error			SLI	LC	AC
Segmental	Consonant substitutions	*n*=	204	97	66
		%	51.3	49.5	55.5
	Vowel substitution	*n*=	10	10	2
		%	2.5	5.1	1.7
Syllabic	Final cluster reduction	*n*=	121	48	34
		%	30.4	24.5	28.6
	Medial cluster reduction	*n*=	24	14	6
		%	6.0	7.1	5.0
	Cluster Creation	*n*=	15	0	1
		%	3.8	0	0.8
	Syllable omission	*n*=	13	12	3
		%	3.3	6.1	2.5
	Metathesis	*n*=	7	4	5
		%	1.8	2.0	4.2
	Final consonant deletion	*n*=	4	11	3
		%	1.0	5.6	2.5
	Total number of errors		398	196	119

When comparing the percentage of correct repetitions among the three groups, little overlap was observed among them. Children with SLI had a range of percentage of correct responses of 29% to 70% (with one outlier who scored 93%), while both LC and AC had a range of 71% to 95%, with one outlier in LC group with a score of 36%. Therefore, only one participant in the SLI group had a score within the range seen in the other two typically developing groups.

One way ANOVAs were conducted to compare the percentage of segmental vs. syllabic errors and consonant substitutions vs. final cluster reduction across groups. The result for the ANOVA for percentage of segmental errors was $F(2,30)=0.273$, $p=.763$, and for syllabic errors was $F(2,30)=0.144$, $p=.866$, indicating there was no difference in the pattern of errors displayed by these groups. The largest proportion of errors was either consonant substitutions or final cluster reduction and the ANOVA examining these two different types of errors was not significant (see

Appendix H I for ANOVAs of all error types). The table in Appendix H shows that only two types of errors were significant, namely cluster creation and final consonant deletion. However, both of them accounted for a minority of error types and there were too few examples to make any analysis meaningful. However, a noteworthy observation from Table 58 was the relatively large number of clusters added or created by children with SLI compared to the two control groups.

5.11 Discussion

This study set out to investigate four main issues. Firstly, whether the nonword repetition test devised in this study can act as a clinical marker of SLI in Arabic, since this is the first investigation of nonword repetition skills in Arabic speaking children with SLI. Secondly, this study endeavoured to shed light on some of the competing theories of the nature of nonword repetition deficits in children with SLI, especially those that attribute these impairments to a central impairment in phonological capacity (Gathercole, 2006; Gathercole & Baddeley, 1990)) or theories that argue for the presence of more influential grammatical factors, such as phonological complexity (van der Lely, 2005) or broader and more basic phonological processing skills (Chiat, 2001; Snowling et al., 1991). Thirdly, the results of the Arabic nonword repetition test will be compared with those obtained from other languages, in order to examine how nonword repetition deficits manifest in different languages and how underlying processes interact with language-specific properties. Finally, this experiment examined the relationship between children's performance on nonword repetition and various psycholinguistic measures obtained from the tests conducted with typically developing children and children with SLI. These measures included receptive vocabulary, sentence comprehension, sentence repetition, expressive language skills, digit span, and nonverbal IQ.

5.11.1 Clinical implications for the study of nonword repetition in Arabic

The results of this first investigation of nonword repetition skills in Gulf Arabic speaking children with SLI show that these children perform significantly worse than their typically developing peers matched on age or language abilities.

Therefore, these results extend the viability of nonword repetition task as a possible clinical marker of SLI to a further language, namely Gulf Arabic. The usefulness of this task may not be constrained to Gulf Arabic; as the design of the test and the stimuli used may render it useful and clinically viable in other varieties of Arabic. The stimuli used in this task consist of early developing sounds that exist in all Arabic dialects and the syllable structures (cv), (cvc), and (cvcc) used are common in most dialects of Arabic (Watson, 2002) . Therefore, this task might be a useful tool in the identification of children at risk of language impairments; especially with the paucity of assessment tools in Arabic (see Shaalan, 2009). Moreover, many studies have found that nonword repetition is less influenced by socioeconomic factors and therefore less prone to bias than other conventional language measures (Campbell, Dollaghan, Needleman & Janosky, 1997; Ebert, Kalanek, Cordero & Kohnert, 2008; Ellis Weismer et al., 2000). The NWR test could avail itself to be used with a wider population than the current sample of children, who mostly come from middle class households. However, larger scale empirical studies are needed to confirm these findings in various varieties of Arabic, including Gulf Arabic.

5.11.2 Discussion of results with respect to different processes underlying nonword repetition

In the following, I discuss the implications of the results of the nonword repetition task for theories that attribute deficits in nonword repetition test to impairments in phonological short-term memory, phonological complexity, and those that argue for deficits in basic phonological processing skills.

5.11.2.1 The phonological short-term memory account of SLI

The findings of significant difference between children with SLI and the other two control groups on this nonword repetition task that comprises only 2 and 3 syllable words organised according to their syllabic structure are not consistent with the phonological short-term memory account of SLI as proposed by Gathercole and Baddeley (1990). According to this account, limitations in the phonological loop, the part of working memory responsible for storing phonological information, are the main cause of deficits in nonword repetition, vocabulary learning, and syntactic development in children with SLI (Baddeley et al., 1998; Gathercole & Baddeley, 1990). This account predicts that children with SLI would score significantly less well than typically developing children on nonwords consisting of three syllables onwards. It also predicts a significant interaction between syllable length and groups in that children with SLI should be differentially affected by syllable length effects. However, despite the presence of a main effect of length in this experiment, indicating that nonword repetition accuracy decreased when syllable length increased from two to three syllables, the nonsignificant interaction of syllable length and groups and the significant group by syllabic complexity interaction show that syllable complexity shows better differentiation among groups than length effects. This is inconsistent with PSTM claim that phonological storage is the central factor in determining NWR skills (Gathercole, 2006). This study found that even on the shortest nonwords used, i.e., bisyllabic nonwords, children with SLI performed significantly less well than their age and language controls when these nonword contained consonant clusters, while most studies of nonword repetitions reported that differentiation between children with SLI and their typically developing controls starts at nonwords of three and four syllable lengths (Dollaghan & Campbell, 1998; Gathercole & Baddeley, 1990; Marton & Schwartz, 2003).

A detailed examination of the interaction between length and cluster effects reveals that the presence or absence of clusters has a stronger effect on children with SLI performance than syllable length. Seventy five percent of all nonwords used in this experiment had consonant clusters in them, and when children with SLI were compared with the language control children on bisyllabic and trisyllabic nonwords with no clusters, the difference was not significant. This is in contrast to the significant difference between the two groups (SLI vs. LC) on nonwords with clusters regardless of syllable length. This weak effect of clusterless nonwords vs. the strong effect of nonwords with clusters is not borne out by the strong phonological short term memory account (Baddeley et al., 1998; Gathercole & Baddeley, 1990), which posits that PSTM is the main factor in determining NWR performance, or even its latest version (Gathercole, 2006), which acknowledges the contribution of other phonological factors (such as prosody, wordlikeness, and lexicality effects). However, this account continues to argue that PSTM has a ‘central’ role in NWR and its contribution exceeds those of the above-mentioned factors (Gathercole, 2006).

It is important however to reiterate that children with SLI performed significantly worse than their age control group on both two and three syllable nonwords and regardless of the presence or absence of clusters. Therefore, the difference between the group of children with SLI and their age controls in terms of capacity limitations in phonological short-term memory is evident and cannot be ignored. However, this experiment shows that their performance is better differentiated from their language controls by phonological complexity, defined here as the presence of medial, final, or medial and final clusters.

5.11.2.2 The phonological complexity account of nonword repetition deficits

The results of the Arabic nonword repetition test show that children with SLI have an increasing difficulty with nonwords as the number of marked structures increases. These findings are consistent with accounts of SLI that attribute weak performance to phonological complexity, such as the Computational Grammatical Complexity (CGC) account (van der Lely, 2005). This linguistic account of SLI argues that deficits in the grammatical aspects of language, such as syntax, morphology, and phonology can explain the type of linguistic problems seen in children with SLI. van der Lely and colleagues (Gallon et al., 2007; Marshall et al., 2002; van der Lely, 2005) argue that grammatical phonology in children with (G)-SLI are characterised by deficits in marked prosodic and syllabic structures, such as unfooted syllables and consonant clusters. The presence of these marked structures increases the complexity of words and nonwords and children with SLI seemed more prone to such deficits in phonologically complex structures than their age or language matched peers. This complexity effect should not be confounded by response time, as one may argue that children with SLI may have difficulty with complex structures, such as consonant clusters, because they take more response time. However, Coady and colleagues (Coady, Evans & Kluender, 2010) showed that children with SLI took the same time to repeat nonwords as their TD peers did.

The results of this experiment show that syllabic complexity plays a more important role than phonological storage in this nonword repetition task in Arabic.

Children with SLI consistently struggle with the task as the number of marked structures increases from no clusters to one and two clusters even in two syllabic nonwords. Their performance on nonwords with complex syllabic structure was significantly less accurate than both their age and language matched controls.

These results replicate the finding of Gallon et al. (2007) and Marshall et al. (2002) where children with (G-) SLI had difficulties repeating even monosyllabic and bisyllabic nonwords when they contained marked syllabic and metrical structures. Both Marshall et al. (2002) and Gallon et al. (2007) found that the performance of children with SLI deteriorated as the number of complex structures increased, similarly to our findings of increasing difficulties as the number of consonant clusters increases from 0 (no cluster) to 1 (medial or final cluster) to 2 (medial and final cluster). These effects of phonological complexity as represented by consonant clusters have been reported in some of the early studies of NWR, such as Gathercole and Baddeley (1989). Moreover, Bishop et al. (1996) found that consonant clusters affected the performance of children with SLI greater than it did in other control groups. When Archibald and Gathercole (2006a) compared the performance of 12 children with SLI on both CNRep (Gathercole & Baddeley, 1996) and NRT (Dollaghan & Campbell, 1998), they showed that children with SLI had difficulties with both tests compared to age controls, however, only on the CNRep, where there were many nonwords with clusters, did they perform significantly less well than their language controls. This shows that a strong phonological storage account, like the one proposed by Gathercole and Baddeley (1990) cannot account for the significant deficits seen in nonwords with consonant clusters. It seems that along with PSTM, there are some important factors that need to be taken into consideration when examining NWR and phonological complexity is one of these factors. In summary, this nonword repetition test in Arabic shows that the impact of phonological complexity on the Arabic NWR test exceeds that of phonological short-term memory

5.11.2.3 The phonological processing account of nonword repetition

While the experiment and the stimuli were designed mainly to investigate the phonological short term memory account and phonological complexity account, the results of this nonword repetition task in Arabic speaking children can be of relevance to discussions of other phonological processes that might have an impact on nonword repetition. Proponents of the phonological processing account of NWR argue that there are various phonological processes implicated in NWR. In the following, some processes relevant to the current experiment are discussed.

Wordlikeness Effects. Results of the present study show that wordlikeness effects have influenced the performance of children with SLI on the Arabic NWR task. The experimental nonwords used in the Arabic NWR test were based on nonfrequent, and therefore less wordlike patterns, while the control stimuli contained more frequent (more wordlike) nonwords. The results of the Arabic NWR test show that the difference in accuracy on low vs. high frequency pattern was not significant ($p=.12$) in the group of children with SLI, while the LC group found the high frequency patterns significantly easier to recall ($p <.001$). Therefore, unlike typically developing children, children with SLI did not benefit from previous linguistic knowledge to form new phonological representations, i.e., their redintegration skills are not as efficient as they are in typically developing children. Therefore, these results support other studies that argue that process of redintegration would give children with better language abilities an advantage over those with language impairments, who would be less efficient at reconstructing traces of nonwords from their long-term phonological representations (Gathercole, 1999; Stokes et al., 2006). Wordlikeness (or morpho-lexical) effects were reported in Italian, where children with SLI scored significantly better on morphological nonwords (nonwords consisting of existing root and suffixes) than nonwords that did not contain any existing morphemes (Casalini et al., 2007). However, when compared to controls, preschoolers (but not first graders) with SLI showed smaller differences between morphological nonwords and nonwords, indicating that age can modulate these morpho-lexical effects in Italian. Overall, the results of this experiment indicate the presence of wordlikeness effects in Gulf-Arabic speaking children with SLI.

Output processes. A common confounding variable in many studies of nonword repetition skills in children with SLI is the effects of output or articulatory processes. Snowling et al. (1991) and Chiat (2001) considered articulatory factors as one of the phonological processes that might be implicated in the poor performance of children with SLI on NWR tasks. Though the current experiment was not designed to test the role of articulatory processes in nonword repetition, the results might be helpful in interpreting the effects of these processes on nonword repetition. Results show that the percentage of consonants correct (PCC), the measurement that is arguably more sensitive to articulatory and phonemic accuracy, revealed no significant difference between the SLI and the language control group, whereas the scoring method based on the percentage of correct nonwords shows a significant difference between the SLI group and language control group. This latter measurement is closely related to the design of the test where nonwords are manipulated based on phonological (syllabic) complexity. Therefore, what distinguishes the SLI group from the LC group is the measure of phonological complexity, but not the PCC that measures articulatory processes. Therefore, it is argued that the children with SLI tested in this study do not have major deficits in their articulatory processes, as opposed to their significant difficulties in tasks that involve phonologically complex structures, such as consonant clusters. The high score of children with SLI on PCC, which is 89.4%, is higher than one expects if these children have deficits in articulatory processes. This lack of evidence for deficits in output processes when children with SLI were compared to language controls is consistent with findings of Gathercole and Baddeley (1990), where they show that the articulation rate of children with SLI was not significantly different from their control children. Moreover, Edwards and Lahey (1998) found that children with SLI did not have clear deficits in their output processes. These two studies seem to concur with the findings of the present experiment on lack of evidence of articulatory deficits in children with SLI in the NWR task.

5.11.3 Cross-linguistic implications of the Arabic NWR

This study adds cross-linguistic evidence for the importance of nonword repetition skills in children with SLI. Most of the evidence available in the NWR literature comes from European languages, such as English (Bishop et al., 1996; Botting & Conti-Ramsden, 2001; Gathercole & Baddeley, 1990); Italian (Casalini et al., 2007); Spanish (Girbau & Schwartz, 2007); Dutch (de Bree et al., 2007); Portuguese (Engel et al., 2008), and Swedish (Hansson et al., 2004; Sahlen et al., 1999). All these languages are of Germanic or Latin origin, while Arabic is a Semitic language that is typologically different from these languages and characterised by the root and pattern nature of word composition. Therefore, nonword repetition seems to work as a sensitive measure to identify children with SLI in a variety of European and non-European languages and can be used to elucidate and compare the competing processes underlying nonword repetition across languages.

The results of this experiment support the findings of another cross-linguistic study of nonword repetition, which did not find evidence for limitations in phonological short-term memory in children with SLI. Stokes et al. (2006) demonstrated that Cantonese-speaking children with SLI did not have significant impairment in nonword repetition, as their performance was not significantly different from typically developing children, a challenging finding for any phonological short-term memory account that attributes deficits in SLI to limited phonological storage. The findings of the current study and that of Stokes et al. (2006) showed weak (Gulf Arabic) or no support (Cantonese) for the claim that limitations in phonological short-term memory are the main cause of NWR deficits in children with SLI. It seems that the presence of consonant clusters in Arabic and the lack of them in Cantonese may explain the performance of children with SLI on nonword repetition task in both languages, though sublexical factors may have a role to play in both languages. The stimuli in both the Cantonese NWR test and the Arabic NWR task consist of simple (early developing) consonants and both languages have a regular stress pattern; however the stimuli in Gulf Arabic differ from Cantonese in that they have consonant clusters and sublexical effects were not controlled. Therefore, these various and differing results obtained from NWR tests in English, Arabic, and Cantonese can be attributed to an interaction between language specific and test-specific parameters. Table 59 summarises some of the parameters involved in four NWR tests in English, Cantonese, and Arabic and how these may have influenced their findings. This table shows that the more marked parameters present in a test/language, the more the NWR task is able to distinguish between children with SLI and their typically developing peers.

Table 59: Various parameters that are involved in some NWR tests in English, Arabic and Cantonese.

Test	Segmental complexity	Syllable Number (PSTM)	(sub) Lexical Effects	Prosodic Complexity	
				Syllabic Complexity	Metrical Complexity
CNRep[*]	+	2-5	+	+	+
Arabic NWR[*]	-	2-3	+	+	NA[b]
NRT [**]	-	1-4	+[a]	-	-
Cantonese NWR[±]	-	1-4	-	-	-

Note. CNRep: The Children's Test of Nonword Repetition test (Gathercole & Baddeley, 1996). NRT: The Nonword Repetition Test (Dollaghan & Campbell, 1998). Cantonese NWR: the test developed by Stokes et al., 2006.
* These two tests showed significant difference between children with SLI and both their age and language controls (in English: Gathercole and Baddeley, (1990) and Archibald and Gathercole, (2006a)) and the current study of Arabic NWR test.
** Archibald and Gathercole (2006a) found there was a significant difference between children with SLI and their age controls, but not between children with SLI and language controls on NRT.
[±] Stokes et al. (2006) found no significant difference between children with SLI and their AC or LC on the Cantonese NWR test.
[a] Although Dollaghan and Campbell (1998) attempted to control for sublexical effects, some of the nonwords they used start with frequent CV combinations, such as /dæ/ and /tei/. See (Stokes et al., 2006).
[b] In the Gulf Arabic NWR test it was not possible to manipulate metrical complexity (stress) separately from syllabic complexity (consonant clusters).

This table shows that the CNRep test (Gathercole & Baddeley, 1996) may be successful in differentiating between children with SLI and their AC and LC peers due to the presence of many marked parameters that tax not only the PSTM skills of these children, but their grammatical phonology and phonological processing skills. Archibald and Gathercole (2006a) found the CNRep, but not NRT (Dollaghan & Campbell, 1998) was able to distinguish between children with SLI and their LC peers. However, both tests showed significant difference between children with SLI and their AC peers. This table suggests that the Arabic NWR test, like CNRep, revealed a significant difference between children with SLI and their AC and LC groups due to the presence of marked parameters. The NRT test (Dollaghan & Campbell, 1996) has fewer marked parameters than the CNRep and the Arabic NWR, but more than the Cantonese NWR test. Therefore, it showed a significant difference between children with SLI and their age controls, but not when they were compared to their language controls (Archibald & Gathercole, 2006a). The

Cantonese NWR (Stokes et al., 2006) test has the fewest marked parameters and therefore this may explain why the group of children with SLI were not significantly different from their AC and LC peers on this task. The only marked structure in the Cantonese NWR test was the number of syllables that ranged from 1-4 syllables and this shows that syllable length as a measure of phonological short term memory may not be sufficient to distinguish between children with SLI and their TD peers on NWR task in Cantonese, and possibly in other languages.

In summary, the results of this experiment, when compared to results from some NWR tests in Cantonese and English suggest that NWR tests with more marked parameters are better able to distinguish between children with SLI and their TD peers.

5.11.4 Correlations between nonword repetition and other language measures

The results of correlations of the Arabic nonword repetition test and other language measures are consistent with studies reported in English and other languages, which showed a good correlation between nonword repetition and other psycholinguistic measures. Due to the small size of the SLI group (*n*=11), the correlations reported in this experiment are for all participants (*n*=33).

Table 56 shows that the score on the Arabic nonword repetition test moderately correlated with digit span (r=.41, p=.01), sentence repetition (r=.40, p=.02), sentence comprehension (r=.37, p=.03), and the expressive language test (r=.50, =.003). The nonword repetition score did not correlate significantly with the receptive vocabulary score (r=.27, p=.12). Nor did it correlate with nonverbal IQ as measured by the Test of Non-verbal Intelligence (TONI-3) (Brown et al., 1997), though this test was conducted with 20 children only.

The moderate correlation between nonword repetition and digit span reported here (r=.41, p=.01) for all the children concurs with other reports of correlation between these two measures of short- term working memory in typically developing children. The correlation coefficient for 8 year old children reported in Gathercole et al. (1992) study was r=.44. This strong correlation is not surprising, since both NWR and digit span tap into verbal short term memory, however NWR is posited as a better indicator of phonological short-term memory as it resembles the task children face when learning new words (Gray, 2003b), especially when nonword repetition test controls for lexicality (wordlikeness) effects. Therefore, nonword repetition has been described as a better identifier of language impairment than digit span (Gray, 2003a).

The nonword repetition task correlates strongly with the Arabic Sentence Repetition (SR) test ($r=.40$, $p=.02$). Sentence repetition requires the child to integrate his phonological, lexical-semantic, and syntactic skills along with his memory and articulatory skills in order to recall sentences, most of these processes are implicated in nonword repetition too. Therefore, it is not surprising that there is such a strong correlation between sentence repetition and nonword repetition. Sentence repetition has been found as one of the best clinical tools to identify children with SLI. This has been true for many languages, such as English (Conti-Ramsden et al. (2001) and Stokes et al. (2006) in Cantonese. This project has demonstrated that Gulf Arabic speaking children with SLI have consistent difficulties with sentence repetition (see chapter 3).

This study reveals a significant correlation between nonword repetition and the Arabic Sentence Comprehension (SC) test ($r=.37$, $p=.03$). Bishop et al. (1999) reported similar level of correlation ($r=.37$, $p<.001$) between nonword repetition, as measured by CNRep (Gathercole & Baddeley, 1996) and the Test of Receptive Grammar (TROG, Bishop, 1989) in children aged 7 to 13 years old. Moreover, Montgomery and colleagues reported a strong correlation between the two constructs in various studies (Montgomery, 1995b; 2002b; Montgomery & Evans, 2009; Montgomery & Windsor, 2007). They used these findings to argue that sentence comprehension requires a significant amount of working memory resources.

Among the language measures studied, the Arabic Expressive Language (EL) test shows the strongest correlation with the nonword repetition task ($r=.50$, $p=.003$). The EL test consists of various tasks that measure various morphosyntactic structures and therefore requires good grammatical abilities. Similarly, the Arabic nonword repetition task is highly laden with consonant clusters and therefore requires good grammatical phonology skills. This is consistent with linguistic accounts of SLI (e.g., the CGC) that expect children with SLI to display deficits in the main components of their grammatical system, which are syntax, morphology, morphosyntax, and phonology. The findings of the current study are consistent with studies that reported a correlation between nonword repetition skills and other measures of expressive grammatical abilities, such as third person singular –s and past tense (Botting & Conti-Ramsden, 2001) and the expressive language score of the Clinical Evaluation of Language Fundamental (CELF-R; Semel et al., 1987) as reported in Montgomery and Windsor (2007).

This experiment found no evidence for a strong correlation between nonword repetition and receptive vocabulary score as measured by the Arabic Picture Vocabulary Test (APVT) (r=.27, p=.12). Most, but not all, studies that investigated the correlation between nonword repetition and vocabulary (receptive or expressive) have found strong correlation between the two (for reviews, see Gathercole, 2006; Baddeley et al., 1998; Baddeley, 2003). According to the phonological short-term memory theory (Gathercole & Baddeley, 1990; Gathercole, 2006; Baddeley et al., 1998), nonword repetition is closely related to the language learning ability of typically and atypically developing children and is considered one of the best predictors of word learning skills of these children. However, Gathercole (2006) explains that the correlation between nonword repetition and word learning is at its strongest at early stages of language development (r=.52-.56 at ages 4, 5, and 6 years) but declined to .27 at 8 years; however it remains significant in the large sample tested. The lack of significant correlation in this experiment could be due to factors such as age of participants or the small number of subjects in this experiment, or due to language-specific factors. However, this is not the first study that has found no correlation between NWR and receptive vocabulary. Bowey (1996) shows there was a lack of correlation between nonword repetition and receptive vocabulary in 238 5-year old typically developing children and therefore questioned the direct involvement of phonological short term memory in vocabulary development. Botting and Conti-Ramsden (2001) studied the correlation between nonword repetition and various measures of language ability at ages 7 and 11 years and found no clear correlation between nonword repetition and receptive and expressive vocabulary measures. This lack of correlation between nonword repetition and vocabulary growth should not be expected if they both are constrained by phonological capacity, especially as one of the major arguments in PSTM account is the important role PSTM plays in learning new words (Baddeley et al., 1998; Gathercole, 2006).

This study reveals that nonword repetition scores in Gulf Arabic correlate with various psycholinguistic measures, such as digit span, sentence repetition, sentence comprehension, and expressive language skills. These patterns of correlations are highly consistent with those reported in many studies in different languages. The lack of correlation between nonword repetition and receptive vocabulary as reported in Gulf Arabic has also been reported in some English studies (e.g., Bowey, 1996 and Conti-Ramsden, 2001). Overall, the results are more consistent with the view that nonword repetition is not a mere measure of phonological short-term memory; rather it is a multi-component measure that taps into various processes, including phonological short term memory (PSTM). This study, however, suggests that PSTM may not be the central factor in determining children's performance on nonword repetition.

5.11.5 Error Analysis

The pattern of errors reported in this study is consistent with most of the studies that analysed errors in NWR tasks. Majority of errors were consonant substitutions, with cluster reduction appearing as a prominent type of error in this study, due to the design of the study where 75% of the stimuli contains either a medial, final, or medial *and* final clusters. Errors of substitution were dominant in most of the studies that analysed the types of errors exhibited by children in NWR (Marton & Schwartz, 2003; Edwards & Lahey, 1998; Montgomery, 1995). The lack of overlap in terms number of errors between the group of children with SLI on one hand and the two control groups (with exception of one outlier in the SLI group and one in LC group) shows that children with SLI presented with significant impairments in their phonological skills.

One type of error that is significantly higher in children with SLI is cluster creation, where children with SLI have 15 errors, while both control groups have a combined total of 1 error. However, it is difficult to reach a conclusive note on this due to the low numbers of this type of error; though the fact that some children with SLI produced more clusters than was required may counter the argument that these children have articulatory/output process deficits. Cluster creation errors in children with SLI were reported by Marshall and colleagues (Marshall et al., 2002; Marshall & van der Lely, 2009) who argue that cluster creation errors indicate there are qualitative differences between children with SLI and their TD peers in their phonological skills.

Overall, examination of the types of errors produced by children with SLI reveals no particular pattern of deficits, although their performance on the overall accuracy does not overlap with typically developing children. The SLI group produces more cluster creation errors than their typically developing controls, however, the lower rate of this kind of error makes it difficult to conclude that this is a definite characteristic of their phonological skills, but this may support the increasing evidence that their output processes are not particularly affected.

5.12 Conclusion and summary

This experiment examines the viability of nonword repetition as a clinical marker of specific language impairment in Gulf Arabic, which has not been studied before. Gulf Arabic speaking children with SLI performed significantly worse than their age and language matched controls on two and three syllable nonwords with different type of clusters. Their performance on the task deteriorates as the number of clusters increases.

Analysis of the data shows that their performance is not consistent with the phonological short-term memory account of SLI (Gathercole, 2006; Gathercole & Baddeley, 1990). Results reveal significant difference in performance of children with SLI even on two syllable nonwords, due to the presence of more phonologically complex structures. Therefore, deficits in phonological short-term memory alone cannot explain all the results of this task. The present findings are consistent with some other studies (e.g., Gallon et al., 2007; Marshall et al., 2002) that found that though phonological short-term memory plays a role, it seems that phonological complexity plays a more important one in children with SLI. The combination of increasing length and complexity poses the highest level of difficulty for children with SLI. Therefore, accounts that argue for deficits in phonological complexity (e.g., van der Lely, 2005) and phonological processing skills (e.g., Snowling et al., 2001, Chiat, 2001) present better explanations of the results, which show that there are various processes influencing the performance of children with SLI, such as the presence of consonant clusters and wordlikeness. Results show that articulatory complexity might not affect the performance of children with SLI. However, further analysis and research is required to test this hypothesis and to investigate other influencing factors, such as the influence of phonotactic probabilities of CV combinations.

