Investigating Speech Output Skills in 3-5 Year Old Arabic-Speaking Children: A Psycholinguistic Approach

Noaf F. Al-kheraiji

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Sheffield
Faculty of Medicine, Dentistry and Health
Department of Human Communication Sciences

March 2017

For My Mother & Father,

Zeinab & Fahad

ACKNOWLEDGMENT

I would like to express my sincere gratitude to the many people who have helped me throughout the years and who made this work possible. I would like to thank Professor Joy Stackhouse for her guidance, generous time, support, and feedback on this project. I thank Dr. Ben Rutter for his enormous support and his feedback throughout this project. I am also appreciative for Dr. Blanca Schaefer for the comments she provided on my work.

I would also like to thank the schools I visited for data collection. I would like to thank the Ministry of Education and the head teachers of the schools. I would also like to thank the class teachers of the schools who have warmly welcomed me and aided in the smooth process of recruiting children. I would also like to thank the children who were a delight to test and were cooperative and patient during testing despite their young age and the long material. I also appreciate their feedback and unexpected comments on the tests, which provided insightful observations into the child’s thoughts and experiences with the tests. I also thank everyone in the department of Human Communication Sciences; the lovely environment and support for one another is remarkable, including all the postgraduate students and especially everyone in room 207.

Furthermore, I will forever be grateful to the role-models and inspirations in my life, my grandmothers, who both passed away during my study at the University of Sheffield, were two of the most exceptional women I have ever known, with their high morals, love for knowledge and strength, they will forever be remembered, may they rest in peace. I thank my grandfather Muthanna, my aunts, uncles, cousins and friends for cheering for me all the way.

Most importantly, I thank my beloved family, the air I breathe, I am forever grateful to my parents Zeinab and Fahad, who this work is dedicated to, and my siblings, Norah, Aziz, Mohammed and Asma. The love and support they have given me is beyond what words can describe. Thank you from the bottom of my heart.

ABSTRACT

This exploratory study aims to describe the profile of speech processing performance across different speech output tasks in typically developing Arabic-speaking children and chart developmental change by looking at cross-sectional data across different age groups. The speech processing demands required to complete the tasks were interpreted within the psycholinguistic speech-processing model of Stackhouse and Wells (1997).

A total of 129 typically developing Saudi Arabic-speaking children were divided into three age groups (3-year olds: 29 children; 4-year olds: 50 children; 5-year olds: 50 children). Children were tested on three speech repetition tasks comprising real words, non-words and syllable sequences. The stimuli of the three tasks were phonetically matched and stimuli items of each task increased in the number of syllables (length). For each task, the children were required to: **a)** repeat each item once i.e. immediate single-word repetition; with responses being scored for repetition accuracy; and **b)** repeat each item multiple times consecutively and at speed i.e. speech motor performance; with behavioural measures of accuracy and consistency used to score productions.

Single repetitions revealed different performance profiles in different age groups; mainly, there were no differences between real word and non-word repetitions, and developmental progress was not evident. With multiple rapid productions, the processing demands of the tasks and age significantly affected children's performance. Generally, the effects of increasing item length was not straightforward; as repetition of short real words and non-words was not necessary better than that of longer items.

The results of this study show that the Arabic-speaking children's speech processing profiles, developmental progression on the speech output tasks and effects of length were not entirely in line with cross-linguistic evidence, on both single repetitions and on multiple rapid productions. These results appear to reflect the unique phonetic and phonological properties of the Arabic language, which could have affected children's performance on the tasks. Therefore, this study has important methodological, theoretical and clinical implications, which will be discussed.

Table of Contents

List of Tables

List of Figures

GLOSSARY & CONVENTIONS

Glossary:

Speech processing: refers to all the skills required to understand and produce speech, including how speech is presented and processed in the brain, and including peripheral skills of hearing and articulation

Lexical knowledge, mental lexicon or mental vocabulary: a store of phonological-semantic complexities of a word.

Underlying representations/lexical representations: is a term used in the present thesis to describe the necessary information required to identify and produce words. It includes the abstract phonological-semantic form of a word that is stored in the mental lexicon, and includes its motoric information.

Phonological processes: used in this thesis in psycholinguistic terms to refer to the underlying cognitive-linguistic operations that are core abilities that give rise to speech.

Phonological error patterns: is used to describe the regularities or "rules" in children's phonology, that could be idiosyncratic to an individual child or that are shared by children of similar age and exposed to a specific language.

Speech motor planning: This is sometimes referred to in the literature as motor programming, however, motor planning in the present study entails the pre-execution stage of assembling and preparing speech movement sequences.

Conventions:

RWR: real word repetition.

NWR: non-word repetition.

SSR: syllable sequence repetition.

/ / : realisation of a target stimuli/token in phonetic transcript.

[ ] : actual spoken realised response.

→ : is realised as / / → [ ].

"….": English translation of a word, for example /kiˈtæ b/ "book".

CV represents structure, where C= consonant and V= vowel.

C: represents gemination or consonant elongation, which is a characteristic of the Arabic language. In Arabic, it is known as Shadda, for example, CVC:V in the Arabic verb /darras/ "he taught".

ˈ : a symbol used to detect syllable stress, for example, /kiˈtæ b/.

INTRODUCTION

It is well recognised that the foundation from which speech and language therapists will plan an appropriate management program for a child with speech difficulty is a comprehensive evaluation of the child's speech difficulty. To implement such a thorough evaluation, various perspectives to describing a speech difficulty should be adopted; combining medical/aetiological, linguistic and psycholinguistic perspectives. This is especially important since there is consensus among specialists that children with speech difficulty form an extremely heterogeneous group.

Generally, clinical decision making on the assessment (and intervention) of a speech difficulty of an Arabic-speaking child resolves around medical, and mainly on descriptive-linguistic perspectives. The clinicians are guided in their decisions by their current state of knowledge on speech difficulties and the available research, that predominantly adopts an articulatory/phonological paradigm (Abou-Elsaad, Baz, & El-Banna, 2009; Amayreh, 2003; Ayyad, Bernhardt, & Stemberger, 2016; Khattab, 2007). On the other hand, a psycholinguistic perspective that investigate the underlying speech processing mechanisms involved in the production of speech in Arabic-speaking children have received extremely little attention in the literature, in both typical and atypical development, and have therefore been limited in informing clinical practice in this population. This approach has however, been influential in conceptualising developmental processes and impairments in many languages, particularly English (Baker, Croot, McLeod, & Paul, 2001; Chiat & Roy, 2007; Dispaldro, Deevy, Altoé, Benelli, & Leonard, 2011; Gathercole, Willis, Baddeley, & Emslie, 1994; Pascoe et al., 2016; RCSLT, 2011; Waring & Knight, 2013).

One of the most influential psycholinguistic models that provide a framework for understanding and explaining the surface linguistic-descriptive information about impaired speech systems is the psycholinguistic speech processing model of Stackhouse and Wells (1997). It captures the key components of psychological processes involved in the perception, storage, planning and production of speech and attempts to identify the level/s in which speech processing is disturbed. The psycholinguistic speech processing model of Stackhouse and Wells is not limited to a clinical population, but can also be used equally well to understanding typical speech processing development. Essentially, Stackhouse and Wells (1997) stress the need to understand typical developmental speech processes before it could be applied in identifying speech processing difficulties in children. Speech impairment in children is defined by reference to normal development, and an understanding of how a particular child differs from one who is developing typically.

The present thesis will attempt to model Arabic-speaking children's speech output development from the psycholinguistic (theoretical) speech processing framework of Stackhouse and Wells. It is hoped that understanding typically developing Arabic speech output processing will inform and help in the detection of underlying processing deficits in atypical speech development. This thesis has an important prospective orientation, with its participating group of Arabic-speaking children, it is a call for research to move in a direction that has remained furtive, since very little is known about the underlying speech processing skills in Arabic-speaking children, to our knowledge, no published study had used a model-based psycholinguistic approach to understanding underlying speech processes in Arabic-speaking children.

The psycholinguistic framework of Stackhouse and Wells hypothesis different levels of speech processing that could be assessed and targeted using a series of tasks. One of the key tenets of assessment within the psycholinguistic model of Stackhouse and Wells is that using a task in isolation is not informative; rather patterns of association and dissociation across more than one task is more informative. The present study intends to target the output speech processing levels of Stackhouse and Wells' model. The present thesis will use speech output tasks and stimuli to tap hypothesizes levels of speech output processing. This study will use repetition tasks at a single-word level, furthermore, it includes repetition at rapid multiple-level i.e. kinematic aspect, therefore, adding a further dimension to this study, as the speech processing demands of the different tasks will shed light on the different levels of processing within the model.

To this end, there is a need to identify the means by which typically developing children progress through the psycholinguistic framework, and the need to refer to normal control data when assessing children with speech disorders. It is hoped that this exploratory study will provide the baseline information for future advances in research on Arabic-speaking children and developing clinical tools.

The study of output speech processing skills in Arabic-speaking children is worth investigating for the following reasons:

1. The Arabic language is the most widely spoken of the Semitic languages, and is spoken by approximately 240 million people. The demographics in Saudi Arabia show that children between the ages 0-5 comprise % 8 of the population (Arabia, 2016). Therefore, investigating speech processing skills in this population allows for the prospect advances in understanding the nature of speech processing in young Arabic-speaking children and would have potential influences on speech and language therapy research and practice.
2. Preschool children aged 3 to 5 years, are the ages targeted in the present thesis, as they are considered the critical ages for language and speech development. Furthermore, the majority of children with speech difficulty are usually referred to speech and language therapy between the ages of 3 and 4 years (Broomfield & Dodd, 2004a).
3. There have been considerable advances in the study of psycholinguistic speech processing skills in typically and atypically developing English-speaking children and many other languages. Psycholinguistic perspectives have been influential in conceptualising developmental processes and impairments in many language, particularly English, while this perspective in the study of Arabic-speaking children's speech has largely been sparse to date.
4. With regard to the theoretical framework of the current study, the psycholinguistic speech processing framework of Stackhouse and Wells (1997) has been published and applied to the investigation of typical and atypical speech in children in languages other than English; such as German (Fox, 2004; Schaefer et al., 2009), French (Wells, Stackhouse, & Vance, 1999) and Portuguese (Vance, 1996). Therefore, applying this framework to the study of Arabic-speaking children's speech output skills will add to the body of cross-linguistic studies and contribute immensely to the literature.
5. If findings from cross-linguistic studies were replicated in Arabic-speaking children, where speech processing demands influences children's performance and developmental changes emerge, then universal trends and the theoretical concept of speech processing would be supported. Nevertheless, it is not suitable to assume a universal order of behaviour and that similar speech processing behaviour and developmental trends will emerge, as Arabic has its own phonetic/phonological properties and, in turn, its own psycholinguistics.

6. The assumed universal order of sound acquisition and sequence of syllable structure (Jakobson, 1968) has been scrutinised over the years by many cross-linguistic studies on phonological development and the emergence of new phonological theories, such as usage–based phonology (Bybee, 2000) and whole-word approaches (Vihman & Keren-Portnoy, 2013). Evidence from the literature on the acquisition of Arabic phonology (e.g., Abdoh, 2011; Dyson & Amayreh, 2000; Khattab & Al-Tamimi, 2013), showed that Arabic-speaking children do not follow a straight forward (simple to complex structure) developmental path; rather children's early words in terms of sound, word length and structure are influenced by the adult phonology. With this evidence in mind, the language properties of the Arabic language could have profound effects on the way children perform on speech processing tasks, and this will be determined by interpreting their results in light of the Arabic stimuli designed for the present study.
7. The use of non-word repetition task could potentially identify processing skills which are free from cultural or linguistic constraints.
8. The literature provides evidence that early lexical representation in young children are represented as whole unites "holistically" and gradually incorporate phonetic detail as the lexicon grows, becoming increasingly segmented (Fergison & Farwell, 1975; Metsala & Walley's, 1998; for a review see Vihman & Keren-Portnoy, 2013). This concept has been supported by studies using tasks of non-word repetition and measures of consistency of repeated productions of words (Vihman & Keren-Portnoy, 2013). Indeed, Arabic-speaking children's performance on a non-word repetition task and behavioural measures of consistency would serve as a window into the developing phonological representations of the Arabic-speaking child.
9. Clinicians resort to varying strategies to carry out their assessment for Arabic-speaking children; they either develop an informal measure or more likely adapt existing assessment tools with norms from English speaking children, i.e. not intended for Arabic speaking children. Providing preliminary norms from Arabic-speaking children on speech output tests will provide a base for comparing children with speech difficulty.

Overview and Organisation of the Thesis

The literature review is divided into three main chapters. Chapter 1 starts with a brief introductory review of literature on the prevalence of speech difficulties with a spotlight on Saudi Arabic-speaking children and the way their difficulties are commonly conceptualised in clinical practice in Saudi Arabia. This leads to the review of the psycholinguistic speech processing approach of Stackhouse and Wells (1997). Chapter 2 of the literature review focuses on the present studies speech output tasks, where the review of the literature is divided into single repetition tasks and speech motor control i.e. rapid consecutive productions tasks. Each of the divisions will include a review of the literature on children's overall performance on tasks, developmental progression and the effects of different length on performance.

Chapter 3 will provide an important overview of the Arabic language, as this thesis is cantered around the Arabic-speaking children and the Arabic tasks designed to investigate their speech processing skills.

Chapter 4 provides an introductory summary of the reviewed literature and outlines the purpose of the present study, its aims and questions. This is followed by the design of the task stimuli and the pilot study.

Chapter 5 provides the methods and procedure for this thesis's main cross-sectional study. Chapter 6 presents the results on single word -level repetitions and includes its discussion. Chapter 7 presents the results on the multiple consecutive repetition level -speech motor performance- and discusses the results. Chapter 8 puts together the general discussion of the results from chapter 6 and 7, and provides study limitations, future direction and conclusion.

CHAPTER 1 LITERATURE REVIEW I: A SPEECH PROCESSING APPROACH TO INVESTIGATING CHILDREN'S SPEECH SKILLS

A psycholinguistic speech processing approach is one theoretical approach that has been applied to children's speech development, and was found useful when exploring problems underlying impaired speech development. To make use of this approach in research and clinical practice, a specific psycholinguistic model should be selected and the proposed information-processing pathways of the model assessed using appropriate tasks. Before going into further detail on the current project's selected model of speech processing, this chapter will first provide the necessary background information on how speech developmental and speech difficulties has been conceptualised in research and clinical practice.

Therefore, this chapter will:

- Provide an overview of the major perspectives in research and clinical practice that have been influential in conceptualising children's speech and informing clinical practice. This overview will reflect on research developments and on current clinical practice in Saudi Arabia; therefore setting the background information behind the rational for adopting the psycholinguistic speech processing approach in the current study (Section 1.1).
- Introduce the psycholinguistic speech processing perspective (Section 1.2)
- Provide a brief overview of speech difficulties in children; this includes its epidemiology with particular attention to the current study's target population i.e. Saudi Arabic children (Section 1.3).
- Review in depth the psycholinguistic speech processing model of Stackhouse and Wells (1997) (Section 1.4).
- Review the principles of assessment within Stackhouse and Wells' psycholinguistic model, where different tasks are hypothesised to trigger different processing levels (Section 1.5).

1.1 Towards Explaining Speech skills in Arabic-Speaking Children

A speech and language therapist is faced with making the important decision of whether a child being assessed has a speech difficulty. The decision is determined with reference to typical development (i.e., whether a child's speech is appropriate for their age); thus the criteria of normality is essential to the clinical assessment of a child's speech. Clinical researchers have endeavoured to provide insights into children's developing speech system and bring them to the attention of practicing clinicians to keep them abreast of theoretical developments and inform clinical practice.

One of the revolutionary and fundamental approaches to understanding children's speech development that has been applied to clinical practice is the descriptive-linguistic approach (Tyler, 2010; Waring & Knight, 2013). The linguistic perspective allows detailed description of a child's overt linguistic behaviour at different levels of analyses (e.g., phonetic, phonological). It is a developmental approach that describes and identifies how a child's speech errors differ from a child who is the same age and developing typically.

Generally, the descriptive-linguistic based approach to understanding speech development and impairment has dominated the Arabic literature. The literature on Arabic-speaking children's speech production has solidly focused on two domains: the phonetic inventories and phonological patterns/processes observed in typically developing children who speak Jordanian, Egyptian and Kuwaiti Arabic (Amayreh & Dyson, 1998; Ammar & Morsi, 2006; Ayyad et al., 2016; Ayyad & Bernhardt, 2009; Dyson & Amayreh, 2000; Saleh, Shoeib, Hegazi & Pakinam, 2007; Shahin, 1995, 2003). However, some studies have further investigated syllable structure (Abdoh, 2011; Ammar, 2002; Salem; 2000) and gemination (Khattab & AlTamimi, 2013). Furthermore, other studies have explored the phonetic and phonological features and errors in children with speech difficulties associated with structural abnormalities such as cleft palate (Abou-Elsaad, Baz, Afsah, & Mansy, 2015; Alawaji, 2014; Shahin, 2006).

Although our knowledge of typical development in Arabic-speaking children at this stage is incomplete, normative studies provide data that are usually used as a baseline for comparison purposes with a child with speech difficulty, and these data are useful for initial assessment and for monitoring progress. The Arabic literature has focused on describing a child's overt speech behaviour, and this has subsequently affected how clinicians conceptualize speech difficulties in Arabic-speaking children; where the investigation of Arabic-speaking children's speech difficulties to date has been firmly grounded on linguistic-based approaches. It is often the case that speech and language therapists in Saudi Arabia assess, analyse and manage speech difficulties in an eclectic way, drawing predominantly on linguistic-based approaches. For example, the Jeddah Institute for Speech and Hearing JISH ("Speech disorders", 2011) recognise children with speech difficulty as having an articulatory (phonetic) difficulty or speech sound errors. Many speech and language therapist use articulation tests designed at their clinics (such as pictures, objects or repeating specific words) to test children's phonetic consonantal inventories at different word position (word initial, medial and final). They usually draw on normative data from Arabic phonetic acquisition studies such as Amayreh & Dyson (1998) that lists the developmental phonetic inventories of children. Generally, efforts to design speech tests with baseline data for Arabic-speaking children have largely focused on phonetic/phonological levels of investigation—for example, the 'Mansoura Arabic Articulation Test' (Abou-Elsaad et al., 2009) and the Qatari project for "baseline data for Arabic acquisition with clinical applications" (Al-Buainain, Shahin, Morsi, Khattab & Al-Tamimi, 2010). Although the descriptive-linguistic approach provides detailed descriptions of children's phonological systems, it does not explain why the system takes the typically developing or impaired forms.

Alternatively, the aetiological or medical perspective, which has a long history in speech and language therapy, considers the integrity of the neurological and anatomical systems in the developing child and aims to explain the underlying causes of speech difficulty when there is an identifiable cause. The perspective starts from a position of pathology rather than normality and has the general assumption that speech difficulty is due to an underlying clinical condition or medical condition. It is important to identify the origins of speech difficulty in a child, if possible (e.g., cleft lip/palate, dysarthria, hearing loss, neurological causes, intellectual or a genetic basis, such as Down syndrome). An aetiological perspective has long been in favour for many speech and language therapists in Saudi Arabia because of the demands of their working environment. In Saudi Arabia, speech and language therapists work in a variety of settings; predominantly, they work in clinical-based settings, such as hospitals or in special needs institutes/associations[1]. Furthermore, clinic-based facilities are the primary placements for students undergoing

[1] A few examples include the Disabled Children's Association, Down Syndrome Charitable Association (DSCA) and Sultan Bin Abdulaziz Humanitarian City.

speech and language therapy training. Within this context, speech and language therapists usually work closely with medical and other professional personnel and are faced with the practical demands of the situation, whereby a diagnostic aetiological label is useful to communicate the need of a child with speech difficulty (e.g., for example, cleft palate, dysarthria or a hearing loss etc.).

However, this perspective is limited by many issues; several authors have questioned the aetiological approach in general as an applicable system in clinical practice (Broomfield & Dodd, 2004b; Fox, Dodd, & Howard, 2002; Stackhouse & Wells, 1997; Waring & Knight, 2013). First, it is argued that one of the main limitations of the approach is the difficulty faced when attempting to differentially diagnose children's speech difficulty with an unknown causal factor. A medical label is not always possible as children with speech difficulty are a heterogeneous population whose difficulties are mostly of unknown origin. Further, some children may present with more than one difficulty, or they may present with one difficulty but do not fall into a specific subgroup (Waring & Knight, 2013). Second, the approach has limited clinical utility when targeting intervention, as it does not describe or predict with any precision the severity or nature of a child's speech difficulty and fails to account for developmental change (Stackhouse & Wells, 1997; Waring and Knight, 2013). It is recognised that even when there is a medical label attached to the child's speech difficulty, such an understanding of the underlying cause contributes to clinical management (e.g., repair of cleft/lip palate velopharyngeal dysfunction), nevertheless, this will have little relevance if it does not alter the therapeutic management of speech and language therapist.

The medical perspective and the descriptive-linguistic perspective complement one another, where the first considers the integrity of the neurological and anatomical systems while the latter describes the language system. The literature on Arabic-speaking children's speech development and difficulties has been limited to these approaches. Although it is unclear to what extent clinicians are influenced by research and implement its findings, these two perspectives have generally dominated clinical management of children's speech difficulties in Saudi Arabia. It should be noted that such knowledge of clinical practice in Saudi Arabia is largely anecdotal, as there are no clinical surveys that report how speech and language therapists manage children with impaired speech; it is fair to say that the dominance of the approaches in the way speech and language therapists in Saudi Arabia conceptualize children's speech difficulties has been influenced by the direction of the Arabic literature, educational, training and working environments in Saudi Arabia.

Nevertheless, describing a developing speech system and a speech disorder in terms of causal factors or linguistic analysis is unsatisfactory[2], as these perspectives do not *explain* the underlying mental operations of speech production and difficulty (Baker et al., 2001; Dodd, 2005a; Stackhouse & Wells, 1997; Waring & Knight, 2013). In contrast, the psycholinguistic perspective embraces the role of explaining speech development and difficulty (i.e. why children make developmental errors and where speech impairment is located).

A psycholinguistic speech processing approach to understanding typical development and impairment has received little attention in the Arabic literature and its application to clinical practice has been close to nil. To the author's knowledge, no published model-based psycholinguistic approach have investigated the underlying mental operations of speech in Arabic-speaking children. It is therefore the intention of the present thesis to investigate typical processing performance in Saudi Arabic-speaking children, in the hopes that understanding normal developmental processes will inform clinical practice.

1.2 The Psycholinguistic Speech Processing Perspective

Psycholinguistics is a branch of linguistics and psychology that aims to explain human linguistic behaviour at a cognitive or psychological level. The term psycholinguistics was first introduced by American psychologist Jacob Robert Kantor in 1936 and later emerged as an academic discipline after an influential seminar at Cornell University in 1951 (Levelt, 2013). A psycholinguistic processing perspective accounts for the underlying cognitive processes required for speech and maps a range of interactive deficits that underlie speech difficulty. Therefore, they go beyond description to *explanation* of speech development and difficulties (Baker et al., 2001; Hewlett, 1990; Stackhouse & Wells, 1997).

The psycholinguistic approach to speech and language development attempts to explain the way typically developing children process speech and language, and therefore formulates hypotheses about the speech process or components that could be impaired.

[2] See Waring and Knight (2013) who provide a comprehensive review and critique of the different perspectives to describing speech difficulties.

To generate a hypothesis about the level/s of breakdown giving rise to speech impairment, psycholinguistics proposes theoretical models that highlight key components of speech processing and the relationship between those components. There are three key components in a psycholinguistic model of speech processing: receptive processing of speech (input), stored or underlying representations and processes involved in the production of speech (output) (Baker et al., 2001).

A number of psycholinguistic models use box-and-arrow diagrams in which the hypothesised levels of speech processing are represented by a “box” and the relationship between the levels are represented by “arrows”. The models differ significantly in their complexity; some have one or two boxes between input and output processes (e.g., Smith, 1973) and others have multiple boxes that outline complex relationships between the different levels of speech processing[3] (e.g., Hewlett, 1990; Hewlett, Gibson & Cohen-McKenzi, 1998; Stackhouse & Wells, 1997). Regardless of the models’ complexity, their application to children is the same, where hypotheses of different levels of processing are systemically tested to locate the level/s of breakdown.

The psycholinguistic approach to investigating speech has been described as a bridge between the aetiological and linguistic-descriptive approaches (Kamhi, 1989; Waring & Knight, 2013). It is viewed as a good approach that ‘attempts to make good some of the shortcomings of other approaches by viewing the children’s speech problems as being derived from a breakdown at one or more levels of input, stored linguistic knowledge, or output’ (Stackhouse & Wills, 1997, p. 7).

Baker et al. (2001) acknowledged that theoretical models of speech processing are deemed to be important in clinical practice, stating that a psycholinguistic approach ‘can have important effects on clinical practice—not only in influencing assessment and intervention procedures, but in reshaping our thinking about the nature of speech impairment’ (p. 686).

Researchers have demonstrated the usefulness of psycholinguistic model-based evaluations in understanding the underlying problems of impaired speech development, regardless of the diagnostic label. Studies have found that groups of children with speech

[3] It is beyond the scope of this thesis to discuss different backgrounds of psycholinguistic models. For a detailed discussion and critique on psycholinguist models, refer to the tutorial by Baker et al. (2001).

difficulties with the same diagnostic label had similar patterns of performance and that identifying these patterns may have diagnostic value (e.g. Bradford and Dodd, 1994; Thoonen, Maassen, Gabreel, Schreuder & Swart, 1997; Williams and Chiat, 1993). Studies have also found different profiles of performance on speech tasks with children given the same diagnostic label (Bryan & Howard, 1992; Pascoe, Stackhouse & Wells, 2005; Stackhouse & Snowing, 1992).

One of the recent models of speech processing that has been widely used as a framework to profile children's performance and investigate both typical and atypical speech development is the speech processing model proposed by Stackhouse and Wells (1997). Numerous studies have used this model to investigate speech skills and employ its key hypothesis that children's speech difficulty could be impaired at different levels of processing. The model has also been successful in understanding the typical development of speech (Pascoe et al., 2016; Stackhouse, Vance, Pascoe, & Wells, 2007). This existing clinically applicable psycholinguistic speech processing model of Stackhouse and Wells (1997) is the focus of the present thesis. Although other psycholinguistic models exist (e.g., Levelt, Roelofs and Meyer; 1999); Stackhouse and Wells' model is favoured due to its perceived advantages. First, the model is considered influential, in so that it has facilitated the understanding of speech processing in both typical and atypical speech development. It is based on a very strong theoretical background and was developed after years of psycholinguistic and cognitive neuropsychological research (Stackhouse & Wells, 1997). Stackhouse and Wells' work has been influenced by previous models developed by Wateson (1987) and Hewlett (1990), where they expanded on Hewlett's (1990) model by extending and adding a wide series of processes, starting from audition to motoric production. Second, researchers have already applied psycholinguistic models to understanding both typical and impaired speech. In fact, all practitioners who work with children with speech and language difficulty draw on psycholinguistics in some way; however, application has been at a surface level and has been disorganised. Unlike other speech processing models, Stackhouse and Wells introduced a systemic speech processing framework to design specific tasks to assess different components of the psycholinguistic model, and to score and compare the performance of different tasks based on the model's theoretical understanding of speech processing in children. Clinicians can then use the information from the assessment tasks to profile an individual child's speech processing skills and deficits as a bases for intervention. Tasks such as picture naming, word repetition and non-word repetition are differentiated within the

Stackhouse and Wells speech processing framework, but this is not always the case with other speech production models; for example, the model proposed by Levelt, Roelofs and Meyer's (1999) does not fully offer complementary contribution to understanding how novel words (non-words) are processed.

It was noted earlier, that the understanding and clinical management of children's speech development and difficulties in Saudi Arabia is predominately medically and linguistically based and have been limited to these approaches. To the author's knowledge, applications of a psycholinguistic speech processing approach to investigate Arabic-speaking children's speech are sparse. Therefore, the use of the psycholinguistic perspective of Stackhouse and Wells to investigate speech development and difficulties can be appealing to those interested in speech development of an Arabic-speaking child and to speech and language therapists working with Arabic-speaking children for a number of reasons. First, in recent years, there has been a growing demand among speech-language therapists in Saudi Arabia to work in school based settings. This demand has been highlighted with the expansion of inclusive schooling/education (See Aldabas, 2015, for a full review). Second, speech and language therapists are constantly challenged when working with children with speech difficulties, as they are faced with the heterogeneity and complexity of caseloads combined with limited assessment resources and norm-referenced data for Arabic-speaking children. Therefore, with the future demands of school-based therapy setting, along with the challenges of diminished diagnostic resources, the importance of introducing the psycholinguistic processing approach is emphasised because, in clinical practice, the psycholinguistic approach can be used with any child, whether the child presents with a known aetiological cause such as dysarthria, cerebral palsy or a speech difficulty of unknown origin. With this approach the speech therapist is not burdened with a diagnostic label, as the approach does not aim to differentiate between diagnostic labels, rather the processing approach could uncover underlying difficulties and individual difference for children given the same diagnostic label, and could locate more than one level of difficulty (Stackhouse and Wells, 1997).