Comparing the findings of the present study with those in other languages, such as English and Cantonese, shows that NWR findings are influenced by two major factors, the typology of the phonological system of a particular language and the design of the nonword repetition test. Therefore, results of nonword repetition tests should not be interpreted as reflecting phonological short-term memory only, but they are highly influenced by the typological differences in phonological systems across languages and the manipulation of other variables that have been found to affect performance on NWR tests, such as metrical structure, articulatory complexity, (sub)lexicality effects, and phonotactic probabilities.

The Arabic nonword repetition test scores correlate with various language measures, such as sentence comprehension, expressive language skills, sentence repetition, and digit-span. These findings are consistent with various studies in other languages. Interestingly, however, this nonword repetition task did not correlate significantly with receptive vocabulary score, while most studies in other languages found strong correlation between these two tasks. Therefore, the lack of strong correlation in this study is attributed to the nature of the nonword repetition task here, which seems to tap phonological complexity more than phonological short-term memory. These results of the first investigation of nonword repetition skills in Gulf Arabic speaking children with SLI are considered preliminary and further studies are needed to confirm these findings.

6. Summary and Conclusion

In this chapter I will summarise my findings, discuss theoretical and clinical implications of the results, and suggests directions for future research.

6.1 Summary of findings

The main aim of this study is to investigate syntactic and phonological complexity in Gulf-Arabic speaking children with SLI and examine their performance on two grammatical tasks with reference to current theories of SLI. Investigation of SLI in Gulf-Arabic can be of important theoretical value due to its syntactic and phonological differences from European languages in general and English in particular. For example, Arabic has a flexible word order, where canonical basic order, such as SVO, exists along other orders that have varying degrees of complexity, such as VSO, OSV, and OVS. These different word orders allow for manipulation of syntactic complexity, with little change in sentence length and therefore little increase in working memory demand. These properties of word orders in Gulf-Arabic are exploited in the experimental study of comprehension of reversible sentences with fronted NP's in order to examine how syntactic complexity influences the performance of children with SLI and compare its role with the role of working memory. Important issues related to working memory can also be studied through the distinctive phonological system of Arabic. The classic nonword repetition task can be seen from a novel perspective exploiting features such as root-and-pattern word composition, the fact that stress is related to the weight of syllable in such a way that syllables with clusters will usually receive stress and the existence of frequent and infrequent vocalic patterns, which will help in controlling wordlikeness. In Arabic it is difficult to have more than three syllables without adding extra morphemes, which meant that the current NWR test has two and three syllable nonwords only. All these phonological properties of Arabic are expected to have some bearing on the results on NWR test and they will elucidate the interaction between the different processes involved in the NWR skills. While the phonological short term memory theory argues that NWR is highly influenced by working memory skills (Gathercole & Baddeley, 1990), there is mounting evidence that there are language specific and test-specific factors that influence the performance of children with SLI on NWR. Therefore, studying NWR in languages with different phonological structures will help in investigating the role of various processes involved in NWR.

Since this investigation is the first to be done with this population and due to lack of normative data on typical and atypical language development of this variety of Arabic, a battery of language tests was developed and conducted with 88 typically developing children and 26 children with SLI, whose identification was verified by comparing their performance to age matched TD children. The tests developed during this project were: the Sentence Comprehension test, the Expressive Language test, the Sentence Repetition test and the Arabic Picture Vocabulary test. These tests are thoroughly explained in Chapter Three and they all possess good psychometric properties with satisfactory levels of reliability and validity. Thus they provide a suitable battery for identification of children with SLI, based on well established diagnostic criteria. A general overview of the performance of children with SLI on these tests shows a similar profile to language deficits seen in other languages, with children with SLI showing more deficits in expressive language skills compared to other language tasks. However, while children in other languages usually show relatively better scores on receptive vocabulary tests compared to other tests, this was not the case in children with SLI in Gulf Arabic, as they had low scores on the receptive vocabulary test compared to other measures. This is possibly attributed to the typological differences in Arabic, where good manipulation of a semantic root is required to derive words from the root.

While these tests provide general information about the profile of children with SLI acquiring Gulf Arabic as a native language, Chapters Four and Five investigate the grammatical abilities of these children more closely.

Chapter Four examines the comprehension of complex grammatical sentences in GA speaking children with SLI by comparing their performance to age and language controls. All participants listened to 54 reversible sentences with three different word orders and agreement cues. These comprised one canonical, unmarked word order (SVO), one word order that involves fronting the object to the initial position of the sentence (OSV_{CL}), and one more complex word order in which both the object and verb are fronted ($OV_{CL}S$). Both fronted word orders have a clitic that is coreferential with the fronted object. Gender agreement (masculine or feminine subject) is systematically manipulated in this experiment. Results of this experiment show that children with SLI scored significantly worse on the test when compared to both TD groups. Though children with SLI show good performance on canonical SVO sentences where their scores are not different from their age and language controls, their scores on the complex sentences that involve fronting is characteristically different. In sentences where only the object moves, their scores are significantly lower than the age control group, and in sentences with movement of both object and verb, their results are significantly different from both age and language controls. These significant difficulties in comprehension of these complex sentences are consistent with many studies that reported that children with SLI have problems understanding sentences with complex structures, such as passive sentences, object relative clauses, and object wh-questions.

Results of the Gulf Arabic nonword repetition test in Chapter Five show that children with SLI are particularly affected by the effects of clusters in nonwords. These children were presented with nonwords of two and three syllable lengths with clusters that ranged from none, to medial only, final only, and medial and final clusters. The scores of children with SLI on the NWR test significantly drop as the number of clusters increases from 0 to 1 and 2.The presence of a significant group by cluster interaction shows that children with SLI are differentially affected by the presence of syllabic complexity as they performed significantly less well than the two TD groups when clusters were present, but when compared on nonwords with no clusters, children with SLI were not significantly different from their language controls. There is no syllable number by group interaction, indicating that syllable length may have some role to play, but its role is not critical to children with SLI in Gulf Arabic. Moreover, results of the NWR test show that children with SLI are not significantly different from their language controls on measures of phoneme accuracy, indicating that they may not have significant deficits in their articulatory processes. However, there was an effect of wordlikeness; while TD children showed sensitivity to presence of common vocalic patterns in Gulf Arabic, children with SLI did not. These results show that although there is evidence supporting deficits in phonological complexity (consonant clusters) and phonological short term memory, the former but not the latter is what differentiates children with SLI from their age and language controls. When children with SLI were compared with language controls on clusterless two and three syllable words, the difference was not significant. These results seem to question the central role played by phonological short term memory and support the mounting evidence that NWR is not a pure measurement of PSTM. It is a multidimensional task that involves various factors, such as phonological (syllabic and metrical) complexity, phonological short term memory, phonotactic probability, wordlikeness, and articulatory processes. Comparisons with other tests of NWR in English and Cantonese support this conclusion.

6.2 Contributions to theory of SLI

The results of assessment tests used throughout the project and the experiments conducted can be of relevance to theories of language deficits in children with SLI. This thesis gives some insights on the nature of SLI in general, and in Gulf Arabic in particular by discussing how children performed on the various language tests developed for this project. Some of the results of children with SLI on these tests concur with those obtained in other languages, while others don't (e.g., results of the receptive vocabulary test). Moreover, the findings of the two experiments can contribute to theories concerning the underlying language deficits in children with SLI.

6.2.1 Implications on the nature of SLI

Comparing the performance of children with SLI on the four language tests can shed light on the impact of typological characteristics of Arabic on the performance of these children. Like most studies of SLI in other languages (see Leonard, 1998), the test battery developed for this project reveals that children with SLI have more deficits in their expressive language skills than they do in receptive language skills. However, the profile of Gulf Arabic speaking children with SLI is characterised by their significantly low scores on the Arabic Picture Vocabulary Test (APVT), which examines receptive vocabulary. Table 31 (page 125) shows that the overall performance of these children on the Arabic Picture Vocabulary Test looks similar to that of the Sentence Repetition test, with both of them having similar levels of difficulty. While English speaking children with SLI typically have significant deficits in sentence repetition (e.g., Conti-Ramsden et al., 2001; Roy & Chiat, 2004; Stokes et al., 2006), they usually present with relative strength in receptive vocabulary (Bishop, 1997; Leonard, 1998). However, Gulf Arabic speaking children with SLI presented with significantly low scores on receptive vocabulary test as well as on sentence repetition. Sentence repetition generally poses significant difficulties for children with SLI due to the many processes involved in this task, such as working memory, syntactic skills, and articulatory processes. Similarly, a multitude of processes are involved in the acquisition of Arabic vocabulary. Arabic is a root and pattern language, where one root is manipulated extensively to produce various items that are semantically related to the root. For example, the root 'k-t-b' (writing) is used to derive the verb 'katab' (he wrote), 'maktab' (office or desk), 'kitaab' (book), 'kaatib' (writer), 'maktaba' (library)…etc. Therefore, a child with good morphological and grammatical skills may identify the word 'maktaba' (library), based on its semantic root, even though she or he may not have encountered it before. In the receptive vocabulary test, it was observed that many TD children had difficulties identifying the picture of the word 'tˤa:biʕ' (post stamp); instead most of them went for the picture of a 'xatim' (rubber stamp), arguably because they were accessing the verb 'y-tˤabbiʕ' ('to rubber stamp'). Moreover, deriving words from roots and patterns requires efficient morphological and phonological skills in order to manipulate these roots and patterns. These manipulations require efficient working memory in order to facilitate long term representations of these derived words. Children with SLI acquiring such languages will be at great disadvantage due to their well-attested deficits in morphosyntax, morphology, phonology, and working memory. There might be other causes for these very low scores on vocabulary that may have to do with parenting style, which is very different from European countries, where children have less exposure to reading books or literacy activities at home. Moreover, in the Gulf, the home language of these children (Gulf Arabic) is not always widely spoken in the community, as most expatriates in the Gulf are not Arabic-speaking and children may not have rich exposure to their language outside their homes. Moreover, most households in Qatar and most other Gulf countries have domestic workers who are mostly non-Arabic speaking and children may spend significant amount of time interacting with these workers, therefore, limiting the quality of input they receive in their native languages. However, more research is required to investigate these factors.

This thesis has shown that the profile shown by GA speaking children with SLI on general language tests can be different from that seen in the two experiments investigating grammatical complexity in this population. Linguistic profiles of children with SLI based on their performance on the four language tests show that they are a very heterogeneous group of children. The following table shows the number of children with SLI who scored within the TD mean (i.e., within 1.5 standard deviations) on the four language tests. It is clear from this table that this group of children is very heterogeneous, with many of them scoring within normal range on some tests, especially the SC test (see Table 33 for more details).

Table 60: The number of children with SLI (n=26) who passed the four language tests.

Test	Number of children who scored within mean of TD children (1.5 SD)
The Sentence Comprehension (SC) test	12
The Expressive Language (EL) test	5
The Sentence Repetition (SR) test	8
The Arabic Picture Vocabulary Test (APVT)	8

This picture, however, is not necessarily reflected in the two experiments examining some aspects of complex grammar in this population. Five out of the 13 children with SLI who participated in the experiment examining comprehension of sentences with fronted NP's had within-normal scores on the Sentence Comprehension test. These children, however, had lower scores on other tests, which lead to including them in the SLI group. These results on the SC test are consistent with most studies of SLI in other languages, where children usually present with very heterogeneous profiles, with many of them scoring within normal range on some linguistic tasks, especially on receptive language tasks (Bishop, 1997; Leonard, 1998). However, on the sentence comprehension experiment, where children were required to comprehend some specific types of complex sentences, these children presented as a relatively homogeneous group. Only three of the children with SLI scored close to the mean of percentage of correct answers of the TD age control group, which was 78%. Children with SLI had scores that ranged between 35% and 74%, with only three children scoring above 70%. The language control group, who matched the SLI group on the Sentence Comprehension test, on the other hand, had five children who scored higher than the mean of the age control children, with four of them scoring 80% and higher. Moreover, their overall performance on the experiment is significantly better than the SLI group. This lack of overlap between the SLI and age control group on this sentence comprehension experiment is not expected if some of these children had normal comprehension skills, as indicated by their scores on the general Sentence Comprehension test. Therefore, part of the heterogeneity of SLI seen in various languages could be due to the nature of the tests used to measure the language skills of these children. These children show clear deficits in their comprehension of fronted NP's, with none of them scoring near the mean of their age controls. Therefore, employing tasks that contain complex grammatical structures, such as comprehension of short sentences with fronted NP's, might provide a better diagnostic tool to distinguish children with SLI from TD peers, as these grammatical structures seem be challenging to the inefficient grammatical system of these children. Other complex grammatical structures that might be implicated in children with SLI in Gulf Arabic are: production of different types of clitics and comprehension and production of relative clauses. These are proposed based on examining the performance of children with SLI on the four language tests developed during this project

Performance on the NWR test was broadly consistent; most children with SLI scored poorer than the mean of the TD group, with one outlier only. This very high performance on the NWR test by one child (ID 11610) is not attributed to selection criteria, which were very conservative. She was one of the children who participated in the sentence comprehension experiment and her score was lower than the mean of the SLI group on that task, even though her score on the general Sentence Comprehension test was within the mean of TD children (see Table 32). Her digit span score was 8.1, while the mean for the AC group was 11.1 and it was 9.1 for the LC group, therefore her high performance of NWR is not explained by superior short term memory. She was diagnosed with SLI based on her score on the Expressive Language test, which was two standard deviations below the mean, and she scored within the mean on the three remaining tests. Therefore, this child presented with good NWR skills, in the presence of poor syntactic and morphosyntactic abilities. This weak link between syntactic abilities and NWR skills was also reported by Norbury et al. (2001) and further supports the notion that these two skills are linked to different genetic factors (Bishop et al., 2006). Moreover, a lack of correlation between NWR and receptive vocabulary has been reported in an English study (Edwards & Lahey, 1998) and in Swedish speaking children with SLI (Sahlen et al. 1999). The current study also produced a low outlying NWR score in the TD group. This individual performed within the range of the SLI group. This poor score is consistent with studies showing that some children with poor nonword repetition skills have normal scores on language tests (Gathercole et al., 2005). Gathercole (2006) tried to explain this by stating that poor phonological short term memory skills can cause SLI only if there are concomitant cognitive risk factors, though this was not clearly explained.

The combined findings of one of the general tests and the sentence comprehension experiment might shed some light on the status of agreement in Arabic speaking children with SLI. The results of the Expressive Language test show that children with SLI did not have significant problems with agreement markers. Moreover, analysis of the performance of these children on the comprehension of sentences with fronted NP's shows that they benefitted from gender agreement and there was no group by agreement interaction. The presence of these relatively good skills on agreement markers is consistent with studies that investigated the performance of Hebrew speaking children with SLI on inflectional morphemes (Dromi et al., 1993; Dromi et al., 2003). Current findings, however, are incongruent with those of Abdalla (2002), who reported that younger Saudi children with SLI (aged between 3-5 years) showed significant deficits in agreement based on spontaneous language samples. However, age differences and differences in elicitation and the task may explain these differences. Therefore, further investigation of this should clarify that status of agreement markers across different ages in Arabic speaking children with SLI.

The findings of both the sentence comprehension experiment and the NWR experiment link SLI in Gulf Arabic to deficits in grammatical complexity, with less evidence supporting the argument that SLI is caused by limitations in working memory. These children seem to have significant deficits in comprehension of reversible sentences with fronted NP's, especially when both the verb and the object are moved. Children with SLI acquiring Gulf Arabic are especially affected by the presence of complex movement, in a manner that seems similar to the results obtained in comprehension of reversible passives in English (van der Lely & Harris, 1990) and in object relative clauses in Hebrew (Friedmann & Novogrodsky, 2004). All these syntactic structures are characterised by complex movement that leads to noncanonical order. In all of these complex syntactic structures children with SLI assign the role of agent to the first NP they encounter in the sentence due to their impaired grammar. In the NWR experiment, grammatical complexity is manifest in the presence of consonant clusters, with nonwords with more consonant clusters posing higher level of difficulties, especially for children with SLI. Moreover, results of the four language tests reveal that Gulf Arabic children with SLI may have significant deficits with other complex grammatical structures, especially production of clitics, relative clauses, and irregular plurals that involve morphophonological manipulation. This is based on general language tests and therefore these conclusions warrant more in-depth examination of these structures and with a larger number of children.

6.2.2 Implications regarding theories of SLI

Both the syntactic and phonological experiments found that limitations in general processing may not adequately account for the performance of children with SLI on the sentence comprehension experiment and the nonword repetition test.

The comprehension of reversible sentences with fronted NP's in Gulf Arabic provides a useful contribution to the study of complex syntactic structures in English and other languages by studying a syntactic structure that has not been investigated in English. Most of the studies so far have looked at structures, such as reversible passives, relative clauses, object wh-questions. However, this study looked at three different reversible sentences, one with SVO sentences, and two types where the object is fronted (OSV and OVS sentences). The results of this sentence comprehension study show that children with SLI, like TD controls, found the SVO sentences easier to understand. However, they had significant problems with OSV sentences where only the object is fronted. Their scores on these sentences were significantly lower than their age controls. On OVS sentences that involve fronting of both objects and verbs, they scored substantially less well than their age and language control children, who matched the SLI group on scores of a general sentence comprehension test. Therefore, this shows that the increase in the number of movements leads to significant decrease in the performance of SLI children. These results are consistent with grammatical accounts that attribute these deficits in sentence comprehension to the presence of grammatical complexity as defined in terms of movement. The results are not congruent with processing accounts that claim limitations in working memory and general processing capacity cause these deficits in sentence comprehension, since both fronted sentences have similar processing loads. Like TD children, children with SLI use Subject-first strategy, however, while TD children employ movement strategies and agreement strategy, children with SLI use Subject first as their default strategy across the three types of sentences, though they also employ agreement in OSV sentences as these sentences have a different configuration (NNV) that may facilitate the use of agreement cues. The performance of children with SLI on this task was qualitatively different from TD groups as they are the only group that scored better on SVO sentences than they did on OSV sentences. Overall, these deficits in children with SLI are mostly caused by the presence of complex word orders, as their performance on agreement did not set them apart from the other two TD groups.

On the NWR task, the results clearly show that though phonological short term memory might have a role to play in NWR skills of children with SLI, it is far from being the main cause of NWR deficits (Gathercole & Baddeley, 1990) or having a central role in NWR deficits (Gathercole, 2006). This study reveals a bigger role for phonological complexity as defined in terms of presence of clusters than the role played by syllable length. The difference between children with SLI and their language controls disappears when these two groups are compared on clusterless words, while this difference is substantial on words with clusters, especially when there are two clusters in a nonword. Moreover, this study found evidence for wordlikeness effects. Children's scores on a frequent vocalic pattern exceeded their performance on the experimental patterns that are less frequent in Gulf Arabic. This supports the mounting evidence that NWR is a multidimensional test and it is not a pure measure of phonological short term memory. This finding is consistent with other studies that have found significant effects of the syllabic complexity (Gallon et al., 2007; Marshall & van der Lely, 2009), metrical complexity (Chiat & Roy, 2007; Marshall et al., 2002; Marshall & van der Lely, 2009; Roy & Chiat, 2004)), phonotactic probabilities and wordlikeness (Dollaghan et al., 1993; Gathercole et al., 1991 Stokes et al., 2006).

6.3 Clinical implications

This thesis has some general clinical implications related to assessment and treatment of children with SLI, and provides more specific insights on working with Arabic speaking children with SLI, especially in Gulf-Arabic speaking countries.

The general language tests conducted with this population demonstrate that children with SLI have very heterogeneous linguistic characteristics. Many of these children showed performance within the TD range on some tests, while they had significant problems with others. Therefore, in order to identify children with SLI more reliably, clinicians and researchers must administer a variety of tests tapping into different areas of language skills. This thesis follows the model presented by Tomblin et al., 1997 and it used a battery of four language tests to measure major linguistic abilities of these children. Using one or two tests only to identify children with SLI, would have resulted in some errors of selection. This becomes clear when the performance of children with SLI, who passed the Sentence Comprehension (SC) test and were included in the SLI group after they met other criteria, is compared with their scores on the NP fronting experiment. These children, who had within average scores on the SC test, did as poorly as the other children with SLI who failed the SC test. Therefore, general language tests, especially if only one or two of them are used, may not be sufficient to identify children with SLI who are known to have a heterogeneous profile on these general tests. However, if these tests are combined with more specific tasks that have been indicated in research in a specific language as a major area of deficits in children with SLI; this may reduce the number of false negatives. For example in English both NWR, tense, object relative clauses and object wh-questions have been found to be particularly difficult for children with SLI and therefore, some of these linguistic structures should be examined more thoroughly along with general language tests to help in identifying children with SLI.

Clinicians and researchers working with children with SLI in Gulf-Arabic can employ the tests developed in this project to help in identification of children with SLI. These tests assess different linguistic skills (sentence comprehension, expressive language, sentence repetition, and vocabulary) and can therefore be of empirical value for clinicians and researchers alike. Analysis of these tests shows that they have good psychometric properties with good levels of reliability and validity. However, the caveat mentioned about the general tests should be heeded too in Gulf-Arabic and clinicians and researchers are advised to supplement these tests with one or more grammatically complex tasks that provide more in depth examination of important linguistic constructs, such as NWR or comprehension of sentences with fronted NP's, as both show good ability to differentiate between children with SLI and their typically developing peers.

The Arabic NWR test developed in this thesis has an extra advantage in that it could be used with other dialects of Arabic with little or no change. The test consists of two and three syllable nonwords with clusters that range from 0-2 and all the sounds are early developing sounds that can be found in most varieties of Arabic. Moreover, in varieties of Arabic where final consonant clusters are not as common as they are in Gulf-Arabic, medial clusters can be used as this study demonstrates that there was no significant difference between the performance of all children on nonwords with medial or final clusters.

Moreover, the results of the two experiments indicate that both comprehension of sentences with fronted NP's and the NWR test can be used to identify children with SLI, whose performance was different not only from their TD age controls, but also from their language controls. These can be of great clinical significance, especially since there is a lack of screening and assessment tools for children with language impairment in Gulf Arabic and other varieties of Arabic. Furthermore, other linguistic structures that children with SLI might have significant problems with are: production of clitics, comprehension of (truncated) passives, relative clauses, and irregular plurals. These were found to pose significant challenges for children with SLI on the language tests developed for this project. However, not all of these structures were well represented in the tests and therefore these findings are preliminary and warrant further investigations in the future. These structures, if found to be challenging for children with SLI, should be considered when planning treatment of language disorders in this population.

6.4 Limitations and directions for future research

This is the first project to investigate SLI in Gulf Arabic and therefore some findings warrant replication with larger samples of children, especially the tests developed during this project. Furthermore, both NWR and comprehension of sentences with complex word order were problematic for children with SLI and they warrant further investigation.

This study involved developing four language tests that were administered with 88 TD children and 26 children with SLI, aged between 4;6 and 9;3 years old. The initial aim was to conduct these tests with 30 children of age groups 5;0-8;11 years old, however it was not possible to do this due to difficulties with data collection and the time frame of this project. Nevertheless, the results of each test show that they all possess good psychometric properties and they could form a basis for a bigger project where these tests are revised and conducted with a representative sample of Qatari children. Moreover, these tests could be administered with children belonging to different age groups (younger than 5 years old and older than 8 years old), though this will probably involve adding more test items. The item analysis and reliability tests identified some items that will require modification or omission in future versions and some will require changes in the pictures. Finally adding some qualitative measures of language, such as taking language samples during testing, may supplement the tests and provide more information on how language deficits are impacting on these children's communication skills. However, this may prove to be time consuming and will require more resources.

The sentence comprehension experiment, which involves listening to canonical and noncanonical sentences where the object is fronted to the initial position of the sentence, may warrant replication. This task appears to differentiate children with SLI from both age and language controls. However, this task may not be suitable for use with children less than 5 years old; if reversible sentences are used as has been the case in this experiment. Al-Akeel (1986) has reported the comprehension of reversible SVO sentences is mastered by the age of five and this study has shown that younger TD children had lower scores on the more complex sentences (OVS sentences that involve fronting of the object and the verb), which means that the comprehension of this complex structure is not mastered by the age of 5. However, this task might prove very useful in older children, since there was no ceiling effect (especially on OVS sentences). Moreover, it would be interesting to add other types of word orders that exist in Gulf Arabic, such as VSO, where sentences do not begin with an argument, like the three types used in the current experiment. The results of Aljenaie and Farghal (2009) experiment indicates that in terms of hierarchy of difficulty, VSO sentences will be somewhere between the canonical SVO sentences, and OSV sentences.

Children with SLI showed significant deficits on the NWR task and their performance was qualitatively different from TD age and language controls, where they were more affected by the presence of clusters than their TD age and control peers. The results show stronger effects of clusters than syllable length. Wordlikeness was found to affect performance of all children as they performed better when the nonwords consisted of regular patterns. It was not possible to control for lexicality effects as 13% of the nonwords are possible words, though may be not in these children's lexicon. Therefore, a smaller version of nonword list could be developed that does not include these CV combinations. Moreover, this short list (e.g., 16 or 24 nonwords) could be contrasted with another one where these CV combinations are used to investigate the effects of phonotactic probabilities in Arabic. Moreover, the almost identical performance of all children on medial only clusters and final only clusters means one of them can be used in the shorter version of the NWR test.

The reported heterogeneity of SLI may warrant further investigation in Arabic and other languages, however, more attention should be paid to the tools used to assess and diagnose children with SLI. Results of the comparisons between some children's scores on the general Sentence Comprehension test, and their scores on the comprehension of complex sentences that involve fronting of object, show that most of the children with SLI who passed the Sentence Comprehension test scored significantly less well than age and language controls on the sentences with fronted NP's. This shows that some of the heterogeneity of SLI might be attributable the nature of general language test. Therefore, the use of tests that target specific linguistic structures known to pose difficulties for children with SLI might be more useful in identification of children with SLI. A good example of these tests is the TROG (Bishop, 1989), which examines the comprehension of many types of complex sentences, or TOPhS (van der Lely & Harris, 1999).

Finally, a more specific definition of grammatical complexity is warranted. Results of both experiments show that there are quantitative and qualitative differences between children with SLI and their TD peers on the two grammatically complex structures investigated in this project, namely; reversible sentences with fronted NP's, and NWR with varying length and clusters. In both experiments, children with SLI were especially affected by the increase in complexity from one movement to two movements (in the sentence comprehension experiment) and the increase in number of clusters in a nonword from 0 to 1 to 2 (in the NWR test). Moreover, in the sentence comprehension experiment, children with SLI showed a clear preference for canonical SVO sentences, while the TD groups did not. In NWR, these children's performance dramatically decreased on nonwords with two clusters, showing increased vulnerability to the presence of more clusters. Moreover, any definition of grammatical complexity should consider the typological properties of each language. This study shows that agreement might not be difficult for Gulf-Arabic speaking children with SLI, while they might have significant difficulties with the receptive vocabulary test, possibly due to morphophonological complexity. Many European languages, especially English, reported opposite trends with children with SLI having more difficulties with agreement and fewer problems with receptive vocabulary. Therefore, this may prove challenging to any theory of SLI that is based on one language. Definitions of grammatical complexity should pay attention to language-specific characteristicss. Moreover, accounts of grammatical complexity should be able to provide a hierarchy of difficulty of the various factors causing this complexity, e.g., whether syntactic movement causes more complexity than anaphoric relationships and whether syllabic complexity is more detrimental than metrical complexity. And the interaction among various complexity factors should be clearly spelled out.

References

Adams, A. M., & Gathercole, S. E. (1996). Phonological working memory and spoken language development in young children. *Quarterly Journal of Experimental Psychology, 49A*, 216-233.

Adams, C. (1990). Syntactic comprehension in children with expressive language impairment. *British Journal of Disorders of Communication, 25*, 149-172.

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.

Aljenaie, K., & Farghal, M. (2009). Comprehension of three word orders in Kuwaiti Arabic child language. *Language Sciences, 31*, 494-506.

Aoun, J., & Benmamoun, E. (1998). Minimality, reconstruction, and PF movement. *Linguistic Inquiry, 29*, 569-597.

Aoun, J., Benmamoun, E., & Sportiche, D. (1994). Agreement and Conjunction in Some Varieties of Arabic. *Linguistic Inquiry, 25*, 195–220.

Aoun, J., Choueiri, L., & Hornstein, N. (2001). Resumption, movement, and derivational economy. *Linguistic Inquiry, 32*, 371-403.

Aram, D., Morris, R., & Hall, N. (1993). Clinical and research congruence in identifying children with specific language impairment. *Journal of Speech and Hearing Research, 36*, 580-591.

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

Archibald, L., & Gathercole, S. (2006b). Short-term and working memory in specific language impairment. *International Journal of Language and Communication Disorders, 41*, 675-694.

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.

Barry, J. G., Yasin, I., & Bishop, D. V. M. (2007). Heritable risk factors associated with language impairments. *Genes, Brain and Behavior, 6*, 66-76.

Bedore, L., & Leonard, L. B. (1998). Specific language impairment and grammatical morphology: A discriminant function analysis. *Journal of Speech, Language, and Hearing Research, 41*, 1185-1192.

Bedore, L., & Leonard, L. B. (2005). Verb inflections and noun phrase morphology in the spontaneous speech of Spanish-speaking children with specific language impairment. *Applied Psycholinguistics, 26*, 195-225.

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

Bishop, D. V. M. (1979). Comprehension in developmental language disorders. *Developmental Medicine and Child Neurology, 21*, 225-238.

Bishop, D. V. M. (1989). *Test for Receptive of Grammar (TROG*), 2nd ed. Manchester: University of Manchester.

Bishop, D. V. M. (1990). Handedness, clumsiness and developmental language disorders. *Neuropsychologia, 28*, 681-690.

Bishop, D. V. M. (1994). Is specific language impairment a valid diagnostic category? Genetic and psycholinguistic evidence. *Philosophical Transactions of the Royal Society Lond. B., 346*, 105-111.

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

Bishop, D. V. M. (2000). Pragmatic language impairment: A correlate of SLI, a distinct subgroup, or part of the autistic continuum. In D. V. M. Bishop & L. B. Leonard (Eds.), *Speech and language impairments in children: Causes, characteristics, intervention and outcome* (pp 99-113). Hove, Sussex, UK: Psychological press.

Bishop, D. V. M. (2004). Specific language impairment: diagnostic dilemmas. In L. Verhoeven & H. van Balkom (Eds.), *Classification of developmental language disorders* (pp. 309-326). Mahweh, New Jersey: Lawrence Erlbaum.

Bishop, D. V. M. (2006). Beyond words: Phonological short-term memory and syntactic impairment in specific language impairment. *Applied Psycholinguistics, 27*, 545-547.

Bishop, D. V. M. (2007). What causes SLI in children? In N. J. Cohen (Ed.), *Encyclopaedia of Language and Literacy Development* (pp. 1-7). London, ON: Canadian Language and Literacy Research Network, retrieved [02/10/2007]from http://www.literacyencyclopedia.ca/pdfs/topic.php?topId=

Bishop, D. V. M., & Adams, C. (1992). Comprehension problems in children with specific language impairment: Literal and inferential meaning. *Journal of Speech and Hearing Research, 35*, 119-129.

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 Behavior, 5*, 158-169.

Bishop, D. V. M., Bishop, S., Bright, P., James, C., Delaney, T., & Tallal, P. (1999). Different origin of auditory and phonological processing problems in children with language impairment: evidence from a twin study. *Journal of Speech, Language, and Hearing Research, 42*, 155-168.

Bishop, D. V. M., Bright, P., James, C., & van der Lely, H. K. J. (2000). Grammatical SLI: A distinct subtype of developmental language disorder. *Applied Psycholinguistics, 21* ,159-181.

Bishop, D. V. M., Carlyon, R. P., Deeks, J. M., & Bishop, S. (1999). Auditory temporal processing impairment: neither necessary nor sufficient for causing language impairment in children. *Journal of Speech, Language, and Hearing Research, 42*, 1295-1310.

Bishop, D. V. M., & McArthur, G. M. (2005). Individual differences in auditory processing in specific language impairment: a follow-up study using event-related potentials and behavioural thresholds. *Cortex, 41*, 327-341.

Bishop, D. V. M., North, T., & Donlan, C. (1995). Genetic basis of specific language impairment: evidence from a twin study. *Developmental Medicine & Child Neurology, 37,* 56-71.

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.

Boersma, P., & Weenink, D. (2004). Praat: doing phonetics by computer (Version 4.2. 28)[Computer program]: Institute of Phonetic Sciences, University of Amsterdam.

Borer, H., & Wexler, K. (1987). *The role of maturation in the explanation of linguistic development*. Dordrecht: Reidel.

Borer, H., & Wexler, K. (1992). Bi-unique relations and the maturation of grammatical principles. *Natural Language and Linguistic Theory, 10*, 147-189.

Borman, G. D., & Benson, J. (2006). *Can brain research and computers improve literacy? A randomized field trial of the Fast ForWord® Language computer-based training program*. WCER Working Paper No. 2006-5, Madison: University of Wisconsin-Madison, Wisconsin Center for Education Research. Retrieved [September 15, 2007] from www.wcer.wisc.edu/ publications/workingPapers/papers.php.

Bortolini, U., Casalini, C., & Leonard, L. B. (1997). Grammatical deficits in Italian-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 40*, 809-820.

Bortolini, U., Caselli, M. C., Deevy, P., & Leonard, L. B. (2002). Specific language impairment in Italian: the first steps in the search for a clinical marker. *International Journal of Language and Communication Disorders, 37*, 77-93.

Bortolini, U., & Leonard, L. B. (1996). Phonology and grammatical morphology in specific language impairment: Accounting for individual variation in English and Italian. *Applied Psycholinguistics, 17*, 85-104.

Bortolini, U., & Leonard, L. B. (2000). Phonology and children with specific language impairment: status of structural constraints in two languages. *Journal of Communication Disorders, 33*, 131-149.

Bortolini, U., Leonard, L. B., & Casalini, C. (1998). Specific language impairment in Italian and English: Evaluating alternative accounts of grammatical deficits. *Language and Cognitive Processes, 13*, 1-20.

Botting, N. (2004). Children's Communication Checklist (CCC) scores in 11-year-old children with communication impairments. *International Journal of Language and Communication Disorders, 39*. 215-227.

Botting, N., & Conti-Ramsden, G. (2001). Non-word repetition and language development in children with specific language impairment (SLI).

International Journal of Language and Communication Disorders, 36, 421-432.

Botting, N., & Conti-Ramsden, G. (2004). Characteristics of children with specific language impairment. In L. Verhoeven & H. van Balkom (Eds.), *Classification of developmental language disorders* (pp. 23-38). Mahweh, New Jersey: Lawrence Erlbaum.

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

Brown, L., Sherbenou, R. J., & Johnsen, S. K. (1997). TONI-3: Test of nonverbal intelligence. *Wood Dale, IL: Stoelting Co*.

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

Bulk, A. (2008). Word order, clitics, and agreement in Arabic information strucutre. In A. Steube (Ed.), *The discourse potential of underspecified structures*. Berlin, New York: Walter de Gruyter.

Burt, L., Holm, A., & Dodd, B. (1999). Phonological awareness skills of 4-year-old British children: an assessment and developmental data. *International Journal of Language & Communication Disorders, 34*, 311 - 335.

Campbell, T., Dollaghan, C., Needleman, H., & Janosky, J. (1997). Reducing bias in language assessment: Processing-dependent measures. *Journal of Speech and Hearing Research*, 40, 519-525.

Casalini, C., Brizzolara, D., Chilosi, A., Cipriani, P., Marcolini, S., Pecini, C., et al. (2007). Non-word repetition in children with specific language impairment: a deficit in phonological working memory or in long-term verbal knowledge? *Cortex*, *43*, 769-776.

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. (2001). Mapping theories of developmental language impairment: Premises predictions and evidence. *Language and Cognitive Processes*, *16*, 113-142.

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.

Chomsky, N. (1981). Principles and parameters in syntactic theory. In N. Hornstein & D. Lightfoot (Eds.), *Explanation in linguistics: The logical problem of language acquisition* (pp. 32-75). London: Longman.

Chomsky, N. (1986). *Knowledge of Language: Its Nature, Origin and Use*. NY: Praeger.

Chomsky, N. (1995). *The minimalist program*. Cambridge Mass: MIT press.

Cinque, G. (1990). *Types of A-dependencies*: MIT Press.

Clahsen, H. (1989). The grammatical characterization of developmental dysphasia. *Linguistics, 27*, 897-920.

Clahsen, H., Bartke, S., & Goellner, S. (1997). Formal features in impaired grammars: A comparison of English and German SLI children. *Journal of Neurolinguistics, 10,* 151-171.

Clahsen, H., & Dalalakis, J. (1999). Tense and agreement in Greek SLI: A case study. *Essex Research Reports in Linguistics*, *24*, 1-25.

Coady, J., Evans, J. L. & Kluender, K. R. (2010). Role of phonotactic frequency in nonword repetition by children with specific language impairments. *International Journal of Language & Communication Disorders*, *45*, 494-509.

Cohen, W., Hodson, A., O'Hare, A., Boyle, J., Durrani, T., McCartney, E., et al. (2005). Effects of computer-based intervention through acoustically modified speech (Fast ForWord) in severe mixed receptive-expressive language impairment: outcomes from a randomized controlled trial. *Journal of Speech, Language, and Hearing Research, 48,* 715-729.

Conti-Ramsden, G. (2003a). Methodological concerns: An introduction. In Y. Levy & J. Schaeffer (Eds.), *Language competence across populations: toward a definition of specific language impairment* (pp. 197-208). Mahwah, New Jersey: Lawrence Erlbaum.

Conti-Ramsden, G. (2003b). 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., & Botting, N. (1999). Classification of children with specific language impairment: longitudinal considerations. *Journal of Speech, Language, and Hearing Research*, *42*, 1195-1204.

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., Simkin, Z., & Pickles, A. (2006). Estimating familial loading in SLI: A comparison of direct assessment versus parental interview. *Journal of Speech, Language, and Hearing Research, 49*, 88-101.

Conti-Ramsden, G., Crutchley, A., & Botting, N. (1997). The extent to which psychometric tests differentiate subgroups of children with SLI. *Journal of Speech, Language, and Hearing Research*, 40, 767-777.

Corriveau, K., Pasquini, E., & Goswami, U. (2007). Basic Auditory Processing Skills and Specific Language Impairment: A New Look at an Old Hypothesis. *Journal of Speech, Language, and Hearing Research, 50*, 647-666.

Crago, M., & Paradis, J. (2003a). Two of a kind? The importance of commonalities and variation across languages and learners? In Y. Levy & J. Schaeffer (Eds.), *Language competence across populations: toward a definition of specific language impairment* (pp. 97-110). Mahwah, New Jersey: Lawrence Erlbaum.

Crago, M. B., & Allen, S. E. M. (1996). Building the case for impairment in linguistic representation. In M. L. Rice (Ed.), *Toward a genetics of language* (pp. 261-289). Mahwah, NJ: Lawrence Erlbaum Associates.

Craig, H. K. (1993). Social skills of children with specific language impairment: Peer relationships. *Language, Speech, and Hearing Services in Schools, 24,* 206-215.

Craig, H. K., & Evans, J. (1993). Pragmatics and SLI: Within-group variations in discourse behaviors. *Journal of Speech and Hearing Research, 36*, 777-789.

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

De Cat, C. (2003). Syntactic manifestations of very early pragmatic competence. *Proceedings of the 27th Annual Boston University Conference on Language Development*, 209-219.

de Jong, J. (1999). *Specific language impairment in Dutch: Inflectional morphology and argument structure.* Unpublished Ph.D., University of Groningen.

de Jong, J. (2003). Specific language impairment and linguistic explanation. In Y. Levy & J. Schaeffer (Eds.), *Language competence across populations: toward a definition of specific language impairment* (pp. 151-170). Mahwah, New Jersey: Lawrence Erlbaum.

de Villiers, J. G. (2003). Defining SLI: a linguistic perspective. In Y. Levy & J. Schaeffer (Eds.), *Language competence across populations: toward a definition of specific language impairment* (pp. 425-447). Mahwah, New Jersey: Lawrence Erlbaum.

Deevy, P., & Leonard, L. B. (2004). The comprehension of Wh-Questions in children with specific language impairment. *Journal of Speech Language and Hearing Research, 47*, 802-815.

Dollaghan, C. (1987). Fast mapping in normal and language-impaired children. *Journal of Speech and Hearing Disorders, 52*, 218-222.

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

Dromi, E., & Berman, R. A. (1982). A morphemic measure of early language development: data from modern Hebrew. *Journal of Child Language, 9*, 403-424.

Dromi, E., Leonard, L. B., Adam, G., & Zadunaisky-Ehrlich, S. (1999). Verb agreement morphology in Hebrew-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 42,* 1414-1431.

Dromi, E., Leonard, L. B., & Blass, A. (2003). Different methodologies yield incongruous results: a study of the spontaneous use of verb forms in Hebrew. In Y. Levy & J. Schaeffer (Eds.), *Language competence across populations: toward a definition of specific language impairment* (pp. 273-289). Mahwah, New Jersey: Lawrence Erlbaum.

Dromi, E., Leonard, L. B., & Shteiman, M. (1993). The grammatical morphology of Hebrew-speaking children with specific language impairment: some competing hypotheses. *Journal of Speech and Hearing Research, 36*, 760-771.

Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). *The British Picture Vocabulary Scale: Second Edition*. Windsor: NFER-Nelson publishing Company Ltd.

Ebert, K. D., Kalanek, J., Cordero, K. N., & Kohnert, K. (2008). Spanish nonword repetition: stimuli development and preliminary results. *Communication Disorders Quarterly, 29*, 67-74.

Edwards, J., & Lahey, M. (1996). Auditory lexical decisions of children with specific language impairment. *Journal of Speech and Hearing Research, 39*, 1263-1273

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.

Elman, J., Bates, E., Johnson, M., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, Mass.: MIT press.

Engel, P. M. J., Santos, F. H., & Gathercole, S. E. (2008). Are Working Memory Measures Free of Socioeconomic Influence? *Journal of Speech, Language, and Hearing Research, 51*, 1580-1587.

Evans, J. L., & MacWhinney, B. (1999). Sentence processing strategies in children with expressive and expressive-receptive Specific Language Impairments. *International Journal of Language and Communication Disorders, 34*, 117-134.

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

Fey, M., & Leonard, L. B. (1983). Pragmatic skills of children with specific language impairment. In T. Gallagher & C. Prutting (Eds.), *Pragmatic assessment and intervention issues in language* (pp. 65-82). Baltimore: University Park Press.

Fisher, S. E., Lai, C. S. L., & Monaco, A. P. (2003). Deciphering the genetic basis of speech and language disorders. *Annual. Review of Neuroscience*, *26*, 57-80..

Fisher, S. E., Vargha-Khadem, F., Watkins, K. E., Monaco, A. P., & Pembrey, M. E. (1998). Localisation of a gene implicated in severe speech and language disorder. *Nature Genetics, 18*, 168-170.

Fodor, J. A. (1983). *Modularity of mind*. Cambridge, MA: MIT Press.

Ford, D. (2009). The influence of word order on Modern Standard Arabic information structure [Electronic Version]. *Graduate Institute of Applied Linguistics Electronic Notes Series*, 3, 1-12. Retrieved January 5, 2010 from http://www.gial.edu/GIALens/issues.htm.

Franck, J., Cronel-Ohayon, S., Chillier, L., Frauenfelder, U. H., & Hamann, C. (2004). Normal and pathological development of subject-verb agreement in speech production: a study on French children. *Journal of Neurolinguistics, 17*, 147-180.

Friedmann, N., Gvion, A., & Novogrodsky, R. (2006). Syntactic movement in agrammatism and S-SLI: Two different impairments. In E. B. A. Belletti, C. Chesi, E. D. Domenico & I. Ferrari (Eds.), *Language acquisition and development* (205-218). Cambridge, UK: Cambridge Scholars Press/CSP.

Friedmann, N., & Novogrodsky, R. (2004). The acquisition of Relative clause comprehension in Hebrew: A study of SLI and normal development. *Journal of Child Language, 31*, 661-681.

Friedmann, N., & Novogrodsky, R. (2007). Is the movement deficit in syntactic SLI related to traces or to thematic role transfer? *Brain and Language, 101*, 50-63.

Frisch, S. A., Pierrehumbert, J. B., & Broe, M. B. (2004). Similarity avoidance and the OCP. *Natural Language & Linguistic Theory, 22*, 179-228.

Frisch, S. A., & Zawaydeh, B. A. (2001). The psychological reality of OCP-Place in Arabic. *Language, 71*, 91-106.

Gallon, N., Harris, J., & van der Lely, H. K. J. (2007). Non-word repetition: An investigation of phonological complexity in children with Grammatical SLI. *Clinical Linguistics and Phonetics, 21*, 435 - 455.

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. 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. (1990). 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. (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.

Gathercole, S. E., Pickering, S. J., Hall, M., & Peaker, S. M. (2001). Dissociable lexical and phonological influences on serial recognition and serial recall. *The Quarterly Journal of Experimental Psychology Section A, 54*, 1-30.

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., 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.

Gathercole, V. C. M., & Thomas, E. M. (2005). Minority language survival: input factors influencing the acquisition of Welsh. In K. T. M. J. Cohen, K. Rolstad, and J. MacSwan (Ed.), *Proceedings of the 4th International Symposium on Bilingualism* (pp. 852-874). Somerville, MA: Cascadilla Press.

Gillam, R. B., Crofford, J. A., Gale, M. A., & Hoffman, L. M. (2001). Language change following computer-assisted language instruction with Fast ForWord or Laureate Learning Systems Software. *American Journal of Speech-Language Pathology, 10*, 231-247.

Gillam, R. B., Frome Loeb, D., & Friel-Patti, S. (2001). Looking back: A summary of five exploratory studies of Fast Forword. *American Journal of Speech-Language Pathology, 10*, 269-273.

Gillam, R. B., & Hoffman, L. M. (2004). Information processing in children with specific language impairment. In L. Verhoeven & H. van Balkom (Eds.), *Classification of developmental language disorders* (pp. 137-157). Mahweh, New Jersey: Lawrence Erlbaum.

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

Goorhuis-Brouwer, S., Coster, F., Nakken, H., & Spelberg, H. L. (2004). Environmental factors in developmental language disorders. In L. Verhoeven & H. van Balkom (Eds.), *Classification of developmental language disorders* (pp. 158-171). Mahweh, New Jersey: Lawrence Erlbaum.

Gopnik, M. (1990). Feature blind grammar and dysphasia. *Nature, 344*, 715.

Gopnik, M., & Crago, M. (1991). Familial aggregation of a developmental language disorder. *Cognition, 39*, 1-50.

Gray, S. (2003a). 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.

Gray, S. (2003b). Word-learning by preschoolers with specific language impairment: what predicts success? *Journal of Speech, Language, and Hearing Research, 46*, 56-67.

Greenberg, J. (1950). The patterning of root morphemes in Semitic. *Word, 5*, 162-181.

Hadley, P. A., & Rice, M. L. (1991). Conversational Responsiveness of Speech- and Language-Impaired Preschoolers. *Journal of Speech and Hearing Research, 34*, 1308-1317.

Hamann, C., Ohayon, S., Dube, S., Frauenfelder, U. H., Rizzi, L., Starke, M., et al. (2003). Aspects of grammatical development in young French children with SLI. *Developmental Science, 6*, 151-158.

Hanson, R. A., & Montgomery, J. W. (2002). Effects of general processing capacity and sustained selective attention on temporal processing performance of children with specific language impairment. *Applied Psycholinguistics, 23*, 75-93.

Hansson, K., Forsberg, J., Löfqvist, A., Mäki-Torkko, E., & Sahlén, B. (2004). Working memory and novel word learning in children with hearing impairment and children with specific language impairment. *International Journal of Language & Communication Disorders, 39*, 401-422.

Haskill, A. M., & Tyler, A. A. (2007). A Comparison of Linguistic Profiles in Subgroups of Children With Specific Language Impairment. *American Journal of Speech-Language Pathology, 16*, 209-221.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science, 298*, 1569-1579.

Haynes, S. N., Richard, D. C. S., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approah to concepts and methods. *Psychological Assessment, 7*, 238-247.

Hill, E. L. (2001). Non-specific nature of specific language impairment: a review of the literature with regard to concomitant motor impairments. *International Journal of Language and Communication Disorders, 36*, 149-171.

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

Holes, C. (2000). *Dialect, Culture, and Society in Eastern Arabia: Glossary* Leiden: Brill.

Hurst, J., Baraitser, M., Auger, E., Graham, F., & Norell, S. (1990). An extended family with an inherited speech disorder. *Developmental Medicine and Child Neurology, 32*, 347-355.

Ibn Manzur. (1290[1981]). *Lisaan al 'arab*. Cairo: Daar Al Ma'aarif.

Jakubowicz, C. (2003). Computational complexity and the acquisition of functional categories by French-speaking children with SLI1. *Linguistics, 41*, 175-211.

Jakubowicz, C., & Nash, L. (2001). Functional categories and syntactic operations in (Ab)normal language acquisition. *Brain and Language, 77*, 321-339.

Jakubowicz, C., Nash, L., Rigaut, C., & Gérard , C.-L. (1998). Determiners and clitic pronouns in French-speaking children with SLI. *Language Acquisition, 7*, 113-160.

Joanisse, M. F. (2004). Specific language impairments in children - Phonology, semantics, and the English past tense. *Current Directions in Psychological Science, 13*, 156-160.

Joanisse, M. F., & Seidenberg, M. (1998). Specific language impairment: a deficit in grammar or processing? *Trends in Cognitive Sciences, 2*, 240-247.

Joanisse, M. F., & Seidenberg, M. S. (2003). Phonology and syntax in specific language impairment: evidence from a connectionist model. *Brain & Language, 86*, 40-56.

Johnston, J. R., & Ellis Weismer, S. (1983). Mental rotation abilities in language-disordered children. *Journal of Speech and Hearing Research, 26*, 397-403.

Just, M., & Carpenter, P. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*, 122-149.

Kail, R. (1994). A method for studying the generalized slowing hypothesis in children with specific language impairment. *Journal of Speech & Hearing Research, 37*, 418-421.

Kamhi, A. G. (1998). Trying to make sense of developmental language disorders. *Language, Speech, and Hearing Services in Schools, 29*, 35-44.

Karmiloff-Smith, A. (1992). *Beyond modularity: a developmental perspective on cognitive science*. Cambridge, Mass.: MIT Press/Bradford Books.

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences, 2*, 389-398.

Karmiloff-Smith, A., Grant, J., Berthoud, I., Davies, M., Howlin, P., & Udwin, O. (1997). Language and Williams syndrome: How intact is "intact"? *Child Development, 68*, 246-262.

Karmiloff-Smith, A., & Thomas, M. S. (2003). Are developmental disorders like cases of adult brain damage? Implications from connectionist modelling *Behavioral and Brain Sciences, 25*, 727-750

Kauffman, N. R. (1995). The Kaufman Speech Praxis Test for Children: Detroit: Wayne State University Press.

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.

Lahey, M. (1990). Who shall be called language disordered? Some reflections and one perspective. *Journal of Speech and Hearing Disorders, 55*, 612-620.

Lahey, M., & Edwards, J. (1996). Why do children with specific language impairment name pictures more slowly than their peers? *Journal of Speech and Hearing Research, 39*, 1081-1098

Lahey, M., Edwards, J., & Munson, B. (2001). Is processing speed related to severity of language impairment? *Journal of Speech, Language, and Hearing Research, 44*, 1354-1361.

Lahey, M., Flax, J., & Schlisselberg, G. (1985). A preliminary investigation of reduplication in children with specific language impairment. *Journal of Speech and Hearing Disorders, 50*, 186-194.

Lai, C. S., Fisher, S. E., Hurst, J., Vargha-Khadem, F., & Monaco, A. (2001). A novel forkhead-domain gene is mutated in a severe speech and language disorder. *Nature*, 413-465.

Le Normand, M. T., Leonard, L. B., & McGregor, K. K. (1993). A cross-linguistic study of article use by children with specific language impairment. *European Journal of Disorders of Communication, 28*, 153-163.

Leonard, L. B. (1989). Language learnability and specific language impairment in children. *Applied Psycholinguistics, 10*, 179-202.

Leonard, L. B. (1992). Specific language impairment in three languages: some cross-linguistic evidence. In P. Fletcher & D. Hall (Eds.), *Specific Speech and Language Disorders in Children* (118-126). London: Whurr Publishers.

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

Leonard, L. B. (2003). Specific language impairment: characterizing the deficit. In Y. Levy & J. Schaeffer (Eds.), *Language competence across populations: toward a definition of specific language impairment* (pp. 209-231). Mahwah, New Jersey: Lawrence Erlbaum.

Leonard, L. B., & Bortolini, U. (1998). Grammatical morphology and the role of weak syllables in the speech of Italian-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 41*, 1363-1374.

Leonard, L. B., Bortolini, U., Caselli, M. C., McGregor, K. K., & Sabbadini, L. (1992). Morphological deficits in children with specific language impairment: The status of features in underlying grammar. *Language Acquisition, 2*, 151-179.

Leonard, L. B., Ellis Weismer, S., Miller, C. A., Francis, D. J., Tomblin, J. B., & Kail, R. V. (2007). Speed of processing, working memory, and language impairment in children. *Journal of Speech, Language, and Hearing Research, 50*, 408-428.

Leonard, L. B., Eyer, J. A., Bedore, L. M., & Grela, B. G. (1997). Three accounts of the grammatical morpheme difficulties of English-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 40*, 741-753.

Leonard, L. B., Miller, C. A., & Gerber, E. (1999). Grammatical morphology and the lexicon in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 42*, 678-689.

Leonard, L. B., & Schwartz, R. G. (1982). Early lexical acquisition in children with specific language impairment. *Journal of Speech and Hearing Research, 25*, 554-564.

Li, A. Y. H., & Aoun, J. (2003). *Essays on the representational and derivational nature of grammar: the diversity of Wh-constructions*: MIT Press.

LoBello, S. G. (1991). A short form of the Wechsler Preschool and Primary Scale of Intelligence-Revised. *Journal of School Psychology, 29*, 229-236.

Mahfoudhi, A. (2002). Agreement Lost, Agreement Regained. A Minimalist Account of Word Order and Agreement Variation in Arabic. *California Linguistic Notes, 27,* 1-28.

Marcus, G. F., & Fisher, S. E. (2003). FOXP2 in focus: what can genes tell us about speech and language? *Trends in Cognitive Sciences, 7*, 257-262.

Marinis, T., & van der Lely, H. K. J. (2007). On-line processing of wh-questions in children with G-SLI and typically developing children. *International Journal of Language & Communication Disorders, 42*, 557-582.