Before going into further detail on Stackhouse and Wells' (1997) psycholinguistic speech processing model, Section 1.3 will provide more information on speech difficulties in general and in Saudi Arabia specifically, as the present thesis is motivated by the clinical application of the psycholinguistic approach to Arabic-speaking children with speech difficulties. The section will be followed by the psycholinguistic framework of Stackhouse and Wells (1997).

1.3 Speech Difficulties in Children

Speech difficulties[4] are estimated to be the most common paediatric communication difficulty, comprising an estimate of 70% of paediatric speech-language therapy caseloads (Mullen & Schooling, 2010). Working with children with speech difficulties provides a privileged insight to understanding the overt and covert nature of children's speech, while they pose a challenge due to their highly complex and heterogeneous nature (Dodd, 1995, 2005; Stackhouse, 1996; Tyler, 2010). While the cause of speech difficulty could be attributed to a known origin such as hearing loss, cleft lip/palate, cerebral palsy or cognitive impairment, in most cases the cause of speech difficulty is unknown (Broomfield & Dodd, 2004a; Shriberg & Kwiatkowski, 1994). A UK incidence study revealed that children with speech difficulty without any co-occurring language, cognitive or physical difficulties form the greatest number of referrals to speech-language services, where the estimated annual incidence was 6.4% (Broomfield & Dodd, 2004a, 2004b). A systematic review of epidemiological studies (Law, Boyle, Harris, Harkness, & Nye, 2000) suggested the prevalence for speech difficulties to be from 2.3% to 24.6% and for combined speech and language difficulties to be from 4.56% to 19.0%.

Clinical population and caseload characteristics provide valuable information that reflects the nature of difficulties in children who receive therapy services. Unfortunately, a review of literature and recognised websites such as the Saudi Society of Speech-Language Pathology and Audiology; the Saudi Commission for Health Specialties; JISH (Jeddah Institute of Speech and Hearing) and The Ministry of Education, did not yield information regarding the prevalence and/or incidence of speech and/or language difficulties in speech-language therapy settings in Saudi Arabia. Furthermore, to our knowledge, published studies showing the prevalence and/or incidence of speech difficulty in Saudi

[4] Speech difficulty is a term used in the present thesis to refer to "*children who experience difficulty acquiring accurate and intelligible speech according to the expected developmental timeline*" (Sosa, 2015, p.24). Different researchers (Dodd, 2005b; Stackhouse & Wells, 1997) and professional bodies such as RSCLT (2015, viewed from official website) favour the term speech difficulties; although the term speech sound disorders (SSD) is widely adopted in the literature as a broad cover term see (Bowen, 2014) for a detailed review on the history and terminology of speech difficulties.
For this present thesis, the term speech difficulties, speech disorders or speech sound disorders (SSD) will be used interchangeably. The term will be used regardless of the origin of the difficulty (i.e. known or unknown causes).

children are sparse. Most of the available studies show the prevalence and/or incidence of speech difficulty in samples of children where the cause of speech difficulty is presumably known (Al-Ghamdi et al., 2002; Al-Sulaiman et al., 2003; Albustanji, Albustanji, Hegazi, & Amayreh, 2014). For example, in a recent study, Albustanji et al. (2014) investigated the prevalence of speech difficulties in a small sample of 80 participants between the ages of 6-15 years old with cleft lip and palate. The study reported that 74% of the sample had a speech difficulty (59 participants). Al-Ghamdi et al., (2002) evaluated 111 children (mean age=6) registered at the Center of Disabled Children in Buraida and Unaizah. The children had cerebral palsy, chromosomal abnormalities or miscellaneous disability. The study found that among other difficulties, 88.3 % of the sample had severe speech difficulty. Similarly, Al-Sulaiman et al., (2003) investigated difficulties associated with cerebral palsy in Saudi children aged 1-3 years old who were referred to the neurology department of King Fahad Hospital in Alkhobar city. The children were evaluated in monthly intervals during a one-year period. According to the study, 52% of the children with cerebral palsy had speech difficulties.

The prevalence and incidence studies thus far were based on children with a known cause of speech difficulty. However, one study investigated the prevalence of speech difficulties of unknown origin in Saudi school-age children (Awaad, 2008). The study included 11015 school-age children between the ages of 8 and 10 years old and the sample was from 62 schools in Jeddah city. The study found that 9.55 % of the children in the sample had speech difficulties. Although these estimates do not include preschool children; it is clear that a large number of Saudi Arabic children do not resolve their speech difficulties before they enter school.

Therefore, based on the available prevalence/incidence studies of speech difficulty in Saudi children and from international figures, one could estimate the number of children with speech difficulties in speech-language therapy settings. The pre-school years are the target sample of the present study, as the literature manifests the preschool ages as the critical and fundamental stages of development where speech and language skills should be developed. Law et al., (2000) indicated that children who do not receive intervention, or who begin intervention in the school years, can continue to have difficulties for at least 28 years. Thus, the pre-school years are crucial for identifying, assessing, and providing intervention to children with communication difficulties including children with speech difficulties.

To this end, researchers and clinicians acknowledge that children with speech difficulties form an extremely heterogeneous group in terms of aetiological involvement, severity, underlying causes, cognitive linguistic involvement and surface speech-error patterns (Dodd, 1995, 2005, 2011; Stackhouse, 1996; Tyler, 2010). Therefore, children with speech difficulty should be managed from different perspectives, including the three major approaches commonly used in research and practice when describing speech difficulty in children, namely, the aetiological (medical) perspective, the descriptive-linguistic perspective, and the psycholinguistic speech processing perspective (Tyler, 2010; Waring & Knight, 2013). The three schemes have different theoretical views on speech difficulties and are generally driven from different academic disciplines.

The following section (1.4) will provide a review of Stackhouse and Wells' (1997) psycholinguistic framework. The review of the framework will include the simple speech processing chain, which is the essence of a psycholinguistic speech processing approach. This will be followed by the framework's complex box-and-arrow speech processing model that introduces the components of the speech processing system. Finally the section will overview the assessment principles of the framework.

1.4 The Psycholinguistic Framework of Stackhouse and Wells

The psycholinguistic framework of Stackhouse and Wells (1997) is a developmental linear model that lists the components underlying speech production and links theory to clinical practice. The aim of the psycholinguistic framework is to provide systemic information about why/where breakdown occur within the speech processing chain of a child with speech difficulties, and what the weakness and strengths of the child's speech are. Nevertheless, Stackhouse and Wells (1997) emphasis that in order to understand and remediate children's speech difficulty, it is vital to identify not only *where* breakdown occurs in the speech processing chain, but also *when* the ability to acquire knowledge and skills normally develop (and *how* it develops). Thus, it incorporates a developmental perspective. The premise of the speech processing framework is therefore: **a)** typical speech development depends on the normal functioning of the speech processing system; **b)** breakdown in one or more of the levels of processing system results in speech difficulty and **c)** speech difficulty can be remediated by targeting the level/s of breakdown in children's speech processing system.

1.4.1 *Simple Speech Processing Chain*

The essence of the psycholinguistic framework of Stackhouse and Wells (1997) is that speech processing involves the hypothetical routing of information from the basic component of input processes, lexical representations and output processing; these components are the key terms used in any psycholinguistic processing model –as noted earlier in section 1.2.3. Figure 1.1 shows a schematic representation of a simple speech processing chain in which Stackhouse and Wells conceptualize the basic components of the speech processing system and the directions of processing. The first component is speech perception (input processing), in which a child receives information through the auditory (or visual) system; the information is then routed up to the central storage of lexical knowledge (lexical representations). The sounds/words are then selected and assembled to generate speech output (output processes). In psycholinguistic terms, information can be processed in a *top-down* direction (speech activity that utilises stored information from the lexical representations) and/or a *bottom-up* direction (speech activity that can be completed without accessing stored linguistic information form the lexical representations). This notion will be delivered in more detail in the subsequent sections on task requirements.

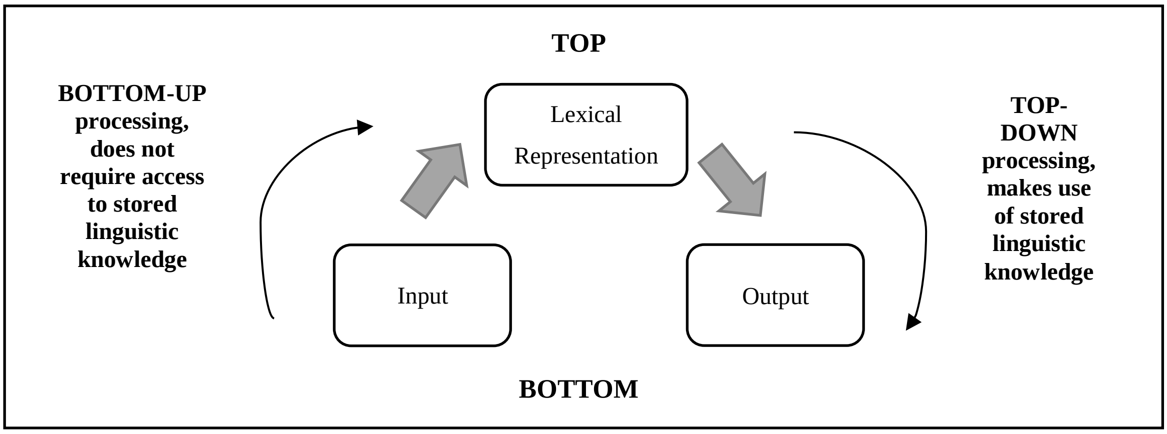

Figure 1.1: The basic components of Stackhouse and Wells' (1997) speech processing model.

1.4.2 *The Speech Processing Box-and-Arrow Model*

Stackhouse and Wells introduced a box-and-arrow model that displays in more detail the levels of speech processing and the processing routes assumed by the framework. It provides a more explicit way for understanding children's speech difficulties from the psycholinguistic perspective, which is helpful when communicating with other professionals or for research purposes. The box-and-arrow model is an information processing model, which is a conventional way of visually representing the speech processing components and the routes between them that are thought to be involved when children process and produce speech. Essentially, the box-and-arrow model takes the simple processing chain, extends and develops it, where the basic levels of input processing, lexical representations and output are built up to include sub-processes involved at each level. The model is presented in Figure 1.2. The plain boxes represent the levels of processing; stored knowledge are represented by the three bold boxes. The shaded boxes are processes hypothesised to occur off-line. Arrows show the root of processing, while the bold arrows represent the flow of information as part of a learning process i.e. off-line processing.

The components of the model in Figure 1.2 are summarised as follow:

Input Processing:

- **Peripheral auditory processing**: the peripheral point of input processing on the left of the model. It represents general auditory ability, not specific to speech and occurs at the ear.

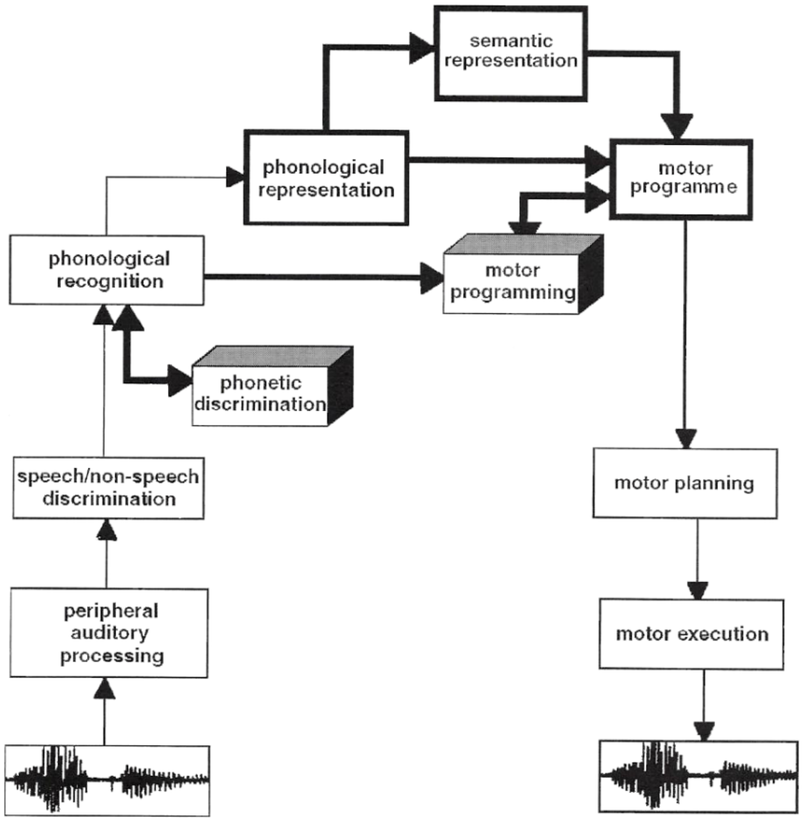

Figure 1.2: The speech processing model proposed by Stackhouse and Wells (1997). Adopted from Stackhouse and Wells (1997).

- **Speech/non-speech discrimination**: a per-linguistic level of processing in which input speech sounds are recognised as speech rather than non-speech/environmental noises before it send for further decoding.
- **Phonological recognition:** the level were the listener recognises the speech signal as belonging to their language and appropriate to the language-specific patters of their language.
- **Phonetic discrimination:** begins in early childhood when the child learns words and starts to learn to contrast segments of different words and when learning new language. It requires the ability to recognise phonetic distinctions that are unfamiliar to the listener.

Lexical Representations

- **Semantic representations:** part of the mental lexicon, where the meaning of a stored word in the mental lexicon is located.

- **Phonological representations:** a central cognitive-linguistic skill, where knowledge about the phonological structure of a word is stored. It does not exist in isolation; rather it includes the meaning of the word.
- **Motor programs:** a series of stored articulatory instructions for producing a word. The motor programs sends articulatory gestures compatible with stored phonological representations.

Output Processing

- **Motor programming:** the level that facilitates the creation of a new motor program rather than relying on pre-existing programs. It is presented in Figure 1.2 as a shaded box where online processing occurs. It is thought of as comporting a store of phonological units that are selected and sequenced/assembled in new combinations. This motor programming device is based on input to create a new motor program. The child's ability to create new motor programs is commonly assessed using the non-word repetition (NWR) task.
- **Motor planning:** after a stored motor program is retrieved or new a motor program is created, the targets are sent to motor planning level. This level assembles the target gestures in correct sequence in real-time, and takes into account the contextual requirements of real-time productions. These contextual requirements such as speed, intonation, pitch, are planned in advance and achieved through neuromuscular activity. It is the level where the motor programs for single words are assembled into a single utterance plan.
- **Motor execution:** occurs at the mouth, or the vocal tract, and includes all the physical organs responsible for producing speech. It is the level where the motor plan is executed to give rise to the speech signal.

1.5 Principles of a Psycholinguistic Assessment

Stackhouse and Wells' (1997) framework provides the theoretical structure for organising tasks so that children's performance on a range of different tasks informs our understanding of children's processing skills, and tasks can be interpreted based on their psycholinguistic speech processing demands. Within Stackhouse and Wells framework (see Figure 1.1 and 1.2), assessment procedures are defined along two dimensions, 1) tasks that are classified as either input or output tasks and 2) tasks that are dependent on linguistic knowledge stored in the lexical representations as well as tasks that are not dependant to some degree on stored lexical representations (top down *v*. bottom up). A comprehensive psycholinguistic approach to assessment focuses on both input and output channels of processing and, with specific tasks, it allows for the detection of speech difficulties at the levels of input, stored representations and output. Tests of auditory skills such as hearing tests and auditory discrimination tasks are one of the first assessments conducted on a child with speech difficulty, these tasks are hypothesised to tap the different levels of input processing, and should determine whether the levels of input are intact. Another level that should be assessed on the psycholinguistic model is the output level. Output levels includes the level of motor execution (at the mouth) and are assessed using tests of oral examination and measures of articulatory skills (sound production in isolation and in sequences). Few examples include the Time by count Test of Diadochokinetic syllable rate (Fletcher, 1978), and the Oral Speech Mechanism Screening Examination (OSMSE-3) (St Louis & Ruscello, 2000). Oral examination of structure and function such as the Fletcher and OSMSE-3 are very common and highly used in speech and language therapy clinics in Saudi Arabia. Furthermore, output tests such as the picture naming tasks are very common in clinics to assess output articulatory skills and phonological processes in children (Abou-Elsaad et al., 2009).

The focus of this thesis is on output tasks i.e., the left side of Stackhouse and Wells' model (See Figure 1.2) where tasks tap both bottom-up and top-down representations. Particularly, the tasks include the repetition of words, non-words and syllable sequences. The use of repetition tasks for this purpose has been influenced by many factors. First, studies of speech skills in Arabic-speaking children mainly cluster around speech activities that look at simple speech processing skills of input, underlying representations and output without considering the different processing levels and different processing demands of speech tasks. For example, with regards to the input side of speech processing, researchers have investigated Arabic-speaking children's speech perception skills using tasks such as discrimination and identification (e.g. Al-Mannai1 & Everatt, 2005; Taibah, 2006; Al-Harbi, 2007; Kishon-Rabin & Rosenhousea, 2000). Further, in relation to phonological representations and the output side, the array of literature on Arabic-speaking children has mostly focused on the phonological system and on speech sound development at the articulatory level (e.g., Amayreh, 1999, 2003; Abou-Elsaad, 2009). Such studies are therefore limited to investigating children's output skills using tasks that only capture one level of speech processing. To address this gap, the present study will use different tasks that are hypothesised to require different processing demands and which will tap different output processing levels within Stackhouse and Wells model (i.e., specifically, repetition tasks). These tasks are hoped to provide novel insights into an Arabic-speaking child's output processing skills. Second, repetition tasks are relatively straightforward to administer and are familiar tasks to clinicians, as the repetition of words in particular are commonly used in clinical practice. Furthermore, repetition tasks have the advantage of being suitable for very young children and children with language or cognitive difficulties (Stackhouse, Vance, Pascoe & Wells, 2007).

Repetition tasks of words, non-words and sounds and sequences of sounds are clearly differentiated within the Stackhouse and Wells framework. However, testing and interpreting responses is very problematic and should be interpreted with caution, this is due to the complex underlying nature of speech tasks as it could tap a number of underlying cognitive and linguistic skills. This is particularly true when testing young children, as their underlying representations and skills develop simultaneously, thus difficulty with one or more of the processing levels will result in reduced accuracy of repetition. When testing children with speech difficulties, an assumption cannot be made that a child has intact underlying skills, and breakdown is on one level without the other. Therefore, the importance of using different tasks to measure performance is emphasised in the investigation of typical development and levels of breakdown in children with speech difficulty. Within this psycholinguistic framework of Stackhouse and Wells (1997), although children's responses on a task informs the understanding of the speech processing skills in children; comparison of children's performance on different tasks that target different levels of processing is considered more informative and can contribute most to the understanding of children's processing skills and how this changes over development. With repetition tasks, direct comparisons could be made and would be most useful if the stimuli used for the repetition tasks are matched in terms of phonetic detail, structure and complexity. Furthermore, task comparisons will be informative if the tasks used are challenging enough to capture children's sensitivity to tasks. Tasks should include stimuli of different lengths to capture any processing or developmental differences in children.

The present thesis focuses on the processes involved in the production of speech namely, the output motor processing levels. The subsequent chapter will therefore review the speech production tasks that are assumed by Stackhouse and Wells' model to tap the different output processing mechanisms covered above, further details on the levels of lexical representations (including motor programs), motor programming, planning and execution will be reviewed along with the tasks. Since, the performance of a child with speech difficulty on speech tasks should be interpreted and compared to the patterns of performance in relation to typically developing peers. Therefore, patterns of performance on speech processing tasks in typically developing children and developmental changes on task performance will also be reviewed.

Summary of Chapter 1:

- There is a universal agreement that children with speech difficulty are heterogeneous in nature and form a large portion of clinical caseloads.
- Researchers have tried to provide a systemic method to describe and classify children's speech difficulties, using aetiological, descriptive-linguistic and psycholinguistic approaches.
- Both the aetiological and linguistic approaches to assessing children with speech difficulty have dominated the literature on Arabic-speaking children's speech output skills and the way speech and language therapist conceptualise speech difficulties in children. Psycholinguistic speech processes has received little attention in the literature in both typically and atypically developing Arabic-speaking children.
- Stackhouse and Wells' (1997) framework is not intended to be a classification system; rather, their framework is a psycholinguistic approach that intends to assess children's underlying speech processing skills. It can be used with any child, and uncover hidden underlying difficulties regardless of the label attached to the child's speech difficulty. The assessment enables therapy to be tailored specifically to a child's needs.
- The principle of the Stackhouse and Wells' (1997) framework is that assessment tasks require different demands in terms of input, lexical representations and output. Different tasks are assumed to tap different speech processing levels. Furthermore, one of the key aspect of the framework is that tasks should not be administered in isolation, as isolated single tests provide minimal information on underlying processing skills. They stress that more than one task should be compared against each other as this will be more informative and would provide a better understanding of children's underlying skills.

- Speech output tasks that include repetition of real words, non-words and sounds/sound sequences are the focus of the present thesis. The three tasks are hypothesised within the model of Stackhouse and Wells to tap different levels of processing.
- In order to apply the psycholinguistic approach on Arabic-speaking children with speech difficulties, it is essential to investigate how typically developing children perform on the tasks, and how children's performance on the tasks change with age, i.e., whether tasks are sensitive to developmental change.

CHAPTER 2 LITERATURE REVIEW II: SPEECH OUTPUT TASKS & CROSS-LINGUISTIC EVIDENCE

Chapter 1 introduced the theoretical background of the psycholinguistic speech processing approach, in particular, Stackhouse and Wells' psycholinguistic speech processing framework (1997). It was understood that different tasks are hypothesised to have different levels of processing within the framework and that children's performance on speech processing tasks should be compared within and between levels in order to have a greater understanding of children's speech processing skills. The previous chapter also stressed that in order to apply speech processing tasks and understand the nature of speech difficulties in children it is essential to study how typically developing children perform on speech processing tasks, therefore, tasks should be investigated for its sensitivity in capturing processing demands and its sensitivity to developmental change.

Therefore, this chapter will:

- Analyse the psycholinguistic properties of repetition tasks that trigger assumed speech output levels of processing i.e. tasks identified previously were real words, non-words and syllables sequences. The tasks will be understood within the current projects adopted Stackhouse and Wells' (1997) model (Section 2.1).
- Since preschool children aged 3 to 5 years are the target of the present study, cross-linguistic psycholinguistic studies on children's performance on repetition tasks on a single-word level will be reviewed (Section 2.2). Both evidence on the effects of processing demands and age will be included.
- Evidence on the effects of length on task performance and developmental changes with length will be reviewed (Section 2.2.2).
- The review of the literature will move to rapid repetitions or rapid consecutive repetition tasks, which capture children's performance at multiple-word level productions real words, non-words and syllable sequences. It will include analysis of the psycholinguistic properties of rapid multiple productions, processing profiles and developmental progression (Section 2.3).

2.1 Speech Repetition Tasks

Any task that requires spoken output (in this case repetition tasks) requires the child to either access pre-existing phonological representations/motor programs which is part of stored lexical representations or to create a new motor program, tax motor planning and execution for speech production. Description of the uses of the tasks and its psycholinguistic analysis will be reviewed below.

2.1.1 *Real Word Repetition RWR*

From a very early age, typically developing children gradually acquire words and build their repertoire. In terms of psycholinguistic models of information processing, the child stores information about individual words in underlying lexical representations. The information stored about the word includes knowledge of a words sound structure (phonological knowledge) and its meaning (semantic knowledge). Therefore, lexical representations are viewed as bodies of knowledge about a word that is built up over time, and as the child develops, the child’s underlying knowledge expands. Chiat (2000) notes that knowledge of a word is a store of phonological-semantic complexities; which is termed mental vocabulary or mental lexicon.

The model proposed by Stackhouse and Wells accounts for the essential central cognitive-linguistic ability to process speech, identified as underlying lexical representations. The young child’s lexical representations are thought as a store of knowledge about a word. This store of early lexical knowledge are described to include thee levels: semantic representations which provide information about the word meaning, the phonological representations which provide information on the word sounds or sound structure, motor programme that provides information on how to say the word, i.e. the articulation of a word. In the speech processing model illustrated in Figure 1 (Chapter1 Section 1.3.2), the three distinct but interconnected components of stored lexical knowledge (lexical representations) of a word can be seen at the top of the model. The components are enclosed in bold as they represent stored knowledge.

Therefore, a task that requires a child to repeat a familiar word involves the child’s ability to repeat a word stored in the mental lexicon. The task would require the child to access lexical representations in long-term memory and activate both phonological (and semantic knowledge, although not necessary with repetition) and the motor program. The motor program has a series of stored articulatory gestural details of a word that will achieve a production compatible with stored phonological representations. The repetition task of a word familiar to the child is designed to target the stored motor program level i.e. whether a child can realize motor programs accurately. Nevertheless, the repetition task involves other levels of processing, including input processing, as the task involves hearing the word first. The child processes the word through peripheral hearing, discriminates the signal as a speech, and recognises the phonological pattern to be specific to Arabic which is then forwarded to the stored lexical representations. The phonological representations and motor program of the word are activated as the word matches an item stored in lexical representations. The motor program is then forwarded to the output processing side of the models which includes motor planning and motor execution (see Figure 1.1). The motor plan is responsible for assembling the articulatory gesture targets of the stored motor program in the correct sequence, and includes the contextual requirements of the production (for example speed of production, rhythm, pitch and intonation). Finally, the instructions from the motor plan are sent to the articulators i.e. the vocal tract, located at the motor execution level. There, the physical organs such as the lungs, vocal folds, tongue, lips and soft palate are co-ordinated for the production of a word. At this level, the motor plan is executed and gives the acoustic signal. It should be highlighted that, the presence of an adult model enables a child to use a non-lexical rout and simply mimic the adult input, using the motor programming level (reviewed below in section 2.1.2). However, at a young age, children’s phonological representations and motor programs are still developing and unspecified; therefore, although there is an accurate adult model for the repetition of a word, a child has underspecified representations that could interfere with accurate production.

2.1.1.1 *Development of Phonological Representation*

The phonological representations are thought of as part of the child’s stored knowledge of the sound structure of a word. For a child to identify a word from spoken input and produce a word accurately, the phonological representations must be accessed. The phonological representations include abstract enough phonological detail/information to recognise the word as distinct from other related items and identify the word and confirm with the stored form. The nature and development of this phonological representations within the lexicon in children has long been investigated. In a classical study that is frequently referred to in the phonological development literature as one of the early studies of word production, Ferguson and Farwell (1975) collected longitudinal data from three 12-month old toddlers acquiring their first 50 words. Ferguson and Farwell suggested that during the early stages of lexical development, children's lexical representations are represented as whole units, where the word or phrase, not the phoneme, is the minimal unit of phonological representations[5]. This came from the observation that some children mastered the phonemes [b] in certain positions of some words however, did not maintain the order in which it was mastered in other words, instead they redistribute in other contexts. This is seen as the "whole-word" system of phonological representations or "holistic" representations in early language development, as they lack segmental detail.

Furthermore, the nature of phonological representations changes with development. There is a point when children transition from the hypothesised holistic phonological representations to a more phonemic, segmental, representation (Metsala & Walley, 1998) According to Metsala and Walley's (1998), lexical reconstructing model, early lexical representations are holistic in nature as there is simply no need to represent words in a more detailed manner. As children's vocabularies grow, however, and as children gain more language experience, the increasing similarity among words in the lexicon creates pressure to form more fine-grained, phonemic representations to allow for accurate word recognition and production.

2.1.2 *Non-Word Repetition NWR*

A non-word (NW) refers to a phototactically legal novel word i.e. made-up word modelled after a native language, while the term non-word repetition NWR refers to a task that requires participants to hear a NW and repeat it immediately. The ability to hear and repeat novel phonetic sequences is one of the most basic and important language abilities and is closely related to language acquisition. Known words would have started its journey into the mental lexicon via a repetition attempt. From the first year of life, children spontaneously mimic words they hear and by the age of 2-years are able to repeat a word on request (Gathercole, 2006), thereby NWR imitates the process by which a new

[5] The phonological development and adult phonology literature include a rich array of evidence supporting "holistic representations" in early phonological development that move to a more segmental-detailed representations in adults - see Vihman & Keren-Portnoy (2013) for a comprehensive review. This view is by no means uncontroversial; studies on children's perceptual skills show that children are sensitive to phonological material, with more segmental and less holistic representations (few examples include, Aslin et al, 1998; Cody & Aslin 2003, 2004; Jusczyk *et al.*, 1999; Swingley & Aslin, 2000, 2002). It is beyond the scope of this thesis to review them; the reader could refer to the citations provided for more information.

word may become a part of the mental lexicon. Stackhouse and Wells' model differentiates between stored knowledge built up over time and by online processing at a given moment in time. Therefore, the model accounts for how the stored motor programs (discussed above as part of stored lexical knowledge) came to existences, by assuming an online motor programming device that creates new motor programs based on input. A child's ability to create a new motor program is tested by a NWR task. According to Stackhouse and Wells (1997), motor programming can be thought of as a box or a repository containing phonological units, where these units are selected and assembled in new combinations. Following the earlier review on the nature of phonological representations, it is suggested that young children with small vocabularies are likely to have phonological units that are holistic in nature and not segment-sized, and with expanding knowledge, their representations would be more segmented, therefore, the model accounts for repeating NWs at a sublexical level. Thus, as the child gains more experience with their language, more options of phonological units are available for combinations.