Marler, J. A., Champlin, C. A., & Gillam, R. B. (2001). Backward and simultaneous masking measured in children with language-learning impairments who received intervention with Fast ForWord or Laureate Learning Systems Software. *American Journal of Speech-Language Pathology, 10*, 258-268.

Marshall, C. R., Marinis, T., & van der Lely, H. (2007). Passive verb morphology: The effect of phonotactics on passive comprehension in typically developing and Grammatical-SLI children. *Lingua, 117*, 1434-1447.

Marshall, C. R. (2004). *The morpho-phonological interface in children with Grammatical-Specific Language impairment.* Unpublished doctoral thesis, University of London.

Marshall, C. R., Ebbels, S., Harris, J., & van der Lely, H. K. J. (2002). Investigating the impact of prosodic complexity on the speech of children with Specific Language Impairment. *UCL Working Papers in Linguistics, 14*, 43-68.

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

Marshall, C. R., & van der Lely, H. K. J. (2006). A challenge to current models of past tense inflection: The impact of phonotactics. *Cognition, 100*, 302-320.

Marshall, C. R., & van der Lely, H. K. J. (2007). Derivational morphology in children with Grammatical-Specific Language Impairment. *Clinical Linguistics and Phonetics, 21*, 71-92.

Marshall, C. R., & van der Lely, H. K. J. (2009). Effects of word position and stress on onset cluster production: Evidence from typical development, specific language impairment, and dyslexia. *Language*, *85*, 39-57.

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.

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

McCarthy, J., & Prince, A. (1995). Prosodic Morphology. In J. Goldsmith (Ed.), *The Handbook of Phonological Theory* (pp. 318-366). Oxford: Blackwell.

Merzenich, M. M., Jenkins, W. M., Johnston, P., Schreiner, C., Miller, S. L., & Tallal, P. (1996). Temporal processing deficits of language-learning impaired children ameliorated by training. *Science, 271*, 77-81.

Miller, C. A., Kail, R., Leonard, L. B., & Tomblin, J. B. (2001). Speed of processing in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 44*, 416-433.

Miller, C. A., & Leonard, L. B. (1998). Deficits in finite verb morphology: Some assumptions in recent accounts of specific language impairment. *Journal of Speech, Language, and Hearing Research*, *41*, 701-707

Miller, C. A., Leonard, L. B., Kail, R. V., Zhang, X., Tomblin, J. B., & Francis, D. J. (2006). Response time in 14-year-olds with language impairment. *Journal of Speech, Language, and Hearing Research, 49,* 712-728.

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

Montgomery, J. W. (1995b). Sentence comprehension in children with specific language impairment: the role of phonological working memory. *Journal of Speech & Hearing Research, 38*, 187-199.

Montgomery, J. W. (2000a). Relation of working memory to off-line and real-time sentence processing in children with specific language impairment. *Applied Psycholinguistics, 21*, 117-148.

Montgomery, J. W. (2000b). Verbal working memory and sentence comprehension in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 43*, 293-308.

Montgomery, J. W. (2002a). 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. (2002b). Information processing and language comprehension in children with specific language impairment. *Topics in Language Disorders, 22*, 62-84.

Montgomery, J. W. (2002c). Understanding the language difficulties of children with specific language impairments: Does verbal working memory matter? *American Journal of Speech-Language Pathology, 11*, 77-91.

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.

Montgomery, J. W. (2005). Effects of input rate and age on the real-time language processing of children with specific language impairment. *International Journal of Language & Communication Disorders, 40*, 171-188.

Montgomery, J. W. (2006). Real-time language processing in school-age children with specific language impairment. *International Journal of Language & Communication Disorders, 41*, 275-291.

Montgomery, J. W., & Evans, J. (2009). Complex sentence comprehension and working memory in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 52*, 269-288.

Montgomery, J. W., & Leonard, L. B. (1998). Real-time inflectional processing by children with specific language impairment: Effects of phonetic substance. *Journal of Speech, Language and Hearing Research, 41*, 1432-1443.

Montgomery, J. W., & Windsor, J. (2007). Examining the language performances of children with and without specific language impairment: contributions of phonological short-term memory and speed of processing. *Journal of Speech, Language, and Hearing Research, 50*, 778-797.

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.

Mustafawi, E. M. (2006). *An optimality theoretic approach to variable consonant alternations in Qatari Arabic* .Unpublished doctoral dissertation, University of Ottawa, Ottawa. Rutgers Optimality Archive (ROA) at http://ruccs.rutgers.edu.

Newbury, D. F., Bishop, D. V. M., & Monaco, A. P. (2005). Genetic influences on language impairment and phonological short-term memory. *Trends in Cognitive Science, 9*, 528-534.

Newbury, D. F., Bonora, E., Lamb, J. A., Fisher, S. E., Lai, C. S., Baird, G., et al. (2002). FOXP2 is not a major susceptibility gene for autism or specific language impairment. *American Journal of Human Genetics, 70*, 1318-1327.

Newmeyer, F. J. (1997). Genetic dysphasia and linguistic theory. *Journal of Neurolinguistics, 10*, 47-73.

Norbury, C. F., Bishop, D. V. M., & Briscoe, J. (2002). Does impaired grammatical comprehension provide evidence for an innate grammar module? *Applied Psycholinguistics, 23*, 247-268.

Norbury, C. F., Tomblin, J. B., & Bishop, D. V. M. (2008). *Understanding developmental language disorders: from theory to practice*: Psychology Press.

O'Brien, E. K., Zhang, X., Nishimura, C., Tomblin, J. B., & Murray, J. C. (2003). Association of specific language impairment (SLI) to the region of 7q31. *American Journal of Human Genetics, 72*, 1536-1543.

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.

Ouhalla, J. & Shlonsky, U (eds.). (2003). *Themes in Arabic and Hebrew Syntax*. Dordrecht: Kluwer.

Owen, A. J., Dromi, E., & Leonard, L. B. (2001). The phonology-morphology interface in the speech of Hebrew-speaking children with specific language impairment. *Journal of Communication Disorders, 34,* 323-337.

Owen, A. J., & Leonard, L. B. (2002). Lexical diversity in the spontaneous speech of children with specific language impairment: application of D. *Journal of Speech, Language, and Hearing Research., 45,* 927-937.

Paradis, J., & Crago, M. (2000). Tense and temporality: Commonalities and differences in French-speaking children with specific language impairment and French second language speakers. *Journal of Speech, Language and Hearing Research, 43,* 834-848.

Paradis, J., & Crago, M. (2001). The morphosyntax of specific language impairment in French: An extended optional default account. *Language Acquisition, 9,* 269-300.

Paradis, J., Crago, M., Genesee, F., & Rice, M. (2003). Bilingual children with specific language impairment: How do they compare with their monolingual peers? *Journal of Speech, Language and Hearing Research, 46*, 1-15.

Paradis, J., Crago, M., Genesee, F., & Rice, M. L. (2003). French-English bilingual children with SLI: how do they compare with their monolingual peers? *Journal of Speech, Language, and Hearing Research, 46*, 113-127.

Paul, R. (1995). *Language disorders from infancy through adolescence: assessment and intervention*. St. Louis: Mosby-Year Book.

Pierrehumbert, J. (1993). Dissimilarity in the Arabic verbal roots, *Proceedings of the North East Linguistics Society (NELS), 23* (pp. 367–381).

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Pinker, S. (1994). *The language instinct: How the mind creates language*. New York, NY and London, UK: William Morrow & Co (New York), Penguin (London).

Pinker, S. (1999). *Words and rules: The ingredients of language*. New York: Basic Books.

Plante, E. (1998). Criteria for SLI: The Stark and Tallal legacy and beyond. *Journal of Speech, Language, and Hearing Research, 41,* 951-957.

Plomin, R., & Kovas, Y. (2005). Generalist genes and learning disabilities. *Psychological Bulletin, 131*, 592-617.

Pokorni, J. L., Worthington, C. K., & Jamison, P. J. (2004). Phonological awareness intervention: A comparison of three programs-Fast ForWord, Earobics, and LiPS. *Journal of Educational Research, 97,* 147-157.

Qafisheh, H. (1997). *NTC's Gulf Arabic-English Dictionary*. Chicago, Illinois: NTC Publishing Group.

Rapin, I., & Allen, D. A. (1987). Developmental dysphasia and autism in preschool children: characteristics and subtypes. In P. F. J. Martin, P.Grunwell, and D. Hall (Ed.), *Proceedings of the first international symposium on specific speech and language disorders in children* (pp. 20-35). London: AFASIC.

Redmond, S. M., & Rice, M. L. (1998). The socioemotional behaviors of children with SLI: Social adaptation or social deviance? *Journal of Speech, Language, and Hearing Research, 41*, 688-700.

Rescorla, L., & Bernstein-Ratner, N. (1996). Phonetic profiles of toddlers with specific expressive language impairment (SLI-E). *Journal of Speech and Hearing Research, 39,* 153-165.

Restrepo, M. A. R., Iacute, A, A., & Gutierrez-Clellen, V. F. (2001). Article use in Spanish-speaking children with Specific Language Impairment. *Journal of Child Language*, *28*, 433-452.

Rice, M. L. (1996). Of language, phenotypes, and genetics: building a cross-disciplinary platform for inquiry. In M. L. Rice (Ed.), *Towards a Genetics of Language* (pp. xi-xxv). Mahwah, NJ: Lawrence Erlbaum Associates.

Rice, M. L. (2000). Grammatical symptoms of specific language impairment. In D. V. M. Bishop & L. B. Leonard (Eds.), *Speech and language impairments in children: Causes, characteristics, intervention and outcome* (pp. 17-34). East Sussex, England: Psychology Press Ltd.

Rice, M. L. (2003). A unified model of specific and general language delay: Grammatical tense as a clinical marker of unexpected variation. In Y. Levy & J. Schaeffer (Eds.), *Language competence across populations: Toward a definition of Specific Language Impairment* (pp. 63-95). Mahwah, NJ: Erlbaum.

Rice, M. L. (2007). Children with specific language impairment: bridging the genetic and developmental perspectives. In E. Hoff & M. Shatz (Eds.), *Blackwell handbook of language development* (pp. 411-431). Oxford: Blackwell.

Rice, M. L., Cleave, P. L., & Oetting, J. B. (2000). The use of syntactic cues in lexical acquisition by children with SLI. Specific Language Impairment. *Journal of Speech, Language, and Hearing Research, 43*, 582-594.

Rice, M. L., Haney, K. R., & Wexler, K. (1998). Family histories of children with SLI who show extended optional infinitives. *Journal of Speech, Language, and Hearing Research, 41*, 419-432.

Rice, M. L., & Oetting, J. B. (1993). Morphological deficits of children with SLI: Evaluation of number marking and agreement. *Journal of Speech and Hearing Research, 36,* 1249-1257.

Rice, M. L., & Wexler, K. (1995). Extended optional infinitive (EOI) account of specific language impairment. In D. MacLaughlin & S. McEwan (Eds.), *Proceedings of the 19th annual Boston University Conference on Language Development, Vol. 2* (pp. 451-462). Somerville, MA: Cascadilla Press.

Rice, M. L., & Wexler, K. (1996a). A phenotype of specific language impairment: extended optional infinitives. In M. L. Rice (Ed.), *Towards a Genetics of Language* (215-237). Mahwah, NJ: Lawrence Erlbaum Associates.

Rice, M. L., & Wexler, K. (1996b). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech and Hearing Research*, *39*, 1239-1257.

Rice, M. L., Wexler, K., & Cleave, P. L. (1995). Specific language impairment as a period of extended optional infinitive. *Journal of Speech and Hearing Research, 38*, 850-863.

Rice, M. L., Wexler, K., & Hershberger, S. (1998). Tense over time: The longitudinal course of tense acquisition in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *41*, 1412-1431.

Rice, M. L., Wexler, K., & Redmond, S. M. (1999). Grammaticality judgments of an extended optional infinitive grammar: evidence from English-speaking children with specific language impairment. *Journal of Speech, Language and Hearing Research*, *42*, 943-961.

Riches, N. G., Tomasello, M., & Conti-Ramsden, G. (2005). Verb learning in children with SLI: Frequency and spacing effects. *Journal of Speech Language and Hearing Research, 48*, 1397-1411.

Rizzi, L. (1997). The fine structure of the left periphery. In C. Hamann (Ed.), *Elements of Grammar* (pp. 281-337). Dordrecht: Kluwer.

Roberts, J., Rescorla, L., Giroux, J., & Stevens, L. (1998). Phonological skills of children with specific expressive language impairment (SLI-E): Outcome at age 3. *Journal of Speech, Language, and Hearing Research*, *41*, 374-384.

Rosen, S. (2003). Auditory processing in dyslexia and specific language impairment: is there a deficit? What is its nature? Does it explain anything? *Journal of Phonetics, 31*, 509-527.

Rouse, C., & Krueger, A. (2004). Putting computerized instruction to the test: a randomized evaluation of a 'scientifically based' reading program. *Economics of Education Review, 23*, 323-338.

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. (1999). Language comprehension and non-word repetition in children with language impairment. *Clinical Linguistics & Phonetics, 13,* 369-380.

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

Schaeffer, J. (2003). Pragmatics and SLI. In Y. Levy & J. Schaeffer (Eds.), *Language competence across populations: toward a definition of specific language impairment* (pp. 135-150). Mahwah, New Jersey: Lawrence Erlbaum.

Schuele, C. M., & Dykes, J. C. (2005). Complex syntax acquisition: A longitudinal case study of a child with specific language impairment. *Clinical Linguistics & Phonetics, 19*, 295-318.

Schutze, C. T., & Wexler, K. (1996). Subject case licensing and English root infinitives. In A. Stringfellow, D. Cahama-Amitay, E. Hughes & A. Zukowski (Eds.), *Proceedings of the 20th annual Boston University Conference on Language Development, Vol. 2* (pp. 670-681). Somerville, MA: Cascadilla Press.

Schwartz, R. G., Leonard, L. B., Messick, C. K., & Chapman, K. L. (1987). The acquisition of object names in children with specific language impairment: Action context and word extension. *Applied Psycholinguistics, 8*, 233-244.

Selkirk, E. O. (1980). The role of prosodic categories in English word stress. *Linguistic Inquiry, 11*, 563–605.

Selkirk, E. O. (1982). *The Syntax of Words*. Cambridge, M.A: MIT Press.

Semel, E., Wiig, E., & Secord, W. (1996). *Clinical Evaluation of Language Fundamentals-3 (CELF 3)*. San Antonio, TX: Psychological Corp.

Shaalan, S. (2009). Considerations for developing and adapting language and literacy assessments in Arabic-speaking countries. In E. Grigorenko (Ed.), *Multicultural Psycheducational Assessment* (pp. 287-314). New York: Springer.

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 2006, University of Newcastle upon-Tyne, U.K, 19-21 July 2006.

Shlonsky, U. (1997). *Clause structure and word order in Hebrew and Arabic: An essay in comparative Semitic syntax*. New York: Oxford University Press.

Shriberg, L. D., Ballard, K. J., Tomblin, J. B., Duffy, J. R., Odell, K. H., & Williams, C. A. (2006). Speech, prosody, and voice characteristics of a mother and daughter With a 7;13 translocation affecting FOXP2. *Journal of Speech, Language, and Hearing Research, 49*, 500-525.

Shriberg, L. D., Tomblin, J. B., & McSweeny, J. L. (1999). Prevalence of speech delay in 6-year-old children and comorbidity with language impairment. *Journal of Speech, Language, and Hearing Research, 42*, 1461-1481.

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.

Sokolov, J. (1989). The development of role assignment in Hebrew. In B. MacWhinney & E. Bates (Eds.), *The crosslinguistic study of sentence proecessing* (pp. 118-157). Cambridge: Cambridge University Press.

Stark, R. E., & Tallal, P. (1981). Selection of children with specific language deficits. *Journal of Speech and Hearing Disorders, 46*, 114-122.

Stavrakaki, S. (2001). Comprehension of reversible relative clauses in Specifically Language Impaired and normally developing Greek children. *Brain and Language, 77*, 419-431.

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, 49*, 219-236.

Stromswold, K. (1998). Genetics of spoken language disorders. *Human Biology, 70*, 297-324.

Stromswold, K. (2001). The heritability of language: A review and metaanalysis of twin, adoption, and linkage studies. *Language*, *77*, 647-724.

Stromswold, K. (2006). Why aren't identical twins linguistically identical? Genetic, prenatal and postnatal factors. *Cognition*, *101*, 333-384.

Tager-Flusberg, H. (2004). Do autism and specific language impairment represent overlapping language disorders? In M. L. Rice & S. F. Warren (Eds.), *Developmental language disorders: from genotypes to etiologies* (pp. 31-52). Mahwah, NJ: Erlbaum.

Tager-Flusberg, H., & Cooper, J. (1999). Present and future possibilities for defining a phenotype for specific language impairment. *Journal of Speech, Language, and Hearing Research, 42*, 1275-1278.

Tallal, P., Hirsch, L. S., Realpe-Bonilla, T., Miller, S., Brzustowicz, L. M., Bartlett, C., et al. (2001). Familial aggregation in specific language impairment. *Journal of Speech, Language, and Hearing Research, 44*, 1172-1182.

Tallal, P., Miller, S. L., Bedi, G., Byma, G., Wang, X., Nagarajan, S. S., et al. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science*, *271*, 81-83.

Tallal, P., & Piercy, M. (1973). Defects of non-verbal auditory perception in children with developmental aphasia. *Nature, 241*, 468-469.

Tallal, P., & Piercy, M. (1974). Developmental aphasia: Rate of auditory processing and selective impairment of consonant perception. *Neuropsychologia*, *12*, 83-93.

Tallal, P., & Piercy, M. (1975). Developmental aphasia: The perception of brief vowels and extended stop consonants. *Neuropsychologia, 13*, 69-74.

Tallal, P., Stark, R. E., Kallman, C., & Mellits, D. E. (1981). A reexamination of some nonverbal perceptual abilities of language-impaired and normal children as a function of age and sensory modality. *Journal of Speech and Hearing Research, 24*, 351-357.

Thal, D., O'Hanlon, L., Clemmons, M., & Fralin, L. (1999). Validity of a parent report measure of vocabulary and syntax for preschool children with language impairment. *Journal of Speech, Language and Hearing Disorders, 42*, 482-496.

The SLI Consortium (SLIC). (2004). Highly significant linkage to the SLI1 locus in an expanded sample of individuals affected by specific language impairment. *American Journal of Human Genetics, 74*, 1225–1238.

Thomas, M., Grant, J., Barham, Z., Gso, M., Laing, E., Lakusta, L., et al. (2001). Past tense formation in Williams syndrome. *Language and Cognitive Processes, 16*, 143-177.

Thordardottir, E. T., & Ellis Weismer, S. (2001). High-frequency verbs and verb diversity in the spontaneous speech of school-age children with specific language impairment. *International Journal of Language & Communication Disorders., 36*, 221-244.

Thordardottir, E. T., & Namazi, M. (2007). Specific language impairment in French-speaking children: Beyond grammatical morphology. *Journal of Speech, Language, and Hearing Research, 50*, 698-715.

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.

Tomblin, J. B., Records, N. L., & Zhang, X. (1996). A system for the diagnosis of specific language impairment in kindergarten children. *Journal of Speech and Hearing Research, 39*, 1284-1294.

Tomblin, J. B., & Zhang, X. (1999). Language patterns and etiology in children with specific language impairment. In H. Tager-Flusberg (Ed.), *Neurodevelopmental disorders* (pp. 361-382). Cambridge, MA: MIT Press.

Tomblin, J. B., Zhang, X., Catts, H., Ellis Weismer, S., & Weiss, A. (2004). Dimensions of individual differences in communication skills among primary grade children. In M. L. Rice & S. F. Warren (Eds.), *Developmental language disorders: From phenotypes to etiologies.* Mahwah, NJ: Erlbaum.

Tsakali, V., & Wexler, K. (2004). *Why children omit clitics in some languages but not in others: New evidence from Greek.* Paper presented at the GALA 2003, LOT.

Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: the declarative/procedural model. *Bilingualism: Language and Cognition, 4*, 105-122.

Ullman, M. T. (2004). Contributions of memory circuits to language: the declarative/procedural model. *Cognition, 92*, 231-270.

Ullman, M. T., & Gopnik, M. (1999). Inflectional morphology in a family with inherited specific language impairment. *Applied Psycholinguistics, 20*, 51-118.

Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex, 41*, 399-433.

van Balkom, H., & Verhoeven, L. (2004). Pragmatic disability in children with specific language impairment. In L. Verhoeven & H. van Balkom (Eds.), *Classification of developmental language disorders* (pp. 283-305). Mahweh, New Jersey: Lawrence Erlbaum.

van der Lely, H. K. J. (1996). Language modularity and grammatically specific language impaired children. In M. Aldridge (Ed.), *Child Language* (pp. 188-201). Avon:UK: Multilingual Matters Ltd.

van der Lely, H. K. J. (1996). Specifically language impaired and normally developing children: Verbal passive vs. adjectival passive sentence interpretation. *Lingua, 98*, 243-272.

van der Lely, H. K. J. (1998). SLI in children: movement, economy and deficits in the computational syntactic system. *Language Acquisition, 72*, 161-192.

van der Lely, H. K. J. (2003). Do heterogeneous SLI deficits need heterogeneous theories? SLI subgroups, G-SLI and the RDDR hypothesis. In Y. Levy & J. S. Schaeffer (Eds.), *Towards a definition of specific language impairment* (pp. 109-134): Lawrence Erlbaum.

van der Lely, H. K. J. (2005). Domain-specific cognitive systems: insight from Grammatical-SLI. *Trends in Cognitive Sciences, 9,* 53-59.

van der Lely, H. K. J., & Battell, J. (2003). Wh-movement in children with grammatical SLI: a test of RDDR Hypothesis. *Language*, *79*, 153-181.

van der Lely, H. K. J., & Christian, V. (2000). Lexical word formation in children with grammatical SLI: a grammar-specific versus an input-processing deficit? *Cognition, 75*, 33-63.

van der Lely, H. K. J., & Harris, J. (1999). *The Test of Phonological Structure (TOPhS)*. London: Available from the author at the Centre for Developmental Language Disorders & Cognitive Neuroscience, University College London.

van der Lely, H. K. J., & Harris, M. (1990). Comprehension of reversible sentences in specifically language-impaired children. *Journal of Speech and Hearing Disorders, 55*, 101-117.

van der Lely, H. K. J., & Stollwerck, L. (1996). A grammatical specific language impairment in children: An autosomal dominant inheritance? *Brain and Language* and Language, *52*, 484-504.

van der Lely, H. K. J., & Stollwerck, L. (1997). Binding theory and grammatical specific language impairment in children. *Cognition*, *62,* 245-290.

van der Lely, H. K. J., & Ullman, M. (1996). The computation and representation of past-tense morphology in specifically language impaired and normally developing Children. In A. Stringfellow, D. Cahama-Amitay, E. Hughes & A. Zukowski (Eds.), *Proceedings of the 20th annual Boston University Conference on Language Development, Vol. 2* (pp. 804-815). Somerville, MA: Cascadilla Press.

Vargha-Khadem, F., Gadian, D. G., Copp, A., & Mishkin, M. (2005). FOXP2 and the neuroanatomy of speech and langauge. *Nature Reviews Neuroscience, 6*, 131-138.

Vargha-Khadem, F., Watkins, K., Alcock, K., Fletcher, P., & Passingham, R. (1995). Praxic and Nonverbal Cognitive Deficits in a Large Family with a Genetically Transmitted Speech and Language Disorder. *PNAS, 92*, 930-933.

Watson, J. C. E. (2002). *The Phonology and Morphology of Arabic*: Oxford University Press, USA.

Wechsler, D. (2002). *WPPSI-III: Wechsler Preschool and Primary Scale of Intelligence, 3rd ed.*: Psychological Corp.

Wexler, K. (1994). Optional Infinitives. In D. Lightfoot & N. Hornstein (Eds.), *Verb Movement* (pp. 305-350). New York: Cambridge University Press.

Wexler, K. (1996). The development of inflection in a biologically based theory of language acquisition. In M. L. Rice (Ed.), *Towards a Genetics of Language* (113-144). Mahwah, NJ: Lawrence Erlbaum Associates

Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua, 106*, 23-79.

Wexler, K. (2000). The Unique checking constraint as the explanation of clitic omission in SLI and normal development. In C. Jakubowicz, L. Nash & K. Wexler (Eds.), *Essays on syntax, morphology and phonology in SLI*. Cambridge, MA: MIT Press.

Wexler, K. (2003). Lenneberg's dream: learning, normal language development, and specific language impairment. In Y. Levy & J. Schaeffer (Eds.), *Language Competence Across Populations: Towards a Definition of SLI* (pp. 11-62). Mahwah, NJ Lawrence Erlbaum.

Wexler, K., Gavarro, A., & Torrens, V. (2004). Feature checking and object clitic omission in child Catalan and Spanish. In R. Bok-Bennema, B. Hollebrandse, B. Kampers Mahne & P. Sleeman (Eds.), *Romance languages and linguistic theory 2002* (pp. 253-268). Amsterdam-Philadelphia: John Benjamins.

Windsor, J. (2002). Contrasting general and process-specific slowing in language impairment. *Topics in Language Disorders, 22*, 49-61.

Windsor, J., Milbrath, R. L., Carney, E. J., & Rakowski, S. E. (2001). General slowing in language impairment: methodological considerations in testing the hypothesis. *Journal of Speech, Language, and Hearing Research, 44*, 446-461.

World Health Organisation (WHO). (1992). *Manual of the International Statistical Classification of Diseases, and Related Health Problems-Tenth Revision (ICD-10)*. Geneva: World Health Organisation.

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (1992). *Preschool Language Scale (PLS-3).* San Antonio, TX: Psychological Corp.

Appendix A: Standardized test items and item analyses

Table A- 1 Cronbach's Alpha values for the Sentence Comprehension test

Item	Cronbach's Alpha if Item Deleted
item2	0.789
item3	0.788
item4	0.791
item5	0.790
item6	0.788
item7	0.789
item8	0.798
item9	0.789
item10	0.787
item11	0.787
item12	0.785
item13	0.786
item14	0.782
item15	0.788
item16	0.788
item17	0.786
item18	0.790
item19	0.786
item20	0.785
item21	0.784
item22	0.772
item23	0.780
item24	0.783
item25	0.784
item26	0.786
item27	0.775
item28	0.775
item29	0.791
item30	0.778
item31	0.780
item32	0.782
item33	0.793
item34	0.789
item35	0.788
item36	0.782
item37	0.790
item38	0.781
item39	0.784
item40	0.795

Cronbach's Alpha	No of items
.790	39

*All children correctly answered item 1 and therefore no Cronbach's Alpha was reported.

Table A- 2 Proportion of correct responses for all items of the Sentence Comprehension test for TD children (n=88).