Therefore, although NWR is not hypothesised to tax stored lexical representations, the motor programming repository is not entirely separate from stored phonological representations/motor programs. Empirical evidence suggests that the accuracy in which a child repeats NWs is closely related with their lexical knowledge and development. In fact, evidence suggests that children as young as 2 years old were influenced by knowledge of their language during NWR tasks. Studies sugest that receptive vocabulary, but not necessary expressive vocabulary, accounts for accurate NWR in English-speaking children ages 2;0 to 4;0 years (Shula Chiat & Roy, 2007; Hoff, Core, & Bridges, 2008). Furthermore, a study on Korean-speaking children ages 3;3-5;8 years also found an association between NWR accuracy and vocabulary knowledge, both expressive and receptive (Lee, Kim, & Yim, 2013). Furthermore, 2-year old English-speaking children's repetition of NWs were influenced by their vocabulary size, lexical and sublexical factors such as neighbourhood density and frequency of phonotactic patterns (Cody & Aslin, 2004; Zamuner *et al.* 2004, Eaton, Newman, Raner & Rowe, 2015).

The present thesis dose not attempt to directly investigate the influence of vocabulary knowledge or lexical/sub-lexical influence on Arabic-speaking children's repetition accuracy. However, these factors will be controlled for to some extend during stimuli design. Nevertheless, it is assumed that as the child gets older and experience with language increases, their repetition accuracy will increase.

Furthermore, different research groups report different sources to support NWR and there is a long history of research into what NWR really measures. Some research groups have focused on phonological short-term memory (or working memory) to explained performance on NWR (Coady & Evans, 2008). Gathercole, Baddeley and colleagues conducted a series of studies using NWR in examining the relation between phonological memory and language development in typically developing children (e.g. Adams & Gathercole 2000, Gathercole & Adams, 1993, 1994; Gathercole, 1995, 2006 Gathercole, Willis, Baddeley, Emslie, 1994,). They have reported significant correlations between NWR and measures of phonological memory, such as digit span, and correlations between measures of phonological memory and vocabulary. They propose that successful NWR is mediated by temporary phonological storage capacity, where a child can retain novel phonological strings in phonological working memory for immediate repetition. Thus, children with more memory resources to hold a novel phonological string would be more successful at repeating a NWR and at lexical acquisition. Gathercole and colleagues generally view NWR as a measure of phonological short-term memory separate from other phonological processes. Therefore, over the years, there has been considerable attention in the literature on the use of NWR as a clinical marker of specific language impairment in children; with mounting evidence that poor NWR is a marker of specific language impairment, as children with language difficulties are less accurate at NWR compared to their typically developing peers, (for an extensive review see Coady & Evans, 2008; Gathercole, 2006). This comes from the notion that NWR involves intact underlying supporting skills such as phonological working memory and that it is closely linked to lexical development, while children with language impairment have phonological memory deficits resulting in poor word learning.

The psycholinguistic model of Stackhouse and Wells (1997) does not account for phonological working memory; nevertheless, they do not disregard a memory load account on the ability to repeat NWs. Within Stackhouse and Wells' model, a child repeating a NW would discriminate and recognise the novel phonological combination, assemble the new phonological units at the level of motor programming, hold the NW in memory long enough to generate a plan for an articulatory output.

Baddeley (2008, 2012) described working memory as a multi-component system serving cognitive tasks and implies simultaneous storage and processing of information over a short period of time. It includes the phonological loop which deals with verbal information and is integral to word learning and NWR. The phonological loop contains the phonological memory capacity and rehearsal process. Therefore, when memorizing a string of digits, it is temporary stored in the loop, and is actively maintained by rehearsal, otherwise the string of digits may decay. Baddeley (2003) suggest that phonological loop is not distinct from language knowledge.

Snowling, and colleagues (Snowling, Chiat, & Hulme, 1991) argue that NWR is not a pure measure of phonological memory, rather, successful NWR requires the accurate perception of the NW, creating at least a transient representation in working memory, segmenting the novel phonological string into speech units, making temporal orders of the units, formulating a motor plan for articulation and then implementing the motor plan. The interaction between the phonological process and memory should be taken into consideration when interpreting NWR. Snowling and colleagues further argue that NWR does not necessary provide a measure of phonological memory that is free of any lexical influences and is "content-free" of the influence of prior stored word knowledge. Children will use their existing lexical knowledge to support NWR. On the other hand, MacDonald and Christiansen's (2002) argue that the distinction between phonological working memory and language knowledge is artificial, and could not exist at all, and that any difference in tasks such as NWR that presumably test phonological working memory could be an artefact of the child's lexicon, and NWR could be a reflection of the properties of the lexicon (such as word-likeliness) and biological constrains (such as precision of underlying phonological representations).

Therefore, a conclusion to be drawn from the literature is that there is strong evidence to support that NWR accuracy is influenced by long-term stored lexical knowledge over and above phonological and memory process.

Another point to consider regarding NWR, that although studies on English-speaking children and other languages have confirmed a relationship between NWR and language experience/development and disorders, Stroke et al. (2006) argue that performance on NWR is not only shaped by children's language experience but also by the phonological structure of the language itself. they found that Cantonese-speaking children ages 4;2 to 5;7 (years; months) performed similarly on a multisyllabic NWR task that followed the phonotactic rules of Cantonese. Stokes *et al*, (2006) suggests that the Cantonese language has a small phonetic inventory; restricted syllable structure with constant stress, this general phonological simplicity of the language allows the children to reconstruct their phonological representations in working memory. Thus, the suitability of NWR as a measure of phonological processing or in differentiating typical form atypical speech/language development cannot apply to different language systems. To our knowledge, NWR tasks have not been developed nor has been investigated in Arabic-speaking children. Developing an Arabic language-specific NWR task allows building and conformation of the relationship between speech/language development and disorders that have been found in other languages. Nevertheless, it is expected that Arabic-speaking children will perform on NWR similarly to English-speaking children, this due to the fact that both language systems have phonetically based language system with varying syllable structures, complexity and stress patterns. Furthermore, Arabic is a distinctive semantic language with many emphatic consonants, and geminate is one of the Arabic language features.

One confounding variable when testing young children who are still in the developing stages of speech production is their speech sound errors. Children as young as 3 years old could fail the repetition tasks due to their speech production limitations (Chiat & Roy, 2007; Stokes & Klee, 2009), where accuracy of NWR is constrained by the maturity of the phonological and articulatory systems. Usually errors in NWR are generally assumed to indicate difficulty in successfully encoding and holding the string of sounds in memory. However, in young children this could not be the case, as they simply may not be able to produce the sounds of the task. Similarly, this could be the case in children with SSD. In order to control for such a confound, studies investigating NWR in very young children as well as studies on children with SSD have used a variety of methods. The methods include, adjusting the scoring criteria to accommodate for the common speech errors seen in children (Roy & Chiat, 2004), or controlling for speech production errors by using an individualised approach, where children's phonemic inventories are a analysed using a standardised articulation assessment and then counted as correct those items in the NWR task that were produced with a consistent error (Chiat & Roy, 2007; Stokes & Klee, 2009). Another method would be designing non-word items that include only early developing sounds such as /b/ and /m/ (Shriberg et al, 2009). Alternatively, some researchers (Hoff, Core & Bridges, 2008) have incorporated real-word repetition as a control variable for speech production ability. In summary of research findings, NWR requires strong representations of underlying speech units, and requires sufficient memory to store and operate on the novel phonological strings (Coady & Evans, 2008). It is widely acknowledged that the ability to repeat a novel word involves complex multistage processes which include speech perception, cognitive (acoustic signal is segmented into smaller speech units, and then stored in memory), and motor process which include motor plan where the speech units are formulated and assembled and articulation.

2.1.3 *Syllable Repetition Task*

Motor execution is the lowest level of the output right-hand of Stackhouse and Wells' model (refer to Figure 1.2) and is the level where the acoustic signal is produced. It involves the vocal tract including all its physical role in speech production (lungs, larynx, oral and nasal cavity). According to Stackhouse and Wells' model, syllable repetitions (as with oral-examination) taps the motor execution level, near the periphery of the output side of the model. It is not a linguistic activity as the sounds do not have to confirm with the child's knowledge of their language. Articulation tests that include sound and syllable imitation are at the lower level of speech processing. They could subtests as in the Diagnostic Evaluation of Articulation and phonology [DEAP] (Dodd, Zhu, Crosbie, Holm, & Ozanne, 2002).Investigating a child's ability to produce speech sounds and sequence a string of speech sounds is a fundamental part of assessment when investigating children with speech difficulty. Children with structural or functional abnormalities, such as cerebral palsy or dysarthria will show difficulty with this task, other subtle speech motor difficulties such as apraxia of speech could also have difficulties with this task.

Thus far, the repetitions tasks were analysed within the psycholinguistic model of Stackhouse and Wells (1997). However, as mentioned previously, according to the framework, interpreting results of tasks in isolation would be misleading, rather, the patterns of children's performance across more than one task is more informative and essential. The following sections will therefore provide evidence from the literature on the effects of repetition tasks, age and increasing number of syllables of stimuli on children's speech processing skills. First, the evidence from psycholinguistic research on children's speech processing skills, at an immediate single-word repetition level will be reviewed, and second, repetitions at a multiple-production level, where rapid consecutive repetitions are required, will be reviewed.

2.2 Single Repetition – Word-Level Repetition

Throughout Chapter 1, it was stressed that performance on tasks are best interpreted when compared across tasks (Stackhouse and Wells, 1997), and that tasks sensitivity to developmental change is essential if it would be applied to children with speech difficulty. To this end, it is important to explore haw typically developing Arabic-speaking children perform and progress within the model.

The studies that investigated typically developing children's performance on repetition tasks will be reviewed below. The reader should recall that this section is interested in both processing performance across tasks and developmental changes on task performance, therefore, the section will first present evidence on children's profiles of performance and then move to developmental change. It is important to clarify that many studies reviewed below include both effects of processing and age and length effects, however, due to the density of information, each effect will be reviewed separately. Furthermore, ages will be presented in (year; month), alternatively, ages between 3;0-3;11 will be presented as 3 year olds, children between 4;0-4;11 will be presented as 4 year olds and children between 5;0-5;11 will be presented as 5 year olds.

2.2.1 *Patterns of processing Performance and Developmental Change*

The literature is mounted with evidence showing that typically developing children show different profiles of performance on different speech repetition tasks (Shula Chiat & Roy, 2007; Dispaldro et al., 2011; Hoff et al., 2008; Roy & Chiat, 2004; Sundström, Samuelsson, & Lyxell, 2014; Torrington Eaton, Newman, Ratner, & Rowe, 2015; Vance, Stackhouse, & Wells, 2005). These studies showed substantial effects of lexical status i.e. RWs versus NWs, where RWs were repeated more accurately than NWs. These findings from different studies suggest that children use top-down processing when repeating RWs, accessing stored lexical knowledge and activating the phonological form of a lexical representation in long-term memory.

Studies on English-speaking children have found that the effects of stored lexical knowledge during repetition tasks were observed in children as young as 2 and 3 years of age. Given children's young ages, and that their exposure and familiarity to words are relatively short-lived; still, repetition tasks of RW and NW were sensitive to processing skills emerging at this key period of language development. For example, Roy & Chiat

(2004) designed a repetition test "The Preschool Repetition Test" aimed towards children as young as 2 years old. The test included 36 items of increasing length; 18 RW items and 18 phonetically matched NW items. A total of 66 typically developing children between the ages of 2 and 4 years (age bands: 2;0-2;11 and 3;0-3;11) were tested on the items to investigate the effects of lexical status on repetition accuracy (RWs vs NWs) (they further investigated effects of age and length, however, this will be included later). Chiat & Roy (2007) later replicated their earlier study on a larger sample size of 315 children, also within the ages of 2 and 4 years (age bands included: 2;0-2;6, 2;6-3;0, 3;0-3;6 and 3;6-4;0). Both studies found that children in each age band repeated RWs more accurately than NWs. Similarly, studies conducted by Hoff et al., (2008) on a sample of 15 2-year old children and replicated by Torrington Eaton et al., (2015) on a larger sample of 86 children, found that their 2 year old participants showed greater accuracy at repeating RWs items than phonetically matched NWs. The studies suggesting that at this young age, there was a beneficial effect of stored lexical knowledge when repeating a word.

On the other hand, Vance et al.'s (2005) study revealed a different profile of performance across repetition tasks in 3-year old children. Vance et al.'s study included children between the ages of 3;1- 3;11, between the ages of 4;1-4;11 and between the ages of 5;1-5;11 (each group contained 20 children)[6]. Children were required to repeat a list of 60 RWs increasing in length and phonetically matched NWs. Children's performance on the tasks were explained within the psycholinguistic speech processing model of Stackhouse and Wells (1997). The study showed that whole-word accuracy scores of RWs were not significantly different from NWs at the age of 3 years (repetition accuracy was 66.7% on RWs vs 64% NWs). Vance et al. suggested that at the age of 3 years, children do not use top-down processing (where existing lexical representations are accessed to support RWR). Rather the availability of an adult input enabled the child to simply "mimic" the word and favour a bottom-up processing i.e. non-lexical rout to repeat familiar words, where the word is perceived through input-processing skills and then a new motor program is created thought motor programming skills are accessed to support repetition on RWs and NWs.

[6] Vance et al.'s (005) study had overall 100 participants, which included school-age children ages 6 and 7 years old. However, due to the nature of the current study which focuses on preschool children; only the 3, 4, and 5 year old's performance are reviewed in detail.

From the age of 4 years, however, a top-down processing rout was in operation, as Vance's et al.'s participants ages 4 and 5 years showed significantly higher accuracy scores when repeating RW's compared to its matched NW's. The beneficial effects of stored lexical representations on speech output processing was seen up to the age of 7 year olds.

Evidence from children speaking languages other than English have also found significant main effects of lexical status on repetition accuracy in preschool children (Dispaldro et al., 2011; Sundström et al., 2014). In a study on Italian-speaking children Dispaldro et al. (2011) investigated the perfroamce of 48 children on a RWR and NWR tasks, children were devided into three age groups (of 3;0, 3;6 and 4;0-years old). The study found that young Italian-speaking children were able rpeat real words more accuratley than no-words. Sundström et al. (2014) investigated the performance of 44 Swiss-speaking children with mean ages of 4;4 and 5;4 (20 children within the age band 4;0-4;11 and 24 children within the age band 5;0-5;11). The study also found significant effects of stored lexical knowledge of real words over non words.

Cross-linguistic evidence reviewed thus fare, shows a similar trend of underlying speech processing skills in typically developing preschool children. They have shown that children as young as 2 years up to 4 years of age are suggested to benefit from existing lexical knowledge over creating new motor programs. The effects were seen regardless of the differences in linguist stimuli presented to children, test construction, administration and scoring methods; for example, whole word accuracy such as Roay and Chiat (2007) or percentage of consonant correct (PCC) such as Sundström et al. (2014).

In contrast, an early study conducted by Williams and Stackhouse (2000) did not show any significant differences in accuracy performance on NWR, RWR tasks groups of children aged 3, 4 and 5-years old, with the performance researched ceiling by 5-years of age. Furthermore, a small scale study on Italian-speaking children (Dispaldro, Leonard, & Deevy, 2013) of preschool children also did not find significant differences between RW and NW tasks (which could have been due to the small sample size and reduced power).

Thus, it is important to explore whether previous findings from cross-linguistic studies will be replicated in an Arabic language context.

It is crucial to the viability of the repetition tasks to establish whether scores on repetition tasks are sensitive to age, and therefore could have the potential of using the tasks as a tool for identifying children whose repetition skills fall below the expected developmental norm. Cross-sectional studies on English-speaking children have shown that repetition accuracy on tasks of RW; and NW's improves with age (Chiat and Roy, 2007; Gthercole & Baddeley, 1990; Lee, Kim and Yim, 2013; Roy and Chiat, 2004; Sundström, Samuelsson and Lyxell, 2014; Vance et al, 2005). For example, Roy and Chiat (2004) and Chiat and Roy (2007) showed that age had a significant effect on repetition where repetition accuracy significantly improved between the ages 2 to 3 years (2;0-2;11 - 3-3;11). In Chiat and Roy (2007) study a clear profile of development was also evident form the age of 2;0< 2;6<3;0. However, form the age of 3;0 there was a lack of age effects on both the RWR and NWR tasks, and no difference between the age bands of 3;0 and 3;6 and 4;0 were found.

The relationship between NWR performance and age has been documented in languages other than English. In the cross-sectional study of Lee, Kim and Yim (2013) typically developing monolingual Korean-speaking children's repetition of NWs significantly improved from 3 to 5-years old (age bands were not clearly specified). Sundström, Samuelsson and Lyxell (2014) found that Swedish-speaking children ages 5-years old performed significantly better than 4-year olds on NWR (5;5 > 4;5).

To our knowledge, no publishes studies compared Arabic-speaking children's performance on processing tasks or documented developmental progress. However, Studies that investigated children's performance on NWR tasks are sparse and typically developing children were part of a small cohort as a control group and compared against children with language difficulties for example (Shaalan, 2010).

2.2.2 *Length Effects on Task Performance*

One of the traditional attributes linked to speech disorders in children, particularly motor speech disorders such as Developmental Verbal Dyspraxia DVD, is that speech errors are more evident as the complexity (i.e. production demands) of a speech task increases (ASHA, 2007). The scenario of increasing task complexity is termed performance load factor (Crary, 1993). Within the psycholinguistic speech processing framework of Stackhouse and Wells (1997), a task with complex stimuli involves the same level of processing as the same task with simple stimuli. However, complex stimuli is placed at a hierarchical level which is above simple stimuli in term of difficulty and is therefore considered to have more processing demands.

Performance load factor has many forms in the literature; one performance load factor that frequently emerges in the literature is increasing length of word/item in number of syllables. Words containing three or more syllables (multisyllabic words) usually contain more and different phonological elements than shorter words (monosyllabic and bisyllabic words). Multisyllabic words contain a higher number of phonemes and levels of stress than words of one or more syllables. They also contain within-word strong/weak syllables and consonant sequences that adjoin syllable borders, consequently, this provides a good insight into speech production skills (James, 2006; James, Van Doorn, & McLeod, 2008).

The unique information yielded by items of increased length is important in clinical practice. Poor productions of multisyllabic items have been tied to impairment, sometimes exclusively, in speech (as well as language and literacy) (Bernhardt & Major, 2005; Dodd, Russell, & Oerlemans, 1993; Lewis, Freebairn, & Taylor, 2000; Pollock, 1991). Studies show that impairment was more evident or more severe when the number of syllables of words increased, particularly with multisyllabic words. Therefore, James (2006) stresses that tests of speech production should not be confined to short monosyllabic or bisyllabic words, otherwise, impairment might be undetected or undermined.

Within the context of typical development, psycholinguistic studies found that increasing processing demands by increasing item length of repetition tasks had significant effects on English-speaking children’s accuracy scores (Shula Chiat & Roy, 2007; Chris Dollaghan & Campbell, 1998; Gathercole et al., 1994; Gray, 2003; Jamie L Metsala & Chisholm, 2010; Roy & Chiat, 2004; Vance et al., 2005; Williams & Stackhouse, 2000).

By using monosyllabic, bisyllabic and trisyllabic items, the studies show that as the number of syllables in an item increased, children's accuracy decreased; these effects were found in repetition tasks of real words, non-words and sequences of sounds.

Psycholinguistic studies by Chiat and Roy (2007), Roy and Chiat (2004) and Vance, et al. (2005) found that length effects were stronger for NWs than for RWs. For RWR tasks, generally, studies did not show consensus on children's performance on different lengths. Roy and Chiat (2004) found that 2 and 3 year old children repeated trisyllabic words as accurately as bisyllabic words and both were poorer at repeating monosyllabic words. However, in their follow-up study on a larger sample of children (Chiat & Roy, 2007), they note that length effects were significant, as accuracy scores were lowest on trisyllabic items and highest on monosyllabic items. Similarly, Vance et al. (2005) showed that item length was a significant factor in repetition accuracy of RWs in children ages 3 and 4 years; where monosyllables were repeated more accurately than bisyllables and both were repeated more accurately than trisyllables. However, by the age of 5, children's performance was at ceiling and no length effects were noted on the RWR task (however, Vance et al. note that 5 year olds still showed lower scores and greater variation for trisyllables compared to shorter lengths).

One the other hand, studies have confirmed that length effects were more evident when using NWs. Children aged 2, 3 and 4 years were least accurate at repeating trisyllabic NWs followed by bisyllabic NWs and were most accurate at monosyllabic NWs (Chiat & Roy, 200; Roy & Chiat, 2004; Vance et al., 2005). Furthermore, Vance et al., found length effects in NWR task was more sensitive to variation in 5 year old children; where the 5 year old children were least accurate on trisyllabic items compared to shorter NWs (no difference were found between mono and bi syllables at this age).

Interestingly, Williams and Stackhouse (2000) found that item length was sensitive to variation in children as young as 3 years old and in children 5 years of age on tasks of RWR, NWR and SSR; although performance was at ceiling at the age of 5, still, the children showed significant differences between item lengths. Williams and Stackhouse's study only included bisyllabic and trisyllabic items on the three phonetically matched tasks.

Languages other than English have also documents the effects of item length in both RW and NW repetition tasks (Dispaldro et al., 2013; Lee et al., 2013; Sundström et al., 2014). Interestingly, the evidence from Italian-speaking children suggests that the effects of length on children's' repetition accuracy could vary depending on the properties of the language. Dispaldro et al.'s 2013 study investigated the effects of length on RW and NW repetition accuracy in 17 Italian-speaking children aged 3;11-5;8; item lengths were bisyllabic, trisyllabic and four-syllabic. They found that children were only significantly more accurate at repeating bisyllabic compared to four-syllabic RWs and NWs. The results could be explained based on the language properties of the Italian, where Italian has a high frequency of trisyllabic and four-syllabic words (34.57% and 31.74%from the main corpus respectively), while trisyllables occur with 14.83% and monosyllables rarely occur with only 0.96% (Mancini & Voghera, 1994); cited in Dispaldro et al., 2013)

It is not clear how Arabic-speaking children would perform on auditory repetition tasks with different item lengths as studies investigating the effects of length on Arabic children's performance are sparse. One study investigated the performance of Qatari Arabic-speaking children on a NWR repetition task (Shaalan, 2010). The typically developing children participating in that study were part of small control cohort to compare against children with language impairment. The study included 11 children with an age range of 5;0-6;9, and 11 children with an age range of 6;3-9;0. The author designed a list of 48 NWs that included both bisyllables and trisyllables and responses were scored as either correct or incorrect. Shaalan (2010) designed the NWs to be language-specific; controlling for phonotactic and morphological rules of Arabic. The study showed that overall bisyllables were repeated more accurately than trisyllables; the percentage of correct repetitions for the younger group (5;0-6;9) was 85.2% on bisyllables and 65.5% on trisyllables, while the older group (6;3-9;0) had higher scores of 92.8% for bisyllables and 76.6% for trisyllables. Although Shaalan did not analyse developmental progression, the difference between the groups is notable. This study suggests that length effects on a NWR task were evident in Arabic-speaking children aged 5 years and older; which is similar to findings from 5 year old English-speaking children (Vance et al., 2005; Williams & Stackhouse, 2000).

However, to our knowledge, no studies have thus far examined the effects of length on RWR, NWR and SSR tasks in preschool Arabic-speaking children and compared performance on monosyllables, bisyllables and trisyllables. This would be particularly interesting for two reasons: first, the Arabic language has its unique phonetic and phonological properties, and second, Arabic-speaking children are influenced by the adult's phonology in terms of acquiring complex syllable structures and the frequency of occurrence of bisyllabic words. (See Chapter 3 on the Arabic language and studies on phonetic and phonological acquisition in Arabic children). Therefore, as with the Italian children, whose developing phonological system was influenced by the rich multisyllabic words of Italian; the Arabic child may gain greater command of longer words over monosyllabic words, which in turn may translate to their performance on NWs.

However, Lee et al. (2013) found that in a sample of 30 Korean-speaking children ages 3;3 to 5;10, bisyllabic NWs were repeated less accurately than trisyllabic NWs. Lee et al. disregard language-specific factors as influential in children’s performance. Rather, they suggest that this is not unusual, as documented by previous researchers on English-speaking 4 to 6 year olds (Gathercole, Willis, Emslie, & Baddeley, 1991) where shorter NWs were repeated less accurately than longer NW’s. Lee et al. suggested that the findings of lower accuracy scores on shorter NWs could be a result of the perceptual saliency of longer NWs, where shorter NWs were harder to perceive accurately. It was suggested that given that it is more difficult for the phonological forms of shorter NWs to be encoded during initial phonological perception shorter non-word stimuli may have more perceptually demanding linguistic features, resulting in reduced accuracy (Alt, 2011; Majerus, Poncelet, Elsen, & Van der Linden, 2006).

In addition, the significant effects of length in repeating items could be due to the memory load required for the task. Indeed, as mentioned in the preceding sections with regard to NWR, that the memory component of NWR has been extensively studied, so much that it is viewed as a measure of phonological memory. The novel string of phonemes must be held in memory long enough to formulate and implement a motor plan. Gathercole and Baddeley (1990, 1995) and Baddeley (2012) assume that immediate recall of words requires the word to be stored in the phonological loop (memory) and maintained by vocal or subvocal rehearsal in real time. Short one-syllabic items could be repeated more quickly, therefore, they are assumed to be easily maintained in the phonological loop. Longer items on the other hand, take longer to rehearse and hence allowing more time for decay, resulting in poor recall and lower score performance.

To summarise, the following length effects were concluded,

- Speech output repetition tasks of RWs, NWs and SSs were sensitive to the increased processing demands of item length in English-speaking children ages 2, 3 and 4 years (monosyllables $>^{\text{greater}}$ bisyllables $>^{\text{greater}}$ trisyllables).
- Speech output repetition tasks are generally more sensitive to increased processing demands in children who are older than 4 years old if trisyllabic items are used.

Children older than 4 years of age may show ceiling accuracy scores with mono and bisyllabic items.

- Furthermore, the effects of length on children's repetition accuracy could be language dependant. As seen in Italian-speaking children, children repeated bisyllabic RWs and NWs as accurately as trisyllabic RWs and NWs.

Table 1.1 summarises studies that investigated speech processing skills in typically developing children using repetition tasks.

Table 2.1: A summary of some studies that compared typically developing preschool children's performance on speech repetition tasks, at a single-word- repetition level.

Study	**Date**	**Participants language**	**Participants** **in (year; month)**
Williams and Stackhouse	2000	English	30 children between 3,4,5; 10 each group
Roy, Chat	2004	English	66 children, age groups 2;0-2;11 & 3;0-3;11
Vance, Stackhouse and Wells	2005	English	100 children age groups 3, 4, 5,6 and 7 years old
Chiat and Roy	2007	English	315 ages 2;0-4;0 2;0-2;6-3;0-3;6 and 4;0
Budd, Hanley and Nozari	2012	English	5, 6, 7, 8 and 11
Lee, Kim and Yim	2013	Korean	60children 3-5 years old
Dispaldro, Leonard and Deevy	2013	Italian	one group of 17 children between the ages of 3;11-5;8
Sundström, Samuelsson and Lyxell	2014	Swiss	44 children, mean age 4 (4) and 5 years olds
Eaton, Newman Ratner and Rowe	2015	English	2 year olds

2.3 Rapid Consecutive Productions: Speech Motor Skills

The theoretical underpinnings and background literature reviewed thus far have focused on speech processing skills at a single-word level (i.e. stimuli repeated once, whether the stimuli were RWs, NWs or SSs). This section addresses the production of stimuli multiple times in succession and at speed, using behavioural measured of accuracy as well as consistency. The present study's inclusion of rapid consecutive (or repeated) productions of different linguistic stimuli (RWs, NWs, SSs), and the use of different behavioural measures, was motivated by a) research studies of Williams and Stackhouse (1998, 2000), that advocated for understanding rapid production tasks form a psycholinguistic perspective; and by b) empirical evidence of the utility of accuracy and consistency of repeated productions in understanding typical and detecting atypical speech behaviour (Dodd, 2005b; Holm, Crosbie, & Dodd, 2007; Macrae & Sosa, 2015; Sosa, 2015; Sosa & Stoel-Gammon, 2006; Sosa & Stoel-Gammon, 2012).