Item	Age Band 1 (n=24)	Age Band 2 (n=23)	Age Band 3 (n=22)	Age Band 4 (n=19)	Average of all TD children
1	1.00	1.00	1.00	1.00	1.00
2	0.79	1.00	1.00	1.00	0.95
3	1.00	1.00	1.00	1.00	1.00
4	0.88	0.91	0.91	0.89	0.90
5	1.00	1.00	1.00	1.00	1.00
6	0.83	0.87	0.86	0.95	0.88
7	0.96	1.00	0.95	1.00	0.98
8	0.79	0.74	0.86	0.74	0.78
9	0.83	0.87	0.91	1.00	0.90
10	1.00	0.91	1.00	0.95	0.97
11	0.79	0.65	0.91	0.95	0.83
12	0.71	0.96	0.95	0.95	0.89
13	0.71	0.83	0.77	1.00	0.83
14	0.67	0.91	0.95	1.00	0.88
15	0.67	0.78	1.00	0.89	0.84
16	0.79	0.61	0.91	0.89	0.80
17	0.67	0.78	0.91	0.95	0.83
18	0.79	0.91	0.95	0.95	0.90
19	0.63	0.74	1.00	0.89	0.81
20	0.71	0.70	0.82	0.74	0.74
21	0.46	0.52	0.55	0.79	0.58
22	0.63	0.52	0.91	1.00	0.76
23	0.79	0.91	0.95	0.95	0.90
24	0.75	1.00	1.00	1.00	0.94
25	0.79	0.83	1.00	0.89	0.88
26	0.58	0.78	0.91	0.89	0.79
27	0.50	0.74	0.77	0.89	0.73
28	0.38	0.57	0.50	0.89	0.58
29	0.54	0.52	0.68	0.79	0.63
30	0.54	0.57	0.77	0.95	0.71
31	0.58	0.61	0.68	0.84	0.68
32	0.71	0.74	0.95	0.95	0.84
33	0.54	0.78	0.45	0.89	0.67
34	0.38	0.57	0.82	0.79	0.64
35	0.42	0.35	0.59	0.63	0.50
36	0.25	0.48	0.82	0.89	0.61
37	0.29	0.35	0.50	0.53	0.42
38	0.54	0.61	0.82	0.84	0.70
39	0.13	0.26	0.55	0.47	0.35
40	0.42	0.43	0.32	0.42	0.40

Note. Age Band 1: 4;6-5;11 years; 2: 6;0-6;11; 3: 7;0-7;11; 4: 8;0-9;4.

***Table A- 3: Proportion of correct responses for all items of the Sentence Comprehension test for children with SLI* (n=26)**

Item	Age Band A (n=5)	Age Band B (n=8)	Age Band C (n=5)	Age Band D (n=8)	Overall mean (n=26)
1	1.00	1.00	1.00	1.00	1.00
2	1.00	1.00	1.00	0.88	0.97
3	0.80	0.88	1.00	1.00	0.92
4	0.80	1.00	1.00	1.00	0.95
5	0.80	1.00	1.00	1.00	0.95
6	0.80	0.63	0.80	1.00	0.81
7	0.60	0.88	1.00	1.00	0.87
8	0.60	0.88	1.00	0.63	0.78
9	0.60	0.88	0.80	0.88	0.79
10	0.40	1.00	1.00	1.00	0.85
11	0.60	0.50	0.60	0.75	0.61
12	0.80	0.63	1.00	0.88	0.83
13	0.60	0.50	0.60	0.63	0.58
14	0.20	0.63	0.80	0.88	0.63
15	0.40	1.00	1.00	1.00	0.85
16	0.40	0.88	0.80	0.50	0.64
17	0.60	0.38	0.40	0.88	0.56
18	0.80	1.00	1.00	0.88	0.92
19	0.80	0.50	0.80	0.75	0.71
20	0.40	0.25	0.60	0.75	0.50
21	0.20	0.50	0.60	0.63	0.48
22	0.00	0.13	0.40	0.75	0.32
23	0.60	0.50	0.60	1.00	0.68
24	0.40	0.75	0.60	1.00	0.69
25	0.80	0.50	0.80	1.00	0.78
26	0.60	0.63	0.20	1.00	0.61
27	0.20	0.50	0.40	0.75	0.46
28	0.20	0.25	0.40	0.50	0.34
29	0.60	0.38	0.60	0.25	0.46
30	0.20	0.25	0.60	0.88	0.48
31	0.00	0.63	0.20	0.63	0.36
32	0.60	0.75	0.80	0.88	0.76
33	0.40	0.88	0.60	0.25	0.53
34	0.80	0.25	0.40	0.50	0.49
35	0.20	0.38	0.20	1.00	0.44
36	0.20	0.38	0.60	0.63	0.45
37	0.20	0.50	0.20	0.75	0.41
38	0.4	0.50	0.40	0.50	0.45
39	0.00	0.13	0.00	0.13	0.06
40	0.2	0.50	0.20	0.25	0.29

Note. Age Band 1: 4;6-5;11 years; 2: 6;0-6;11; 3: 7;0-7;11; 4: 8;0-9;4.

Table A- 4: Cronbach's Alpha values for the Expressive Langauge test (n=68)

Item	Cronbach's Alpha if Item Deleted
1	0.933
2	0.933
3	0.932
4	0.932
5	0.932
6	0.931
7	0.933
8	0.931
9	0.932
10	0.931
11	0.932
12	0.932
13	0.933
14	0.931
15	0.932
16	0.932
17	0.932
18	0.933
19	0.932
20	0.933
21	0.933
22	0.931
23	0.933
24	0.933
25	0.932
26	0.931
27	0.932
28	0.932
29	0.932
30	0.933
31	0.932
32	0.931
33	0.932
34	0.931
35	0.930
36	0.931
37	0.931
38	0.931

Item	Cronbach's Alpha if Item Deleted
39	0.930
40	0.933
41	0.933
42	0.932
43	0.932
44	0.932
45	0.931
46	0.932
47	0.933
48	0.931
49	0.932
50	0.933
51	0.931
52	0.933
53	0.932
54	0.931
55	0.931
56	0.932
57	0.933
58	0.931
59	0.934
60	0.932
61	0.932
62	0.933
63	0.931
64	0.934
65	0.931
66	0.933
67	0.932
68	0.932

Reliability Statistics

Cronbach's Alpha	N of Items
.933	68

Table A- 5: Proportion of correct responses by test items for the Expressive Language test for TD participants

Item	Age Band 1 (n=24)	Age Band 2 (n=23)	Age Band 3 (n=21)	Age Band 4 (n=18)	Overall TD children (86)
1.	1.00	0.96	1.00	1.00	0.99
2.	1.00	1.00	1.00	1.00	1.00
3.	0.88	1.00	1.00	1.00	0.97
4.	1.00	1.00	1.00	1.00	1.00
5.	0.58	0.65	0.62	0.78	0.66
6.	0.83	0.96	1.00	1.00	0.95
7.	0.96	1.00	1.00	1.00	0.99
8.	0.71	0.87	0.86	0.89	0.83
9.	0.96	1.00	1.00	1.00	0.99
10.	0.58	0.87	1.00	0.94	0.85
11.	0.50	0.70	0.71	0.94	0.71
12.	0.42	0.35	0.43	0.50	0.42
13.	0.21	0.39	0.62	0.72	0.49
14.	0.79	0.65	0.71	0.78	0.73
15.	0.46	0.78	0.76	0.72	0.68
16.	0.88	0.83	0.95	0.94	0.90
17.	0.42	0.78	0.67	0.83	0.67
18.	0.08	0.00	0.05	0.06	0.05
19.	0.13	0.09	0.24	0.39	0.21
20.	0.67	0.70	0.71	0.83	0.73
21.	0.88	0.91	0.86	0.94	0.90
22.	0.83	0.91	1.00	1.00	0.94
23.	1.00	1.00	0.90	1.00	0.98
24.	0.83	1.00	0.90	1.00	0.93
25.	0.79	0.78	0.76	0.89	0.81
26.	0.79	0.91	0.95	1.00	0.91
27.	0.67	0.78	0.86	0.94	0.81
28.	0.58	0.61	0.86	0.78	0.71
29.	0.96	1.00	1.00	1.00	0.99
30.	1.00	1.00	0.95	1.00	0.99
31.	1.00	1.00	1.00	1.00	1.00
32.	0.21	0.30	0.52	0.94	0.50
33.	1.00	1.00	1.00	1.00	1.00
34.	0.50	0.87	0.86	1.00	0.81
35.	0.46	0.83	0.86	1.00	0.79
36.	0.29	0.52	0.76	0.83	0.60
37.	0.67	0.96	0.95	1.00	0.89
38.	0.58	0.87	0.90	1.00	0.84
39.	0.46	0.91	0.95	1.00	0.83
40.	1.00	1.00	1.00	1.00	1.00
41.	0.88	1.00	0.95	1.00	0.96

Item	Age Band 1 (n=24)	Age Band 2 (n=23)	Age Band 3 (n=21)	Age Band 4 (n=18)	Overall TD children (86)
42.	0.92	1.00	0.95	0.94	0.95
43.	0.88	1.00	1.00	0.94	0.95
44.	0.92	1.00	1.00	1.00	0.98
45.	0.17	0.52	0.67	0.89	0.56
46.	0.63	0.83	0.86	0.94	0.81
47.	0.00	0.04	0.05	0.17	0.06
48.	0.17	0.52	0.48	0.83	0.50
49.	0.38	0.30	0.48	0.83	0.50
50.	0.33	0.52	0.71	0.94	0.63
51.	0.46	0.57	0.71	0.78	0.63
52.	0.71	0.70	0.86	0.78	0.76
53.	0.67	0.78	0.62	0.78	0.71
54.	0.38	0.65	0.71	0.78	0.63
55.	0.79	0.83	0.86	0.94	0.85
56.	0.63	0.78	0.81	0.89	0.78
57.	0.38	0.52	0.62	0.44	0.49
58.	0.54	0.70	0.86	0.78	0.72
59.	0.38	0.35	0.48	0.22	0.36
60.	0.63	0.87	0.90	0.78	0.79
61.	0.92	0.78	0.95	0.94	0.90
62.	0.96	0.91	0.90	0.94	0.93
63.	0.21	0.43	0.48	0.72	0.46
64.	0.21	0.30	0.29	0.50	0.32
65.	0.08	0.43	0.48	0.56	0.39
66.	0.21	0.57	0.38	0.56	0.43
67.	0.96	1.00	0.86	0.94	0.94
68.	0.33	0.39	0.71	0.67	0.53

Note. Age Band 1: 4;6-5;11 years; Age Band 2: 6;0-6;11 ; Age Band 3: 7;0-7;11 ; Age Band 4: 8;0-9;4.

Table A- 6: Proportion of correct responses by test items for the Expressive Language test for children with SLI

Item	Age Band A (n=5)	Age Band B (n=8)	Age Band C (n=5)	Age Band D (n=8)	Overall mean (n=26)
1.	1.00	0.88	1.00	0.88	0.94
2.	1.00	1.00	1.00	1.00	1.00
3.	0.80	0.75	1.00	0.88	0.86
4.	0.80	0.75	1.00	1.00	0.89
5.	0.40	0.25	0.40	0.38	0.36
6.	0.40	0.50	1.00	1.00	0.73
7.	1.00	1.00	0.80	1.00	0.95
8.	0.00	0.13	0.60	1.00	0.43
9.	0.80	0.75	1.00	0.88	0.86
10.	0.00	0.63	0.80	1.00	0.61
11.	0.20	0.50	0.20	0.63	0.38
12.	0.00	0.00	0.20	0.25	0.11
13.	0.40	0.13	0.00	0.50	0.26
14.	0.00	0.25	0.20	0.63	0.27
15.	0.40	0.25	0.00	0.75	0.35
16.	0.20	0.38	0.80	0.88	0.56
17.	0.20	0.13	0.80	0.88	0.50
18.	0.00	0.00	0.00	0.13	0.03
19.	0.00	0.00	0.00	0.13	0.03
20.	0.00	0.50	0.60	0.50	0.40
21.	0.60	0.63	0.60	0.88	0.68
22.	0.20	0.50	1.00	1.00	0.68
23.	1.00	0.63	1.00	1.00	0.91
24.	1.00	0.75	1.00	1.00	0.94
25.	0.80	0.25	0.40	0.63	0.52
26.	0.40	0.25	0.60	1.00	0.56
27.	0.40	0.50	0.40	0.63	0.48
28.	0.00	0.50	0.40	0.63	0.38
29.	1.00	0.75	0.60	0.88	0.81
30.	0.80	0.75	0.80	0.88	0.81
31.	0.60	0.75	1.00	1.00	0.84
32.	0.00	0.00	0.20	0.63	0.21
33.	0.40	0.88	1.00	0.88	0.79
34.	0.20	0.63	0.40	1.00	0.56
35.	0.00	0.13	0.40	0.75	0.32
36.	0.00	0.13	0.20	0.50	0.21
37.	0.40	0.63	0.80	1.00	0.71
38.	0.00	0.25	0.80	0.88	0.48
39.	0.20	0.38	0.40	0.88	0.46
40.	0.80	1.00	1.00	1.00	0.95
41.	0.80	0.88	1.00	0.88	0.89
42.	0.80	0.50	1.00	0.63	0.73

Item	Age Band 1 (n=24)	Age Band 2 (n=23)	Age Band 3 (n=21)	Age Band 4 (n=18)	Overall TD children (86)
43.	0.40	0.63	0.80	0.88	0.68
44.	0.60	0.63	1.00	0.88	0.78
45.	0.00	0.00	0.60	0.63	0.31
46.	0.20	0.38	0.60	0.75	0.48
47.	0.00	0.00	0.00	0.00	0.00
48.	0.00	0.00	0.00	0.50	0.13
49.	0.20	0.25	0.00	0.13	0.14
50.	0.00	0.50	0.60	0.50	0.40
51.	0.00	0.00	0.20	0.63	0.21
52.	0.40	0.25	0.80	0.88	0.58
53.	0.20	0.25	0.20	0.75	0.35
54.	0.00	0.13	0.00	0.38	0.13
55.	0.20	0.25	0.40	0.50	0.34
56.	0.40	0.13	0.40	0.63	0.39
57.	0.00	0.38	0.20	0.50	0.27
58.	0.00	0.25	0.40	0.63	0.32
59.	0.20	0.13	0.00	0.13	0.11
60.	0.20	0.38	0.40	0.88	0.46
61.	1.00	0.50	0.20	0.88	0.64
62.	1.00	0.75	0.60	0.88	0.81
63.	0.00	0.13	0.00	0.38	0.13
64.	0.20	0.50	0.00	0.38	0.27
65.	0.00	0.00	0.40	0.50	0.23
66.	0.40	0.13	0.00	0.13	0.16
67.	0.40	0.63	0.60	0.75	0.59
68.	0.00	0.25	0.40	0.50	0.29

Note. Age Band 1: 4;6-5;11 years; Age Band 2: 6;0-6;11 ; Age Band 3: 7;0-7;11 ; Age Band 4: 8;0-9;4.

Table A-7: Cronbach's Alpha for the Sentence Repetition test

Item*	Cronbach's Alpha if Item Deleted
2	0.893
5	0.894
6	0.893
7	0.894
8	0.895
9	0.893
10	0.893
11	0.890
12	0.894
13	0.893
14	0.890
15	0.890
16	0.891
17	0.894
18	0.890
19	0.893
20	0.891
21	0.896
22	0.890
23	0.887
24	0.889
25	0.888
26	0.887
27	0.887
28	0.890
29	0.890
30	0.885
31	0.885
32	0.891
33	0.890
34	0.887
35	0.892
36	0.888
37	0.893
38	0.891
39	0.892
40	0.893
41	0.893

*No variation was observed in items 1,3, and 4.
Overall Cronbach's alpa=.894, n=38.

Table A-8: Proportions of correct responses by test items for the Sentence Repetition test for typically developing participants

Item	Age Band 1 (n=24)	Age Band 2 (n=23)	Age Band 3 (n=21)	Age Band 4 (n=18)	Mean (n=86)
1.	1.00	1.00	1.00	1.00	1.00
2.	0.92	0.99	0.98	0.98	0.97
3.	0.99	1.00	1.00	1.00	1.00
4.	1.00	1.00	1.00	1.00	1.00
5.	0.97	1.00	0.98	0.98	0.98
6.	0.93	0.96	1.00	1.00	0.97
7.	0.94	1.00	1.00	1.00	0.99
8.	0.96	0.97	1.00	1.00	0.98
9.	0.83	0.93	0.92	0.98	0.91
10.	0.96	0.99	0.98	1.00	0.98
11.	0.76	0.88	0.92	0.91	0.87
12.	0.94	1.00	0.98	0.96	0.97
13.	0.94	0.94	0.95	0.98	0.95
14.	0.54	0.65	0.73	0.91	0.71
15.	0.79	0.90	0.97	0.93	0.90
16.	0.76	0.78	0.87	0.91	0.83
17.	0.97	0.99	0.97	0.98	0.98
18.	0.76	0.88	0.93	0.91	0.87
19.	0.89	0.93	0.93	0.98	0.93
20.	0.78	0.94	0.93	0.96	0.90
21.	0.82	0.84	0.82	0.91	0.85
22.	0.76	0.91	0.95	0.98	0.90
23.	0.54	0.68	0.85	0.96	0.76
24.	0.64	0.72	0.85	0.85	0.77
25.	0.65	0.81	0.90	0.80	0.79
26.	0.33	0.54	0.73	0.80	0.60
27.	0.08	0.22	0.17	0.52	0.25
28.	0.17	0.26	0.43	0.50	0.34
29.	0.14	0.35	0.35	0.44	0.32
30.	0.29	0.62	0.73	0.76	0.60
31.	0.25	0.33	0.42	0.70	0.43
32.	0.03	0.01	0.07	0.17	0.07
33.	0.03	0.10	0.18	0.48	0.20
34.	0.36	0.64	0.75	0.63	0.59
35.	0.31	0.55	0.58	0.52	0.49
36.	0.06	0.10	0.08	0.20	0.11
37.	0.00	0.00	0.00	0.06	0.01
38.	0.00	0.01	0.02	0.11	0.04
39.	0.00	0.01	0.02	0.13	0.04
40.	0.00	0.00	0.00	0.06	0.01
41.	0.01	0.01	0.00	0.04	0.02

Table A- 9: Proportions of correct responses by test items for the Sentence Repetition test for children with SLI

Item	Age Band 1 (n=5)	Age Band 2 (n=8)	Age Band 3 (n=5)	Age Band 4 (n=8)	Mean (n=26)
1.	1.00	1.00	1.00	1.00	1.00
2.	0.93	0.88	0.87	1.00	0.92
3.	1.00	1.00	1.00	1.00	1.00
4.	1.00	1.00	1.00	1.00	1.00
5.	0.73	0.88	1.00	0.96	0.89
6.	0.60	0.75	0.80	1.00	0.79
7.	0.47	0.71	1.00	1.00	0.79
8.	1.00	1.00	1.00	1.00	1.00
9.	0.40	0.54	0.67	0.92	0.63
10.	0.67	0.79	0.87	1.00	0.83
11.	0.53	0.67	0.80	0.79	0.70
12.	0.93	1.00	1.00	1.00	0.98
13.	0.67	0.75	0.87	0.96	0.81
14.	0.00	0.21	0.47	0.54	0.30
15.	0.47	0.50	0.40	0.79	0.54
16.	0.53	0.63	0.80	0.79	0.69
17.	0.53	0.79	0.93	0.92	0.79
18.	0.27	0.46	0.80	0.92	0.61
19.	0.27	0.46	0.93	0.88	0.63
20.	0.47	0.58	0.80	0.83	0.67
21.	0.53	0.63	0.60	0.83	0.65
22.	0.20	0.58	0.73	0.79	0.58
23.	0.20	0.29	0.73	0.79	0.50
24.	0.00	0.33	0.40	0.67	0.35
25.	0.13	0.33	0.60	0.83	0.48
26.	0.07	0.08	0.13	0.54	0.21
27.	0.00	0.00	0.00	0.08	0.02
28.	0.00	0.08	0.07	0.13	0.07
29.	0.00	0.00	0.20	0.17	0.09
30.	0.00	0.21	0.47	0.50	0.29
31.	0.00	0.17	0.13	0.33	0.16
32.	0.00	0.00	0.00	0.08	0.02
33.	0.00	0.00	0.00	0.04	0.01
34.	0.00	0.08	0.27	0.29	0.16
35.	0.00	0.13	0.20	0.29	0.15
36.	0.00	0.00	0.00	0.17	0.04
37.	0.00	0.00	0.00	0.00	0.00
38.	0.00	0.00	0.00	0.00	0.00
39.	0.00	0.00	0.00	0.04	0.01
40.	0.00	0.00	0.00	0.00	0.00
41.	0.00	0.00	0.00	0.00	0.00

Table A- 10: Arabic Picture Vocabulary Test (APVT): Proportion of correct responses for the 8 year old TD group (n=18) on the APVT items.

Item	proportion of correct responses	Item	proportion of correct responses	Item	Proportion of correct responses	Item	proportion of correct responses
1.	1.00	39.	0.78	77.	0.94	115.	0.72
2.	1.00	40.	1.00	78.	0.94	116.	0.89
3.	1.00	41.	1.00	79.	0.89	117.	0.72
4.	1.00	42.	0.61	80.	0.39	118.	0.11
5.	1.00	43.	0.67	81.	0.72	119.	0.39
6.	1.00	44.	0.94	82.	0.33	120.	0.72
7.	1.00	45.	1.00	83.	0.78	121.	0.61
8.	1.00	46.	0.94	84.	0.78	122.	0.50
9.	1.00	47.	0.89	85.	0.78	123.	0.11
10.	1.00	48.	1.00	86.	0.83	124.	0.50
11.	1.00	49.	1.00	87.	0.39	125.	0.28
12.	1.00	50.	0.83	88.	0.78	126.	0.17
13.	1.00	51.	1.00	89.	0.89	127.	0.33
14.	0.94	52.	0.94	90.	0.22	128.	0.94
15.	1.00	53.	0.94	91.	0.67	129.	0.28
16.	1.00	54.	0.44	92.	0.06	130.	0.39
17.	0.94	55.	0.72	93.	0.72	131.	0.17
18.	1.00	56.	0.50	94.	0.94	132.	0.72
19.	1.00	57.	0.28	95.	0.83		
20.	1.00	58.	0.22	96.	0.72		
21.	1.00	59.	0.94	97.	0.56		
22.	1.00	60.	1.00	98.	0.44		
23.	0.83	61.	0.94	99.	0.94		
24.	0.67	62.	0.94	100	0.94		
25.	0.56	63.	1.00	101	0.44		
26.	1.00	64.	0.56	102	0.94		
27.	0.94	65.	0.72	103	0.22		
28.	0.22	66.	0.50	104	0.33		
29.	1.00	67.	0.61	105	0.50		
30.	0.94	68.	0.61	106	0.61		
31.	0.94	69.	0.22	107	0.56		
32.	0.78	70.	0.89	108	0.44		
33.	1.00	71.	0.94	109	0.44		
34.	0.67	72.	0.83	110	1.00		
35.	0.61	73.	1.00	111	0.06		
36.	1.00	74.	0.67	112	0.94		
37.	0.94	75.	0.22	113	0.78		
38.	1.00	76.	0.67	114	0.61		

Table A- 11: Arabic Picture Vocabulary Test (APVT): Proportion of correct responses for TD children on the first 60 items of the test.

Item	Age Band A (n=22)	number of participants who completed this item	Age Band B (n=22)	number of participants who completed the item	Age Band C (n=19)	number of participants who completed the item	Age Band D (n=18)	number of participants who completed the item	Mean
1	1.00	22	1.00	22	1.00	19	1.00	18	1.00
2	1.00	22	1.00	22	1.00	19	1.00	18	1.00
3	1.00	22	1.00	22	1.00	19	1.00	18	1.00
4	1.00	22	1.00	22	1.00	19	1.00	18	1.00
5	1.00	22	1.00	22	1.00	19	1.00	18	1.00
6	1.00	22	1.00	22	1.00	19	1.00	18	1.00
7	1.00	22	1.00	22	1.00	19	1.00	18	1.00
8	0.95	22	0.95	22	1.00	19	1.00	18	0.98
9	1.00	22	1.00	22	1.00	19	1.00	18	1.00
10	0.59	22	0.91	22	0.95	19	1.00	18	0.86
11	1.00	22	1.00	22	1.00	19	1.00	18	1.00
12	1.00	22	1.00	22	1.00	19	1.00	18	1.00
13	1.00	22	0.95	22	1.00	19	1.00	18	0.99
14	0.95	22	0.95	22	1.00	19	0.94	18	0.96
15	0.95	22	0.95	22	1.00	19	1.00	18	0.98
16	1.00	22	1.00	22	1.00	19	1.00	18	1.00
17	0.82	22	0.82	22	0.84	19	0.94	18	0.86
18	0.77	22	0.95	22	1.00	19	1.00	18	0.93
19	1.00	22	1.00	22	1.00	19	1.00	18	1.00
20	0.91	22	1.00	22	1.00	19	1.00	18	0.98
21	0.95	22	1.00	22	1.00	19	1.00	18	0.99
22	0.95	22	1.00	22	1.00	19	1.00	18	0.99
23	0.59	22	0.82	22	0.79	19	0.83	18	0.76
24	0.50	22	0.68	22	0.84	19	0.67	18	0.67
25	0.36	22	0.50	22	0.89	19	0.56	18	0.58
26	0.91	22	1.00	22	1.00	19	1.00	18	0.98
27	0.55	22	0.68	22	0.89	19	0.94	18	0.77
28	0.45	22	0.23	22	0.37	19	0.22	18	0.32
29	1.00	22	0.91	22	1.00	19	1.00	18	0.98
30	0.64	22	0.82	22	0.89	19	0.94	18	0.82
31	0.91	22	1.00	22	1.00	19	0.94	18	0.96
32	0.68	22	0.68	22	0.89	19	0.78	18	0.76
33	0.82	22	0.95	22	0.84	19	1.00	18	0.90
34	0.55	22	0.55	22	0.89	19	0.67	18	0.66
35	0.77	22	0.82	22	0.68	19	0.61	18	0.72
36	0.59	22	0.41	22	0.95	19	1.00	18	0.74

Item	Age Band A (n=22)	number of participants who completed this item	Age Band B (n=22)	number of participants who completed the item	Age Band C (n=19)	number of participants who completed the item	Age Band D (n=18)	number of participants who completed the item	Mean
37	0.64	22	0.95	22	0.95	19	0.94	18	0.87
38	0.95	22	1.00	22	1.00	19	1.00	18	0.99
39	0.55	22	0.73	22	0.79	19	0.78	18	0.71
40	0.95	22	1.00	22	1.00	19	1.00	18	0.99
41	0.95	22	1.00	22	1.00	19	1.00	18	0.99
42	0.14	22	0.23	22	0.79	19	0.61	18	0.44
43	0.59	22	0.59	22	0.58	19	0.67	18	0.61
44	0.68	22	0.68	22	1.00	19	0.94	18	0.83
45	0.95	22	0.91	22	0.95	19	1.00	18	0.95
46	0.55	22	0.64	22	0.79	19	0.94	18	0.73
47	0.50	22	0.77	22	0.79	19	0.89	18	0.74
48	0.95	22	1.00	22	1.00	19	1.00	18	0.99
49	0.82	22	0.68	22	0.79	19	1.00	18	0.82
50	0.14	22	0.41	22	0.53	19	0.83	18	0.48
51	0.59	22	0.82	22	0.95	19	1.00	18	0.84
52	0.55	22	0.68	22	0.63	19	0.94	18	0.70
53	0.73	22	0.82	22	0.89	19	0.94	18	0.85
54	0.09	22	0.45	22	0.32	19	0.44	18	0.33
55	0.36	22	0.55	22	0.63	19	0.72	18	0.57
56	0.18	22	0.18	22	0.47	19	0.50	18	0.33
57	0.05	22	0.09	22	0.21	19	0.28	18	0.16
58	0.27	22	0.09	22	0.11	19	0.22	18	0.17
59	0.91	22	0.91	22	0.95	19	0.94	18	0.93
60	0.73	22	0.86	22	0.95	19	1.00	18	0.88

Table A- 12: Arabic Picture Vocabulary Test (APVT): Proportion of correct responses for the children with SLI (n=26) on the first 60 items of the test.

Item	Age Band A (n=5)	number of participants who completed this item	Age Band B (n=8)	number of participants who completed the item	Age Band C (n=5)	number of participants who completed the item	Age Band D (n=8)	number of participants who completed the item	Mean
1	1.00	5	1.00	8	1.00	5	1.00	8	1.00
2	1.00	5	1.00	8	1.00	5	1.00	8	1.00
3	1.00	5	0.88	8	1.00	5	1.00	8	0.97
4	1.00	5	1.00	8	1.00	5	1.00	8	1.00
5	0.80	5	1.00	8	1.00	5	1.00	8	0.95
6	1.00	5	1.00	8	1.00	5	1.00	8	1.00
7	0.60	5	0.88	8	1.00	5	1.00	8	0.87
8	1.00	5	0.88	8	1.00	5	1.00	8	0.97
9	1.00	5	1.00	8	1.00	5	1.00	8	1.00
10	0.20	5	0.63	8	0.60	5	0.75	8	0.54
11	1.00	5	1.00	8	1.00	5	1.00	8	1.00
12	1.00	5	1.00	8	1.00	5	1.00	8	1.00
13	0.80	5	0.63	8	1.00	5	1.00	8	0.86
14	0.80	5	0.75	8	0.80	5	1.00	8	0.84
15	0.80	5	0.88	8	1.00	5	1.00	8	0.92
16	0.80	5	0.88	8	0.80	5	1.00	8	0.87
17	0.80	5	0.38	8	0.40	5	0.88	8	0.61
18	0.20	5	0.50	8	0.60	5	1.00	8	0.58
19	1.00	5	0.88	8	1.00	5	1.00	8	0.97
20	1.00	5	0.88	8	1.00	5	1.00	8	0.97
21	1.00	5	0.75	8	0.80	5	1.00	8	0.89
22	1.00	5	0.88	8	1.00	5	1.00	8	0.97
23	0.00	5	0.25	8	0.20	5	0.88	8	0.33
24	0.40	5	0.38	8	0.00	5	0.25	8	0.26
25	0.40	5	0.29	7	0.40	5	0.88	8	0.49
26	1.00	5	0.71	7	0.80	5	0.88	8	0.85
27	0.40	5	0.29	7	0.60	5	0.75	8	0.51
28	0.60	5	0.14	7	0.00	5	0.25	8	0.25
29	0.60	5	1.00	7	1.00	5	1.00	8	0.90
30	0.80	5	0.29	7	0.80	5	1.00	8	0.72
31	1.00	5	0.86	7	1.00	5	0.88	8	0.93
32	0.80	5	0.57	7	0.60	5	0.63	8	0.65
33	0.20	5	1.00	7	0.80	5	1.00	8	0.75
34	0.20	5	0.14	7	0.20	5	0.88	8	0.35
35	0.20	5	0.14	7	0.60	5	0.88	8	0.45
36	0.60	5	0.14	7	0.60	5	0.50	8	0.46

Item	Age Band A (n=5)	number of participants who completed this item	Age Band B (n=8)	number of participants who completed the item	Age Band C (n=5)	number of participants who completed the item	Age Band D (n=8)	number of participants who completed the item	Mean
37	0.50	4	0.40	5	1.00	5	0.88	8	0.69
38	0.50	4	0.80	5	0.80	5	1.00	8	0.78
39	0.50	4	0.60	5	0.60	5	0.88	8	0.64
40	0.50	4	1.00	5	1.00	5	1.00	8	0.88
41	0.50	4	0.80	5	1.00	5	1.00	8	0.83
42	0.50	4	0.20	5	0.00	5	0.25	8	0.24
43	0.50	4	0.40	5	0.80	5	0.88	8	0.64
44	0.50	4	0.80	5	1.00	5	0.88	8	0.79
45	0.50	4	0.80	5	0.80	5	1.00	8	0.78
46	0.50	4	0.80	5	0.60	5	0.63	8	0.63
47	0.50	4	0.20	5	0.40	5	1.00	8	0.53
48	0.50	4	0.60	5	1.00	5	1.00	8	0.78
49	0.33	3	0.60	5	1.00	5	0.88	8	0.70
50	0.67	3	0.40	5	0.40	5	0.50	8	0.49
51	0.33	3	0.40	5	0.60	5	0.75	8	0.52
52	0.00	3	0.60	5	0.40	5	0.88	8	0.47
53	0.33	3	0.60	5	0.60	5	0.88	8	0.60
54	0.00	3	0.00	5	0.00	5	0.25	8	0.06
55	0.67	3	0.40	5	0.60	5	0.50	8	0.54
56	0.33	3	0.20	5	0.40	5	0.25	8	0.30
57	0.00	3	0.20	5	0.40	5	0.13	8	0.18
58	0.00	3	0.20	5	0.20	5	0.00	8	0.10
59	0.33	3	0.80	5	0.80	5	0.75	8	0.67
60	0.33	3	0.60	5	0.80	5	0.75	8	0.62

Appendix B: Familiarity rating of verbs and animals used in the Comprehension of sentences with fronted NP's

Table B-1: Familiarity ratings for the verbs

Verb	Familiarity Rating	Gloss
1. y-arfɪs	4.5	kick
2. y-dʒɪr	3.3	pull
3. y-xarɪʕ	4.4	scare
4. y-ʁasɪl	4.8	wash
5. y-tˤɪg	4.6	Hit
6. y-maʃʃɪtˤ	4.3	Comb
7. y-ħɪk	4.4	Scratch
8. y-ɪlħeg	4.3	Chase
9. y-bu:s	3.5	Kiss
10. y-ʔakkɪl	4.6	Feed
11. y-arsɪm	3.8	Draw
12. y-fattɪʃ	3.8	Search

* y- marks third person masculine singular (it is usually the default gender marker in Arabic

Table B-2: Familiarity rating for each pair of animals used in the sentence comprehension experiment

Animal	Familiarity Rating	Gloss	Animal	Familiarity Rating	Gloss
1a- tʃalb	4.61	dog	1b- Fi:l	3.88	elephant
2a- ʔasad	3.63	lion	2b- Dub	3.40	bear
3a- χaru:f	4.21	sheep	3b- ħsˤa:n	4.00	horse
4a- Naħlə	4.33	bee	4b- Fara:ʃə	4.42	butterfly
5a- Batˤtˤə	4.10	duck	5b- gatˤwə	4.33	cat
6- ħayyə	3.63	snake	6b- Bgarə	4.39	cow
7- Silħafa:t	3.29	turtle	7b- Sɪndʒa:b	2.90	squirrel
8a- ħma:r	3.90	donkey	8b- ħamamə	4.00	pigeon
9a- Du:da	2.90	worm	9b- ðˤifdaʕ	3.50	frog
10a- Tinni:n	3.40	dragon	10b- Naʕʕa:mə	3.50	ostrich
11a- Fa:r	3.90	mouse	11b- Namlə	4.70	ant
12a-Tɪmsa:ħ	3.70	crocodile	12b- Zara:fə	3.70	giraffe

Appendix C: A List of the 54 sentences used in the experiment

Type	greement	Gender				
		Subj	Obj			
SVO	SVO1	m	m	Il-fi:l	ya-rfɪs	ɪl-tʃalb
				the-elephant M	3M-kicks	the dog M
				The elephant kicks the dog		
	SVO1	m	m	Id-dub	y-dʒɪr	ɪl-ʔasad
				the-bear M	3M-pulls	the-lion M
				The bear pulls the lion		
	SVO1	m	m	Il-χaru:f	y-χarɪʕ	lɪ-ħsˤa:n
				the-sheep M	3M-scares	the-horse M
				The sheep scares the horse		
	SVO1	f	f	Il-fara:shə	t-aʁsɪl	ɪl-naħlə
				the-butterfly F	3F-washes	the-bee F
				The butterfly washes the bee.		
	SVO1	f	f	ɪl-batˤtˤə	t-tˤɪg	ɪl-gatˤwə
				the-duck F	3F-hits	the-cat F
				The duck hits the cat.		
	SVO1	f	f	Il-ħayyə	t-mashɪtˤ	l-ɪbgarə
				the-snake F	3F-combs	the-cow F
				The snake combs the cow.		
SVO						
	SVO2	m	f	Is-sɪnja:b	y-ħɪk	ɪl-sɪlħifa:t
				the-squirrel F	3MS-scratches	the-turtle F
				The squirrel scratches the turtle.		
	SVO2	m	f	lɪ-ħmaar	Y-ɪlħag	Il-batˤri:g
				the-donkey M	3M-chases	the-pigeon F
				The donkey chases the pigeon.		
	SVO2	m	f	Iðˤ-ðˤɪfdaʕ	y-bu:s	χl-du:də
				the-frog M	3M-kisses	the-worm F
				The frog kisses the worm.		
	SVO2	m	f	It-tɪnni:n	y-ʔakkɪl	ɪl-naʕʕa:mə
				the-dragon N	3M-feeds	the-ostrich F
				The dragon feeds the ostrich		
	SVO2	m	f	Il-fa:r	y-arsɪm	ɪn-namlə
				the-mouse M	3M-draws	the ant F
				The mouse draws the ant		
	SVO2	m	f	It-tɪmsa:ħ	y-fattɪʃ	Iz-zara:fə
				the-crocodile M	3M-searches	the-giraffe F
				The crocodile searches the giraffe		
SVO	SVO3	f	m	Il-sɪlħɪfa:t	t-ħɪk	ɪs-sɪndʒa:b
				the-turtle F	3F-scratches	the-squirrel M
				The turtle scratches the squirrel		

	SVO3	f	m	Il-ħama:mə	t-ɪlħaq	lɪ-ħma:r
				the-pigeon F	3F-chases	the-donkey M
				The pigeon chases the donkey		
	SVO3	f	m	Il-du:də	t-bu:s	ɪðˤ-ðˤɪfdaʔ
				the-worm F	3FS-kisses	the-frog M
				The worm kisses the frog		
	SVO3	f	m	Il-naʔʔa:mə	t-?akkɪl	ɪt-tɪnni:n
				the-ostrich F	3F-feeds	the-dragon M
				The ostrich feeds the dragon		
	SVO	f	m	In-namlə	t-arsɪm	ɪl-fa:r
				the ant F	3F-draws	the-mouse M
				The ant draws the mouse		
	SVO3	f	m	ez-zara:fə	t-fatɪʃ	ɪt-tɪmsa:ħ
				the-giraffe F	3F-searches	the-crocodile M
				The giraffe searches the crocodile		
OSV						
	OSV1	m	m	Il-fi:l	l-tʃalb	ya-rfɪs-a
				The-elephant M	the dog M	3M-kiks-him CL
				The elephant, the dog kicks him		
	OSV1	m	m	Id-dub	ɪl-ʔasad	y-dʒɪr-a
				The-bear M	the-lion M	3M-pulls-him CL
				The bear, the lion pulls him.		
	OSV1	m	m	Il-χaruf	lɪ-ħsˤan	y-χarʕ-a
				the-sheep M	the-horse M	3M-scares-him CL
				The sheep, the horse scares him		
	OSV1	f	f	Il-fara:ʃə	ɪl-naħlə	t-ʁasɪl-ha
				The-butterfly F	the-bee F	3F-washes-her CL
				The butterly, the bee washes her		
	OSV1	f	f	Il-batˤtˤə	ɪl-gatˤwə	t-tˤɪg-ha
				the-duck F	the-cat F	3FS-hits-her CL
				The duck, the cat hit her		
	OSV1	f	f	Il-ħayyə	l-ɪbgara	t-maʃɪtˤ-ha
				The-snake F	the-cow F	3FS-combs- her CL
				The snake, the cow combs her		
	OSV2	m	f	Il-sɪlħɪfa:t	ɪs-sɪndʒa:b	y-ħɪk-ha
				The-turtle F	the-squirrel M	3M-scratches-her CL
				The turtle, the squirrel scratches her		
	OSV2	m	f	Il-ħama:mə	lɪ-7mar	y-ɪlħaq-ha
				The-pigeon F	the-donkey M	3M-chases-her CL
				The pigeon, the donkey chases her		
	OSV2	m	f	Il-du:da	iðˤ-ðˤifdaʔ	y-bu:s-ha
				The-worm F	the-frog M	3M-kisses-her CL
				The worm, the frog kisses her		
	OSV2	m	f	Il-naʕʕa:mə	ɪt-tɪnni:n	y-ʔakkɪl-ha
				The-ostrich F	the-dragon M	3M-feeds-her CL

				The ostrich, the dragon feeds her		
	OSV2	m	f	In-namlə	ɪl-fa:r	y-arsɪm-ha
				the ant F	the-mouse M	3M-draws-her CLI
				The ant, the mouse draws her..		
	OSV2	m	f	Iz-zara:fə	et-tɪmsaħ	y-fatɪʃ-ha
				The-giraffe F	the-crocodile M	3M-searches-her CL
				The giraffe, the crocodile searches her		
	OSV3	f	m	Is-sɪndʒa:b	ɪl-sɪlħɪfa:t	t-ħɪk-a
				the-squirrel M	the-turtle F	3F-scratches him CL
				The squirrel, the turtle scratches him		
	OSV3	f	m	lɪ-ħmar	ɪl-ħama:mə	t-ɪlħɪg-a
				the-donkey M	the-pigeon F	3F-chases him CLI
				The donkey, the pigeon chases him		
	OSV3	f	m	Iðˤ-ðˤɪfdaʔ	ɪl-du:də	t-bu:s-a
				the-frog M	the-worm F	3F-kisses-him CL
				The frog, the worm kisses him		
	OSV3	f	m	It-tɪnni:n	el-naʕʕamə	t-ʔakl-a
				the-dragon	the-ostrich F	3F-feeds-him CL
				The dragon, the ostrich fees him		
	OSV3	f	m	Il-fa:r	ɪn-namlə	t-arsɪm-a
				the-mouse M	the ant F	3F-draws-him CL
				The mouse, the ant draws him		
	OSV3	f	m	It-tɪmsa:ħ	ɪz-zara:fə	t-fatʃ-a
				the-crocodile M	the-giraffe F	3F-searches-him CLI
				The crocodile, the giraffe searches him		
OVS	OVS1	m	m	Il-tʃalb	ya-rfɪs-a	il-fi:l
				the dog M	3M-kicks-him CL	the-elephant
				The dog, the elephant kicks him		
	OVS1	m	m	Id-dub	y-dʒɪr-a	Il-ʔasad
				the-bear M	3M-pulls-him CL	the-lion M
				The bear, the lion pulls him		
	OVS1	m	m	lɪ-ħsan	y-xarʔ-a	ɪl-xarou:f
				the-horse M	3MS-scares-him CL	the-sheep M
				The horse, the sheep scares him		
	OVS1	f	f	Il-fara:ʃə	t-ʁasɪl-ha	ɪl-naħlə
				the-butterfly F	3F-washes-her CL	the-bee
				The butterfly, the bee washer her		
	OVS1	f	f	Il-gatˤwə	t-tɪg-ha	ɪl-batˤtˤə
				the-cat M	3F-hits- her CL	the-duck
				The cat, the duck hits him		
	OVS1	f	f	Il-ħayyə	t-maʃɪt-ha	l-ɪbgarə
				the-snake F	3FS-combs- her CL	the-cow F
				The snake, the cow combs her.		
	OVS2	m	f	Il-sɪlħɪfa:t	y-ħɪk-ha	ɪs-sɪndʒa:b
				the-turtle F	3M-scratches-her CL	the-squirrel M
				The turtle, the squirrel scratches her		

OVS2	m	f	Il-ħama:mə	y-ɪlħaq-ha	lɪ-ħma:r
			the-pigeon F	3M-chases-her CL	the-donkey M
			The pigeon, the donkey chases her		
OVS2	m	f	Il-du:də	y-bu:s-ha	iðˤ-ðˤIfdaʔ
			the-worm F	3M-kisses- her CL	the-frog M
			The worm, the frog kisses her		
OVS2	m	f	Il-naʕʕa:mə	y-ʔakkɪl-ha	It-tInni:n
			the-ostrich M	3M-feeds- her CL	the-dragon M
			The ostrich, the dragon feeds her.		
OVS2	m	f	In-namlə	y-arsɪm-ha	ɪl-fa:r
			the ant F	3M-draws-her CL	the-mouse M
			The ant, the mouse draws her		
OVS2	m	f	Iz-zara:fə	y-fatɪʃ-ha	It-tɪmsa:ħ
			the-giraffe F	3M-searches-her CL	the-crocodile M
			The giraffe, the crocodile searches her		
OVS3	f	m	Is-sɪndʒa:b	t-ħɪk-a	Il-sɪlħɪfa:t
			the-squirrel M	3FS-scratches-him CL	the-turtle F
			The squirrel, the turtule scratches him		
OVS3	f	m	lɪ-ħma:r	t-ɪlħɪg-a	ɪl-ħama:mə
			the-donkey M	3FS-chases-him CL	the-pigeon F
			The donkey, the pigeon chases him		
OVS3	f	m	Iðˤ-ðˤɪfdaʔ	t-bu:s-a	ɪl-du:da
			The-frog M	3FS-kisses- him CL	the-worm F
			The frog, the worm kisses him.		
OVS3	f	m	It-tɪnni:n	t-?akl-a	Il-naʕʕa:mə
			The-dragon	3FS-feeds- him CL	the-ostrich F
			The dragon, the ostrich feeds him.		
OVS3	f	m	Il-fa:r	t-arsɪm-a	ɪn-namlə
			The-mouse	3FS-draws- him CL	the ant F
			The mouse, the ant draws him		
OVS3	f	m	It-tɪmsaħ	t-fatʃ-a	ɪz-zara:fə
			The-crocodile M	3FS-searches- him CL	the-giraffe F
			The crocodile, the giraffe searches him		

Appendix D: The nonword stimuli organised according to nonroots, number of syllables, patterns, and syllable types

Non-root	Pattern	No of Syllables	Syllable Type No Cluster	Medial Cluster	Final Cluster	Medial+Final
/STL/		2 syllables	CV.CVC	CVC.CVC	CV.CVCC	CVC.CVCC
	a-u		sa.tul	das.tul	sa.tulb	das.tulb
		3 syllables	CV.CV.CVC	CVC.CV.CVC	CV.CV.CVCC	CV.CVC.CVCC
	a-u-a		da.su.tal	das.bu.tal	da.su.talb	da.sum.talb
/KDF/	u-i	2	ku.dif	suk.dif	ku.difs	suk.difs
	u-i-a	3	su.ki.daf	suk.bi.daf	su.ki.dafs	Su.kim.dafs
/DLS/	a-u	2	da.lus	kad.lus	Da.lusb	kad.lusb
	a-u-a	3	ka.du.las	kad.mu.las	ka.du.lasb	ka.dum.lasb
/ SBN/	a-u	2	sa.bun	das.bun	sa.bunf	kas.bunf
	u-i-a	3	ku.si.ban	kus.mi.ban	ku.si.banf	ku.sib.banf
/ DNF/	a-u	2	da.nuf	sad.nuf	da.nufd	Sad.nufd
	a-u-a	3	sa.du.naf	sad.lu.naf	sa.du.nafd	sa.dun.nafd
/ KMS/	a-u	2	daka.mus	dak.mus	ka.musd	dak.musd
	u-i-a	3	du.ki.mas	duk.li.mas	du.ki.masd	du.kim.masd
/DFL/	a-a	2	da.fal	kad.fal	da.falb	kad.f.alb
	a-a-a	3	ka.da.fal	kad.ba.fal	ka.da.falb	ka.dam.falb

Note. Fullstops indicate syllable boundary.

Appendix E: Nonword Repetition Test

Table E-1: Nonword Repetition Test: List A

Item	Nonword		Nonword
1	ka.du.las	31	da.nufd
2	ka.da.fal	32	sa.tulb
3	su.ki.dafs	33	du.kim.masd
4	sa.tul	34	su.kim.dafs
5	da.falb	35	kad.mu.las
6	ku.si.banf	36	sa.dun.nafd
7	dak.mus	37	das.bu.tal
8	das.bun	38	ka.du.lasb
9	sad.nufd	39	ka.musd
10	suk.difs	40	ku.dif
11	ka.dum.lasb	41	ku.sib.banf
12	duk.li.mas	42	su.ki.daf
13	kas.bunf	43	du.ki.mas
14	ka.dam.falb	44	da.nuf
15	da.su.tal	45	sa.bun
16	sad.lu.naf	46	da.sum.talb
17	dak.musd	47	da.fal
18	ka.da.falb	48	da.lusb
19	sad.nuf	49	kad.ba.fal
20	das.tul	50	ka.mus
21	kad.lus	51	Suk.bi.daf
22	ku.difs	52	Sa.bunf
23	das.tulb	53	Suk.dif
24	sa.du.nafd	54	kad.f.alb
25	da.lus	55	Du.ki.masd
26	ku.si.ban	56	kus.mi.ban
27	sa.du.naf		
28	kad.lusb		
29	kad.fal		
30	Da.su.talb		

Trial items: 1-Kal 2- Meek 3-Difel 4-Fedeleb

Table E-2: Nonword Repetition Test: List B

Item	Nonword		Nonword
1	duk.li.mas	31	ka.musd
2	das.bun	32	ku.sib.banf
3	su.ki.dafs	33	ka.du.lasb
4	da.falb	34	ku.dif
5	das.tul	35	ka.mus
6	kus.mi.ban	36	da.lus
7	ka.dum.lasb	37	ku.si.banf
8	du.kim.masd	38	kad.f.alb
9	su.kim.dafs	39	da.sum.talb
10	ka.da.falb	40	suk.bi.daf
11	kad.mu.las	41	ka.da.fal
12	sa.tulb	42	sad.lu.naf
13	sa.dun.nafd	43	da.su.tal
14	kad.lus	44	suk.difs
15	kad.fal	45	dak.mus
16	sa.du.naf	46	kas.bunf
17	kad.lusb	47	sa.tul
18	dak.musd	48	sad.nuf
19	das.tulb	49	ka.du.las
20	su.ki.daf	50	sad.nufd
21	kad.ba.fal	51	du.ki.masd
22	ku.si.ban	52	das.bu.tal
23	da.nuf	53	ku.difs
24	du.ki.mas	54	ka.dam.falb
25	suk.dif	55	da.nufd
26	da.fal	56	sa.bunf
27	da.lusb		
28	sa.bun		
29	sa.du.nafd		
30	da.su.talb		

Appendix F: Pairwise comnparisons of types of clusters for the nonword repetition test

(I) cluster	(J) Cluster	Mean Difference (I-J)	Significance [a]
1	2	10.09*	.000
	3	13.36*	.000
	4	33.36*	.000
2	1	-10.09*	.000
	3	3.27	.178
	4	23.27*	.000
3	1	-13.36*	.000
	2	-3.27	.178
	4	20.00*	.000
4	1	-33.36*	.000
	2	-23.27*	.000
	3	-20.00*	.000

* The mean difference is significant at the .05 level.
[a] Adjustment for multiple comparisons: Bonferroni.

Note. 1=nonwords with no clusters, 2=nonwords with medial clusters only, 3=nonwords with final clusters only, and 4=nonwords with medial and final clusters.

Appendix G: Mulitple comparison with Bonferroni correction for the different types of cluster

Bonferroni

Dependent Variable	(I) groups	(J) groups	Mean Difference (I-J)	Sig.
Average score of No cluster words	SLI	Lang Control	-10.6364	.153
		Age Control	-21.2727*	.001
	Lang Control	SLI	10.6364	.153
		Age Control	-10.6364	.153
	Age Control	SLI	21.2727*	.001
		Lang Control	10.6364	.153
Average score of Medial cluster words	SLI	Lang Control	-18.2727	.061
		Age Control	-26.7273*	.004
	Lang Control	SLI	18.2727	.061
		Age Control	-8.4545	.799
	Age Control	SLI	26.7273*	.004
		Lang Control	8.4545	.799
Average score of Final cluster words	SLI	Lang Control	-29.5455*	.001
		Age Control	-39.4545*	.000
	Lang Control	SLI	29.5455*	.001
		Age Control	-9.9091	.516
	Age Control	SLI	39.4545*	.000
		Lang Control	9.9091	.516
Average scores of M and F cluster words	SLI	Lang Control	-26.4545*	.019
		Age Control	-34.9091*	.002
	Lang Control	SLI	26.4545*	.019
		Age Control	-8.4545	1.000
	Age Control	SLI	34.9091*	.002
		Lang Control	8.4545	1.000

*. The mean difference is significant at the .05 level.

Appendix H : Descriptive statistics for types of errors in NWR

	Type of error		SLI	LC	AC
Structural errors	Final cluster reduction	*N*=	121	48	34
		M	20.17	4.36	3.09
		SD	4.07	4.15	2.77
		Range	2-17	1-15	0-8
	Medial cluster reduction	*N*=	24	14	6
		M	4.00	2.33	0.55
		SD	2.04	1.63	0.52
		Range	0-7	0-4	0-1
	Syllable omission	*N*=	13	12	3
		M	2.17	1.09	0.27
		SD	1.72	1.51	0.65
		Range	0-5	0-5	0-2
	Metathesis	*N*=	7	4	5
		M	2.33	0.36	1
		SD	0.55	0.50	0.71
		Range	1-2	0-1	0-2
	Final consonant deletion	*N*=	4	11	3
		M	2.0	1.38	0.75
		SD	0.58	0.52	0
		Range	1-2	1-2	0-1
	Cluster Creation	*N*=	15	0	1
		M	2.5		0.09
		SD	1.5		0.30
		Range	0-5		0-1
	Total Structural errors	*N*=	184	89	51
		%	46	43	45
Segmental errors,	Consonant Substitutions	*N*=	204	97	66
		M	34.0	8.82	6.0
		SD	9.07	4.90	4.34
		Range	4-31	2-19	1-12
	Vowel substitution	*N*=	10	10	2
		M	2.86	2.50	2.0
		SD	1.03	1.29	0.0
		Range	0-3	1-4	2-2
	Total non Structural errors,	*N*=	214	107	68
		%	54	57	55
Total errors *N*=			398	196	119

Appendix I: AVOVAs for various types of error patterns on the NWR test for the three group (SLI, LC, and AC)

		df	F	Sig.
medial cluster reduction	Between Groups	2.00	0.12	0.89
	Within Groups	30.00		
	Total	32.00		
final cluster reduction	Between Groups	2.00	0.96	0.39
	Within Groups	30.00		
	Total	32.00		
syllable omission	Between Groups	2.00	0.23	0.79
	Within Groups	30.00		
	Total	32.00		
cluster creation (medial+final)	Between Groups	2.00	4.09	0.027*
	Within Groups	30.00		
	Total	32.00		
metathesis	Between Groups	2.00	1.11	0.34
	Within Groups	30.00		
	Total	32.00		
Final consonant deletion	Between Groups	2.00	3.56	0.041*
	Within Groups	30.00		
	Total	32.00		
consonant substitution	Between Groups	2.00	0.31	0.74
	Within Groups	30.00		
	Total	32.00		
vowel substitution	Between Groups	2.00	0.46	0.64
	Within Groups	30.00		
	Total	32.00		

Appendix J: Articulation Screener

	Phoneme	Stimulus	Initial	Medial	Final	Gloss
1.	/b/	/ba:sˤ/	b		sˤ	Bus
2.	/t/	/təmər/	t	m	r	Date
3.	/tˤ/	/ tˤyya:rə?	tˤ	b	l	Airplane
4.	/d/	/di:tʃ/	d	i	tʃ	Rooster
5.	*/dˤ/	/dˤɪfdaʕ/	dˤ	F	ʕ	Frog
6.	/k/	/keik/	k	e	l	Cake
7.	/g/	/galb/	g	l	b	Heart
8.	/q/	/qaləm/	q	l	m	Pen
9.	/ ðˤ/	/ ɪðˤfir/	I	ðˤ	r	Finger nail
10.	/m /	/maree dˤ/	m	r	dˤ	Patient
11.	/n/	/namɪl/	n	m	l	Ants
12.	/ tʃ /	/ tʃa:y/	tʃ	a	y	tea
13.	/dʒ/	/ dʒɪsɪr/	dʒ	s	r	bridge
14.	/f/	/fa:r/	f	a	r	mouse
15.	/θ/	/θalʒ/	θ	l	ʒ	ice
16.	/ð/	/ ðɪbba:n/	ð	b	n	Fly(insect)
17.	/ðˤ/	/ ðˤɪl/	ðˤ	i	l	shadow
18.	/s/	/sari:r/	s	r	r	bed
19.	/sˤ/	/ sˤa:ru:χ/	sˤ	r	x	rocket
20.	/z/	/zɪr/	z	i	r	button
21.	/ʃ/	/ ʃa:rɪʕ/	ʃ	r	ʕ	street
22.	/x/	/χeitˤ/	x	ei	tˤ	thread
23.	/ʁ/	/ ʁeim/	ʁ	ei	m	cloud
24.	/ħ/	/ ħɪ sˤa:n/	ħ	sˤ	n	horse
25.	/ʕ/	/ ʕanab/	ʕ	n	b	grapes
26.	/h/	/hadɪyyə/	h	d	ye	gift
27.	/l/	/la:ʕɪb/	l	ʕ	b	player
28.	/r/	/ra:s/	r		s	head
29.	/w/	/warag/	w	r	g	paper
30.	/y/	/youm/	Y		m	day

* The substitution of /ðˤ/ for /dˤ/ is acceptable in Gulf Arabic.