Rapid consecutive productions are most commonly known in the literature and by clinicians as maximum repetition rate; that is, articulatory diadochokinesis or speech diadochokinetic (DDK) tasks. This task is usually part of the assessment category of maximum performance tasks[7]—which includes maximum phonation duration and maximum repetition rate (Rvachew & Brosseau- Lapré, 2012). DDK tasks are one of the most widely used assessments in clinical practice that examine speech motor functions of speed, precision, and coordination of sequenced movements of the articulators of a child with speech difficulty. It is well established that the assessment of speech motor control[8] is an important part of the assessment process of a child's speech difficulty. This assessment determines whether there is a motor component to a child's speech disorder and, if so, the extent to which a motor component contributes to a child's speech difficulty (Gadesmann & Miller, 2008). Many published tests such as the Diagnostic Evaluation of Articulation and Phonology *(*DEAP) (Dodd et al., 2002) and Oral Speech Mechanism Screening Examination (OSMSE-3) (St Louis & Ruscello, 2000) include the DDK task as a subtest to assess speech motor functions. Typically, DDK tasks of speech motor functions used in research and clinical practice are based on the rapid repetition of

[7] Maximum performance task is an activity that examines how an individual performs when exerting as much energy as possible during that activity.

[8] Speech motor control refers to "the systems and strategies that regulate the production of speech, including the planning and preparation of movements and the execution of movement plans to result in muscle contractions and structural displacements" (Kent, 2000)

monosyllables, such as /pə/ (e.g. /pə/, /pə/, /pə/), and sequences of nonsense syllables, such as /pə tə kə /—with a focus on alternating articulatory movements (Fletcher, 1978; Gadesmann & Miller, 2008; Henry, 1990; S. Rvachew & Brosseau-Lapré, 2012). Hence, rapid consecutive productions or speech DDK, has been defined as “to rapidly start and stop the movement of the articulators and to execute repetitive, alternating, sequential movements typically associated with speech articulation” (Johanson, 1980, as cited in (Cohen, Waters, & Hewlett, 1998). A more accurate description of DDK performance is described by Fletcher (1978) as: “The study of motor control integrity in bodily functions through performance in rapidly alternating movements, e.g., pronation and supination of the hand and side to side motions of the tongue. In speech, the term has been extended to include syllable repetition at a maximum rate of utterance” (p. 2).

Thus, DDK tasks usually indicate the speed in which a child can move his/her oral articulators in a task that approximates normal speech, but is not affected by the imponderable phonological complications that affects conversation. This indicates the predominant conception that DDK tasks assess neuromotor components of speech difficulty and rather overlooks other underlying linguistic/processing difficulty (Wilcox, Morris, Speaker, & Catts, 1996) . However, speech production is a highly complex motor behaviour requiring rapid and coordinative control of orofacial articulators, at the same time they interact with cognitive and linguistic skills (Kent, 1992, 2000; Krishnan et al., 2013; Nip, Green, & Marx, 2009; Smith & Goffman, 2004). These speech motor control skills were traditionally distinguished from underlying phonological/linguistic processes, however, experimental research provide evidence of the influence of higher levels of linguistic processing to lower levels of motor implementation.

Williams and Stackhouse (Williams & Stackhouse, 1998, 2000) recognised that DDK stimuli design commonly used syllables and syllable sequence repetition that target the neuromotor function integrity. They addressed the issue by designing DDK tasks in a systemic way so that children’s performance can be understood from the theoretical psycholinguistic speech processing framework of Stackhouse and Well (1997); where different tasks are assumed to tap different demands (See Chapter 1). Williams and Stackhouse (2000) state that “if DDK tasks are to be a valid assessment tool, a stronger developmental perspective is required and careful attention has to be given to the design of tasks, the stimuli used and how scoring procedures are employed” (p.272). Williams and Stackhouse designed their DDK tasks to contain the common repetition task of nonsense two and three syllable sequences (e.g. /pə pə /and /pə tə kə/), furthermore, they included phonetically – consonant - matched real words (e.g. paper and patacake) and nonwords (e.g. /paɪpɪ/ and / pɒtɪkəʊk/) (See stimuli list in Stackhouse et al., 2007). Therefore, with the psycholinguistic framework, repetition of syllable sequences, real and nonwords are clearly differentiated and the different levels of speech processing that give rise to a speech signal can be analysed and compared across tasks. In addition to the task stimuli design, Williams and Stackhouse's study add a further dimension to the investigation of responses on DDK tasks. A fundamental outcome from Williams and Stackhouse's studies was the inclusion of consistency scores of DDK tasks along with scoring accuracy and rate; as accuracy and rate are the measure usually employed with of DDK tasks. William and Stackhouse (2000) emphasis the need to use the supplementary measures of consistency stating that: "Assessments of DDK performance, therefore, need to include consistency measures as well as the more common measures of accuracy and rate" (p.287).

The importance of including consistency scores of rapid consecutive repetitions, as well accuracy, stems from the notion that, depending on the nature of speech difficulty in children, performance on a task could be different with a single repetition versus repeated repetitions (regardless of speed). For example, if a child is asked to repeat a word three times, the child may produce it inaccurately but the same way each time (consistently); while other children would repeat the word inaccurately and inconsistently. Inconsistent speech lacks systemic error patterns and children's inconsistent speech errors are unpredictable, even their families find them unintelligible and often refer them for assessment at an earlier age of three years old, rather than the usual age of four years old (Dodd, 2005b; Holm et al., 2007). Inconsistent speech errors that persist or show highly inconsistent levels and patterns that are characterized by multiple error types i.e. unpredictable variation of the same phone sequence, could reflect a deficit in planning (degraded phonological plan for word production) of the speech processing chain. This results in productions with broad articulatory parameters, and it is hypothesised that children with a deficit in planning may have to create a new phonological plan each time they produce a particular word (Dodd, 2005b). Therefore, the information gathered with repeated productions provides more insight into the underlying nature of speech difficulty than single repetitions. Furthermore, from a clinical perspective, inconsistency could be considered a positive sign of change. Grunwell (1992) studied categories of inconsistency to describe the changes observed in a child's speech system and argues that inconsistency is considered positive when there is evidence that the productions are moving towards the accurate adult target. These findings illustrate the need to investigating typically developing children's task performance on measures of both accuracy and consistency to determine the degree and the age when productions are moving closer to accurate productions.

Many studies have since followed Williams and Stackhouse and included behavioural measures of consistency as well as accuracy in their diadochokinetic tasks such as (Hof, Wijnen, & Dejonckere, 2009; Preston & Edwards, 2009), although on adult participants. Clinically, the Nuffield Centre Dyspraxia Program (Williams & Stephens, 2004) has utilized the Stackhouse and Wells (1997) psycholinguistic assessment framework to measure rapid consecutive repetitions in the diagnosis of DVD. With the different tasks and measures, both the level of speech motor control and linguistic domains could be investigated developmentally.

Further, it was noted earlier that DDK tasks indicate the speed in which a child can move their oral articulators in a task that approximates normal speech. Therefore, most available DDK tasks include the rate by which a child can repeat syllables and sequences of syllables i.e., the speed of repetitions. Speech rate can indicate the presence of neurological impairment, such as dysarthria, and can also be used to evaluate changes over time in both developmental and acquired speech motor disorders (Kent, Kent, & Rosenbek, 1987; Rvachew, Ohberg & Savage, 2006; Thoonen, Maassen, Wit, Gabreëls, & Schreuder, 1996; Wit, Maassen, Gabreels, & Thoonen, 1993; Wren, Roulstone, & Miller, 2012). Moreover, studies have shown that children with speech difficulty can be subgrouped based on their performance on maximum performance speech motor tasks. For example, Thoonen and colleagues validated and cross-validated the ability of maximum performance tasks to differentiate between children with speech difficulty (Thoonen, Maassen, Gabreëls & Schreuder, 1999; Thoonen, Maassen, Wit, Gabreëls and Schreuder, 1996). Their study revealed that the performances of children aged 6 to 10 years of age on general maximum performance speech motor tasks distinguished between children diagnosed with dysarthria, Developmental Verbal Dyspraxia (DVD), speech difficulties (the authors termed this as unspecified speech difficulty) and typical speech development. The children were required to perform two maximum performance tasks that included the maximum prolongation of sounds and a DDK task in which the children repeated repeat as fast as possible monosyllabic sounds /pə pə / and sequences of alternating sounds /pə tə kə/. The study found that children with dysarthria could be differentiated from typically developing children and those with DVD on vowel prolongation and the DDK rate when repeating monosyllabic sequences. Further, children with DVD differed in the prolongation of fricatives and the DDK trisyllabic repetition rate compared to typically developing children. However, despite the success of the procedure in identifying the different groups of children, an overlap was observed in the performance of the three groups. Theoonen et al. (1999) reported that the performance of the non-specific speech disorder group on monosyllabic repetition and maximum phonation duration tasks approached that of the DVD group, and the DVD group approached the dysarthric group in tasks of trisyllabic repetition rates. In some cases, the DVD group had similar levels of performance as the dysarthria group, especially with slow monosyllabic repetition rates. In fact, some children with dysarthria scored higher than typically developing children on the DDK rate task. Further, in Theoonen et al.'s (1996) study, the children with DVD had difficulty sequencing DDK trisyllabic sequences compared to typically developing children. This highlights the difficulty with scoring rate of repetitions when the accuracy of productions are affected. Thoonen et al. requires the accurate production of syllables to score rate responses. However, this is not entirely possible with young children and children with speech difficulties. In fact, the accuracy of the pronunciations of a target sequence is possibly one of the most challenging issues when analysing DDK rate data. In an attempt to address this issue, a study conducted by Yaruss and Logan (2002) instructed young, typical preschool children (mean age 4;7, year; month) to rapidly and accurately repeat the trisyllabic syllable sequences /pə tə kə/ or the words /patticake/. The authors then calculated the percentage of misarticulated consonants in each DDK trial by dividing the number of consonant errors by the total number of correctly attempted consonants. However, this method is not straightforward, and it is unclear how this procedure would be effective in calculating rate. Other studies such as Robinson & Klee (1987) include calculations of the errors and inaccuracies of syllable repetitions in a DDK task, but they did not clarify whether the rates were calculated from all the repetitions or only from the accurate production of syllables. Likewise, standardised tests such as the OSMSE-3 (Louis & Ruscello, 2000) include a DDK task involving the repetition of syllable sequences. The rate of repetitions is calculated using a stopwatch and the accuracy of the repetition is also scored, but the test does not provide instructions on whether to time inaccurate repetitions.

The challenges observed in calculating the rate of repetitions in young children and children with speech difficulties who are bound to produce speech errors could limit the usefulness of measuring rate. For example, Canning and Ross (1974) reported that voicing errors produced by children reduced their speed during a DDK task. With these challenges, some researchers have shifted their focus from the rate of a DDK task to the frequency and type of errors produced as potential valuable measures of children's oral motor development. Researchers such as Yaruss and Logan (2002) and Williams and Stackhouse (1998, 2000) emphasised the importance of using more useful assessment measures such as the accuracy and consistency of DDK performance. In fact, there is evidence questioning the usefulness of rate measures in detecting developmental change in preschool children. Williams and Stackhouse (2000) found that consistency and accuracy measures of DDK were more sensitive measures than the rate of production in detecting developmental change in preschool children aged 3 to 5 years of age. They found that there was no developmental progress in the rate of production in children aged between 3 and 4 years or 4 and 5 years, even with increased item length. Further, Williams and Stackhouse (2000) reported high levels of variability in the rate of production within different age groups, as each child approached the tasks differently. The authors stressed that developmental sensitivity of DDK measures must be determined if they were to be a valid assessment tools. Therefore, the challenges observed when measuring and interpreting the rate of production in preschool children and the research findings on the developmental sensitivity of accuracy and consistency measures in DDK task highlights the importance of using other measures to explore children's performance.

To this end, the present thesis will focus on the accuracy and consistency of productions during DDK tasks. These measures, as reviewed previously, could hold promise in capturing developmental change in typically developing children and identifying speech motor difficulties in children with speech difficulties.

Before going further to review the literature, the theoretical understanding of rapid productions described above will be highlighted.

With reference to the speech processing model of Stackhouse and Wells (1997) (reviewed in Chapter 1 Section 1.3) single repetitions and repeated repetitions of an item multiple times (whether the item was a RW, NW or SS) requires the same route through the speech processing mode (presented in Figure l.2). Rapid productions, however, requires a higher processing load than single repetitions, and is placed above each task in the hierarchy of task difficulty. In The Compendium of Auditory and Speech Tasks (Stackhouse, Vance, Pascoe, et al., 2007) it is suggested that to repeat a familiar word multiple times at speed requires the child to access lexical representations (or possibly bypass the representations) and make use of existing motor programs and trigger that motor program both accurately and consistently at speed. To repeat a non-word multiple times at speed, the child should create a new motor program, and trigger that program repeatedly at speed, sustaining that production. As for repeating sounds and sound sequences, the child is required to accurately and consistently produce syllable sequences at the lowest level of execution.

Furthermore, as mentioned previously, all speech output processing tasks will tax motor planning skills, whether the item was a word, non-word or a syllable sequence and whether items were repeated once or multiple times. The level of planning is where the neuromuscular activity of the output is planned in advance (assembled into a single utterance). This pre-execution level is influenced by many factors including the contextual requirements such as the speed of production. Therefore, rapid consecutive repetitions of an item has a richer contextual requirements, where the processing load on the planning stage increases with the requirement of multiple productions and speed. According to Stackhouse and Wells (1997) motor planning skills can also be tapped though assessments that elicit connected speech such as describing a picture or sentence repetition tasks. In the present study, rapid consecutive repetitions of the same item are thought to, at least partly, reflect the planning stage of processing. At this level of motor planning, the child's stored motor programs of a single word, the new motor program for a single non-word or abstract syllable sequences of sounds thought to occur at the execution level, have to be assembled repeatedly and rapidly into a single plan to produce the string of consecutive productions. With these consecutive repetitions, the ability to trigger the motor program or motor programming levels and plan the correct production and sequence in real time is measured by accuracy and consistency. Additionally, the execution level, where the physical organs of speech give rise to the speech signal, is also a level that is challenged with rapid productions of a linguistic unit.

Motor planning is a potential source that contributes to the accurate production of a novel word. Learning to say a new word is a basic language skill that is linked to motor abilities. Not only does the new encoded speech signal require successful perception, phonological encoding, storage, it also requires transformation of the novel phonological string into a series of coordinated, timed oromotor execution movements in real time. Using non-words stimuli contain novel speech movement sequences allowing for the examination of speech motor processing. The literature has focused on NWR within the scope of underlying language process, lexical access, phonological processing and phonological memory. However, there is less attention on other underlying skills such as phonological assembly and articulation processes i.e. speech motor planning that contribute to successful NWR.

Therefore, rapid speech output tasks of different tasks may provide clues about the underlying nature of a speech deficit in a child, especially with the inclusions of accuracy and consistency measure. Speech difficulty in children could be linked to the planning and/or execution of sequences of motor movement required for speech. These levels of motor planning and execution could be assessed through the repeated repetitions of items. A small number of children with speech difficulty (an estimate of approximately 5%) present with motor speech disorders such as developmental verbal dyspraxia or Dysarthria (RCSLT, 2011) (McNeil, 2003; Ozanne, 1995; Shriberg,1994; Shriberg, Aram, Kwiatkowski, 1997). For example, dysarthria results from a breakdown at the level of motor execution and involves more pervasive effects and abnormalities on muscle tone and weakness. Furthermore, although there is still debate on the underlying causes of CAS, it is generally identified by a number of speech motor characteristics including difficulty sequencing articulatory movement and syllables, and unusual and inconsistent speech errors of both consonants and vowels (ASHA, 2007; Deger and Ziegler 2002, Ogar *et al.* 2006 Ziegler, 2008).

The present thesis will replicate the key methodological procedures used by Williams and Stackhouse (2000), however with a different linguistic community, i.e. on an Arabic speaking population, using the psycholinguistic approach to investigate children's accuracy, consistency performance on a range of speech output repetition tasks. The tasks include real-word repetition, non-word repetition and syllable sequence repetition. The emphasis on the following sections will therefore be on the following: a) the tasks used to assess DDKs, and b) the performance measures of DDKs. In order for the DDK tool to be a clinical valid and applicable tool for Arabic speaking children, a stronger developmental perspective is required. Therefore, the tasks and measures will be reviewed in more detail separately, based on two criteria: on their developmental sensitivity (developmental change), and on validity as a tool to be applied in clinical populations.

Generally, while accuracy is the performance measure of single repetitions, the inclusion of consecutive repetitions allows for other scopes of behavioural measures, such as rapid repeated repetitions for the investigation of planning and execution levels of speech processing. Below is a review of available normative data for English-speaking children.

It is important to clarify when inconsistent productions are clinically significant, given that typically developing children produce inconsistent productions during the early stages of childhood.

2.3.1 *Performance across Tasks and Developmental Change*

In Williams and Stackhouse's (2000) study, single repetition of RWs, NWs and SSs did not capture different profiles of performance in 30 typically developing children ages 3 (3;0-3;11), 4 (4;0-4;11) and 5 (5;0-5;1) years. However, when the children were challenged with higher speech motor processing demands, children ages 3 and 4 years showed different profiles of performance on the tasks and developmental sensitivity was noted. In their study, after children repeated the items once, the children were required to repeat the stimuli items of RWs, NWs and SSs five consecutive times at speed i.e. diadochokinetic tasks. Using an analysis of variance for quantitative analysis, findings from their study suggest that consecutive repetitions of RWs and NWs were no different from consecutive repetitions of SSs. However, children aged 3 and 4 years were significantly less accurate on consecutive productions of NWs compared to RWs. Furthermore, accuracy had become a developmentally sensitive measure for this age group, as 3-year olds were least accurate compared to 4 year olds. In contrast, by the age of 5 years, the discrepancy between the tasks no longer existed and children were able to produce all the tasks consecutively with the same degree of accuracy. Williams and Stackhouse suggested that from a speech processing perspective, typically developing 3 and 4 year old children were able to assemble a new motor program for NWs or access stored motor programs for RWs, but the children found it more challenging to create a new motor program for NWR accurately on five consecutive occasions.

Word variability, *intra-word variability* or *inconsistency* is defined as variability in repeated productions of the same word or token within the same context, for example the English word *cat* as [kæt], [dæt], and [dæ] (example from Macrae, Tyler and Lewis, 2014) . The term variability and inconsistency have been used in the literature to refer to the same observed phenomenon; however some researchers and their teams prefer to differentiate between the two terms (for example, Dodd, 1995; Dodd & Bradford, 2000; Holm, Crosbie and Dodd ,2007; Sosa & Stoel-Gammon, 2006; Sosa, 2015; Stoel-Gammon, 2004;). For example, according to Holm, Crosbie and Dodd (2007) variability in productions is a term used to describe the developmental speech phenomenon observed in young, typically developing children. They state that variability is "attributed to factors described in normal acquisition and use of speech - e.g. phonetic context, pragmatic influence, maturation or cognitive-linguistic influences" (p 468). In contrast, *inconsistency* is a term used to describe the speech of children with speech difficulties, in which a child produces multiple error types with repeated productions, these errors are unpredictable variations that that are both segmental (phonemic) or structural (vowel-consonant sequence within a syllable) (Bradford & Dodd, 1996; Holm *et al.*, 2007). An example from Bradford and Dodd (1996) would be: [drvkm kinv, fokum timv, bwokjum kinv] for vacuum. Word variability or intra-words variability is the focus of the present thesis. For the present thesis, the term consistency and inconsistency will be used throughout the thesis to describe the repeated productions in both typically and atypically developing children.

2.3.2 *Developmental Inconsistency*

Furthermore, as mentioned above, William and Stackhouse's study emphasised the importance of including score of consistency when measuring tasks of diadochokinesis. To our knowledge, no study has detailed children's consistency performance on rapid repetitions, with the exception of Williams and Stackhouse (2000). They found that, first, at the age of 3, consistency scores across the tasks were significantly higher than accuracy scores i.e. children produced an item inaccurately but the same way on each of the five productions, and there were no differences across tasks. From the age of 4, accuracy and consistency scores were in line and no differences were noted; however, at this age NW scores on both accuracy and consistency were lower than RWs. By the age of 5 years, accuracy and consistency were stable across all the tasks with scores close to ceiling, with no differences across tasks. Furthermore, Williams and Stackhouse's study found that developmental progression on consistency scores were evident: children aged 5 years old were highly consistent reaching ceiling (91.6% consistent), followed by 4 year olds (89.7%) and the least consistent were the 3 year-olds (84.5%). Overall, children at the age of 3 years old tend to be more consistent when repeating items on more than one occasion even when inaccurate; from around the age of 4 years old, accuracy and consistency came into line. Nevertheless, children's inconstancies remained minimal even at the age of 3 years old and consistency increased with age. Furthermore, the results suggest that consistency was developmentally sensitive at least between the ages of 3-4 years old.

Findings from Williams and Stackhouse study on rapid consecutive production were in line with empirical evidence that scored accuracy and consistency of multiple productions of items but not consecutively in a speech motor task. Studies have shown that typically developing children exhibit some form of inconsistency when producing words which significantly decrease with age (Burt, Holm, & Dodd, 1999; Holm et al., 2007; Iuzzini & Forrest, 2011; Macrae, 2013; McLeod & Hewett, 2008; Anna V Sosa, 2015; Anna Vogel Sosa & Stoel-Gammon, 2006). Holm et al. (2007), reported inconsistency of 12% for children between the ages of 3;6 and 3;11 and only 13% for children between the ages of 3;0 and 3;5. The younger children were least consistent, and by the age 4.5, children were highly consistency (>95%). Sosa (2015) however, replicated Holm et al's (2007) study and reported higher inconsistency levels in children ages 3-years old, than that reported in Holm et al. Sosa (2015) investigated intra-word inconsistency in typically developing 33 children ages 2; 6 to 3;11 (years; months) with age bands of 2; 6-2;11, 3;0-3;5 and 3;6-3;11. Sosa reported high levels of inconsistency in the youngest age groups. Results revealed that inconsistency decreased with age; however, the older children ages 3;6-3;11 still showed considerable inconsistency with an average of 57% inconsistent productions. Iuzzini and Forrest (2011) also found decreasing word inconsistency with increasing age in children with typical speech development, ages 3;0 to 5;8 (age bands: 3, 4, 5 year olds). Inconsistency reflected the proportion of target words that were produced variably across three or more productions. Inconsistency was highest in 3-year olds, and by the age of 5-years old, children's word consistency was 80 % (while segmental level consistency was 100% consistent).

Common findings can be extracted from studies investigating typical inconsistency despite different procedures (spontaneous vs imitation/naming; rapid consecutive repetitions vs productions within the same context but not necessary consecutive and number of repetitions), scoring methods applied and purposes. First, word inconsistency appears to be prevalent in the speech of young typically developing children, this inconsistency lasts for up to 3 years of age, and by the age of 4 children's productions are highly stable with consistent error patterns. Second, even though children's productions are inaccurate, they still remain more consistent. Therefore, as accuracy of productions increase with age, word consistency also increases, with productions being relatively stable by the age of 4 years old.

The underlying causes of typically developing young children's inconsistency are not fully understood and may be different across children. However, possible explanations for typically young children's inconsistencies could be that 1) the child's underlying phonological representations are "fuzzy" and incomplete; the child is still developing representations, and the lack of sufficient detail causes the child to produce a phoneme or word variably form utterance to another; 2) the child's phonological representations contain the information for a correct articulation of a word, however due to inadequate articulatory ability i.e. output motor execution skills are not fully developed, preventing the child from producing correct surface form consistently (Betz and Stoel-Gammon, 2005). The following sections will discuss the two possible sources of inconsistency.

2.3.2.1 *Underlying Phonological Representations*

Chapter 2 Section 2.1.1.1 reviewed the proposal that that the underlying representations in typically developing children with small vocabularies are holistic in nature (Metsala & Walley, 1998) and may be underspecified (Bernhardt & Stemberger, 1998; Menn & Matthei, 1992) and that young children's representations become more segmental in nature as the child's vocabularies grow. This holistic underlying representation in young typically developing children has been attributed to the observed inconsistency in young children's production of words.

One important finding from the classical study of Ferguson and Farwell (1975) was that as the children in their study (1-year olds) gradually acquire new sounds and increased the number of their phonetic repertoire, they also showed high inconsistencies of the words they produced. A notable example from their study was data from a girl aged 1:4 (years; months) who produced the word "*pen*" in ten different forms during a single 30-mint session. Ferguson *et al*. (1975) proposed that the variation in the production is the child's attempt to organize the target word's segmental properties, which is a reflection of the general knowledge of the articulatory features of the target word (i.e. nasality, bilabial closure, alveolar closure and voiclessness). Other instances of variation in production include phonetic segments inconsistency, for example, a child many have mastered the phoneme /b/ in the initial position of a set of words, however, does not produce the phoneme /b/ initially in other words i.e. shows inconsistency in initial /b/ in other contexts (Ferguson, 1986). Based on the observations of inconsistency in production, they proposed the "whole-word" system of phonological representations in early language development or what Ferguson (1986) calls prosodic variability (inconsistency). This "holist" phonological representations and lack of segmental detail could be one of the sources of high inconsistency in children's productions. The inconsistency in production seems to peak during critical stages of language acquisition (Sosa and Stole-Gammon, 2006). Sosa and Stole-Gammon (2006) studied word inconsistency in a longitudinal study that followed four children between the ages of 1;0 and 2;0 year. The study used Ingham's (2002) proportion of whole world inconsistency measure, and found that word inconsistency peaked during developmental change, specifically during the early stages of lexical acquisition when the children acquired the first 150–200 words and started combining two words. From a phonological developmental perspective, this peak in inconsistency is indicative of the reorganization and emergent systematicity of the phonological representations that suggests the emergence of phonemic representations (Bernhardt & Stemberger, 1998; Sosa and Stole-Gammon, 2006; Vihman, 1996). If lack of segmental detail in children's early phonological representations is hypothesized to be a source of inconsistency in young children's productions then it would be expected that as the child gains more segmental detail to their representations, the less inconsistent they might be (Sosa, 2015). If this is the case, then studying word inconsistency during different stages of development could inform our understanding of underlying phonological representations during those periods. To date, no existing studies have actually investigated inconsistency /consistency in typically developing Arabic-speaking children.

Inconsistency in typical development has been shown to vary as a function of child-specific characteristics, including both age and vocabulary. Studies have found a high predictive relationship between expressive vocabulary and word inconstancy in children up to the age of 4 years old (Macrae, 2013; Macrae & Sosa, 2015; Anna V Sosa & Stoel-Gammon, 2012) were children with smaller vocabularies predicted higher inconsistency. Although the present study does not intend to investigate vocabulary knowledge, these studies point to the importance a child's vocabulary size when measuring consistency.

2.3.2.2 *Speech Motor Development*

Another source that could contribute to inconsistency of repeated productions of linguistic units in children is immaturity of the speech motor control system (Kent, 1992). Speech motor control refers to the system and strategies that regulate the production of speech. Children's speech production skills of their ambient language develops over an extended period of time and seem to significantly lag behind many of the child's associated cognitive and perceptual skills of their language. Infants with normal hearing and with no cognitive or physical impairments begin to produce consonant-vowel syllable or speech-like sounds at around 7 to 10 months; more commonly known as babbling. It is during this period that infants reflect control over the mandible, labial and lingual subsystems for producing speech, although these structures have already been used for several months for non-speech vegetative and vocal activities such as crying. With gradual speech development, children typically master the sounds of their ambient language at around 8 years old, suggesting that children's control over their neuromuscular systems for speech production is markedly limited. Although perceptually acceptable speech sound productions are observed at young ages, the child's control over their speech motor system takes a relatively different time course.

Typically developing children show a dramatic difference in their speech motor control skills compared to adults. Kinematic studies on the development of motor control for speech production have showed that children have slower speech rates and are quantitatively less consistent in their articulatory movement patterns compared to adults (Green, Moor, Hishigawa & Steeve, 2000; Green, Moor & Reilly, 2002; Smith & Goffman, 1998; Smith & Zelaznik, 2004; Walsh & Smith, 2002). There is a gradual transition to faster and more stable articulatory movements with age, and speech movement consistency with repeated productions continue to increase up to adolescence (Smith & Goffman, 1998) and children reach the adult like mature speech motor control after 14-16 years (Smith & Zelaznik, 2004; Walsh & Smith, 2002).

Kent (2006) defined Speech production as "a complex motor act that requires the coordination of respiratory, laryngeal and articulatory subsystems involving over 100 muscles belonging to 5 different structures and functional classes". At the same time, this coordinative control interacts with cognitive and linguistic networks; that is there are two levels of maturation that interact with each other.

Given the complexity of speech production and the many levels required to achieve a speech goal, how do children develop this motor skill? In this section, children's motor control development for speech production will be reviewed in a hierarchy fashion. Given the many subsystems involved in speech production this study will predominately focus on children's development of the articulatory (orofacial) system, and its interaction with intrinsic linguistic processing. Thus, for the purposes of this essay, speech motor control will be discussed exclusively in relation to speech output.

Kinematic work by Goffman and colleagues (Goffman &Smith, 1999; Goffman, Gerken & Lucchesi, 2007) however, suggests that the relationship between motor control movement variability and speech/word accuracy and consistency/inconstancy is not straightforward. High movement variability of articulator exists in young children when compared to adults even when children produced words accurately (Goffman and Smith, 1999).