Appendix K: Apraxia and Oral-Motor Screener

Child's name:____________________________ D.O.B: ______________________________

School:_________________________________ Date:_________________________________

A. Engage child in a conversation for 5 minutes and notice the following:

1- Intelligibiliy: ☐ Intelligible ☐ Not intelligible

2- Grimaces or unusual oral motor movements: ☐ Not observed ☐ Observed

Other observations: (e.g. difficulties controlling saliva): ______________________

B. Ask the child to perform the following actions:

1- To pucker up (as if kissing something): ☐ Can ☐ Cannot

2- To spread his lips (as in smiling): ☐ Can ☐ Cannot

3- To alternate puckering and smiling: ☐ Can ☐ Cannot

B. Ask the child to perform repeat the following words

	Word (English)		
CVC	mouz (banana)	☐Can	☐ Cannot
$C_1V_1C_2V_2$	haki (take)	☐Can	☐ Cannot
Blend Synthesis	swa:r (bracelet)	☐Can	☐ Cannot
/r/ blends	fra:ʃ (bed)	☐Can	☐ Cannot
C_FVC_B	di:k (rooster)	☐Can	☐ Cannot
C_BVC_F	ga:t	☐Can	☐ Cannot
CVCVC	wagɪf (standing)	☐Can	☐ Cannot
Polysyllabic Synthesis	ka:mɪra (camera)	☐Can	☐ Cannot

Appendix L: The Arabic Language Test

Record Form

Child's Name/ID:…………………………….. Gender: M / F

School: …………………………………………………Grade………….

Examiner:……………………………………………………….……………..

	Year	Month	Day
Test Date			
Birth Date			
Chronological Age			

	Raw Score		
Sentence Comprehension A	/22		
Sentence Comprehension B	/18		
	Total SC	**/40**	
Expressive Language A	/24		
Expressive Language B	/44		
	Total WS	**/68**	
Sentence Repetition A	/54		
Sentence Repetition B	/69		
	Total SR	**/123**	
Total Score			/231

Appendix M: Sentence Comprehension (A)

Manual	Repetition
SC(A)	Not Allowed

Instructions: Before presenting any picture say to the child: 'I want you to point to………..'. Place a circle on the number representing the child's answer. Then circle 1 for correct answers, 0 for incorrect one, and NR for no response.

Practice A: "ɪlli y-tˤi:r'? 3 2 **1** Practice B: "ɪl-tʃalb ɪl-kɪbi:r " 3 **2** 1
' that 3MS-fly the-dog the-big'
'flying'? 'the big dog'

		Score		
1- ɪl-walad y-ɪsbaħ The-boy 3MS-swim The boy is swimming	3 **2** 1	0	1	NR
2- ɪl-dʊb f-ɪl ʕaraba:nə The bear in-the wagon The bear is in the əwagon	3 2 **1**	0	1	NR
3- ɪl-fa:r taħt ɪl kɪrsɪ The-mouse under the chair The mouse is under the chair	3 **2** 1	0	1	NR
4- χalasˤt ɪl-ʔakɪl Finished the-food I finished the food	**3** 2 1	0	1	NR
5- ɪl-bɪnt gaʕdə t-ɪsbaħ The-girl Aux(is) 3FS-swim The girl is swimming	3 **2** 1	0	1	NR
6- ɪl-rayyal fɪtɪħ ɪl-ba:b The-man opened the-door The man opened the door	3 2 **1**	0	1	NR
7- ɪl-walad y-sˤi:ħ ʕaʃa:n tˤayya:rt-ə ɪnkasrat The-boy 3MS-cry because airplane-his broke The boy cries because his airplane broke down	3 **2** 1	0	1	NR
8- ɪl-bɪnt χaðat ʃwayyat ward ħag ʊm-ha The-girl took some flower for mother-her The girl took some flowers for her mother	3 **2** 1	0	1	NR
9- ɪhya tɪsʕad wə-hwa y-ɪlʕab mɪrdʒeiħə She 3FS-climb and-he 3MS-play swing She is climbing and he is playing on the swing	3 **2** 1	0	1	NR
10- ɪl-rayyal ɪllɪ gaʕɪd taħt lɪ-ʃyarə y-ɪlbas kabbous The-man who sitting under the-tree 3MS-wear hat The mean who is sitting under the tree is wearing a hat	**3** 2 1	0	1	NR

	Score			
11- Warri-nɪ wein y-ɪlʕab ɪl-walad ku:rə Show-me where 3MS-play the-boy ball Show me where the boy play football	3 2 **1**	0	1	NR
12- ɪl-mara ɪmsɪkat smɪtʃə kbi:rə The woman caught fish big The woman caught a big fish	**3** 2 1	0	1	NR
13- ɪl-tʃalb ɪl-rʊma:dɪ da:χɪl ɪl- sˤandu:g The dog the-grey inside the-box The grey dog in the box	3 2 **1**	0	1	NR
14- ɪl-bɪnt mʊb ga:ʕdə t-arsɪm The-girl not Aux-is 3FS-draw The girl is not drawing	3 2 **1**	0	1	NR
15- Laʔ la t-almɪs No! don't 2MS-touch No! Don't touch	3 **2** 1	0	1	NR
16- ɪhwa ra:ħ y-a:kɪl ɪt-tɪffa:ħə He will 3MS-eat the-apple He will eat the apple	3 **2** 1	0	1	NR
17- ɪhwa dʒa:hɪz ʕaʃa:n y-rouħ y-argɪd He ready because 3MS-go 3MS-sleep He is ready to go to sleep	3 **2** 1	0	1	NR
18- ɪhya t-agdɪr t-amsɪk ɪl-kɪta:b She 3FS-can 3FS-catch the-book She can reach the book	**3** 2 1	0	1	NR
19- ɪl-bɪnt t-ɪlħag-ha gatˤwat-ha The-girl 3FS-follow-her cat-hers The girl, the cat follows her.	**3** 2 1	0	1	NR
20- ɪl-ʊm warrat ɪl-gatwə ħaq ɪl-tʃalb The-mother showed the-cat to the-dog The mother showed the cat to the dog	3 2 **1**	0	1	NR
21- ɪl-walad y-ɪsbaħ The-boy 3MS-swim The boy is swimming	**3** 2 1	0	1	NR
22- ɪl-walad ʃa:f ɪl-bɪnt ɪllɪ gaʕdə t-ʃi:l matˤrɪga The-boy saw the-girl that Aux-is 3FS-carry hammer The boy saw the girl was carrying a hammer	3 2 **1**	0	1	NR

Appendix N: Sentence Comprehension (B)

Manual	Repetition
SC(B)	Not Allowed

Instructions: Same as in SC-A

	Score			
23- ɪl-walad mʊb gaʕɪd y-ɪsˤʕad The-boy not Aux-is 3MS-climb The boy is not climbing	4 3 **<u>2</u>** 1	0	1	NR
24- ɪl-bɪnt ʕɪnd-ha tʃalb kɪbi:r w mnaggatˤ ʔabyaðˤwʔaswad The girl has-her dog big and spotty white and black The girl has big spotty (with black and white) dog	4 3 2 **<u>1</u>**	0	1	NR
25- ɪl-walad ɪlli gaʕɪd taħt lɪ-ʃyarə ɪl-kɪbi:rə y-akɪl mouza The boy that sitting under the-tree the-big 3MS-eat banana The boy who is sitting under the big tree is eating a banana	**<u>4</u>** 3 2 1	0	1	NR
26- ɪl-walad gaʕɪd y-dɪz ɪl-ya:hɪl The boy Aux-is 3MS-push the-child The boy is pushing the child	4 **<u>3</u>** 2 1	0	1	NR
27- ɪl-bɪnt tʕadɪl ɪl-χeimə w-ɪl rayyal y-gatˤɪʕ ɪl-χaʃab The-girl fix the-tent and-the man 3MS-chop the-wood The girl is fixing the tent and the man is chopping the wood	4 3 2 **<u>1</u>**	0	1	NR
28- ɪl-mara ɪlli ʃayla lɪ-bnayya tˤayyaħat ʃantˤat-ha The-woman who carry the-girl dropped bag-her The woman who is carrying the girl dropped her bag	4 3 **<u>2</u>** 1	0	1	NR
29- ɪl-batˤtˤə gaʕdə t-ɪmʃɪ sˤoub ɪl-bɪnt The-duck aux-is 3FS-walk towards the-girl The duck is walking towards the girl	**<u>4</u>** 3 2 1	0	1	NR
30- ʔawal tˤɪflei:n wagfi:n f-ɪl tˤabour bas ɪl tˤɪfl ɪl-θalɪθ gaʕɪd y-ɪlʕab First two children standing in-the line but the child the third aux-is 3MS-play The first two children are standing but the third one is playing	**<u>4</u>** 3 2 1	0	1	NR
31- ɪl-ɪbu gaʕɪd y-warɪ ɪl-ya:hɪl ɪl-tʃalb The-father is showing the-child the-dog The father is showing the child the dog	4 3 2 **<u>1</u>**	0	1	NR
32- ɪl-ʊm gaʕdə t-ɪsʔal leiʃ ma t-ɪlbas dʒakeit-ɪk The-mother Aux-is 3FS-ask why don't 2MS-wear jacket-your The mother is asking 'why don't you wear your jacket'	4 3 **<u>2</u>** 1	0	1	NR
33- ɪl-walad gaʕɪd y-ɪnlɪħɪg The-boy aux-is 3MS-being chased The boy is being chased	**<u>4</u>** 3 2 1	0	1	NR

	Score			
34- ɪhya ra:ħ t-saʕɪd ʊm-ha t-sawɪ ɪl-ʕaʃə She will 3FS-help mother-her 3FS-make the dinner She will help her mother prepare the dinner	4 **<u>3</u>** 2 1	0	1	NR
35- ɪl-bɪnt gaʕdə t-ɪlbɪs dʒakeit-ha ɪl-dʒɪdid maʕa ɪn-ha ma: tɪħta:dʒ-ə The-girl Aux-is 3FS-wear jacket-her the-new although that-she not need-it The girl is wearing her new jacket eventhough she does not need it	4 **<u>3</u>** 2 1	0	1	NR
36- ɪhya ʃrɪbˤat ɪl-ħalib gabɪlˤ ma: t-a:kɪl ɪl-baskout She drank the-milk before not 3FS-eat the-biscuit She drank the milk before eating the biscuits	4 3 2 **<u>1</u>**	0	1	NR
37- ɪl-walad gaʕɪd y-sˤi:ħ lɪ-annə gasˤ sˤɪbʕ-ə The-boy Aux-is 3MS-cry for-he cut finger-his The boy is crying because he cut his finger.	4 **<u>3</u>** 2 1	0	1	NR
38- ɪl-bɪnt t-goul wein χaʃeit-u ɪl-hadɪyə The-girl 3FS-say where hid-you the-present The gir is saying: where did you hide the present	**<u>4</u>** 3 2 1	0	1	NR
39- ɪl-bɪnt ra:dʒʕə ɪl-beit mɪn ɪl-madrɪsə The-girl returning the-home from the-school The girl is going home from the school	4 **<u>3</u>** 2 1	0	1	NR
40- ɪl-walad ra:ħ y-ʔakɪl ɪl-tʃalb The-boy will 3MS-feed the-dog He will eat the apple	4 3 **<u>2</u>** 1	0	1	NR

Appendix O: Expressive Language Test (A)

Manual	Repetition
EL(A)	Allowed once only

Instructions: Say to the child: “I will show you some pictures and I will say something about these pictures. I want you to complete what I say. Is this clear? Let’s try some”. Point to the picture in practice 1a): ‘ɪl walad gaʕɪd y-ɪlʕab’ (the boy is playing), then point to the picture in practice 1 b: ‘‘ɪl walad gaʕid” (the boy is………..).” If the child does not answer in 10 seconds, point to the picture in Practice 1b and say: ‘look ‘‘‘ɪl walad gaʕid ya-kɪl’ (the boy is eating)”. Continue until the child understands the instructions.

Practice 1a:
‘ɪl-walad gaʕid y-ɪlʕab
The-boy Aux-is 3MS-play
The boy is playing

1b:
ɪl-walad gaʕɪd….(y-akɪl)
The-boy Aux-is (3MS eat)
The boy is (eating)

Practice 2a:
Ha:ðə ɪl-sˤaħan fih fara:wlə waħdə
This the-plate in-it strawberry one
This plate has one strawberry

2b:
Ha:ðə ɪl-sˤaħan fih χams…(farawla:t)
This the-plate in-it five…(strawberries)
This plate has five (strawberries)

Scoring: Circle 1 for correct answers, 0 for incorrect one, and NR for no response.

	Score		
1- A. Haðɪ ɪl-lɪʕbə barrə ɪl-sˤandoug This the-toy outside the-box This toy is outside the box B. Haðɪ ɪl-lɪʕbə ……….. ɪl-sˤandoug (daχɪl) This the-toy ………… the-box (outside) This toy is ……………..the box (outside)	0	1	NR
2- A. ɪhni fi: bɪnt gaʕdə targɪd Here there girl Aux-is 3FS-sleep Here, the girl is sleeping B. ɪhni fi: bɪnt gaʕdə……(t-ɪlʕab) Here there girl Aux-is ……..(3FS-play) Here, the girl is ……………..(playing)	0	1	NR
3- A. ɪhni fi: batˤtˤə waħdə Here there duck one Here, there is one duck B. ɪhni fi: θalaθ ……….(batˤtˤa:t) Here there three…………(ducks) Here, there are three……..(ducks)	0	1	NR

4- A. Maryam t-aʕtˤɪ hadɪyyə ħag ɪχɪt-ha Maryam 3FS-give present to sister-her Maryam gives a present to her sister B. Maryam t-aʕtˤɪ hadɪyyə ħag …….(ʊχu-ha) Maryam 3FS-give present to…….. (brother-her) Maryam gives a present to……. (her brother)	0	1	NR
5- A. Haðɪ ʕaraba:nt-ə This (F) wagon-his This is his wagon. B. Haðə............ (seikel-ha) This (M)………(bike-her) This is ……..(her bike)	0	1	NR
6- A. ɪl-walad y-sˤɪ:d sɪmatʃ The-boy 3MS-catch fish The boy catches fish B. ɪl-ʔwla:d………………(y-ɪlʕɪb-un (kura)) The-boys………………(3M-play-Plural (ball)) The boys………….(play (bal))	0	1	NR
7- A. Mɪnu gaʕɪd y-argɪd? ɪhwa gaʕɪd y-argɪd Who Aux-is sleep? He Aux-is 3MS-sleep Who is sleeping? He is sleeping. B. Mɪnu gaʕɪd y-ʔaʃɪr?..........(ɪhya gaʕɪdə t-ʔʃɪr) Who Aux-is sleep? ………(She Aux-is 3FS-point) Who is sleeping? ………...(She is pointing)	0	1	NR
8- A. ɪl-ʔays kri:m ha:ðə ħag ɪl-walad Ice Cream this for the-boy This ice cream is for the boy B. ha:ðɪ ɪl-ʔalʕa:b………(ħag ɪl-bɪnt) These the-toys ………..(for the-girl) These toys …….(are for the girl).	0	1	NR
9- A. ɪhwa y-ʔakkɪl nafs-a He feed self-him He feeds himself B. Hɪya t-labbɪs………….(nafs-hə) She 3FS-dress……….(self-her) She dresses…………..(her self)	0	1	NR
10- A. Ha:ðɪ nɪfeiχə, ʔms ɪl-walad nɪfaχ nɪfeiχə This balloon, yesterday the-boy blew balloon This is a balloon, yesterday the boy blew a balloon B. Haðɪ ku:rə, ʔms ʃɪnu sawwa ɪl-walad…..(ʃa:t ɪl-ku:rə) This ball, yesterday what di the-boy …..(hit the-ball) This is a ball, what did the boy do yesterday?....(hit the ball)	0	1	NR
11- A. ɪhni: ka:n fi: warda, ɪl-walad ʃala-ha Here was there flower, the-boy took-it (FS) Here, there was a flower, the boy took it. B. ɪhni: ka:n fi: warda:t, ɪl-bɪnt……(ʃalat-hum) (FP) Here was there flowers, the-girl…...(took-them)	0	1	NR

	Score		
12- A. Ha:ðə ɪl-walad gaʕɪd y-tɪzaɦlag, Ha:ðə ɪl-walad ra:ħ y-tɪzaɦlag This the-boy Aux-is 3MS-skate, this the boy will 3MS-skate This boy is skating, this boy will skate. B. Ha:ðə ɪl-walad gaʕɪd y-lawwɪn, Ha:ðə ɪl-walad…(ra:ħ y-lawɪn) This the-boy Aux-is 3MS-paint, this the boy.(will 3MS-paint) This boy is painting, this boy…..(will paint)	0	1	NR
13- A. Ha:ðɪ ɪl-bɪnt gaʕdə t-ɪrkab ħsˤa:n This the-girl Aux-is 3FS-ride horse This girl is riding a horse. B. Yaʕni ha:ða lɪ-ħsˤa:n ɪlli ɪl-bɪnt ………(ra:kbit-ah) So This the-horse that the-girl…… (rode-it) So, this is the horse that the girl (rode)	0	1	NR
14- A. Mɪnu gaʕɪd yɪqraʔ kɪta:b ? ɪhwa gaʕɪd yɪqraʔ kɪta:b Who Aux-is 3MS-read book ? He Aux-is 3MS-read book Who is reading a book? He is reading a book B. Mɪnu gaʕɪd y-akɪlʔ…… (ɪhma gaʕɪd-in yakl-u:n) Who Aux-is 3MS-eat?......(They are-plural eat-plural) Who is eating? ………….(They are eating)	0	1	NR
15- A. ɪl-rayyal y-goul ħag ɪl-bɪnt "ʃufɪ ana asˤalɪħ ɪl-χeimə The-man 3MS-say to the-girl" see I fix the-tent. The man says to the girl, I fix the tent. B. W-ɪnti ………………(tgatˤʕ-in (ɪl-χaʃb) And-you……………..(cut-2F (the-wood)) And you …………….(cut the wood)	0	1	NR
16- A. ɪhni: fi tʃalb sari:ʕ Here there dog fast Here there is a fast dog B. W-hni: fi………………(tʃalb ein batˤiʔin) And-here there ………… (two dogs slow-Dual) And here, there are ………(two slow dogs) 18. W-hni: fi………………(θalɪθ tʃla:b sari:ʔə) And-here there ………… (three dogs slow) And here, there are ………(three slow dogs)	0	1	NR
19. A. ɪhni: fi sayyarə ħamrə Here there car red Here there is a red car B. W-hni: fi……………….(sayyartein ħʊmʊr) And-here there ………… (two cars red) And here, there are ………(two slow dogs) 20. W-hni: fi………(θaləθ sayyara:t/sɪya:yi:r ħamrə/ħʊmʊr) And-here there … (three cars red) And here, there are ………(three red cars)	0	1	NR
	Score		

	0	1	NR
21 A. ɪl-gatˤe t-tˤayɪħ ɪl-sˤuə. ɪl-ɪbu y-goul: 'ɪnta kɪsart ɪl-sˤuə. the-cat -3FSdrops the-picture. the-father 3MS- says: 'You broke the picture Here the cat drops the picture. The father says: 'you broke the picture' B. ɪl-walad y-goul : la:!........(ma: kɪsart ɪl-sˤuə) the-boy 3MS-say: No!......(not broke the-picture) They boy says: No!..........(I did not break the picture)			
22. A. ɪhni, ɪl-bɪnt t-ʔakɪl ɪl-ʕasˤfur lɪ-sˤʁi:r Here, the-girl 3FS-feed the-bird the-little Here, the girl feeds the little bird. B. A. ɪhni, ɪl-bɪnt t-ʔakɪl………(ɪl-batˤtˤə lɪ-kbi:ʁə) Here, the-girl 3FS-feed………..(the-duck(F) the-big (F)) Here, the girl feeds……………..(the big duck)	0	1	NR
23 A. ɪl-zar:fə rqʊbat-ha tˤawi:lə The-giraffe (F) neck-her long The giraffe has a long neck B. w-ɪl sɪlħɪfa:t rqʊbat-ha………(gɪsˤi:rə) And-the turtle-F neck-her……….(short (F)) And the turtlen's neck is………..(short)	0	1	NR
24 A. ɪl-walad ha:ðə y-goul a:nə ʕɪndi flus wa:yɪd The-boy this 3MS-say I have money a lot This boy says: 'I have a lot of money' B. ɪl-walad ha:ðə zaʕla:n, y-goul a:nə….. (ma: ʕɪndi flus) The-boy this upset , 3MS-say I……….(not have money) This boy is upset, he says……………...(I don't have money)	0	1	NR

Appendix P: Expressive Language Test (B)

Manual	Repetition
EL(B)	Allowed once only

Instructions and scoring: Same as in EL(A)

A. possessive Pronouns Practice: ɪl-walad ʕində ku:rə yɪdi:də, yaʕnɪ ɪl-ku:rʕ ha:ðɪ……(mal-ta) The boy has ball (F) new(F), so the-ball(F) this……..(his-(F)) The boy has a new ball, so this ball is …………….(his)			
25. ɪl-bɪnt ʕind-ha kɪtab, yaʕni ɪl-kɪta:b haðə……..(mal-ha) The-girl has-F book(M), so the-book(M) this(M)….(her-F) The girl has a book, so the book is …….(hers)	0	1	NR
26. ʊhma ʕind-hʊm radu: yɪdi:d, yaʕni ɪl-radu: haðə……..(mal- hʊm) They have-Pl radio(M), so the-radio (M) this(M)….(theirs) They have a new radio, so this readio is….. …….(theirs)	0	1	NR
27. ɪl-rayyal ɪʃtarə naðˤðˤa:rə yɪdi:də, yaʕni haðɪ ɪl-naðˤðˤa:rə ..(mal- tə) The-man bought glasses(F) new(F), so this(F) the-glasses….(his(F)) The man bought new glasses, so these glasses are………………….(his)	0	1	NR
28. ɪl-walad y-gu:l, ʃu:f: haðə ɪl-kabbous ma:ltɪ, w haðə ɪl-kabbous…(malik). The-boy 3MS-say, see: this(M) the-hat mine, and this the-hat(M) (yours(M)) The boys says, see: this hat is mine and this hat is …………………….(yours)	0	1	NR
B. Possessive Nouns No practice			
29. Haðə χa:lɪd, ħag mɪnu: ɪl-kalb haðə…………….(ħag χa:la:ɪd) This (M) Khalid, for who the-dog this……………..(for Khalid) This is Khalid, for whom is this dog?...................(for Khalid)	0	1	NR
30. Haði Mni:rə , ħag mɪnu: ɪl-dʒu:tɪ haðə…………….(ħag Mni:rə) This (F) Mnira, for who the-shoe this……………..(for Mnira) This is Mnira, for whom is this shoe?...................(for Mnira)	0	1	NR
C. Regular Plural **Practice:** Haðə tɪlɪfoun waħɪd, w haðeilə arbaʕ…….(tɪlɪfoun-a:t) This(M) telephone(M) one(M), and these four….(telephones) Thi is one telephone, and these are four……(telephones)			
31. Haðɪ tˤa:wlə waħdə, w haðeilə χams…….(tˤa:wl-a:t) This(F) table(F) one(F), and these five……...(tables) This is one table, and these are five…………(tables)	0	1	NR
32. Haðə mʊdarrɪs waħɪd, w haðeilə θalaθ…….(mʊdarrɪs-i:n) This(M) teacher(M) one(M), and these three….(teachers) This is one teacher, and these are three……(teachers)	0	1	NR
33. Haðɪ sa:ʕə waħdə, w haðeilə arbaʕ…….(sa:ʕ-a:t) This(F) clock(F) one(F), and these four……...(clocks) This is one clock, and these are four…………(clocks)	0	1	NR
D. Irregular Plural **Practice:** Haðɪ lɪʕbə waħdə,, w haðeilə θalaθ…(ʔlʕa:b) This(M) toy(F) one(F), and these three….(toys) This is one toys, and these are three……...(toys)			
34. Haðə galˤam waħɪd, w haðeilə θalaθ…(aglˤa:m) This(M) pen (M) one(M), and these three….(pens) This is one pen, and these are three…………(pens)	0	1	NR