The use of acoustic or kinematic methods have been used extensively to investigate stability/consistency of speech control. However, few studies of speech motor control have actually used behavioural methods to investigate consistency of repeated productions in children (Williams & Stackhouse (1998, 2000) and Habgood (2000). (in which this current thesis replicates) Williams and Stackhouse (2000) instigated 30 typically developing children ages 3, 4 and 5years old on DDK tasks and measured accuracy as well as consistency of repeated productions. They found that children's consistency of response improved significantly between 3 and 4 years of age. Children aged 5 years olds were highly consistent reaching ceiling (91.6% consistent), followed by 4-year olds (89.7%) and the least consistent were the 3 year-olds (84.5%). The majority of 3-year olds responses were consistent even when inaccurate.

There is considerable overlap between motoric-based and linguistic-based influences that are proposed to be potential sources of intraword inconsistency. It was mentioned previously that word inconsistency peaks during critical stages of vocabulary growth. Theoretically, the peaks in inconsistency in repeated productions supports the dynamic system theory. This theory accounts for real-time changes in motor behaviour in infants (Thelen & Bates, 2003). This theory was extended to account for long-term changes in motor development. A dynamic systems view of development "considers the origins and functions of inconsistency as absolutely central for understanding change" (Thelen & Smith, 1994, p. 145). Variability is associated with transitions between developmental stages and is viewed as a "*potential driving force of development and a potential indicator of ongoing processes*" (van Geert & van Dijk, 2002, p. 341). For example, newborn infants lying on their backs, perform highly coordinated alternating leg kicks. At about 1 month of age, coordination between the legs becomes highly variable. This inconsistency leads to new forms of coordination between the legs, for example, simultaneous kicking (Thelen & Smith, 1994). Thus, peaks in inconsistency of repeated production seen in young children is a sign of maturing system and developmental change. According to van Geert & van Dijk (2002) inconsistency is a characteristic of development of biological and psychological system, including speech-language development.

It was established earlier in this chapter that hearing and repeating novel phonetic sequences is an important human skill that is crucial in learning new words and begins in infancy spanning throughout lifetime and a child's ability to repeat novel phonetic sequences is the cornerstone of language acquisition and involves multistage process (Gathercole, 2006; Stackhouse & Wells, 1997).

Further studies have investigated developmental changes in response of the motor system to complex increasing the demand on the motor system. Smith and Goffman (2004) and Smith (2006) suggest that there is a “top-down” where linguistic goal influence physiological measures and “bottom-up” where the motor system influences language processing. Kinematic studies provide evidence that oral movement features are influenced by the linguist context in which a word is produced, and have provided evidence that there is a close relation between language processing and formulation and execution of motor commands. Sadagopan and Smith (2008) investigated duration and stability of the lower lip for children ages 5 to 12 years and young adults when repeating the sentence “Buy bobby a puppy” alone and when embedded in a longer utterances. The study found that children who are in the path of developing mature speech motor control up to the age of 12, show increased speech movement inconsistency when they repeat a sentence as language processing demands increase and also showed slower rate with the embedded sentence, where young adults variability decrease with age and there speed increase when repeating embedded more complex sentences. The authors suggest that adults and children older than 12 are planning their utterances in chunks and by 12 years there is evidence of using mature motor planning strategies.

The factors that allow for increased rates of speech with age are not fully understood, but have been attributed to gains in biologic factors (e.g., anatomic growth, neurologic and neuromuscular maturation) and learned skills that support rapid spoken language production (motor learning; semantic, lexical, and phonologic access; and motor programming and planning). Investigating how these variables and speaking rate increase with age will have implications for understanding the contributions of biologic and spoken language processing factors on speaking rate development.

As mentioned above, speech motor control, and is dependent on cognitive, linguistic, and motor workloads (Green & Nip, 2010). The processing demands on spoken language production that are imposed by different speaking tasks may vary depending on factors including attention, utterance familiarity (word frequency and phonotactic probability), utterance length, and syntactic complexity. Children speak faster during simple speech tasks, such as the repetitions of simple syllables, than during more demanding speaking tasks, such as conversational speech (Haselager et al., 1991). The relations between speaking rate and task demands suggest that children speak slower than adults, in part, because their articulator movement speeds are slowed by their reduced capacity to formulate spoken language. Therefore, the influence of speaking task demands on children's rate of speech would be supported by the observation that children's speeds of articulator movement are faster for low-demand speaking tasks, but slower for high-demand speaking tasks across development.

The findings by Williams and Stackhouse (2000) support what other research teams have found regarding consistency in preschool children. Overall, evidence thus far suggests that even with more challenging tasks such as DDK, when consecutive rapid productions are required of a token, consistency of repeated productions was high and inconsistency remained minimal in overall responses and decreased with age. Importantly, children's responses were consistent even when inaccurate.

The source of inconsistency in productions in typical and atypical speech development is attributed to both motor-based and linguistic sources. Typically developing children show increased consistency in repeated productions as the motor system matures and as phonological representations become more detailed (segmental). On the other hand children with motor planning difficulties or difficulties establishing or accessing phonological representations may continue to show inconsistency as they get older (Sosa, 2015).

In summary, speech inconsistency is prevalent in the speech of typically developing children and is seen form the age of 2 years old up to the age of 4 years. Therefore, caution should be made when interpreting intra-word inconsistency as indicative of specific subtypes of speech disorder. An important note to consider when interpreting the findings from studies measuring accuracy and consistency is that 3-year old children showed more inconsistency in their performance on processing tasks compared to older groups, rending the reliability of the tasks with these age groups questionable.

Key points from Chapter 2

Summary on Tasks Involving Single Repetitions:

- The literature reviewed has provided evidence suggesting that, in general, children ages 3 to 5 years old show different profiles of performance on auditory repetition tasks. With studies showing that, RWs were repeated more accurately than NWs. Furthermore, children's performance on repetition speech tasks was sensitive to age.

- In contrast, Williams and Stackhouse's (2000) studying 30 typically developing children ages 3, 4 and 5 years old did not show differences in performance across tasks of RWR, NWR and SSR tasks.
- A conclusion to be drawn from the literature is that there is strong evidence to support that NWR accuracy is influenced by both long-term stored lexical knowledge and phonological memory.
- To our knowledge, no published results have been found in Arabic-speaking children comparing their performance on different repetition tasks using a box-and-arrow model of speech processing.

Summary on Task Involving Rapid Multiple Productions - Speech Motor Takes:

- Children with a speech motor difficulty such as Developmental Verbal Apraxia show inconsistent productions and are challenged with motor speech tasks. Furthermore, inconsistency is used as differential diagnostic marker for subtypes of SSD (Dodd, 1995, 2005; Bradford & Dodd, 1996).
- Inconsistency is a characteristic found in typically developing children. This inconsistency in word production peaks between the ages of 1;0-2;0 with the emergence of two-word utterances (Sosa & Stoel-Gammon, 2012) and decrease with age. It is evidence of maturing phonological and motor systems (Burt, Holm, & Dodd, 1999; Holm, Crosbie, & Dodd, 2007; McLeod & Hewett, 2008).

CHAPTER 3 PHONETICS & PHONOLOGY OF ARABIC

This chapter provides an overview of the Arabic phonetic and phonological system. This chapter covers aspects of the Arabic language that are relevant to the current project and is by no means exhaustive. This chapter will provide an overview of the linguistic system of the Arabic language (more precisely, the Saudi Arabic phonetic system) and review relevant studies on phonological acquisition on Arabic-speaking children. The International Phonetic Alphabet (IPA) system will be used throughout this thesis to transcribe Arabic words.

3.1 General Background

Arabic is a language that belongs to the Afro-Asiatic Semitic family, South-Western subgroup (Arabian branch) and is the most widespread of the living Semitic languages. It is one of the world's major languages and it is estimated that about 240 million people speak Arabic as their first language (https://www.ethnologue.com/). Arabic is the official language of more than 20 countries, from Western-Asia (Middle East) to North-Africa (Holes, 1995).

The Arabic language has three different varieties (Amayreh, 2003): Traditional Classical Arabic, Modern Standard Arabic (also called Educated Spoken Arabic) and a large number of regional and social dialects/ colloquial. The varieties of the Arabic language is considered by some authors to be a "a spectrum or better still a continuum which has at one extreme the purest Classical Arabic and at the other, the purest type of colloquial Arabic" p.87 (Bakalla, 1984)

Arabic dialects are spoken on a daily basis while Standard Arabic is confined to formal written and spoken occasions, such as literate activities, educational settings and the media (e.g. the news and most radio and televised programs). Therefore, the Arabic child grows learning their regional Arabic dialect as their mother tongue, and learns Standard Arabic through television (many animated cartoon series use standard Arabic) and later formally at school.

Standard Arabic is unified and is the descendant of Traditional Classical Arabic (Holes, 1995). Regional Arabic dialects are extracts of Standard Arabic and resemble it greatly. Many regional dialects of Saudi Arabia resemble Standard Arabic due its geographic location in the heart of the Arabia Peninsula, however, they do differ from Standard Arabic in terms of phonology, morphology, syntax and lexicon (Ingham, 1994; Versteegh, 1997), for example, /r ɑ ħ/ “he left or gone”, /leʃ/ “why” in Saudi Arabic dialect is produced as [ðahab] and [limaða] in Standard Arabic.

3.2 Consonants of Arabic

Arabic has 28 consonant phonemes; many of the consonants are shared with English. However, Arabic is rich in emphatic consonants which do not occur in English. They include pharyngeal fricative /ħ/ /ʕ/, uvular fricatives /x/ /ʁ/, uvular stop /q/ and pharyngealized consonants, such as / sˤ/, /dˤ/, /tˤ/, and /ðˤ/. The emphatic coronals (/sˤ/, /dˤ/, /tˤ/, and /ðˤ/) cause assimilation of emphasis to adjacent non-emphatic coronal consonants (Mace, 1998; Versteegh, 1997). There are additional consonants that occur in some dialects of Arabic such as Jordanian and Lebanese, they are mainly used in loan words from other languages such as English and French, and are sometimes inconsistent (Dyson & & Amayreh, 2007; G. Khattab, 2007), however, this is not the case with Saudi Arabic dialects. Table 3.1 shows the Standard Arabic connotes (which are also the consonants of the Saudi Arabic dialects).

Table 3.1: IPA-type chart of the Arabic consonant inventory.

Arabic Consonants												
		Labial	**Emphatic**		**Plain**		**Palato-alveolar**	**Palatal**	**Velar**	**Uvular**	**Pharyngeal / Epiglottal**	**Glottal**
			Dental	**Alveolar**	**Alveolar**	**Dental**						
Nasal		m			n							
Stop	**voiceless**		tˤ		t				k	q		ʔ
	voiced	b	dˤ		d				g			
Fricative	**voiceless**	f		sˤ	s	θ	ʃ		x		ħ	h
	voiced		ðˤ		z	ð			ɣ~ʁ		ʕ	
Affricate (voiced)							d͡ʒ					
Approximant		w		l				j				
Trill				r								

In Saudi Arabic dialects, as in many other Arabic dialects, it is socio-phonetically acceptable to replace single speech sounds commonly used in standard Arabic. For example, /q / in Standard Arabic is realised as → [g] in Saudi Arabic dialects, as in the word /qɪrd/ “monkey” → [gɪrd], /qama r/ “moon” → [gamar] and /qalam/ “pen” → [galam]. Other dialectal sound replacements include (/θ/ → [s],[t]), (/ð/→[z],[d]); (/ðˤ/→ [dˤ]) (Amayreh & Dyson, 2000; Ingham, 1994).

3.3 Arabic Vowels

Unlike many languages of the word, the Arabic vowel system is impoverished compared to its consonant system. The number of distinct vowel qualities in Arabic is low compared to its rich consonant inventory. Arabic consists of three main vowels; the high front unrounded vowel /i/, the high back rounded /u/ and the low front unrounded /a/, and their corresponding three shorter vowels / ɪ, ʊ, a/. There are two diphthongs in Arabic /aj/ and /aw/. Vowels alternate depending on the phonetic context and allophones may occur; most commonly /a, aː/ are approximated to [æ] in the environment of most Arabic consonants. Also, /a, aː/ is retracted to low back unrounded [ɑ] when it occurs next to emphatic consonants, for example, /b ɑ: sˤ / 'bus' and /tˤɑˈwiːl/ 'tall' (Mace, 1998; Thelwll et al. 1999; Holes, 2004).

The vowels system of English and other languages differ considerably between dialects; the same is true for vowels in different Arabic varieties. Whereas the three main vowels exist in all Arabic varieties, other vowels can occur, such as in Jordanian Arabic (Dyson & Amayreh, 2007), Lebanese Arabic (Khattab, 2007), Kuwaiti Arabic (Ayyad & Bernhardt, 2009) and Egyptian Arabic (Kotby et al., 2011). To our knowledge, there do not seem to be many references to the Saudi Arabic vowel system compared to Standard Arabic. However, it is known that in addition to the vowels / i, u, aː/, other vowels appear mainly in Saudi Arabic dialects. They include the high-mid front unrounded vowel /e/, the high and low-mid back rounded vowel /o / with its shorter counterparts / ɛ, ɔ / (Ingham, 1994). See Table 3.1 below for a summary of Arabic vowels, particularly Saudi Arabic vowels. It is beyond the scope of this paper to discuss the vowel system in different Arabic varieties, and the reader could refer to the citations mentioned in this section for further information.

3.4 Word and Syllable Structure

Words in Arabic are generally one to three syllables long, however four or more syllable long words could be found in stable loanwords of foreign origin (e.g., / b æ l æ k ɔ n æ/ 'balcony'). Stress is predictable in Arabic and automatic (Holes, 2004), the stress in a word occurs in the penultimate syllable. If the syllable is heavy in a word (contains three elements or double articulation, such as CVVC or CVCCV) that syllable is stressed.

Table 3.2: The Arabic Vowels.

		Front		Back	
		long	short	short	long
close		/iː/			/uː/
close near			/ɪ/	/ʊ/	
Mid	close	/eː/			/oː/
	open		/ɛ/	/ɔ/	
open near		/ æ/			
open		/aː/	/a/		/ɑː/

3.5 The Acquisition of Arabic

3.5.1 *Arabic speech development*

The literature on Arabic-speaking children's phonological development is sparse and to date there are no available published papers on the phonological development of Saudi Arabic-speaking children. However, information on the acquisition of Arabic phonology is available in other Arabic dialects; they cluster mainly around dialects of Jordanian and Palestinian (e.g., Amayreh & Dyson, 1998, Dyson & Amayreh, 2000; Amayreh, 2003), Egyptian Arabic (as in Omar, 1973, Ammar and Morsi, 2006 and Saleh *et al*., 2007) and Kuwaiti Arabic (Ayyad, 2011) and Qatari Arabic (Al-Buainain et al., 2012).

With regards to the acquisition of consonants, Amayreh and Dyson (1998) and Amayreh (2003) investigated the acquisition of Arabic consonants in different word positions in monolingual Jordanian Arabic-speaking children aged 2;0 to 6;4. They investigated the levels of customary production (i.e., at least 50% accuracy in at least two word positions), acquisition age (i.e., 75% accuracy in all word positions) as well as mastery (i.e., 90% accuracy of consonants in all word positions). Amayreh and Dyson grouped Arabic consonant acquisition into three development stages: early (2;0 to 3;10), intermediate (4;0 to 6;4) and late (after the age of 6;4). Generally, the children acquired consonant stops at an early stage than fricatives and front consonants were acquired before back consonants. Furthermore, non-emphatic consonants occurred before emphatic consonants. Interestingly, back consonants /k, ħ/, which occurs frequently in Arabic, were acquired by children at a very early stage of speech development. With regard to development stages, the earlier stage is characterised by the acquisition of 10 consonants, while the intermediate stage is characterised by the acquisition of fricative, affricates and liquids.

Finally, the late stages are suggested to be the stage were the Arabic-speaking children acquire the consonants that were not acquired during the early stages. Table 3.3 summarises the ages of acquisition from Amayreh and Dyson (1998) and Amayreh (2003) studies.

Table 3.3: Consonant Acquisition Developmental Stages in Jordanian Arabic-speaking children (Amayreh & Dyson, 1998; Amayreh, 2003).

	Development Stages		
	Early (2;0 to 3;10)	**Intermediate (4;0 to 6;4)**	**Late (> 6;4)**
Consonants	/ b, t, d, k, f, ħ, m, n, l,w/	/s, h, j, ʁ, x, ʃ, r/	/d͡ʒ, tˤ, dˤ, q, ʔ, sˤ, z, θ, ð, ʕ, ðˤ/

One of the most recent studies on Arabic speaking children's phonological development was that conducted by Ayyad (2011) on 80 Kuwaiti Arabic-speaking children ages 3;8 to 5;2. Children were divided into groups, a young group comprising 43 children ages 3;8 - 4;5 and an older group comprising 4;6-5;2.

For the young age group, the following consonants were acquired at different word positions:

- Acquired by 90% or more of the children: plosives /b, b:, d, t, tˤ k, g, q: /, fricatives /ðˤ, ħ, h, x:/ , affricates /tʃ/, nasals /m, n/, liquid /r/ and approximates /w,j/.
- Acquired by 89-75% of the children: plosives / tˤ, q: /, fricatives /ð, ʃ, s:ˤ, ʁ, x: , ʕ, / and approximates /l/.
- Not acquired at the level of 75% by the children: fricatives /s, sˤ, z, θ/ affricate /d͡ʒ/ and trill /r/.

For the older age group, the following consonants were acquired at different word positions:

- Acquired by 90% of the children in all word positions: plosives / b b:, d, t, tˤ k, q:, g, ʔ/, fricatives /f, ðˤ, ħ, h, x:, ʃ, ʁ /, affricates /tʃ/, nasals /m, n/, liquids /r, l/ and approximates /w,j, j:/.
- Acquired by 75-89% of the children in all word positions: fricatives /θ ð, ʕ/ affricate /tʃ/ and liquid /l/.
- Not acquired at the level of 75% by the children at the age of 4: /s, s:ˤ, r/.

Table 3.4 provides a summary of the consonant inventory of children speaking Arabic (Jordanian, Egyptian and Kuwaiti).

Table 3.4: A Summary of Age of acquisition of the Arabic consonants.

Arabic Dialect	Age of acquisition	Consonant inventory
Jordanian	Younger than 4	Plosives /b, t,d,k/, fricatives /f, ħ/, nasals /m, n/, approximants /l,w/.
	4 -6;4	Plosives /b, t,d,k/, fricatives / f, ħ, s, ʃ, x, ʁ/, trill, tap/flap.
	6;6- 8;4	Plosives /b, t,d,k, tˤ, dˤ, q, ʔ/, fricatives / f, ħ, ʃ, x , sˤ, ð , θ, z, s, ðˤ, ʁ, ʕ/, trill, tap/flap., affricates /d͡ʒ /.
Egyptian	3 -4	Mastery production: plosives /t, k/, fricatives /f, ħ, x , h/, nasals /m, n/ approximates /l, w,j/. customary production: plosives /b, d, tˤ, dˤ, g /, fricatives /ʁ, ʕ, sˤ, z, zˤ s, ʕ/.
	4 -5	Mastery production: plosives /t, k/, fricatives /f, ħ, x , h, ʕ/, nasals /m, n/ approximates /l, w,j/. customary production: plosives /b, d, tˤ, dˤ, g /, fricatives /ʁ, s, sˤ, z, zˤ/
Kuwaiti	3;8-4;5	Plosives /b, b:, t, t:, d, k, q:, ʔ , g, tˤ/, fricatives /f, ħ, h, ʃ, x, sˤ, ðˤ/, affricates /tʃ/, nasals /m, n/, approximates /w, j, j:/ trills.
	4;6-5;2	Plosives / b, b:, t, t:, d, k, q:, ʔ , g, tˤ/ fricatives / f, ħ, h, ʃ, x, , ðˤ/ affricates /tʃ/, nasals /m, n/, approximates /w, j, j:/ trills.

A number of studies have investigated the phonological processes (structural and systemic) used by Arabic-speaking children. Al-awaji (2014) summarised these simplifying processes investigated by different researcher as follow:

“1- Stopping: Dyson and Amayreh (2000), Ayyad (2011) and Al-Buainain et al.(2012) reported a high number of mismatches for the following consonants /s, ð , θ, d͡ʒ/ across word positions.

2- Fronting: Amayreh and Dyson (2000) found the fronting process occurring only for stops, whereas Ayyad (2011) reported frequent occurring of fronting but only for velar and uvular stops, whereas fronting of fricatives occurs only on a very few occasions. Furthermore, Al- Buainain et al. (2012) reported the process of velar and palatal fronting in her data (i.e. /t/→ [k] , /g/ → [d], /ʃ/→ [s])

3- Dentalisation: Dyson and Amayreh (2000) and Al- Buainain et al.(2012) found the following fricatives were frequently affected by dentalisation (/s, sˤ,/ → [θ], /z/→[ð]). The same was reported in Ayyad’s (2011) study with the addition of the affricate.

4- Glottal stop: the glottal stop has been reported frequently in Jordanian studies (e.g. Dyson and Amayreh, 2000) as a replacement for uvular stop; however this is a feature of the dialect. On the other hand, such replacement was uncommon in the Kuwaiti study done by Ayaad (2011). In Al-Buainain’s et al.’s study, the pharyngeal was replaced by glottal stop in children up to age 3; 2 (i.e. /ʕ/→ [ʔ]).

5- De-pharyngealisation (de-emphasis): this simplification indicates the loss of the secondary articulation (e.g. /dˤ/ → [d]) which has been reported frequently in Jordanian children (Dyson and Amayreh, 2000) and persists up to the age of 4;4, whereas Kuwaiti children 35 were using the secondary articulation by the age of four. Al-Buainain et al. (2012) did not report the occurrence of this process in her data.

6- Voicing: Jordanian studies did not report voicing as a part of phonological processes; however, Ammar and Morsi (2006) reported voicing process in Cairine children aged three to five years old. In Kuwait, Ayyad (2011) found devoicing in the younger and the older age groups as well as the at risk group. Al-Buainain et al. (2012) reported devoicing in her data in Qatari children aged 2; 2 to 3; 0" (pp.34).

3.5.2 *Word and Syllable structure*

Studies investigating the acquisition patterns of Arabic phonology and syllable structure (Abdoh, 2011; Ammar, 2002; Ammar, 1999; Salem, 2000); suggest that Arabic-speaking children do not follow the assumed universal order of syllable structure acquisition. The Arabic-speaking child is exposed to a wide range of word shapes, the most commonly occurring shapes CVCVC, CVCCVC and CVVCVC. The most common syllable length was the bisyllables, which are much more common than monosyllables, with the majority of nouns having a bisyllabic or trisyllabic structure (Watson, 2002). Abdoh's (2011) studied early word- shapes of twenty-two Saudi Arabic speaking children (Hijazi dialect of the western region) ages 1;0-1;9. Relevant findings from the study found that the most frequent word shape/length was bisyllables compromising 60.9% of children's productions, followed by monosyllables with 38.2% of productions and then trisyllables at 0.9%. When combining the frequency of word lengths, the following frequency of syllable structure was noted: CVCV at 29.1% > CVC > CVCCV > CV:CV > CV at 10%. Other acquisition studies also suggest that Arabic-speaking children acquire a wide range of complex syllable structures form an early age, and by 3-year of age, 90% Egyptian Arabic-speaking children had acquired most syllable structures (Ammar, 2002). Cluster reduction of complex syllable structures such as CVCC was noted up to the age of 4-years old; and by that age children master the CVCC structure (Ammar,1999; Ammar and Morsi, 2006).

CHAPTER 4 RESEARCH AIMS AND TEST BATTERY DEVELOPMENT

4.1 Introduction

The literature reviewed in the previous chapters introduced the psycholinguistic speech processing model of Stackhouse and Wells (1997), where the general essence of the approach in research and practice is to locate, the child's speech processing strengths and weakness. Tasks are used to illustrate the hypothetical processing levels of the model. Within the model (refer to Chapter 1 Section 1.3) three tasks were identified that are assumed to involve top-down and bottom-up processing. A repetition task involving sounds and sequences of syllables assesses the lower execution level of sound production, it is a bottom-up processing activity, where no prior stored knowledge is involved and it is least influenced by linguistic factors. The repetition of non-words would assess a child's motor programming skills, where it is assumed that it does not require access to stored lexical knowledge; at least to some extent (See literature review, Chapter 2 Section 2.1.2). These tasks are distinguished from a real word repetition task, in which it assumes top-down processing, involving access to stored lexical representations (See literature review, Chapter 2 Section 2.1.1). Therefore, the tasks are placed within a continuum based on the demands made on stored linguistic knowledge. Within this psycholinguistic approach, it was emphasised that interpreting results from one task in isolation may be misleading, and that matched stimuli from different tasks provides more information about a child's underlying processing capabilities Furthermore, Stackhouse and Wells stress the need to investigate how typically developing children progress through the hypothesised levels and investigate developmental progression on the tasks. Thus, the present study aims to assess and compare preschool Arabic-speaking children's underlying speech processing levels assumed to be activated by repetition tasks, and to chart developmental progression on task performance. It is not clear haw Arabic-speaking children will perform on the speech output tasks and how age will influence their performance, it is important to explore whether previous findings from psycholinguistic literature will be held constant in an Arabic context.

When drawing on psycholinguistic research into typical speech development, the literature (Chapter 2) provides evidence of a complex pattern of speech processing performance for different age groups when immediate single-word repetitions are elicited. It was given that:

- Preschool children up to the age of 5 years, show substantial effects of processing demands, where RWs were repeated more accurately than NWs (Chiat & Roy, 2007; Dispaldro, et al., 2011; Eaton, et al., 2015; Hoff, et al., 2008; Roy & Chiat, 2004; Sundström, et al., 2014; Vance et al, 2005). Furthermore, Vance et al, did not find differences in children's performance on repetition tasks. However, findings reported in Williams and Stackhouse's study (2000) were an exception, where their study did not document differences in speech processing output demands of RWR, NWR and SSR tasks.
- Furthermore, the literature provided evidence that generally, performance on speech output tasks improves between the ages of 3 and 4 years, and between 4 and 5 years (Sundström, et al., 2014; Vance et al, 2005; Williams & Stackhouse, 2000).
- Auditory repetition tasks of both RWs and NWs share several skills, which include perception, lexical and phonological knowledge, motor planning and execution. However, NWR task has been viewed as a measure of short-term phonological memory (Baddeley, 2012; Gathercole & Baddeley, 1990) in which reduced accuracy scores of immediate repetition of NWs compared to RWs in young children has been explained within its concept.

Therefore, based on the speech processing model, and on the evidence from psycholinguistic research, it was predicted that if top-down processing influenced performance in RWR then young children whose exposure to language is short-lived and limited would not benefit from stored lexical knowledge. However, children 4 years and older would start to benefit from stored knowledge when repeating RWs over NWs. The repetition of SS would also be privileged, as it requires a simple level of input recognition, planning and execution, with little linguistic demands of stress and syllable boundaries; therefore, accuracy would be predicted to increase compared to the other tasks. Furthermore, it was predicted that NWR would require more memory capacity compared to RWR and SSR, therefore, scores of NWR task would reduce as it requires the most memory capacity, while SSR would require the least memory capacity due its simple structure.

Furthermore, the speech processing model of Stackhouse and Wells (1997) assumes that increasing task difficulty such as increasing number of syllables of items or complexity will involve the same level of processing, however, it will increase processing load. This study, therefore aims to explore Arabic-speaking children's performance on the tasks when the load of the tasks assumed level of processing increases, by presenting items of increasing length: including monosyllabic, bisyllabic and trisyllabic items. Studies found that performance declines systemically with length. This decline is dependent on lexical status RW, NW and SS and it is dependent on age. It was given that:

- Item length of RWs, NWs and SS affect repetition accuracy scores in preschool children ages 3, 4 and 5 (Chiat & Roy, 2007; Dispaldro et al., 2009, 2011, 2012; Gathercole et al, 1995; Metsala & Chisholm, 2010; Roy & Chiat, 2004; Sundström, et al., 2014; Vance et al., 2000; Williams & Stackhouse, 2000). Studies generally found that repetition accuracy was greatest for monosyllabic items and lowest for trisyllabic items and that item length effects were stronger for RW than for NW, where the effects of length on accuracy of repetition was seen in NWs more than RWs in 3 to 5 year old children.
- Within the template-based theoretical approach to representations of phonological knowledge, phonological structure of words are represented as language-specific phonotactic templates. The segmental and phonological pattern of the adult language (ambient language shapes) and the exposure to frequent prosodic structures in a language influences children's template shapes (Vihman & Corft, 2007). The literature on Arabic-speaking children supports this theoretical hypothesis; particularly, early acquisition studies has shown that the properties of the Arabic prosodic structure influence Arabic children's language acquisition, in terms of the acquisition of bisyllables and complex syllable structures.
- The findings from the literature generally suggest that including multisyllabic stimuli items may be more sensitive to individual variation and could enable speech errors to be detected and examined. This is particularly true for NWR with children as old as 5 years
- Within a phonological memory account (Baddeley, 2012) short-items would be maintained and held in memory for a shorter time, and produced more quickly and accurately than longer items, which requires a longer time held in memory and longer time to decay.