35. Haðə kɪrsɪ waħɪd, w haðeilə arbaʕ…(kara:sɪ) This(M) chair (M) one(M), and these four….(chairs) This is one chair, and these are four…..……(chairs)	0	1	NR
36. Haðə seikal waħɪd, w haðeilə θalaθ…(sɪya:kɪl) This(M) bike (M) one(M), and these three….(bikes) This is one bike, and these are three……..……(bikes)	0	1	NR
E. Dual **Practice:** ɪhni: fi: kɪta:b, w-ɪhni: fi:…….(kɪta:b-ein) Here, there book, and here there….(two books) Here, there is one book and here there are…..(two books)			
37. ɪhni: fi: ħsˤa:n , w-ɪhni: fi:…….(ħsˤa:n -ein) Here, there horse , and here there….(two horses) Here, there is one horse and here there are…..(two horses)	0	1	NR
38. ɪhni: fi: ʃɪrtˤɪ , w-ɪhni: fi:…….(ʃɪrtˤɪ-yein) Here, there policeman , and here there….(two policemen) Here, there is one policeman and here there are…..(two policemen)	0	1	NR
39. ɪhni: fi: bei:ðˤə , w-ɪhni: fi:…….(bei:ðˤt-ein) Here, there egg , and here there….(two eggs) Here, there is one egg and here there are…..(two eggs)	0	1	NR
F. Present Gender/Number Markers **Practice:** ɪhni: ɪl-ʕasfu:r y-a:kɪl, w-ɪhni: ɪl-ʕasfu:r………….. (y-tˤi:r) Here, the bird 3MS-eat, and here the-bird……….. (flies) Here, there is one bird, and here the bird………….. (flies)			
40. ɪhni: ɪl-bɪnt t-arkɪðˤ, w-ɪhni: ɪl-ʔawla:d……(y- arkɪðˤ-u:n) Here, the-girl 3FS-run, and here boys………… (M-run-Pl) Here, the girl runs, and here the boys………….. (run)	0	1	NR
41. ɪhni: ɪl-walad y-gu:l, ɪnta tnitˤ foug. Hni y-gu:l, ɪnti…(t-nitˤ-i:n) Here, the-boy 3MS-say you(2SM) jump up.Here 3MS-say, you(2SF)..(jump-2F) Here the boy says you jump up, her he says, you (jump)	0	1	NR
42. ɪhni: ɪl-bint t-gu:l ɪntay t-dɪzi:n lɪ-ʃdʒarə. Hni: ɪl-walad y-gu:l ɪntaw..(t-dɪz-u:n) Here, the-girl 3FS-say you(F) 2F-push the-tree. Here, the-boy 3MS say you (Pl)….(2-push-Pl) Here, the girl says you push the tree. Here, the boys says you (push)	0	1	NR
G. Construct State **Practice:** Haðə ʔkɪl ɪl-walad. Lɪʕbat mɪnu ha:y….(Lɪʕbat ɪl-bɪnt) This food the-boy. Toy who this……...(Toy the-girl) This is the boy’s food. Who’s toy is this (the girl’s toy)			
43. Haðə ba:b ɪs-sayyarə, w-haðˤə…..(ta:yɪr ɪs-sei:kal) This door the-car ,and-this …...(tyre the-bike) This is the car’s door, and this is….……...(the bike’s tyre)	0	1	NR
44. Haðɪ kara:sɪ ɪl-sˤaf, w-haðə ……….(tˤa:wlat ɪl-muʕalɪm)ɪ This chairs the-class, and-this …...(desk the-teacher) This is the class chairs, and this is the ….……...(the teacher’s desk)	0	1	NR
H. Derivation of Nouns **Practice:** ɪl-rayya:l ɪllɪ y-darrɪs f-il saf, n-sami:-h….(mudarrɪs) The-man who 3MS-teach in-the class 1Pl-call-him…(teacher) The man who teaches in the classroom we call……(a teacher)			
45. ɪl-walad ɪllɪ y-ɪlʕab f-ɪl fari:q n-sami:-h….(la:ʕɪb) The boy who 3MS-play in-the team 1Pl-call-him…(player) The boy who plays in a team, we call……….………..(a player)	0	1	NR
46. ɪl-rayya:l ɪllɪ y-itˤbaχ, n-sami:-h….(tabba:χ) The-man who 3MS-cook, 1Pl-call-him…(cook) The man who cook, we call…………(a cook)	0	1	NR

I. Derivation of Adjectives **Practice:** ɪl-mara t-gu:l ɪdeinɪk ʕalei-hʊm wʊsˤaχ. Yaʕnɪ qasˤsd-ha ɪdein ɪl-walad (wasˤχa) The-woman 3FS-sayy hands-your on-them dirt. So means-her hand the-boy (dirty) The woman says your hands have dirt on them, so she means the boy's hands are.... (dirty)			
47. Mariam t-gu:l 'χa:lɪd ʕɪndak <u>ħaðˤ</u> wa:yɪd ħɪlu. Yaʕnɪ qasˤsd-ha Khalid walad…….(maħðˤu:ðˤ) Mariam 3FS-say 'Khalid have-2M luck very nice, so means-her Khalid boy……….(lucky) Mariam says: Khalid you have a very nice luck, so she means Khalid is a …..boy (lucky)	0	1	NR
48. ɪl-mudarris y-gou:l fi: wayɪd <u>ɪzʕa:dʒ</u> f-ɪl saf, yaʕnɪ qasˤd-a haðeilə ɪl-ʔawla:d ……..(muzʕɪdʒi:n) The-teacher 3MS-says there a lot noise in-the classroom, this means-him these the-boys……….(noisy) The teacher says there is a lot of noise in the class, this means the boys are……(noisy)	0	1	NR
49. Hni: ɪl-mudarris-ə t-gou:l intau <u>ti-fawwaqtau</u> ʔala ɪl-sˤaf..fi: yaʕnɪ qasˤd-a haðeilə ɪl-bana:t ……..(mitfawqaat) The-teacher 3FS-says you excelled over the class...., this means-him these The girls……….(excellent) The teachers says you excelled over the class, so this means these Girls are………(excellent)	0	1	NR
J. Past Tense **Practice:** ɪl-tʃalb ga:ʕɪd y-nɪtˤ. Ihnɪ ɪl-tʃalb χallasˤ. Goul-li ʃɪnu: sawwə ɪl-tʃalb………..(natˤ) The-dog Aux-is 3MS-jump. Here the-dog finished…Tell-me what did the-dog (jumped) The dog is running, here it has finished. Tell me what the dog did……(jumped)			
50. l-awla:d y-ɪsmɪʕu-:n musi:qa. Alħi:n χalʊsˤau. Goul-li ʃɪnu: sawwə ɪl-awla:d ………..(smaʕau) The-children 3M-listen-Pl music. Now finished. Tell-mw what did The-boys………….(listened) The children are listening to music. Now, the finished. Tell me what the boys did………..(listened)	0	1	NR
51. ɪl-bɪnt gaʕdə t-ɪktɪb ɪl-wa:dʒɪb ħag-ha. ɪl-bɪnt χallɪsˤat. ʃɪnu: sawwat ɪl-bɪnt awwal……..(ktɪbat) The-girl Aux-is 3FS-write homework for-her. The-girl finished.What did the-girl first……………(wrote) The girl is writing her homework. What did she do first………………………....(wrote)	0	1	NR
52. ɪhni ɪl-walad y-dɪz ɪl-kartoun. ɪhni ɪl-walad χallas. ʃɪnu: sawwa ɪl-walad awwal……..(daz) Here the-boy 3MS-push the box. Here the-boy finished. What did the-boy first……………(pushed) Here the-boy pushes he box. Here the boy is finished. What did the boy do first?..............(pushed)	0	1	NR

K. Clitic Pronouns **Practice:** ɪl-tʃalb y-nɪtˤ ʕala ɪl-su:r. Yaʕni haðə ɪl-su:r ɪlli ɪl-tʃalb…..(natˤʕale-ih) The-dog 3MS-jump on thee-fence Means this the-fence that the dog….(jumped on-it) The dog is jumping on the fence. So, this is the fence that the dog..……(jumped on)				
53.	ɪl-awla:d gaʕdi:n y-ɪlʕib-oun lɪʕbə. Yaʕni haði ɪl- lɪʕbə ɪlli ɪl-awla:d ……..(lʕabau fi-ha) The-children Aux-are 3M-play-pl game. Means this the-game that the-children…….(played in-it) The children are playing a game. So, this is the game that the children ……………... (played)	0	1	NR
54.	ɪl-rayya:l y-ɪgʕɪd taħt ɪl-ʃadʒarə. Yaʕni haði ɪl-ʃadʒarə ɪlli ɪl-rayya:l y-ɪgʕɪd ……..(taħt-ha) The-man 3MS-sit under the-tree. So, this the-tree that the-man 3MS-sits……(under-it) The man sits under the tree. So, this is the tree the man sits…..(under)	0	1	NR
55.	ɪl-awla:d y-a:kl-oun pizza. Yaʕni haði ɪl pizza ɪlli ɪl-awla:d……….. (kalou-ha) The-children 3-eat-Pl pizza. So, this the-pizza that the-children…….. (ate-it) The children 3M-eat-Pl pizza. So, this is the pizza that the children ………..(ate)	0	1	NR
56.	ɪl-bɪnt tarsɪm louħə, yaʕni haði ɪl- louħə ɪlli ɪl-bɪnt……… (rsɪmat-ha) The girl is drawing a picture. So this the-picture that the girl……. (drew-it) The girl is drawing a picture. So, this is the picture that the girl…… (drew)	0	1	NR
57.	ɪl-walad lɪga galˤam w-ʕatˤah ħag sˤaħb-ah w-galah: ɪmsɪk haðə…… (galˤam-ik) The-boy found pen and-gave for friend-his(M) and-said: hold this ……(pen-your 2M) The boy found a pen and gave it to his friend and said: hold it! This is…..(your pen)	0	1	NR
58.	ɪl-walad ʃa:f sayyarat ɪl-ʔawla:d w-gal ʃu:fʊ haðɪ sayyart-ɪ w- haðɪ ……….(sayyart-kʊm) The-boy saw car the-boys and-said see this car-mineand this………..(car-your) The boy saw the boys' car and he said: see! This is my car and this is …….(your car)	0	1	NR
59.	haðə χa:lɪd w-haðɪ ku:rt-ə. Haðɪ Maryam w-haðə...(ħsa:n-ha) This Khalid and-this ball-his. This Maryama and-this…(horse-her) This is Khalid and this is his ball.This is Maryam and this is…(her horse)	0	1	NR
60.	haðɪ Suad w-haðɪ ballouna:t-ha. Haðə Fahad w-haðə....(qɪtˤa:r-ə) This is Suad and-this balloon-her. This Fahad and-this…(train-his) This is Saud and this is her balloon. This is Fahad and this is...(his train)	0	1	NR

L. Demonstrative Pronouns **No Practice**			
61. ɪl-bɪnt t-gou:l m-abi ha:ðɪ ɪl-ʔlʕa:b. Abi…….(haðeilə) The-girl 3FS-say not-want this the-toys. I want… (those) The girl says I don't want these toys, I want those (toys).	0	1	NR
62. ɪl-bɪnt t-gou:l m-abi ha:ðɪ ɪl-ballu:nə……….(haði:k) The-girl 3FS-say not-want this the-balloon. I want… (that) The girl says I don't want this balloon. I want………(that (one)).	0	1	NR
L. Demonstrative Pronouns **No Practice**			
63. Ha:ðɪ ɪl-bɪnt sari:ʕ-ə, bas ha:ðɪ ɪl-bɪnt…..(ʔasraʕ) This the-girl fast-F , but this the-girl…..(faster) This girl is fast, but this girl is (faster). 64. W-ha:ðɪ ɪl-bɪnt ……………(ʔasraʕ waħdə) And-this the-girl ……………(fastest one-F) And this girl is……………….(the fastest).	0	1	NR
65. Ha:ðɪ ɪs-su:rə ħɪlwə, bas ha:ðɪ ɪs-su:rə (ʔħla) This the-picture beautiful, but this picture (more beautiful) This picture is beautiful, bu this picture is…(more beautifu) 66. W-ha:ðɪ ɪs-su:rə……………….(ʔħlə waħdə) And-this the-picture ……………(most beautiful) And this picture is the ……………(most beautiful)	0	1	NR
L. Demonstrative Pronouns **Practice:** Ma-ħħad sa:ʔad ɪl-bɪnt t-abnɪ ɪl-sˤaru:χ…ɪhya banat ɪl-sˤaru:χ….(b-ru:ħ-ha) No-one helped the-girl 3FS-build the rocket…she built the rocket..(by-self-her) No one helped he girl builde the rocket, she built the rocket……….(herself)			
67. Ma-ħħad sa:ʔad haðeil-ə ɪl-waladein. ʊhma lbɪsau….(b-rouħ-hʊm) No-one helped these boys-Dual . They got dressed..(by-self-their) No one helped these two boys. They got dressed…(themselves). 68. ɪl-walad y-tˤalɪʕ f-ɪl manðˤarə. ɪʃ gaʕɪd y-ʃu:f? ɪl-walad y-ʃu:f……. (rouħ-ə) The-boy 3MS-look i-the mirror. What is see? The-boy 3MS-see……. (self-him) The boy looks in the mirror. What does he see? The boy sees….. (himself).	0	1	NR

Appendix Q: Sentences Repetition-A

Sentence Repetition: Not allowed

Instructions:
I will say some sentences and I want you to repeat them exactly the way I say them. For example, if I say: Practice 1: "ʊχu:y y-tˤalɪʕ ɪl-tɪlfɪzyoun" (my brother watches TV), you have to say it exactly like me.
Practice 2: Say: 'wein ra:ħ ɪl-walad' (where did the boy go).
Scoring: 3 for correct answers with no mistakes, 2 when there is one error, 1 for 2-3 error, 0 for more than 3 errors. NR=No response.

Sentence	Score				
1. ʃu:f ha:ðɘ See this-M See this?	3	2	1	0	NR
2. ɪħnz ra:ħ n-sa:fɪr We will 1P-travel We will travel	3	2	1	0	NR
3. ʃɪnu: ha:ðɘ? What this? What is this?	3	2	1	0	NR
4. ʔa:na ʔagdar ʔʃi:l-ɘ I can carry-it I can cary it	3	2	1	0	NR
5. ʃɪnu: ʔakal ɪl-walad? What ate the-boy? What did the boy eat?	3	2	1	0	NR
6. ʔa:nɘ t-ˤɪħt w-ʕawwart ru:ħɪ I 1-fell and-hurt myself I fell and hurt myself	3	2	1	0	NR
7. mɪn wei:n dʒib-t ɪl-dʒɪwa:tɪ ɪl-qadi:mɘ From where get the-shoes the-old Where did you get the old shoes from?	3	2	1	0	NR
8. ru:ħ-ʊ lɪʕb-u barrɘ! go-2Pl play-2Pl outside! Go play outside!	3	2	1	0	NR
9. tabɪ tɪlbas ɪl-dʒakeɪt ɪl-ʔaħmar ha:ðɘ? want-2M 2-wear the-jacket the-red this? Do you want to wear this red jacket?	3	2	1	0	NR
10. ɪl-bana:t ftaħ-aʊ ba:b ɪl-sayyarɘ the-girls opened-3Pl door the-car The girls opened the car door	3	2	1	0	NR
11. Weɪn ra:ħ n-ru:ħ n-ɪlʕab w-n-ʔakɪl? Where will 1Pl-go 1Pl-play and-1Pl-eat? Where will we go to play and eat?	3	2	1	0	NR
12. ʔa:na ma-abi ʔaʃtɪʁɪl b-ru:ħɪ	3	2	1	0	NR

I not-want work by-self-my I don’t want to work by myself					
13. ʔa:naʔ t-ˤalıʕ barrə ʕaʃa:n ʔʃım hawa I 1 M-going outside so 1-smell air I am going outside to smell (fresh) air	3	2	1	0	NR
14. ʔa:na gaʕıd aħıtˤ tˤʊma:tˤ w-χya:r ʕalə ıl-χʊbız ma:lı I Aux-am 1-put tomato and-cucumber on the-bread mine I am putting tomato and bread on my bread	3	2	1	0	NR
15. Ma: ra:ħ t-sˤi:r tˤıwi:l ʔıða ma-kalt Not will 2-become tall If not-eat You won’t become tall if you don’t eat	3	2	1	0	NR
16. ʔabı ʔalbas dʒıwatı ıl-riyaðə haðeı:lə Want 1-wear shoes the-sport these I want to wear these sport shoes	3	2	1	0	NR
17. Weı:n ıl-ʃırtˤı ıllı mısak ıl-ħara:mı? Where the-policeman who caught the-thief? Where is the policeman who caught the thief?	3	2	1	0	NR
18. ıða: χallasˤ-tı kıl ʃay ra:ħ ʔ-aʕti-tʃ ħal:wə If finish-2F every thing will 1-give-you candy If you finish everything, I will give you candy.	3	2	1	0	NR

Appendix R: Sentences Repetition-B

Instructions and Scoring: same as in SR-A

Sentence	Score				
1. Mıta ðˤayyaʕ ıl-wʃalad ıl-lıʕba ıl-sˤafrə When lost the-boy the-toy the-yellow When did the boy lose the yellow toy?	3	2	1	0	NR
2. ıl-ʃaħınə ka:n y-ılħag-ha ıl-ba:sˤ The-truck(F) was 3MS-follow-it the-bus(M) The truck, the bus was following it	3	2	1	0	NR
3. ıl-awla:d ma-gdaraʊ ya-kl-u:n ıl-ħalawıya:t The-boys not-could 3M-eat-Pl the-sweets The boys could not eat the sweets	3	2	1	0	NR
4. ıl-ʔarnab ma-ınħatˤ daχıl ıl-qafasˤ The-rabbit not-was put in the-cage The rabbit was not put inside the cage	3	2	1	0	NR
5. ʔablˤat ıl-sˤaf ıl-ra:bıʕ χallısˤat dars ıl-ʕʊlu:m The female teacher the-class the-fourth finished-F class the-science The fourth class (female) teacher finished the science class	3	2	1	0	NR
6. ıl-walad ıllı rfısa lı-ħˤa:n tˤa: ħ da:χıl ıl-ħʊfra the-boy who hit-him the horse fell in the ditch The boy whom the horse hit fell in the ditch	3	2	1	0	NR
7. ıl-ħa:rıs mısak ıl-ku:rə w-ıl dʒımhu:r sˤaffag l-ə The goalie caught the-ball and-the-fans clapped for-him The goalie caught the ball and the fans clapped for him	3	2	1	0	NR
8. ıl-walad ra:ħ y-ıʃtıri ʕasˤi:r bʊrtʊqa:l maʕa ınna taʔʔaχχar ʕala ıl-madrısə The-boy went 3MS-buying juice ornage eventhough was late for the School The boy went to buy orange juice eventhough he was late for school	3	2	1	0	NR
9. ıl-walad ıʔʃtara kıta:b hag sˤadi:q-a ıllı y-ħıb ıl-qısˤasˤ ıl-bu:lisıyyə The-boy bought book for friend-his who 3MS-like the-stories the-police The boy bought a book for his friend who likes detective stories	3	2	1	0	NR
10. ıl-walad ma kallam ıl-mudarrıs ıllı y-sˤaħıħ ʕala:ma:t ıl-sˤaf ıl-sa:dıs The-boy not talked the-teacher who 3MS-marks grades the-class the sixth The boy did not talk to the teacher who is marking the sixth year grades	3	2	1	0	NR
11. ɪl-kʊtʊb w-ɪl ʔglˤa:m tɪbarraʕ fi:hʊm tˤlla:b ɪl-sˤaf ɪl-sa:dɪs The-books and-the pens donated in-them students the-sixth the-class The books and the pens were donated by the students of the sixth class.	3	2	1	0	NR
11. ʔa:na ma-abi ʔaʃtıʁıl b-ru:ħı1 I not-want work by-self-my I don't want to work by myself	3	2	1	0	NR
12. ɪl-bɪnt ɪlli taskɪn wara bei:t-nʔə maʕa:y b-nafs ɪl-madrisə The-girl who lives behind house-our with me in-same the-school The girl who lives behind our house is with me in the same school	3	2	1	0	NR

13. lau ɪl-χaddamə sawwat kei:k w-baskout ka:n ɪn-ʔaklau mɪn zɪma:n If the-maid made cake and-biscuits was Passive-eaten from longtime If the maid had made cake and biscuit they would have eaten long time ago.	3	2	1	0	NR
14. ɪl-tˤɪllab ktɪb-au rɪsa: ə ħag sˤa:ħibh-ʊm ɪlli sa:far f-ɪl fasˤl ɪl-ʔawwal The-students wrote-Pl letter to friend-them who travelled in-the term the-first The students wrote a letter to their friend who travelled in the first term	3	2	1	0	NR
15. ɪl-walad ɪlli y-su:g ɪl-sayyarə ɪl-zargə ħa:tˤ naðˤðˤa:rə soudə The-boy who drives the-car the-blue wearing glasses black The who drives the blue car is wearing black glasses	3	2	1	0	NR
16. Weɪ:n ıl-ʃırtˤı ıllı mısak ıl-ħara:mı? Where the-policeman who caught the-thief? Where is the policeman who caught the thief?	3	2	1	0	NR
17. ıl-ʔawla:d ʃallau ɪl-ʔalʕa:b w-ratɪbou-ha w-ħatˤtˤou-ha f-ɪl maχzan The-boys picked the-toys w-arranged-them and put-them in-the store The children picked the toys, arranged them and put them in the store	3	2	1	0	NR
18. gabʊlˤ ma: y-ɪtrɪk ɪl-ʔawlad ɪl-sˤaf ɪn-tˤalˤab mɪnhʊm tasli:m ɪl-wa:dʒɪb Before that 3MS-leave the boys the-class passive-asked from-them handing the-homework Before the boys left the classroom, they were asked to hand the homework.	3	2	1	0	NR
19. ɪl-walad ɪlli ma: ħaðˤar ɪl-tamri:n mu: masmu:ħ-lah y-ɪlʕab maʕa ɪl-fari:q mʊddat ɪsbu:ʕ The boy who did not attend the training not allowed 3MS-play with the Team for one week. The boy who did not the training is not allowed to play with the team	3	2	1	0	NR
20. ɪl-mʊdarrɪs ħatˤ tɪsaʕ kʊtub ʕɪlmɪyya dʒɪdi:də maħdʒu:zə ħagna f-ɪl maktaba The-teacher put nine books scientific new reserved for-us In-the library The teacher put 9 new scientific books reserved for us in the library	3	2	1	0	NR
21. mʊdarrɪs ɪl-rɪya:ðˤɪyya:t dʒa:b ɪl-mɪsˤa:tˤɪr w-ʕad-ha w-raqqam-ha w-wazaʕ-ha ʕala ɪl-sˤaf Teacher the-maths brought rulers and counted them and marked them and handed them out to the class The maths teacher brought rulers, counted them, marked them, and handed them to the class	3	2	1	0	NR
22. ɪ ɪl-mʊdarrɪs ɪlli f-ɪl sˤaf ɪl-θa:ni qa:l inah raħ yɪsqi ɪl-zarʕ fɪ ɪdʒa:zat ɪl-sei:f	3	2	1	0	NR

Appendix S: Arabic Picture Vocabulary Test (APVT)

Record Form

Name:…………………………….... Gender: M / F

School: ……………………………………Grade:…………………..

Examiner:……………………........

	Year	Month	Day
Test Date			
Birth Date			
Chronological Age			

Instructions:

Establish rapport with the child in a short conversation. Explain how this test goes by saying: “We will have a look at this picture book and I want you to point to the picture I am talking about”. Start with practice 1 and 2 by saying: “I want you to point to. ‘dʒu:tɪ’ (shoe)”. Encourage the child if h/she does not point and correct him if necessary. Praise him for trying regardless of accuracy. For older children, you can accept answering in number of item instead of pointing.

Practice 1: Point to....... ‘dʒu:tɪ’ (shoe)”.

Practice 2: Point to....... ‘smɪtʃə’ (fish)”.

Scoring:

Put ☑ when the child answers correctly and if the child is incorrect, put a (/) on the item number and write the number of the picture the child chose. To calculate raw score subtract the number of errors from the number of last item in the ceiling group.

Basal: Always start at item 1. Ceiling: you can stop if there are 8 incorrect items in one group. If you start a group, you need to complete it even if child reaches ceiling.

Ceiling Item	
Minus Errors	-
Raw Score	=

	Group 1		Group 2		Group 3		Group 4
.1	yɪʃrab drink (v.)	.13	maχammə broom	.25	mʊmarɪðˤə nurse	.37	ygɪ:s measure (v.)
.2	bei:bI baby	.14	rqʊbə neck	.26	yʃɪg tear(v.)	.38	ħaʃi:ʃ grass
.3	bgarə Cow	.15	ʔɪsˤbaʕ finger	.27	tˤawu:s peacock	.39	gʊfʊl lock
.4	ʕeiːn Eye	.16	wardə flower	.28	tˤa:bɪʕ (post)stamp	.40	χaʃabɪ wooden
.5	yɪrkiðˤ Run	.17	yarfɪs kick	.29	ynɪtˤ jump	.41	sˤaffa:rə whistle
.6	bei:t House	.18	ʕɪʃ nest	.30	batˤi:χ warda	.42	mɪnga:r beak
.7	tɪqra read (v.)	.19	da:ʔɪrə le cir	.31	maʕdʒu:n playdough	.43	kaʕab heel
.8	tˤabʊl drum	.20	ħayyə snake	.32	yħa:sɪb pay(v.)	.44	dʒouz ɪl-hɪnd coconut
.9	sei:kel bike	.21	ħmaːr donkey	.33	ðei:l tail	.45	maki:nət χɪya:tˤə sewing machine
.10	tʃu:f see (v.)	.22	sˤabu:n soap	.34	mɪntɪfɪχ blown(adj.)	.46	qalʕə castle
.11	ba:sˤ bus	.23	ʃwayyə little	.35	bʊrʁɪ screw	.47	yfatɪʃ search(v.)
.12	gatʃˤwə cat	.24	tˤabi:b doctor	.36	yqa:bɪl meet(v.)	.48	ʃamʕə candle

	Group 5		Group 6		Group 7		Group 8
.49	yɪsˤʕad climb(v.)	.61	sˤayya:d hunter	.73	taħfʊr dig	.85	qɪmmə summit
.50	qaðɪ judge	.62	fɪtˤɪr mushroom	.74	Va:n van	.86	tʊqabbɪl interview
.51	χaya:lɪ fictional	.63	tɪfħasˤ examine	.75	yħaðɪr lecture(v.)	.87	ʕʊzlə isolation
.52	tɪdʒu:rɪ safe(n.)	.64	ykaʃɪr grin	.76	ħɪntʃ chin	.88	qamħ wheat
.53	kanʁar kangaroo	.65	bayðˤa:wɪ oval	.77	mʊtafadʒɪʔə surprised	.89	tʊʁa:dɪr leave(v.)
.54	ʕʊmlə coin	.66	taʁtˤɪs dive(v.)	.78	mʊhaððab polite	.90	dʒɪðɪʕ trunk
.55	bʊrdʒ tower	.67	ba:zɪllə peas	.79	tasˤa:dʊm clash(n.)	.91	tahmɪs whispter
.56	fauðˤawɪ messy(adj.)	.68	ru:mansI romantic	.80	naħi:f thin	.92	da:ʔɪrɪ ring road
.57	dʒarsu:n waiter	.69	ħʊqnə syringe	.81	ðˤabʕ hynea	.93	gi:ta:r guitar
.58	Mi:nə port	.70	yʊʃawwɪh distort	.82	mʊtafa:ʔɪl optimistic	.94	mʊftarɪs predator
.59	Dainasu:r dinasaur	.71	masnaʕ factory	.83	taʕa:wʊn cooperation	.95	faχu:r pround
	ħalazoun snail	.72	tɪlɪskoub telescope	.84	hʊdhʊd hadhoud	.96	mʊnhaka exhausted

	Group 9		Group 10		Group 11		
.97	sra:dʒ lantern	.109	taħtası	.121	aħfa:d grandchildren		
.98	madʒara galaxy	.110	ħımðˤıyya:t citrus	.122	waʕl ibex		
.99	La:ma Lama	.111	maʕbad temple	.123	Zara:dıyyə plier		
.100	ʃa:ʔık thorny	.112	rıdaʔ gown	.124	ʕarfadʒ Arfaj		
.101	qamʕ funnel	.113	Tawbi:χ scolding	.125	Leiθ lion		
.102	taqtˤıf pluck	.114	mʊtasawwıl begger	.126	zumbʊrʊk spring		
.103	dʊwa:t ink pen	.115	mʊtadʒahım grumpy	.127	Sa:ksfoun saxphone		
.104	Qarnabi:tˤ cauliflower	.116	hırrə kitten	.128	Faras ıl-baħar Sea horse		
.105	rısʁ wrist	.117	mʊtaha:lik decaying	.129	yʊdallıs cheat		
.106	yaʕwı howl	.118	yaʕdu jog	.130	mıtru: al-anfa:q underground		
.107	mʊstanqaʕ swamp	.119	yansıdʒ weave	.131	tʃılu: chillo		
.108	mʊtawa:zı parallel	.120	yaltahım devour	.132	χanzı:r barrı boar		