It was therefore, predicted that, short monosyllabic items would be produced more accurately than longer items of RW's NWs and SSs, however, it was predicted based on the literature, that 5 year old children would perform at ceiling, and it will be the trisyllabic NWs that would be able to capture variation in performance. However, it was could be that the Arabic-speaking children would be influenced by the properties of the Arabic language. It could be that the complex structure and the dominate bisyllabic structure of the Arabic language would influence children's performance on the different item lengths. On the other hand, it was predicted that longer trisyllabic items will still be less accurate, based on processing demands and memory load. Therefore, the concept of the phonological storage account was explored using stimuli of increasing length including one, two and three syllabic items. Particularly, the short-term memory or "working memory" could account for children's performance on stimuli of increasing length. It was predicted that the longer items would increase memory load and therefore, reduce performance accuracy. Performance would decline systemically with word length and that would require greater recall while.

With reference to the speech processing model, the repetition of an item once will give different processing information than multiple productions. Furthermore, rapid productions will increase the processing load and tap further processing levels. Generally, within the model, repeating an item will involve the same route through the model whether the repetition was a single repetition or multiple repetitions. However, multiple, fast repetitions will increase the demands on planning and execution levels of processing. The literature reviewed previously, it was mentioned that children with speech difficulty could have an underlying speech motor component to their difficulty. Repeating items more than once would therefore reveal underlying difficulty.

In the review of the literature, it was given that,

- Studies of Williams and Stackhouse (1998, 2000) have shown that measures of accuracy (with respect to an adult model) and consistency (with respect to a child's own speech system) did not differ between tasks of RWs, NWs and SSs in 3 year old children; however, different performance on RW and NWs was present in 4 year olds and was no longer evident from the age of 5 years, were scores were at ceiling. Furthermore, accuracy scores were developmentally sensitive between 3,4 and 5 year olds, and the tasks were developmentally sensitive between 3 and 4 years of age on consistency of rapid consecutive productions.
- Inconsistent productions is found at relatively low levels in 3-year old children and consistency would increase with age (Burt et al, 1999; Holm et al, 2007; Sosa,

2015, William et al, 2000). Inconsistent productions is no longer an evident marker in the speech of 4-year old children.

- Young children are suggested to have their individual underlying lexical representations, which are generally, still developing and are proposed to be incomplete and inaccurate in terms of phonological representations or articulatory, motor programs specifications. Therefore, increasing consistency is a sign of maturation.
- Increasing length of syllables affected children's performance, where trisyllables were the least accurate and consistent.

Based on the evidence form the literature, it was predicted that Arabic-speaking children participating in this study would replicate findings from other studies. It was predicted that accuracy and consistency would be greatest for RWs than NWs in children aged 4 years. Furthermore, given that accuracy and consistency are developmentally sensitive markers for children below the age of 5-years, it is predicted that Arabic-speaking children would replicate the findings from English-speaking children. Accuracy would increase between the ages, while consistency would increase between 3 and 4 year olds; from the age of 4 and above it was predicted that children would be highly consistent on the tasks.

To meet the aims, the first step was to devise matching stimuli for the repetition tasks taking into consideration the properties of the Arabic language, which will be introduced below. The following section will summarise the aims and specify the research questions.

4.1.1 *Research Aims and Questions*

The present study is a psycholinguistic speech processing study designed to provide cross-sectional data on preschool Arabic-speaking children's speech output processing skills and describe the profile of performance across different tasks. More precisely, the study adopts the speech processing model proposed by Stackhouse and Wells (1997) as the theoretical cornerstone of the present investigation. The aims of this project are therefore:

1- To investigate typically developing Arabic-speaking children's performance across the designed speech output/motor repetition tasks (real word, non-word and syllable sequence)

2- To chart developmental change on the tasks.

3-To investigate the effects of stimuli length on children's performance on the repetition tasks, and explore developmental changes on tasks with different lengths.

Children's performance on the tasks will be investigated using:

1) Single repetitions as measured by whole word accuracy. The following questions will be addressed on single repetitions of stimuli:

Question 1: How do typically developing Arabic-speaking children perform on the speech output processing tasks (real-words, non-words and syllable sequence) as measured by accuracy of repetition?

a) Is children's performance affected by the test conditions / speech processing demands of tasks (real words vs non-words vs syllable sequences)?

b) Are there age-related (developmental) changes on performance on the tasks (real word, non-word and syllable sequence repetition)? i.e. are the tasks developmentally sensitive?

Question 2: How do children perform on accuracy of repetition when stimuli length increases?

a) Does the stimuli length of the test conditions (monosyllabic vs bisyllablic vs multisyllabic stimuli) affect children's repetition accuracy?

b) Are there age-related (developmental) changes in children's accuracy performance on the tasks when stimuli length increases?

2) Rapid multiple productions as measured by accuracy and consistency. The following questions will be addressed on rapid multiple productions/speech motor task:

Question 3: How do TD Arabic-speaking children perform on the – rapid consecutive repetition/speech motor - tasks (real-words vs non-words vs syllable sequence) using measures of accuracy and consistency?

a) Are accuracy and consistency of responses affected by the test conditions (real word vs non-word vs syllable sequence)?

b) Are there age-related (developmental) changes in children's accuracy, consistency performance on the DDK tasks (real word, non-word and syllable sequence repetition)? i.e., are the tasks developmentally sensitive when accuracy and consistency of repetitions are measured?

Question 4: how do children perform on accuracy and consistency of –rapid productions-tasks when stimuli length increase?

a) Does stimuli length of the test conditions (monosyllabic vs bisyllablic vs multisyllabic stimuli) affect accuracy and consistency of responses?

b) Are there age-related (developmental) changes in children's accuracy and consistency performance when length of stimuli increases? i.e. is stimuli length developmentally sensitive when measured by accuracy and consistency

Objectives

The main purpose for adopting the psycholinguistic approach as the major part of the current investigation was to:

- Enhance our understanding on the underlying speech processes involved in the typical development of Arabic-speaking children.
- This understanding will help establish a baseline for future research on developing assessment tools for children with speech difficulties.

4.2 The Test Battery

The section aims to design speech output processing tasks and develop linguistic stimuli that will be used for the speech output tasks. Before going into further detail about the design of the tasks, stimuli and scoring measures, an outline of the tasks and measures are provided below.

The tasks were as follow:

- Real word repetition task (RWR)
- Non-word repetition task (NWR)
- Syllable sequence repetition task (SSR)

The behaviour measures were as follow:

- Accuracy of one repetition or single repetition (A1)
- Accuracy of rapid consecutive repetitions (Ar)
- Consistency rapid consecutive repetitions (C)

4.2.1 *Psycholinguistic Analysis of Repetition Tasks*

From the psycholinguistic approach reviewed in Chapter 1, Section 1.3, it is clear that assessment tasks require different demands in terms of input, access to stored representations and output. The present thesis is interested in comparing and contrasting the speech output/motor processing routes of repetition tasks. In Chapter 2 the psycholinguistic processing demands of the tasks were analysed in detail based on Stackhouse and Wells's (1997) model. This chapter will therefore focus on providing a brief overview of the repetition tasks and their speech processing demands based on the task requirement i.e. single repetition or multiple rapid repetition.

To have a greater understanding of the processing mechanisms involved in the tasks, it is helpful to refer to Figure 1.1 in Chapter 1, which shows Stackhouse and Wells (1997) speech processing model; the model provides a more explicit look at the assumed levels of processing of each task and their processing routes.

4.2.1.1 *Real Word Repetition (RWR) Task*

Recall from the literature reviewed (Chapter 2 Section 2.1.1) that there are three components to lexical representations: semantic representations, phonological representations, motor programs and the link between them. For a child to produce a familiar word (for example /kitab/ “book”), a child must have stored phonological representations of the word and realizations of its motor program where the child would produce a word compatible with stored phonological representations.

With the RWR task, the child is required to repeat a list of real words after the examiner. Psycholinguistic analysis of the real word repetition task requires the child to recognise the word from input as a familiar Arabic word, accesses an existing phonological representation and motor program stored in their lexical representations to repeat a word once. Forward the motor program to motor planning, where it is then executed and gives the acoustic signal, and repetition accuracy is measured. To repeat a word rapidly and consecutively, the child must not only access and trigger phonological representations and motor programs repeatedly and at speed, but also tap into motor planning and execution skills. The planning level is where words are assembled and productions are influenced by the context requirement such as speed of the spoken word/utterance, while the execution level is where the physical organs must rapidly execute that plan and produce the target acoustic signal. For rapid productions, both accuracy and consistency of productions will be measured

For young children, their phonological representations and motor programs are not fully developed and are imprecise, thus their production of words would be inaccurate. However, with the repetition task, the child repeats a word after the adult model, where the child may update and revise their phonological representations and motor programs based on the adult model, thus increasing their accuracy scores. Furthermore, there are cases when a child is unfamiliar with the word, or the word is one which the child has not mastered. In this case, the child may treat the word as a novel word, and rather than accessing stored motor programs to repeat the word, the child may use input-processing skills to perceive the word, then create a new motor program for output repetition (creating new motor programs will be discussed below). To control for such a confound and ensure that the child has stored lexical representations of a word, the children participating in this study had to complete a picture naming task prior to administering any of the other tasks.

4.2.1.2 *Picture naming*

The picture naming task is a task that requires access to the mental lexicon. This task is critical and should be administered prior to the real-word repetition task to ensured that the child has stored knowledge of the words in memory (the word is stored within their lexical representations) and was familiar with it, and therefore, the word was not treated as a non-word during the real-word repetition task. This prerequisite task is particularly important to the present study, since, to our knowledge, no database exists for Arabic-speaking children on familiarity, frequency or age of acquisition (See Section below for further detail). Therefore, this task allows the researcher to determine whether a word is in fact stored within a child's lexical representations (it does not however, determine the frequency of the word within the child's stored word knowledge).

In this prerequisite task, the child is shown a picture of a target word (for example, a picture of a "book"; /kitab/), the child recognises the picture through the visual system, and the visual stimulus triggers accesses to the semantic representation of the picture (Figure 1). The child's semantic knowledge of the word is linked to stored phonological knowledge of the picture, and thus the semantic representations should activate the child's stored phonological structure of the word. However, according to Stackhouse and Wells (1997), access to phonological representation is not necessary required for recognition of the picture, and thus the motor program, which specifies the stored gestures required for producing the name of picture, can be accessed directly from semantic representations, leading to automated production. This automatic access to stored motor programs is usually thought to occur in older children in which a word is highly frequent.

4.2.1.3 *Non-word Repetition (NWR) Task*

The NWR task addresses whether a child could accurately repeat a speech signal without reference to stored lexical representations. The use of non-words targets a child's online speech processing system as opposed to stored knowledge that has been built up over time. In this task, the child creates a new motor program i.e. motor programming without relying on a pre-existing program (Figure 1.1).

To repeat a non-word accurately, the child hears the non-word through peripheral auditory processing, then discriminates whether the heard signal was a speech or non-speech signal. The non-word although unknown to the child, should confirm to the phonological patterns of (in this case) Arabic (for example /kutib/), and therefore the level of phonological recognition is activated. The child then searches the lexicon and scans stored phonological representations. On finding no entry to the stimulus (example /kutib/), the child has to activate motor programing and select new combinations of phonemes and structures.

After the new motor program has been created, the next stage of processing is the motor planning stage, where the gestural targets have to be assembled correctly in real-time and then forwarded for motor execution. As with the real-word repetition, this stage will show the repetition of the non-word once, while motor planning for the non-word is better illustrated with the rapid consecutive repetitions.

In summary, this task requires the child to repeat a non-word after the examiner. Psycholinguistic analysis of this task is different from the real word repetition task. The child is unfamiliar with the word; it does not exist within their lexical representation as the child has not heard or spoken them before. The aim of the task is to assemble a new motor program, plan and execute the signal to repeat the non-word once, and then with the same process, need motor planning skills to repeat the non-word rapidly and consecutively.

4.2.1.4 *Syllable Sequence Repetition (SSR) Task*

This is a non-lexical task where the child repeats a sound or a sequence of alternating sounds, for example /k æ / or /k æ t æ /. Psycholinguistic analysis of this task requires the child to access his/her motor planning skills to assemble the syllables .The task also assesses the child's articulatory skills at the level of motor execution. Although this is a speech task where the child must hear and recognise the signal as consisting of speech sounds, Stackhouse and Wells (1997) place the syllable repetition task on the lower end of output as they have equal stress patterns and vowels that do not comply with either real-word or non-words, they are therefore illegal. The task, taps motor plans and execution to repeat the syllables once, and as with RWR and NWR, the ability to repeat the syllables rapidly and consecutively taps planning skills.
SSR contain non-meaningful syllables that are more likely to be less contaminated with linguistic factors such as RW or NW

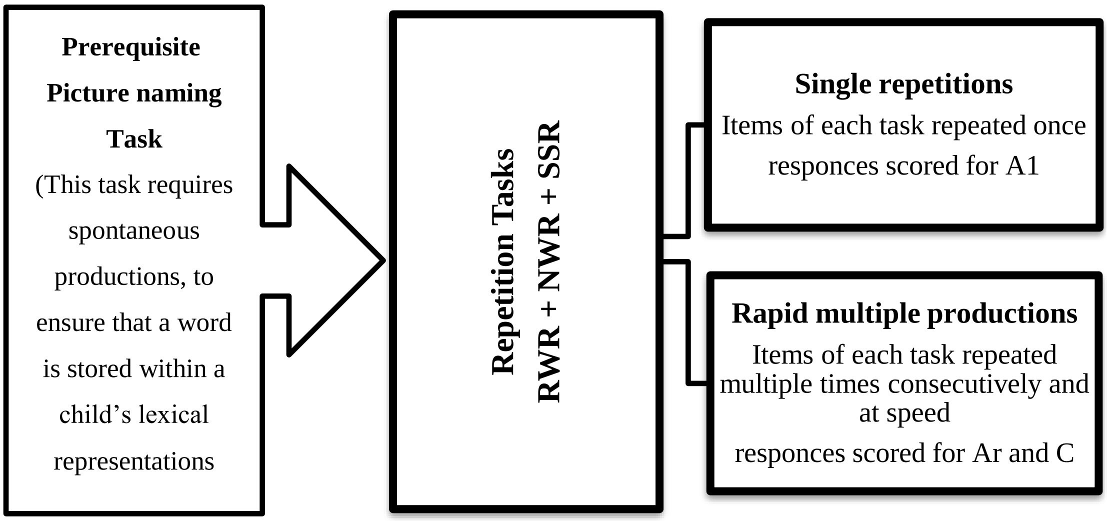

Figure 4.1 Overview of the test battery. This includes speech output tasks and behavioural measures.

Note: RWR= real word repetition, NWR= non-word repetition, SSR= syllable sequence repetition. A1= whole word accuracy of single repetition. Ar= whole word accuracy of multiple repetitions; C = consistency of consecutive repetitions.

4.2.2 *Task Stimuli*

In the previous section, tasks were designed and described based on the psycholinguistic framework of Stackhouse and Wells (1997). This section will summarise how stimuli were developed and matched across the tasks; as in order for comparisons to be made across processing demands, the tasks must involve matched stimuli, presentation, and scoring.

When devising the lists of linguistic stimuli, certain criteria were taken into account; these included developmental factors and linguistic factors (lexical and phonological factors) (Pascoe, Stackhouse and Wells, 2006; Stackhouse & Wells, 1997).

4.2.2.1 *Real words*

4.2.2.1.1 ***Developmental and Lexical Factors***

One task required of children in this study was the repetition of real words; where a linguistic input activates a child's access to the motor program of the word in stored lexical representations. Therefore, the selected word list should be within the receptive and expressive vocabulary of children as young as 3-years old (one of the age groups targeted in this thesis). Furthermore, the stimuli words should be able to be presented by a picture i.e. visually transparent. As mentioned previously, a picture naming task should precede the real word repetition task to ensure that the target words are actually stored in the child's representations. Thus, the target selection of words were nouns and unambiguous objects, so that a word is produced by the child spontaneously when shown a picture. Furthermore, it was necessary to select words/pictures that did not have common lexical replacements in the Saudi Arabic language, for example a /xæ ruf/ "goat/lamb" could be known to a very young child as [ʁæ næ m æ] or [tajs]; a /ba tˤ.tˤ.anjæ / "blanket" commonly known as [lɛħæf] or [ʁa tˤa].

To our knowledge, appropriate measures of vocabulary knowledge and age of acquisition for young Arabic-speaking children is lacking. Available resources for familiarity and word frequency are based on adult usage (for example, Buckwalter, 2011), which are not considered appropriate measures for young children. Therefore, to develop the Arabic stimuli suitable for children with limited vocabularies, two approaches were followed. The first approach was to look at cross-sectional and/or longitudinal Arabic acquisition studies that investigate different aspects of children's speech development and draw early words that were produced by the children. Studies such as Abdoh (2011) and Khattab &

Al-Tammi (2013) investigated different aspects of early developing words from Arabic-speaking children. The studies analysed spontaneous speech samples from recordings of children within the ages of 1;0-2;0 years. From the data provided by the studies, early words produced by the young children were selected (regardless of whether they were produced accurately), which included /b æ b/ 'door'; /d ʊ b/ 'bear'; /ʃ æ ms/ 'sun'; /s æm ækə/ 'fish'. The drawn words were nouns and could also be illustrated pictorially.

Second, popular nursery books that are widely available in bookstores, and books used in nurseries and preschools were studied. Additionally, the picture-naming Arabic articulation test for children aged 3;6-5;10 years was used for word selection (Abou-Elsaad, Baz and El-Banna, 2009). The test includes 58 pictures that elicit spontaneous productions of single-words with Arabic consonants represented in the initial, medial and final word positions. The test is originally developed for Jordanian Arabic-speaking children, therefore dialectal variation on items were present, examples include the item / ʔ æ l æ m/ 'pen' which is produced in the Saudi Arabian dialect [ɡ æ l æ m], /b æl æ ħ/ 'date' is produced [t æ mr]. In addition, sociolinguistic variations were noted in the word list, for example the words /jæs.min/ 'jasmine' and /mæ.hæl.læ.bej.jæ/ 'pudding' are words more commonly used in the Levantine area of the Arabic world. Therefore, only words that were suitable for the Saudi Arabic-speaking child were selected, and words such as 'pen' and 'date' were selected and adjusted to the linguistic system of Saudi Arabic.

4.2.2.1.2 ***Phonological/phonetic factors***

The present project is not an articulation test study and does not intend to investigate the phonetic inventories or phonological error patters of Arabic-speaking children. This study is a speech processing study which aimed at investigating different levels of output/motor processing using repetition tasks. Both single repetition and Diadochokinetic (DDK) motor productions are central to this thesis, therefore, the selection of stimuli included words with different articulatory places, where children would repeat items in a series alternating sounds. The sounds within words represent alternating movements from front to back and vice versa, for example, the English word 'buttercup' used for spoken DDK tasks (Williams & Stackhouse, 2000) has three places of articulation, bilabial, alveolar and velar, thus, moving from front to back sounds /b→ t→ k/. The Arabic stimuli words selected had to involve alternating articulatory places, for example, back to front in the bisyllabic item /k iˈ t æ b/ and front to back in trisyllabic item /b u r t u q æ l/. Manner and voicing of items were not controlled for in this study, nor were the position of liquid consonants.

4.2.2.1.3 *Word Frequency and word density*

Lexical factors of word frequency and neighbourhood density influence production consistency and accuracy (Garlock, Walley, & Metsala, 2001; Sosa & Stoel-Gammon, 2012) (refer to Chapter 2 Section 2.1.2). Ideally, the words selected should be controlled for frequency and neighbourhood density, however, this was challenging when considering word frequency in Arabic, as research in this area seems to be lacking, and it is difficult to place the words in hierarchy for frequency of use. Nevertheless, general assumptions could be made; in general, it is expected that words used in the daily activities of the child are the most frequent and common, thus words such as “door, book, pen, cup, cake, button, dates” could be more frequent than “grape, orange” and less frequently used could be words such as “elephant, fish”. The four-syllabic word ‘balcony’ would function as a very least common word. Generally, it is challenging to pinpoint what is frequent or not in a child’s mental lexicon, as individual differences in word frequency occur. For example, a child wearing glasses would use the word ‘eyeglasses’ with high frequency compared to peers. Word frequency and neighbourhood density could be confounding variables to the present study. However, as mentioned in the literature review frequency of word type has been investigated in the Arabic language with interesting findings.

4.2.2.1.4 *Word length, structure and complexity*

According to studies on English-speaking children increasing stimuli length affects accuracy of responses and captures developmental change, with monosyllables being the most accurate and tri-syllables being the least accurate (See Chapter 2). However, this does not necessary apply to Arabic-speaking children, as the properties of the Arabic language structure are different from English (refer to Chapter 3). The Arabic children’s early words are influenced by the adult phonology in terms of predominance of bisyllables and the acquisition of complex syllable structures. The inclusion of items of different lengths would therefore shed light on the effects of the child’s ambient language on speech processing skills.

The number of syllables were taken into account when selecting stimuli words. The stimuli words were divided into three sections, representing single syllabic items, two syllabic items and multisyllabic items. Ideally, it is best to maintain the structure and stress pattern across test items, however, this was a challenging task, therefore the stress pattern of the stimuli words selected were not similar; some words had equal syllable stress and other words were stressed syllables before the last syllable. Furthermore, the structure of the syllables varied to different degrees within each item length.

Overall, the preliminary word list included 11 monosyllables, 13 bisyllables and 11 multisyllables (See **Appendix 1** for the full preliminary list of words and their syllable structure). A few examples are presented in the Table 4.1 below:

Table 4.1: Examples from the preliminary selection of words.

Word length	Meaning in English	Word structure	Arabic words in IPA
Monosyllables	Door	CVC	b æ b
	Dog	CVCC	k ɛ lb
Bisyllables	Book	CVCVC	k iˈ t æ: b
	Lightbulb	CVC.CV	ˈl æ m.b æ
Multisyllables	Orange	CVC.CVCVC	b u r. t u ˈq æ: l
	Apple	CVC:VCV	t u ˈf .fæ ħ æ

Note: C= consonant, V= vowel, ˈ = stressed syllable, C: = (CC) geminate structure.

The stimuli words selected do not include initial consonant cluster (as this is mainly not present in the Arabic language). The monosyllabic words were the only list of items with a final consonant cluster, they included: (/ʃ æ ms/ ‘sun’, /k æ lb/ ‘dog’, /t æ mr/ ‘date’, /g i rd/ ‘monkey’).

Furthermore, there was one word within the list which is commonly assimilated to the following sound ‘regressive assimilation’; the word /læ mb æ:/ is usually pronounced [l æ nb æ:] in the adult model. Finally, five of the trisyllabic words had a geminate structure (/qu:ˈb.bæʕæ/ ‘hat’, /næˈðˤ.ðˤɑræ/ ‘eyeglasses’, /θ æˈl.læ dʒæ/ ‘fridge’,/t uˈf.fæ ħæ/ ‘apple’, / mæ xæˈd.da / ‘pillow’. There was also one multisyllabic word which is actually a loan-word from English / b æ l æ k ɔ n æ/ ‘balcony’.

An advantage of the selected word list is that the words are nouns and most of them are also produced in Standard Arabic (the literate classical Arabic used for formal reading and writing); where the words have the same form in Saudi Arabic dialect and Standard Arabic. However, within the list of multisyllabic words, there was one exception; the foursyllabic word /bæ læ kɔ næ/ ‘balcony’ could be produced in Standard Arabic as [ʃ u rf æ]; in either case, both words are used in the Saudi Arabic dialect.

4.2.2.2 *Creating non-words*

To tap into the level of motor programming of the speech processing chain, non-word stimuli were created. The non-words developed had to be phonologically legal in Arabic, and most importantly, matched phonetically to the real words to allow for comparisons between the levels of processing demands. Thus, to avoid developing any illegal non-words in Arabic, and to create phonetically matched stimuli, the non-words were derived from the real words.

For the present study, the following criteria were used (Pascoe, Stackhouse and Wells, 2006; Stackhouse and Wells, 1997):

- Non-words stimuli were matched phonetically to the real words in structure e.g. CVC or CVCC; and the stress pattern of the non-words were matched to its counterpart real word.
- All the consonant phonemes were maintained and occurred in the same position in both real words and non-words. Only the vowels of the non-words were changed. Matching stimuli on consonants also allows for controlling for developmental speech production errors confound; so if a child fails to accurately produce a real word due to developmental speech errors, they will also fail to produce its counterpart non-word, and not due to difficulties with processing the non-word.
- When changing the vowels for the non-words, the general theme rule was that Saudi Arabic front vowels were substituted for other Saudi Arabic close, mid and open front vowels (/i/ and /e/ to [æ] and /æ / to /[i]). The back vowels were substituted for other high and low back vowel (/u/ to [ɔ] and /o/ to [ɔ] (See Table 4.2 and refer to Chapter 3 Section 3.3 for an overview of the Saudi Arabic vowel system). Vowel length was also maintained when changing to other vowels.

When creating the non-words, the goal was to achieve a systemic vowel change throughout the list of words using the general theme rule listed above. However, it was challenging to follow a regular rule with some of the word, as in many cases changing the vowel of a word either a) created an English word, b) created a phonetically or acoustically similar Arabic word used in a different Arabic accent variety, or b) created another Arabic word. Therefore, in these cases, vowels were changed as appropriate while maintaining vowels length.

Table 4.2 provides examples of the non-words and the process for creating them (See Appendix 2 for the full list).

Table 4.2: Examples of the preliminary non-word list and how they were developed.

Word length	Arabic words	Vowel alternation	The non-word
Monosyllables	b æ b	Near-open front long vowel /æ:/ →to front high long vowel [i:]; [bib], this non-word could be an English word, or a sound describing a car horning, thus the long back vowel was used [o]	b o b
	k ɛ lb	Mid-open mid-front short vowel /ɛ/→ near close mid-front short vowel /ɪ /	k ɪ lb
Bisyllables	k i ˈ t æ: b	When changing the front close vowel /i/ to near-open front vowel [æ:] and the /æ:/ in the second syllable to [i] a real-word was created “writer” thus the vowel /i/ in the first syllable was changed to a back close long vowel [u].	k u ˈ t i: b

Furthermore, the stimuli items of the non-words were controlled for lexicality effects of syllables within the bi and tri items. There were only three non-words that contained a final syllable that was a word: two bisyllabic non-words, one contained the final syllable /tˤi r/ “go fly” and the other /m i l/ “mile”, and from the trisyllabic non-words there was the final syllable /qi:l/ “was said” (a word used in Standard Arabic). Nevertheless, these words are usually used by adults and may not be in the lexicon of young children.

4.2.2.3 *Creating the syllable sequences*

The syllable sequences stimuli were developed to measure the lower execution level of speech processing. To match the stimuli of the syllable sequences task to the real world and non-word tasks, the criteria used when creating the syllable stimuli was similar to the non-words in terms of matching consonants.

The following criteria were used:

- The stimuli were created by maintaining the consonants of the real word and only changing the vowel. However, unlike the stimulus of the non-words, the syllable sequences stimuli did not include different vowels within an item; only one vowel was used and maintained throughout an item, as the task intended to target the lower level of motor execution.

- The syllables included simple structures, therefore, requiring less demand on speech processing. All the monosyllables contained a CV structure, bisyllables contained a CVCV structure and trisyllables contained a CVCVCV structure (with the exception of the four-syllabic item /bæ læ kæ næ/, with a CVCVCVCV). Furthermore, there were no stress patterns in any of the items, therefore, reducing linguistic contamination.

Table 4.3 provides a few examples of syllable sequence stimuli, and how they were derived from real words in terms of consonants.

Table 4.3: Examples from the preliminary items of syllables and syllable sequences.

Word length	Arabic words	Syllables	Structure
Monosyllables	b æ b	b æ	CV
	k ɛ lb	k æ	
Bisyllables	g i ˈtˤa: r	g æ tˤæ	CV CV
	k i' t æ: b	kæ tæ	
Multisyllables	b u r. t u ˈq æ: l	bæ tæ qæ	CV CV CV
	t u ˈf .fæ ħ æ	tæ fæ ħæ	

Summary of the Preliminary Test Items:

In total, there were 34 pictures (for the prerequisite picture naming task), a total of 34 items for the real word repetition task, 34 items for the non-word repetition task and 34 items for the syllable sequence repetition task. Items of each task included 10 monosyllabic items, 13 bisyllabic items and 11 multisyllabic items.

See Appendix 3 for the full list of preliminary stimuli for the repetition tasks

4.3 Pilot Study: Modifications to Stimuli

After the development of the tasks and stimuli items, the tasks and the bank of items were piloted on an opportunistic small sample of typically developing Saudi Arabic-speaking children living in Sheffield, UK. The pilot study therefore aimed to:

1- Test children's responses to the selected pictures.

2- Test the stimuli items designed for the non-word and syllable sequence repetition tasks.

The pilot study will not include any statistical analysis, scoring procedures or task performance comparisons. This section only aims to present children's responses to the stimuli and the necessary modifications applied to them.

4.3.1 *Participants*

Efforts were made to recruit children from the ages of 3-5 years old for the pilot study; however, that proved to be challenging. E-mail letters were sent through the Saudi-club to their list of members inviting children from the ages of 3-5 to participate in the pilot. The children whose parents responded and agreed for their children to participate in the pilot were bilingual with English as their functional language. This is because many of the Saudi families involved in the Saudi-club were students studying at different universities, their children usually attended local nurseries, and therefore were exposed extensively to English for long periods of time. As a result, the children were unable to name many of the pictures in Arabic; rather they would be able to name certain items in Arabic while unable to name other items except in English. For example, one child who participated in the pilot was 3;5 (years, months) and attended full time nursery school in Sheffield. During the picture naming task, he was unable to name many of the pictures in Arabic, even when verbal cues where provided. Thus, children within the age range of 3-5 whose parents agreed for them to participate in the pilot proved to be inappropriate for the needs of the pilot study. Therefore, older children were recruited, as they were bilingual and fluent in both Arabic and English languages, and would be able to provide feedback on the test items.

Ten Saudi Arabic children (N=10) were recruited to the pilot study from a weekend Arabic primary school in Sheffield. There were 5 girls and 5 boys between the ages of 6; 00 to 7; 00; mean age = 6; 4; and S.D. = 3.4. The children were fluent in both Arabic and

English, with Arabic as their mother tongue. The children attend a mainstream primary English school in Sheffield during the week, and during the weekend they attend an Arabic primary school. The children were in the first year of the Arabic primary school, and they were at the end of the first semester. One advantage of this sample is that it includes children from different regions of Saudi Arabia, including Riyadh, the capital in the central province of Saudi Arabia, the eastern, western and southern regions; the participant's demographics are provided in Table 4.4. Hence, when children were asked to name the pictures or repeat the stimuli non-words or syllable sequences, the children provided responses that reflected different upbringing and regional backgrounds, therefore responses to the stimuli were enriched with children’s general experiences and regional dialects.

Table 4.4: Participant demographics for the pilot study

Participant ID	Gender	Age (years; months)	City
1- CH	F	7;00	(Eastern Region) Alhassa
2- JN	F	6;3	(Western Region) Jeddah
3- RAN	F	6;4	(Western Region) Taif
4- LA	F	6;2	(Central Region) Riyadh
5- HSS	M	6;4	(Western Region) Taif
6- AHM	M	6;3	(South Region) Abha
7- JD	M	6;1	(Western Region) Jeddah
8- YQ	F	6;00	(Western Region) Jeddah
9- NW	M	6;7	(Central Region) Riyadh
10- ZE	M	6;5	(South Region) Abha
	Mean	6;4	
	S.D.	3.4	

Since the school does not include children with special needs, the children included in the study were known to be typically developing within a mainstream school. The class teacher reported that the children's language skills were typical compared to their peers and reported that the children had no known history of literacy difficulties, visual or hearing impairment or known speech, language communication difficulty or attention deficits. Although the class teacher reported that the children did not have significant speech difficulties, four of the boys recruited to the study had varying speech sound errors. The pilot study was approved by the University of Sheffield’s Ethics Review Procedure (See Appendix 4). The head teacher returned a signed consent form to recruit children and test them at their weekend school, and parents of the children returned a signed consent form prior to testing.

4.3.2 *Procedure*

Each child was tested individually in a quiet room in the school during school hours, and parents were invited to attend if they wished. Each child was seen twice, each session lasted for approximately 15-20 minutes. The tests were carried out by the researcher, who is a trained Speech-Language Therapist. The order of presentation of the repetition tasks was randomised, and also, the order of presentation for different syllable lengths was randomised. Children's responses were audio-recorded using a MERANTZ, model number PMD670, with a microphone EM-8. The audio recordings were then used to transcribe and score the accuracy of responses.

4.3.3 *Tasks and Stimuli*

The following tasks were administered:

- Picture naming task (34 pictures)
- Repetition tasks of non-words + syllable sequences. The preliminary stimuli of non-words and syllable sequences listed in Appendix 3 were presented.

4.3.4 *Performance on the Test Items*

4.3.4.1 *Picture Naming Task*

There were 34 pictures and an additional 4 pictures as practice items. All children who participated in the pilot study were able to name the pictures representing monosyllabic words spontaneously and without prompting. On the other hand, children responded differently to naming four pictures representing items of bisyllabic and multisyllabic words. For the four items the children either produced a lexical replacement to name the picture or the children needed more prompting and verbal cues to name the target word of the picture, Table 4.5 provides a summary of children's responses to the four pictures. Form the table provided, it could be noted that children used lexical replacements for three words, where six of the children lexically replaced /sˤur sˤæ:r/ 'cockroach', two children replaced /fil fil/ 'pepper', one child replaced /ħæb ħæb/ 'watermelon'. Lexical replacements were due to dialectal variations and/or individual variations i.e. different words used in different households, and were not due to difficulties identifying the coloured pictures. Children were able to produce the target word instead of the lexical replacement when prompted, for example, two children used /ʃɛdˤ.dˤa/ as a lexical replacement to the target word /fil fil/, nevertheless, they were able to produce the target word with verbal prompting such as ‘what else do you call it?’, the other child was able to name the item after providing a CV initial phonetic cue. However, there was one child from the eastern province of Saudi Arabia, participant (1-CH), who named the picture of a ‘watermelon’ to /d͡ʒʊħ/ as a lexical replacement to the target word /ħæb ħæb/ and was not able to produce a different lexicon for the picture even with prompting, remarking that she is not familiar with word / ħæb ħæb/ and has never used it before. This is also a common lexical replacement in other regions of Saudi Arabia, such as the central area. Lexical replacements are very common between Arabic dialects in general, and since the children included in the pilot study were from different geographical areas of Saudi Arabia, including the Central, western and eastern provinces, lexical replacements due to dialect differences were noted and flagged.

Table 4.5: Summary of children’s productions of five target words in the picture naming task

Meaning in English	**Arabic word in IPA**	**number of children**	**children’s responses to the picture**	**Type of response**
Cockroach	/sˤur sˤ æ:r/	six children	[sˤɑ rur], [sˤɑr sˤu:r/ or named other insects	Lexical “dialectal” replacements, named other insects and needed prompting to identify the target word
Pepper	/fil fil/	two children	[ħ ɛ b. ħr] [ʃ̠ ɛdˤ.dˤ a]	Lexical replacements
Watermelon	/ħ æb ħ æb/	two children	[d͡ʒ̠ ʊ ħ] [bɑdˤix]	Lexical replacements
Balcony	/bæ l æ k ɔ n æ/	five children	[tˤɑ g æ] “window” or prompting for a response	difficulty naming the picture

Furthermore, five children had difficulty naming the word ‘balcony’ /b æ l æ k ɔ n æ/ even with prompting and verbal/phonetic cues, this could be due to the words relatively low frequency, and not all children are exposed or are familiar to a balcony.

The children in this pilot study had a mean age of 6;3 years old; however, the main study will include a younger age group, and the word list was designed to be tested on children as young as 3 and 4 years old. This means that younger children are likely to have lower vocabulary sizes compared to the older children in this pilot study. Children named some pictures differently based on their dialect, and needed prompting to produce the target word. Younger children with smaller vocabulary sizes and less linguistic experience may not be familiar with lexical replacements; this is a major confounding variable as one of the aims of the picture naming task was to confirm that the word is stored within the child's lexical representation. As a result, the five words were excluded from the list of words. This resulted in a list of words that included 10 monosyllabic words, 10 bisyllabic words and 10 trisyllabic words.

4.3.4.2 *Repetition of Non-words*

During testing, it was evident that eight of the non-words either resembled an accent variation in Arabic or the non-word was phonetically close to the real word. The items are therefore, less likely to be perceived as a non-word, and amendments to their vowels was necessary, Table 4.6 shows the non-words before and after amendments.
The first list of items shows that there were three items in which the participants and the researcher found that their phonetic realizations resembled the real word and its pronunciation was in close proximity to the real word, they include, /d ɔ b/, /m ʊ z/ and /g i l i m/. The second list of items included /kilb/, /g i m ir /, /l i mbi /, /mixi di/ and /kinibi/; the items were in close phonetic proximity to the Levantine Arabic accent. The phonetic proximity of the items to real-words and accent variations was not evident until testing the children and repetitions of the items were required. This resulted in changing the vowels of the non-words.

Table 4.6: Non-words used for the pilot study and the amendments made.

Phonetic realization close to real-word		**Close to an Arabic accent variation**	
before change	after change	before change	after change
d ɔ b	d æ b	k ɪ l b	k u l b
m ʊ z	m ɛ z	g i m ir	g u m ur
g i l i m	g u l u m	l i mb i	l u mb u
		mi xi di	mu xu du
		ki ni b i	ku nu bu

4.3.5 *Revision of the Test Battery Stimuli*

The amendments to the items list are provided in Table 4.7. Four words and their matched non-words and syllable sequence were excluded from the test list. The vowels of eight non-words were changed. Appendix 5 and 6 shows the final list of stimuli and pictures used. They include 30 real words, 30 non-words and 30 syllable sequence and 4 practice items.

Table 4.7 The initial stimuli items and the items after amendments, RW: real word, NWR: non-word, EXL: item excluded from list.

meaning of the word in English	Arabic word in IPA	Task	Final Change IPA	Change due to
Cockroach	sˤur sˤ æ:r	RW	EXL	dialect Lexical replacement
Pepper	fil fil	RW	EXL	dialect Lexical replacement
Watermelon	ħ æb ħ æb	RW	EXL	dialect Lexical replacement
Balcony	b æ l æ k ɔ n æ	RW	EXL	less familiar with older children
	k ɪ l b	NW	k u l b	realization similar to an accent
	g i m ir	NW	g u m ur	realization similar to an accent
	l i mb i	NW	l u mb u	realization similar to an accent
	mi xi di	NW	mu xu du	realization similar to an accent
	ki ni b i	NW	ku nu b u	realization similar to an accent
	d ɔ b	NW	m ɛ z	phonetically close to the real word - high word-likeness
	m ʊ z	NW	d æ b	phonetically close to the real word –high word-likeness
	g i l i m	NW	g u l u m	phonetically close to the real word- high word-likeness

The following chapters will present the main cross-sectional study on typically developing Arabic speaking children.

CROSS-SECTIONAL STUDY - SPEECH OUTPUT PROCESSING IN TYPICALLY DEVELOPING CHILDREN

CHAPTER 5 PARTICIPANTS, METHODS & PROCEDURE

The previous chapter focused on designing speech output/motor tasks and their stimuli. This chapter focuses on the aim of the present study, where the effects of processing demands, age and stimuli length are investigated in typically developing Saudi Arabic-speaking children.

This chapter includes the design of the present study, the participants, methods and data preparation for analysis.

5.1 Study Design

To address the research questions a cross-sectional study design was adopted. There were three age groups of children 3, 4 and 5 year olds. The study investigated typically developing Arabic-speaking children's performance on the three speech output/motor processing tasks as measured by accuracy and consistency. Within each age group, children's performance on the tasks was first compared; and then performance on the tasks were compared between the groups. Speech processing skills was therefore investigated form a psycholinguistic and developmental perspective.

5.2 Participants

Data was collected and analysed from 129 children between the ages of 3;0 and 5;10 years old. Children were divided into three age groups comprising group 1: 3-year olds, group 2: 4-year olds and group 3: 5-year olds. Participants information are summarised in Table 5.1.

Table 5.1: Participants information, including the mean age of each group (years; months) and standard deviations (s.d)

Age groups	Mean (s.d)	Minimum	Maximum	N(total: 129)	Gender
3 year olds	3;6 (2.96)	3;00	3;11	29	13 boys, 16 girls
4 year olds	4;5 (3.06)	4;00	4;11	50	22 boys, 28 girls
5 year olds	5;5 (2.63)	5;00	5;10	50	26 boys, 24 girls

Participants were recruited from five mainstream kindergarten/nursery schools in Riyadh city, Saudi Arabia. The five schools approached by the researcher were located at different neighbourhoods throughout the city of Riyadh; this was done to minimise the possible effects of socioeconomic status and educational biases; although it should be noted that studies showed that neighbourhood socioeconomic status did not appear to affect children's performance on repetition tasks, particularly non-word repetition tasks (Balladares, Marshall, & Griffiths, 2016; S. Chiat & Polišenská, 2016; Shula Chiat & Roy, 2007; Ellis Weismer et al., 2000; Roy & Chiat, 2004). A formal letter of permission from the Saudi Ministry of Education and signed consent forms from the head teachers of the five schools to recruit children were provided prior to data collection. Participant questionnaires, information sheets and consent forms were distributed to the children's parents/caregivers via the class teachers and returned to them (see Appendix 7 for an example of the questionnaire; forms were translated to Arabic). The class teachers were instructed to distribute the forms to parents whose children have no record of speech or language difficulties, no developmental difficulties and no hearing difficulties; this was an important initial step as many of the kindergartens in Saudi Arabia are inclusive schooling systems, where children with hearing aids, Down syndrome are included within the mainstream system. Children within the age band of 5;0-5;11 (years; month) were found in Kindergarten classes-3 (KG-3)[9], children within the age band of 4;0-4;11 (years; month) were found in KG-2 classes, and children within the age band of 3;0-3;11 (years; month) were found in KG-1. Originally, 140 preschool children between the ages of 3;0 and 5;11 (years; months) returned signed consents of them taking part in the study. However, the number reduced to 129 after the inclusion criteria.

5.2.1 *Inclusion criteria*

The children had to meet certain selection criteria before they were included in the study, the criteria were as follows:

- Children were Saudi Arabic and Arabic had to be the main spoken language at home; the child was accepted as a participant if they had English as a second language. The minimal criteria was set, as it should be noted, that kindergarten children's educational system in Saudi Arabia includes English classes; thus, all the children in the targeted schools had at least formal English lessons during school hours, which includes the alphabet, stories and simple nursery rhymes.
- No history of hearing loss, neurological or developmental disorders such as autism, ADHD.

[9] KG: kindergarten. KG-3 is the final year of kindergarten before the first year of primary school. Children are usually around 6 years old when they enter 1st grad primary school.

- No history of speech, language and/or communication difficulties.

To ensure that the criteria were met, the parents/caregivers of children completed a questionnaire (see Appendix.7). The questionnaire gathered information regarding the children's speech/language background and development, health, physical development, vision and hearing. Furthermore, an input test and a language test were administered. The input test was a picture card Aural Re/Habilitation input kit developed by JISH –Jeddah Institute for Speech and Hearing- for children with hearing loss (hearing aid users/cochlear implants). The kit is used in some clinics to assess input skills informally. For example, the kit includes identification of voiced/voiceless contrasts, emphatic vs non-emphatic sounds etc. Discrimination skills are also assessed informally using the picture cards. Furthermore, the preschool-language scale PLS-3 (Zimmerman, Steiner,& Pond, 1992) was administered to assess receptive and expressive language skills. However, it should be emphasised that the norms provided in the test were developed for English-speaking children, and are not standardized for Arabic-speaking children. Nonetheless, due to the lack of language tests in Arabic, clinicians usually use this test to informally assess an Arabic-speaking child's language skills. Therefore, to establish a baseline of children's general language skills for the present study and avoid undetected language difficulties, instructions from the PLS-3 were translated and delivered in Arabic during testing. As a general guideline, the children were only excluded if they obviously lagged behind on the test norms; more than 2-3 standard deviation below the mean.

Five children were excluded from the study at an early stage of the project. Upon returning the questionnaire, two children aged 4 and 5-years old, were excluded from the study, as they had moderate developmental delays (growth and general development) that had not been reported at the consent distribution stage. In addition, one 5 year old child and one 4 year old, were excluded from the study after they were observed to have a language difficulty during an informal conversation and performed poorly on the language screen (PLS-3). In addition, one 5-year old child was excluded from the study, as during two sessions of the screening stages, the child was not attentive throughout the sessions; the class teacher later reported that she was concerned regarding the child's achievement and his lag behind his peers.

Furthermore, six children aged 3-years old were excluded from the study for varying reasons. Two children refused to leave the classroom and their teacher. One child refused to cooperate during testing and elected not to participate, even with the use of different reinforcements and activities. One child did not show for the entire semester although his parents returned a signed consent form. Finally, three of the young children were reported to have speech/language difficulties, as their parents/carers had concerns regarding their child's intelligibility and language development and will seek future help form a speech/language therapist.

Ethical approval for the study was granted by the Ethics Committee of the Department of Human Communications Sciences at the University of Sheffield (see Appendix 8.).

5.3 Material

An overview of the tasks administered are provided in Figure 5.1. The final amended stimuli list, pictures and three practice items were administered (see Appendix 5 and 6).

Picture naming task	• 30 pictures, representing 10 monosyllables, 10 bisyllables and 10 trisyllables
Practice items	• 3 pictures (1 monosyllabic, 2 bisyllabic, 1 trisyllabic) • 4 real-word repetition 3 non-word repetition and 3 syllable sequences
Real Word Repetition (RWR)	• 10 Monosyllabic RWs • 10 Bisyllabic RWs • 10 Trisyllabic RWs
Non-Word Repetition (NWR)	• 10 Monosyllabic NWs • 10 Bisyllablic NWs • 10 Trisyllablic NWs
Syllable Sequence Repetition (SSR)	• 10 Monosyllabic SSs • 10 Bisyllabic SSs • 10 Trisyllablic SSs

Figure 5.1: Overview of the test battery

Table 5.2 includes examples of different word lengths, its pictures and their non-words and syllable sequences, (See Appendix 5 for the full list of stimuli of the repetition tasks).

Table 5.2: Examples of stimuli with different lengths

Picture Naming Task	Repetition Task		
	RW-Real word	**NW-Non words**	**SS-Syllables**
	/k ɛ lb/	/k u l b/	/k æ/
	/k i' t æ: b/	/k u 't i b/	/k æ t æ/
	/bur.tu ˈqɑ l/	/b æ r.t æ ˈq i: l/	/b æ t æ q æ/

5.4 Procedure

Each child was tested individually in a quiet room in the school during school hours in the presence of a teacher or an observation room. Each child participated in three to five sessions, which lasted approximately 15-30 minutes. Some younger children ages 3-years old, required shorter sessions that did not exceed 10-15 minutes, as they became inattentive and distracted. The tests were carried out by the researcher, who is a trained speech and language therapist. The order of presentation of the repetition tasks was randomised. In addition, the order of presentation for the different syllable lengths was randomized to avoid fatigue effects. Each child's responses were audio-recorded using an Olympus DS-40 Digital recorder. For this main study, each child's audio recordings were entered into the *Praat* computer software (Boersma &Weenink, 2014). This was particularly useful for the analyses and transcription of fast consecutive productions, as it allowed the researcher to hear each production separately and transcribe each responses with more precision. (See scoring responses below).

5.4.1 *Tasks*

5.4.1.1 *Picture naming task and practice items*

Prior to administering the repetition tasks, the child had to complete a picture naming task. The examiner presented the pictures to the child and was asked to name them, thus spontaneous naming were required of the child without modelling when shown the picture. The examiner only gave descriptive verbal prompt when the child did not know the word. Prompting is defined here as a verbal cue, or question to elicit naming a picture which results in spontaneous production. For example, to prompt / ʕ i g æ l / "men's head wear" the tester could say "daddy wears /ʃ i m æ ʁ/and … " or "what does daddy wear on his head" (pointing to head)?" If the child still did not know the word after prompting, the tester produced the initial consonant, if the child still did not know the picture, the initial CV structure would be next then the second if the child still did not know the word. If the child still did not know the word after prompting, the tester named the picture and asked the child to repeat after her, then checked later to ensure that the child could name the picture without cues.

After completing the picture-naming task, three practice items were presented, they included three pictures, three words and non-words and syllable sequences (representing mono, bi and trisyallbic items). The practice items were presented to ensure the clarity of the instructions and general procedures during practice the child was instructed to repeat the item five times as fast as possible, however without interfering with the intelligibility of the production. If the child stops before the completing the five repetitions, the child was instructed to continue further. However, if the child still failed to produce an item five consecutive times, the repetitions were accepted and no further feedback was required. In addition, feedback was provided if the child produced the items slowly or too fast as to it interfering with intelligibility. Feedback was provided on the practice items, however no further help was provided for the main test items.

5.4.1.2 *Real word repetition*

For the real word repetition task, the examiner named a word without reference to a picture and asked the child to repeat after her. The child's response was scored for the baseline. After the child's imitation of the target word, the child was asked to repeat the word five consecutive times as fast as they could. All responses were audio recorded for transcription and scoring.

5.4.1.3 *Non-word repetition*

In the non-word repetition task the examiner explained to the child that the words he/she will hear are “silly” words and not real words, the child was also instructed to repeat the non-word exactly as they hear them. The non-word repetition task followed the same procedure as the real word task, where the child repeats the word after the tester once and then five times at speed. The examiner produced the syllable stress patterns of the non-word without modifications. All responses were audio recorded for transcription and scoring.

5.4.1.4 *Syllable sequence repetition*

The syllable repetition task also followed the same procedure of the real word and non-word repetition tasks, where the child repeats after the tester once, and then repeats the target five times at speed. As explained above, the child’s responses on all of the tasks were audio recorded. All responses were audio recorded for transcription and scoring.

5.4.2 *Scoring Responses*

Each child repeated each item once after the examiner; this was the child’s first attempt at repetition and will be treated as the child’s baseline single repetition. Then the child repeated each item rapidly five consecutive times. Therefore, each task involved two measurements of accuracy and a measurement for consistency. As mentioned previously, due to the nature of the rapid repetition task, children’s audio recordings were entered into Praat software in the present study to allow precise and accurate transcription of individual productions, especially on the rapid consecutive takes. Figure 5.2 is a spectrogram representation of a 5 year old child’s rapid consecutive production. of the item /bab/ “door”.

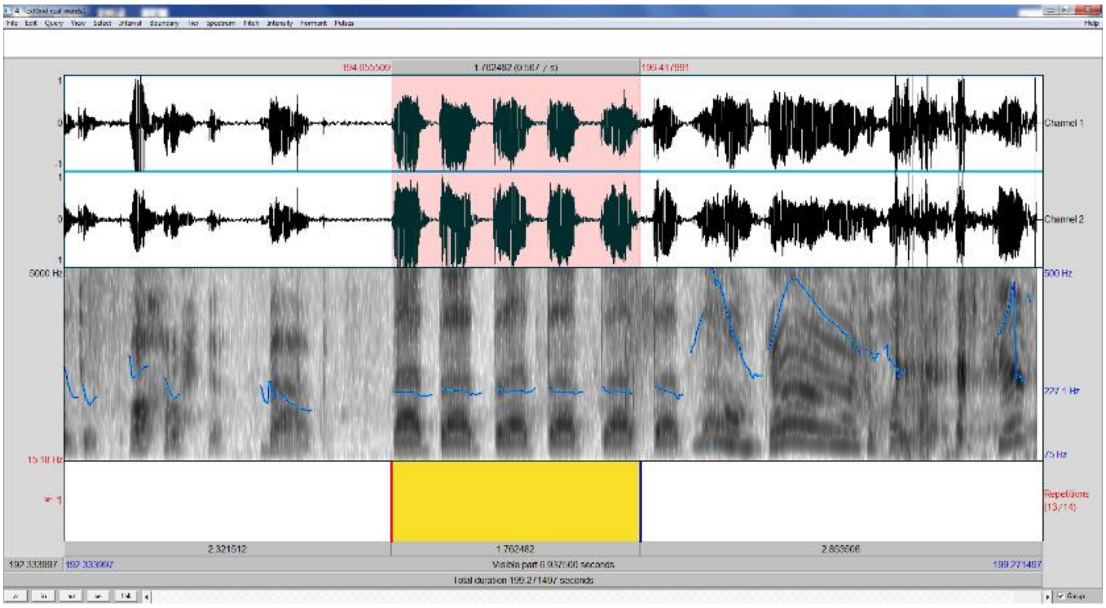

Figure 5.2: Image captured from Praat software showing a spectrogram and waveform of a child's consecutive production of the monosyllabic word /bab/.

5.4.2.1 *Transcription*

The author transcribed children's responses using headphones. The child's single repetition and each of the consecutive repetitions were transcribed by the author using IPA broad phonemic transcription.

Diacritics were used during transcription to better reflect the child's actual productions, such as nasalization's, retractions etc. However, diacritic differences did not constitute to inaccuracy nor did they constitute to inconsistency if they occurred anywhere during repetition. For example, if a child produced an aspirated alveolar plosive /tʰ/ on the item [tʰ æmr] as the baseline repetition, however, produced an unaspirated alveolar plosive [t ⁼æmr] on one or more of the consecutive repetitions, the child would still be considered as accurate and consistent. Another example of diacritic differences includes voicing versus devoicing: if a child for example devoiced the velar /g / on the baseline repetition or on any of the consecutive repetitions [g̊ æ m æ ɾ], it was scored as accurate and consistent. This also applies to other diacritic differences including for example, dentalizations, weak versus strong articulations, no audible release of a plosive consonant versus audible release (See Table 5.2 for examples). However, in instances were changes to phonemic realizations occur, then accuracy and consistency scores will be affected. For example, if a child devoices the voiced velar /g/ completely and this resulted in the voiceless velar /k/, or for example, if a child did not pharyngealize the voiceless alveolar stop / tˤ / as in the word /giˈtˤa: r/ “train” resulting in an alveolar stop /t/, then scoring responses will be affected.

Table 5.3: The examples included in this table are all considered accurate and consistent productions.

	Target	Repetitions				
		1	2	3	4	5
ʰ Aspiration vs ⁼ unaspirated alveolar plosive	/ t ʰ æ mr /	[t ʰ æmr]	[t ⁼ æmr]	[t ⁼ æmr]	[t ʰ æmr]	[t ʰ æmr]
Dentalization	/[dʊb/ or /[ʃæms/	[dʊb] [ʃæms]	[dʊb] [ʃæms]	[dʊb] [ʃæms]	[d̪ʊb] [ʃæm s̪]	[d̪ʊb] [ʃæm s̪]

Furthermore, changes in any of the following phonemes place/manner of articulation while maintaining voicing did not affect accuracy or consistency scores, the phonemes were:

- Voiced postalveolar / ʃ / to voiced retroflex / ʂ / and vis versa.
- Voiced velar /x / to voiced uvular / χ /, or voiceless velar / ɣ / to voiceless uvular / ʁ /
- Post alveolar voiced palato-alveolar affricate /d͡ʒ/ to voiced palato-alveolar sibilant/ʒ/
- Voiceless palato-alveolar sibilant /ʃ/Voiceless palato-alveolar affricate /t͡ʃ/
- Alveolar trill / rˤ~r / to tap / ɾ / or approximant /ɹ /
- Emphatics: pharyngealized voiced dental fricative /ðˤ / to pharyngealized voiced alveolar stop /dˤ]/.

During transcription, the discrepancies between the phonemes were very miner, and these minor changes in place do not yield any change of meaning in the Arabic language, therefore they did not affect accuracy or consistency scores. Similarly, studies such Sosa (2015) and Vance et al. (2005), did not account for minor diacritic changes when scoring children’s responses.

5.4.2.2 *Behavioural Measures and Scoring*

It was noted previously that the present thesis focuses on measures of accuracy and consistency, while the measurement of rate in a DDK task was excluded due to key factors informing this decision (See Chapter 2 Section 2.3). However, it is useful to note some of the issues when measuring DDK rate in preschool children—aged 3, 4 and 5 years old. One of the primary concerns when asking a child to repeat items as fast as possible, is the challenges they encounter during the task, which in turn affect their rate of repetition. For example, Canning & Rose (1974) reported that children younger than 5 years old had difficulty producing trisyllabic syllable sequences in a DDK task; as the task was considered to be too abstract for the young child. Canning & Rose also reported that speech errors such as voicing errors produced by children when repeating items slightly reduced their speed on the DDK task. Further, Wit, Maassen, Gabreëls and Thoonen (1993) stated that "many young children have difficulties with the unnaturalness of maximum performance tests" (p 453). Therefore, the issue of speed of productions makes it reasonable to question the usefulness of measuring rate in young children and whether it would provide useful information.

Accuracy of production means the correct adult –like realization of speech sounds. There are many available single-word standardized articulation/phonology tests for English speaking children (for example, DEAP and HAAP-3) The goal of the single-word tests is to compare the child's performance to normative data of the child's same age, and assign severity level of a child's speech difficulty (Rvachew & Brosseau-Lapré, 2012). Many articulation tests and other productive phonology tests use two-way scoring (correct or incorrect) and most tests score the consonants of the target word. The two-way scoring is a word based analysis, which scores the whole word as phonetically correct or incorrect compared to the adult model (whole word correct). This whole word correct has been used in the literature to measure real words and non-words elicited by repetition, word naming tasks or spontaneous speech. Particularly, this scoring method has been used in psycholinguistic assessments to capture differences in children's performance on different tasks (Newbold, Stackhouse, & Wells, 2013). For example, Vance, et al. 2005 compared typically developing 3 to 7 year old children's performance on picture naming, real word and non-word repetition tasks using the word correct measure to calculate the accuracy of productions, the results showed significant differences in children's performance on the three tasks and the pattern and improvement in the task performance significantly changed as children got older. Within a psycholinguistic framework, typically developing children as well as children with speech difficulty may be able to repeat non-words more accurately than real words, this occurs due to immature or insufficient speech processing skills, leading to inaccurate motor programmers. If a child has accurate existing motor programmers however is unable to create a new motor programe (motor programming deficit) then the child will be able to repeat real words more accurately than non-words. In Vance, et al.'s (2005) study, children at the age of 3 showed no difference in accuracy between real word and non-word repetition which could be due to using a bottom-up approach in processing speech without the use of lexical representations. 4-year old child performed better in real word repetition than non-word repetition. The 4 year old children used their existing lexical representations with an advantage of imitating an adult model and had the opportunity to update their immature motor programming

Williams and Stackhouse (2000) investigated accuracy of children's productions on different DDK tasks in a cross-sectional study; tasks included real word, non-word and syllable repetitions. Accuracy of whole word was scored (correct or incorrect responses). Accuracy was scored for the first repetition and scored for five consecutive repetitions. Results on five repetitions showed that children ages 3 and 4-years scored better in real word repetition and the lowest scores were on the non-word repetition, the 5-year old children performed equally across tasks. The results aid the investigation done by Vance et al. where accuracy of different tasks is a sensitive measure in detecting developmental change. The Nuffield Centre Dyspraxia Program (Williams & Stephens, 2004) is an assessment test that also uses a psycholinguistic approach to measure children's DDK and diagnose CAS; accuracy of production is scored on children's performance on different DDK tasks (real words non-word and syllable repetitions) the test scores responses as either correct or incorrect and compares children performance across tasks. Other standardized tests of speech motor performance; such as Apraxia Profile (AP; Hickman, 1997), the Oral Speech Mechanism Screening Examination, Third Edition (OSMSE–3; St.Louis&Ruscello, 2000), and the Verbal Motor Production Assessment for Children (VMPAC; Hayden & Square, 1999); measure accuracy of different aspects of the child's performance such as phonetic accuracy and syllable sequencing, however scoring varies across tests, and measuring accuracy of DDK performance is not straightforward. For example, OSMSE measures accuracy of syllable repetitions and scores them as correct or incorrect, however it is not clear whether accuracy is scored on every repetition and then a proportion is calculated or repetition accuracy is scored as whole.

With regard to consistency measures, although many studies refer to inconsistency of productions, no standardized measures across studies exist. Some studies that have developed a measure of consistency in their studies include (Dodd, 1995; Sheriberg et al, 1997; Ingram, 2002; Tyler, Lewis & Welch, 2003; Betz & Stoel-Gammon, 2005).

In Dodd's (1995) study; Dodd developed a measure of whole-word inconsistency for her study in order to classify children with speech disorders into subgroups, there were 25 words, where three productions of each word was elicited. The production of a word was labelled as variable (inconsistent) if at least two of the three productions were different. If a child showed variable productions of at least ten out of the 25 words; the child was classified as having an inconsistent speech disorder. Furthermore, The standardised test of articulation and phonology (DEAP) developed by Dodd and colleagues (Dodd et al., 2002) includes an inconsistency assessment for children ages 3;0 to 6;11. Children are asked to name pictures of 25 words, the child names each three times, each trail separated by an activity. The child scores zero if all three productions are the same and scores 1 if one production is different. The percentage is then calculated. If inconsistency was 40% or more then recalculations are made by extracting any production of words that are developmentally appropriate, thus the final score would include inconsistency as a disorder.

Betz and Stoel-Gammon (2005) described three formulas to calculate the different types of error consistency; the formulas provide a framework for researchers interested in investigating consistency as a predictor of change over time or in studies that aim to differentiate types of phonological disorder. The formulas calculate the number of errors at a whole word level and not at the phoneme level. If a child's production of a target word contained one phoneme error, the whole word was considered erroneous, and individual phoneme errors made on the same word are not calculated.

Consistency measures from Dodd's (1995) study and DEAP are not a DDK measure since they measure a child's consistency on separate activities and do not include the severity level of the inconsistency. Also a child's productions of target words are mainly elicited by picture naming and not repetition. Williams and Stackhouse (2000) study and the Nuffield dyspraxia program (2004) provide a good model for measuring consistency of DDK performance and comparing it to accuracy of productions. First, Children's consistency is measured in different DDK tasks including sound level, syllable sequences, real word and non-word repetition levels. Second, consistency scores include the rating of the child's inconsistency. Investigation children's DDK at different speech processing levels and using scoring measure such as consistency and accuracy are more sensitive in capturing developmental change and could aid in the differential diagnosis of children with speech difficulties.

The present study scored the stimuli as follow:

5.4.2.2.1 ***Accuracy***

The child's accuracy of productions was based on the adult-like realization of consonants, where the 'proportion of whole-word accuracy' was measured.

- ***a- Accuracy of first attempt of the target compared to the adult model (A1):*** the child's first attempt at repeating the target item (real-word, non-word or syllable sequences) was scored for whole-word accuracy, where an item was scored as either *accurate=1* or *inaccurate=0* compared to the adult model. Only consonant accuracy was scored compared to the adult model, and any speech sound errors related to consonants was scored as incorrect, thus, speech sound errors such as substitutions, omissions etc. and typical phonological process found in typical developing child's speech are scored as incorrect. However, although only consonant accuracy was scored, vowel changes that led to the lexicalization of a non-word was scored as incorrect (lexicalizations were also scored independently when it occurred although, see below).

- ***a- Accuracy of five repetitions of the target compared to the adult model (Ar):*** The child's attempt at repeating an item (real word, non-word and syllable sequence) five consecutive times was scored for accuracy. To calculate the total proportion Ar score for an item; each of the five repetitions (abbreviated as: R1, R2, R3, R4 and R5) was scored for whole-word consonant accuracy, where a score of =1 indicates that the token was accurate and matched the adult model, and incorrect = 0 indicating that the token was inaccurate. To calculate the Ar score of an item, the following formula was used:

$$\text{Ar for an item} = \frac{total\ accuracy\ score\quad i.e\ R1+R2+R3+R4+R5}{number\ of\ repetitions}$$

Transcribed examples from a boy ages 3;5 (years; months) are provided in Table 5 . Examples from the table include item 1 /bæb/, which is produced accurately when repeated after the examiner once, thus A1 for the items was scored 1. When the item was repeated consecutively, all five repetitions were accurate, therefore Ar score will be: $\frac{total\ accuracy\ score}{number\ of\ repetitions}$, $\frac{5}{5}$ =1. Another example is Item number 3 / k ɛ lb / shows that the first/baseline production was accurate, therefore A1=1; while only three productions of the consecutive repetitions were accurate (R3 and R4 repetitions were inaccurate), yielding an Ar proportion score of: $\frac{total\ accuracy\ score}{number\ of\ repetitions}\frac{3}{5}$ = 0.6.

5.4.2.2.2 *Consistency*

Consistency of five repetitions compared to the child's own model (C): consistency was based on the "child's own model" i.e. based on the child's own speech sound system and whether the child produces a target item consistently compared to his/her first production regardless of accuracy of productions. Consistency of repetitions refers to the 'proportion of whole-word consistency', which computes the number of similar productions of a target item divided by the total number of repetitions. Therefore, the more consistent the child is in their productions the higher their consistency scores[10]. The following formula was used:

$$\text{C for an item} = \frac{total\ of\ similar\ repetitions}{number\ of\ repetitions}$$

From Table 5, both examples 1 and 4 represent target items where each production was produced the same regardless of accuracy. The target word /bæb/ the same in all five repetitions (correctly), while produced the target item /z ɪ r/ the same in all five repetitions (incorrectly) both yielding a total consistent score of $\frac{total\ of\ similar\ repetitions = 5}{number\ of\ repetitions = 5}$ = 1.

5.4.2.2.3 *Repeating items less than five times:*

If a child repeats an item less than five times, they were still scored for Ar and C. The same formula was used to calculate the proportion score of Ar and C. For example, if a child produced an item /kɛlb/ (See Table 5.) four times instead of five:

Ar score would be = $\frac{total\ accuracy\ score}{number\ of\ repetitions}\frac{2}{4}$ = 0.5

[10] Computing consistency scores was similar To Marquardt, Jacks and Davis (2004), however they used token variability, were they computed the number of different variant productions instead of similar productions.

C score will be $\frac{total\ of\ similar\ repetitions = 2}{number\ of\ repetitions = 4}$ = 0.5

5.4.2.2.4 ***Computing Scores***

Computed scores is *the average sum of the proportion scores*. From the stimuli design section described previously, it is was noted that each task (real-words, non-words and syllable sequences) comprised of monosyllabic items, disyllabic items and multisyllabic items. Therefore, the average proportion scores were calculated for the stimuli lengths in each task, and then the total average proportion was calculated for the whole task regardless of length. Analysis was therefore carried out on the total average proportion scores of each task (regardless of length), and on the average proportion scores of the stimuli length of each tasks.

The scoring sheets used in this study are provided in Appendix 9a and b.

Table 5.4: Transcribed examples of real word repetitions from a boy aged 3; 5 (year; month); scoring of accuracy of baseline and accuracy and consistency of consecutive repetitions.

#	**Adult Model**	**Child's single repetition**	**Rapid Consecutive repetitions/productions**					**A1**	**Ar**	**C**
			R1	R2	R3	R4	R5			
1	b æ b	b æ b	b æ b	b æ b	b æ b	b æ b	b æ b	1	5/5=1	5/5=1
2	ʃ æ ms	ʃæmç	tʃæmç	çæm	tʃæmç	ʃæmç	ʃæmç	0	0/5=0	4/5=0.8
3	k ɛ lb	k ɛ lb	k ɛ lb	k ɛ lb	k ɛ b	k ɛ b	k ɛ lb	1	3/5=0.6	3/5=0.6
4	z ɪ r	ðɪr	ðɪr	ðɪr	ðɪr	ðɪr	ðɪr	0	0/5=0	5/5=1
5	g i ˈtˤɑ:r	g i ˈtˤɑ:r	kitʷ	kiˈtɑr	kitɑɹ	˳g̥itˤɑ:r	giˈtˤɑ:r	1	2/5=0.4	2/5=0.4

Note: A1= whole-word Accuracy of single-immediate repetition Ar = accuracy of multiple repetitions; C = consistency of consecutive repetitions.

5.5 Data preparation and analysis

Data was entered into an SPSS Software version 20 for statistical analysis. There were three age groups of children (3 year olds, 4 year olds and 5 year olds), 3 test conditions (RWR, NWR, SSR) and their 9 sub-test conditions (monosyllables, bisyllables and trisyllables). Children's performance on the tests and subtests were measured for Accuracy of one-single repetition (A1), Accuracy of consecutive repetitions (Ar), Consistency of consecutive repetitions (C). The primary dependant variables of interest were scores of Accuracy on first repetition (A1), scores of Accuracy on consecutive repetitions (Ar) and consistency of consecutive repetitions (C). The independent variables were the test conditions (RWs versus NWs versus SSs), the sub-test conditions of length of stimuli (monosyllables versus bisyllables versus trisyllables) and age.

5.5.1 *Sample Characteristics*

Before analysis was carried out, the data was examined for normal distribution and homogeneity of variances. First, distribution of the data was investigated for all the dependant variables individually in each age group. Numerical values were investigated for the Shapiro-Wilk's normality test[11], and Z-values for Skewenss and kurtosis were calculated[12] (Doaen, Seqard, L.E., 2011); a summary of the values are provided in Appendix 10. The values of Shapiro-Wilk's show that for most tasks, children aged 5-years old were highly negatively skewed on accuracy scores. On consistency measure, the same effects were seen for 4 and 5-year old groups on the SSR tasks, while the monosyllabic items show negatively skewness in all age groups. Second, homogeneity of variance was investigated for both within group (repeated measure) and between the groups. Mauchly's Test of sphericity was used to confirm homogeneity of variance for the repeated measure (main effects of within-subject test conditions), χ^2 (df) =value; and the alpha level was set at .05. When sphericity assumption was violated, the Greenhouse and Geisser correction (ε) was used to estimate sphericity and the F-ratio, as this is a more conservative correction to Type 1 error than other corrections such as the Huynh-Feldt correction (Field, 2009). Mauchlys test was only reported when the assumption of sphiricity was violated and Greenhouse-Geisser estimates were used instead.

Levenes Sphericity test was used to investigate equality of variances between the groups. The test was significant between groups in the test conditions, with alpha levels > .05. The unequal variances were actually expected as negatively Skewed data on five-year old scores in most test conditions could have resulted in the unequal variances. Also, 3 year old children seem to have larger variances in general compared to the 4 and 5-year old children. In addition, there were unequal sample sizes; 3 year olds have a smaller sample size than 4 and 5-year olds, this could also adversely interfere with the assumptions of equal variances on the Levens test.

[11] The null hypothesis for the Shapiro-Wilk's test is that the data are normally distributed, and the null hypothesis is rejected at p <.05.
[12] Skeweness and Kurtosis Z values should be within ±1.96

Although negatively skewed data was present and unequal variances assumed, the data was analysed using *parametric* test ANOVA, in which it assumes the normal distribution of variances i.e equal standard deviations. An option to correct the problems of normality and assumptions of homogeneity of variances is transforming the data; however, the data was not transformed for a number of reasons. The data entered for statistical analysis are proportions ranging from 0 up to 1, which is driven from a count and expressed as decimal fractions. It is very common to find skewed data when using proportions, with data piled against 0 or 1, which can also cause unequal variances. Another point to consider is the ceiling effect seen in the older age group of children especially in monosyllable sequence repetition task, with a mean score 1 and Standard deviation of 0 (no error) thus there is no implication of transformation, since there is actually no deviation from the mean. Furthermore, in most tasks only the 5-year old group showed skewed data, therefore transforming data of one group will result in transforming data of other groups with normal distributions. Generally, ANOVA is often suggested to be robust, where *F* ratio performs the same and controls for Type 1 error with skewed distributions, with authors claiming that transforming the data to meet normality is seldom worth the effort (Field, 2009). Another point to consider, is.. Also, if comparisons were to be made with other studies of English speaking children, it would be fair to use the same methods used in analyses (as they also had skewed data).

5.5.2 *Statistical Analysis*

Data was analysed using a mixed ANOVA design with 3 (between group) × 3 (repeated measures) variables. First, for each of the dependant measures (**A1, Ar, C**) analysis was carried out to compare children's performance on the three test conditions (RWR, NWR, SSR), using 3 between group (3 vs 4 vs 5-year olds) × 3 repeated measure (RWR vs NWR vs SSR) design. Second, for each of the dependant measures (**A1, Ar, C**) analysis was carried out to investigate the effects of stimulus length (monosyllabic vs bisyllabic vs trisyllabic items) of each test condition (RWR, NWR, SSR) using 3 between group (3 vs 4 vs 5-year olds) × 3 repeated measure (monosyllabic vs bisyllabic vs trisyllabic) design. Follow-up testing were conducted using *post-hoc* comparisons, with effects being significant at an alpha level of > .05. For the repeated measure (within-subject conditions), the *Bonferroni post-hoc* comparisons were used and reported. The *Bonferroni post-hoc* comparisons are generally the most robust technique when violating sphericity, this is especially true when controlling for Type-1 error rate as it is conservative (Field, 2009). For the one-way (between age-group) comparisons, the *Games-Howell pair-wise* comparisons were used and reported, as it is the preferred method with unequal variances between the groups and unequal sample sizes (Field, 2009).

5.5.3 *Reliability*

Inter-rater reliability was conducted to examine the extent to which the scores obtained by children are consistent and objective across examiners. A sample of 10 children in each age group (3, 4 and 5 year olds) were selected and a qualified Speech and language therapist carried out the reliability scoring. The rater was a native Saudi-Arabic-speaking speech and language therapist with background experience and training on transcribing typical and atypical speech. The scoring procedure was explained and examples of responses were reviewed with the rater, along with training samples. Tasks of real word, non-word, syllable sequences were scored for accuracy and consistency. The average measure correlations were calculated. According to Hammond (2006), reliabilities above .70 are desired if a test is to be used as a research tool. However, a minimum requirement of .55 is also often cited as appropriate for assessments administered in experimental group studies (e.g., Rost, 2007). The general percentage of agreement was 90% on all tasks, and there were generally high level correlations indicating high inter-rater reliability. Inter-rater correlations were as follows: for the 3 year old children, correlations for both accuracy and consistency on the three repetition tasks were between .72 and .88. Correlations for the 4 year old sample on accuracy and consistency on the three repetition tasks were between .89 and .98, while the 5 year old sample had high correlations on all accuracy and consistency scores of the three tasks, where correlations were above .95 indicating that this age group could presumably be easier to score due to their high accuracy and consistency levels.

CHAPTER 6 RESULTS AND DISCUSSION I: PERFORMANCE ON SINGLE REPETITIONS

Children's overall performance on single repetitions will be presented and discussed in this chapter (Section 6.1 and 6.2). This will be followed by presenting the results and the discussion of length effects on children's performance (Section 6.3 and 6.4). Finally, this chapter will conclude with a general discussion (6.5).

6.1 Results 1: Accuracy of Immediate Single Repetitions

This section answers the first of the research questions outlined previously in the methods chapter (Chapter 4 Section 4.1.1):

Research questions addressed:

Question 1: How do typically developing Arabic-speaking children perform on the speech output processing tasks (real-words, non-words and syllable sequence) as measured by accuracy of repetition?

a) Is children's performance affected by the test conditions /or speech processing demands of tasks (real words vs nonwords vs syllable sequences)?

b) Are there age-related (developmental) changes on performance on the tasks (real word, non-word and syllable sequence repetition)? i.e. are the tasks developmentally sensitive?

6.1.1 *Accuracy of Immediate Repetition (A1): Performance Across Test Conditions (RW's, NW's, SS's) and Developmental Change*

Means and standard deviations of the total average proportion of A1 scores are presented in Table 6.1. Inspection of the data shows that children in each age group were least accurate when repeating NWs and were most accurate when repeating SS. Also, there was a developmental change in accuracy scores. Where accuracy scores of each test condition improved with age. For all the three test conditions, children aged 5 years performed at ceiling. Figure 6.1 confirms this observation. From the table and figure, it could be seen that the younger age groups, especially the 3-year old children had large variances and confidence intervals compared to the other older group.

Table 6.1: Means, standard deviation (s.d) and ranges of (A1) for the three test conditions in each age group. Post-hoc comparisons of within and between groups.

		Age group					
Tasks		**3 year olds**	**3 < 4**	**4 year olds**	**4 < 5**	**5 year olds**	**3 < 5**
Real words RW	Mean (s.d)	**.78 (.21)**	*p*=.1	**.85 (.10)**	*p*=.27	**.96 (.04)**	* *p*= .04
	Min-Max (Range)	.37- .97		.66- 1		.90- 1	
	RW vs NW	*P=.1*		*p*=.1		*P=.25*	
Non-words NW	Mean (s.d)	**.71 (.2)**	*p*=.16	**.84 (.11)**	*p*=.24	**.95 (.04)**	**p*=.003
	Min-Max (Range)	.37- .93		.63- .97		.90- 1	
	NW vs SS	* *p*= .05		**p*=.03		*P=.5*	
Syllable sequences SS	Mean (s.d)	**.81(.1)**	*p*=.37	**.89 (.1)**	*p*=.28	**.97 (.04)**	* *p*=.01
	Min-Max (Range)	.60- 1		.66- 1		.90- 1	
	RW vs SS	*p=.43*		*P=.08*		*P=.1*	

Note: *p value: ** significant at p < .01, * significant at p <.05, ^ marginal significance at .06*

This means that children within one age group showed large variation in their performance on the tasks.

Further analysis was carried out to test the effects of test condition and age on accuracy performance (A1), using 3 (age group)× 3 (test condition, RWR, NWR, SSR) ANOVA. Mauchly's test indicated that the assumption of sphericity was met for the main effects of test condition, χ^2 (2) =5.9, P> .05.

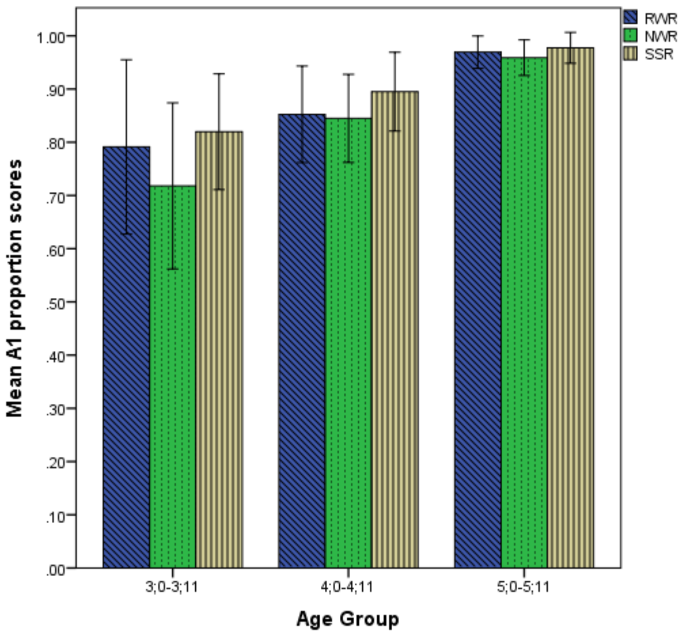

Figure 6.1 Mean proportion scores of Accuracy of baseline repetition (A1) by age group and test condition. Error bars represent 95% Confidence interval

Results showed that there was a significant main effect of test condition F (2, 202) =29.1, *P*<.001 and age F(2,101) = 22.4, *p< .001* . There was also a significant interaction between test condition and age F (4, 202) = 5.9, p<.001. This indicates that children's accuracy on the test conditions differed with age. The main effect of test condition was further analysed using Bonferroni post-hoc comparisons (See Table 6.1 below for all mean differences and post-hoc comparisons). Results showed that the 3-year old age group showed significant differences only between the NWR task and the SSR task. Although the mean proportion accuracy of the NWR task is clearly lower than the RWR task, this difference failed to reach significance. The insignificant difference could be due to the large standard deviations and confidence intervals of the tasks overlapping. The mean difference between RWR and NWR = .073; CI (95%) was between - .066 and + .21, thus CI included a zero which indicating no difference between the means and therefore an insignificant result. The 4-year old children showed similar A1 results to the 3-year old group, there was only a significant difference between the SSR and NWR task. The 5-year old children reached ceiling on all the tasks, and performed equally well, as there was no significant difference between the tasks.

To investigate developmental changes on each of the tasks, the Games-Howell post-hoc comparison was conducted. Unexpectedly, the only significant age group difference was between the 3-year olds and the 5-year olds (See Table 6.1). Although there were mean differences between the groups on the tasks, the large standard variations and CI overlapped resulting in an insignificant difference on the tasks between the 3 and 4-year olds, and between the 4 and 5-year olds.

In summary, there were only significant differences between the syllable sequence repetition task and the non-word repetition task in 3 and 4 year old children. Developmental change was only evident between the 3 and 5 year olds. The following processing and developmental trend was observed:

***Sig**

a) **Processing demands**: 3 & 4 year olds: SSR = RWR = NWR

5 year olds: SSR = RWR = NWR

b) **Developmental change**: only significant between 3 and 5 year olds

6.2 Discussion of Research Question 1

The findings presented above are discussed in light of the hypothetical speech processing model of Stackhouse and Wells (1997) and the literature reviewed in chapters 2. As with the presentation of the results, the discussion will focus on a) each age group's performance on the tasks (i.e. children's performance at different levels of speech processing) and b) the tasks' sensitivity to developmental progression.

6.2.1 *Profiles of Performance and Developmental Progress*

Generally, the present findings partially supported the predictions; where two unexpected findings emerged from this study. First, the Arabic-speaking children who were 3 and 4 years old showed processing profiles that did not replicate numerous reported cross-linguistic findings (Chiat & Roy, 2007; Dispaldro, et al., 2011; Torrington Eaton, et al., 2015; Hoff, et al., 2008; Roy & Chiat, 2004; Sundström, et al., 2014; Vance et al, 2005; Williams & Stackhouse, 2000). However, performance of the 5 year olds did replicate findings from Williams and Stackhouse's (2000) study. Second, developmental progression between the ages 3 and 4 years old were not observed in the present findings, and age differences only occurred between 3 and 5 year old children; this also did not replicate cross-linguistic processing studies (Sundström, et al., 2014; Vance et al, 2005; Williams & Stackhouse, 2000).

6.2.1.1 *Profile of Processing Performance – Task Comparisons*

Ssurprisingly, both the 3 and 4 year old children did not show different profiles of performance across the RWR and NWR tasks, as measured by whole-word accuracy, nor were there significant differences in their performance between SSR and RWR tasks. Nevertheless, the children showed greater facility at repeating SSs compared to NWs. When comparing the processing demands of the different tasks within the context of the processing model (Stackhouse & Wells (1997), the results suggest an interesting pattern of speech processing performance with the stimuli presented at a single repetition level.

First, it seems that beneficial effects of stored representations/motor programs assumed to be accessed during RWR over the creation of new motor programs accessed during NWR were not observed in this study with the stimuli presented to the children. Second, children were able to repeat sounds and sound sequences as accurately as words; this suggests that there were no advantages of using higher levels of stored phonological representations/motor programs over lower levels of motor execution and vice versa. Third, when the children were challenged with creating a new motor program that has higher levels of linguistic demands (such as syllable stress and vowel changes) they were significantly less accurate than when repeating isolated and sequential sounds (which were less contaminated by linguistic demands); this pattern observed in the 3 and 4 year olds supports the speech processing models hypothesised level of processing, where it is assumed that syllable repetition tasks are located at the lowest level of output processing and should not be placed with the higher levels of non-word processing.

The pattern of performance observed in the 3 and 4 year old children, were not entirely in line with cross-linguistic evidence. Starting with the youngest age group (3;0-3;11 year old participants); empirical evidence showed different profiles of speech processing performance emerging in this age groups. The children's failure to show significant differences between the RWs and NWs on whole word accuracy scores did not replicate many speech processing studies on children within the same age range (e.g. Chiat & Roy, 2007; Dispaldro et al., 2011). However, the youngest children's performance did replicated findings by Vance et al. (2005) on English-speaking children aged between 3;0-3;11 years. The findings therefore support Vance et al.'s notion that young 3 year old children possibly process both RWs and NWs similarly, favouring a bottom-up processing route; where the speech input signal from the adult model was used for both tasks to recreate a motor program, and children did not use stored phonological representations/motor programs to support RWR. See Table 6.2 for an overview of scores from 3 year old children[13].

Table 6.2: Score of 3 year old children on real word repetition RWR and non-word repetition NWR from three studies (Arabic, Italian and English).

Language of participants	**Scoring method**	**Number of items**	**RWR**	**NWR**
Arabic (the present study)	Proportion of WWA	30	0.78	0.71
English (Vance et al., 2005)	Percentage of WWA	60	67%	64%
Italian (Dispaldro et al., 2011)	PPC	16	94%	91%

Note: Percentage of phonemes correct =PCC. Whole-word accuracy=WWA.

[13] Onley studies using percentages/proportions of accuracy scores were included in the table. Studies such as Chiat and Roy (2007), Roy and Chiat (2004) and others using number of items produced accurately instead of percentages/proportions were not included in the table.

However, there is reason to suggest that the Arabic-speaking children in this study repeated RWs using an interaction of both top-down processing (using existing phonological representations/motor programs) and bottom-up processing (without using prior stored lexical knowledge). In support of this proposition is the observed overall mean proportion scores and ranges of the RWR and NWR tasks. Clearly, the 3 year old's mean accuracy score on the RWR task (0.78) were higher than their score on the NWR task (0.71). However, this difference failed to reach statistical significance most likely due to the large stretch of score ranges on both tasks (stretching from 0.37 to 0.90+). This large individual variation lead the author to speculate that, it is possible that some children used stored phonological representations/motor programs to support RWR, while other children created new motor programs for the same task. It is also possible that the same child used an interaction of both stored lexical knowledge to repeat some words while created new motor programs to repeat other words. This could be due to children's young age, a child's knowledge of their language is limited and still developing and it is the age were children's vocabularies are expanding rapidly. Furthermore, the wide age ranges of this group meant that children's experience with their language differs widely and in turn could have affected their performance differently. (See Section 6.5 on general discussion of Chapter 6).

When turning to children between the ages of 4;0-4;11 years old, the results were particularity unexpected. Based on cross-linguistic evidence, it was predicted that children at least older than 3-years would benefit from stored phonological knowledge/motor programs for RWR over creating new motor programs for NWR. However, contrary to predictions, children's performance on the tasks did not replicate findings from studies such as Sundström et al. (2014) on Swedish-speaking children and Vance et al. (2005) on English-speaking children. The studies, found that children within the age band (4;0-4;11) repeated RWs more accurately than NWs. The Arabic-speaking children in this study, however, did not show significant differences in their performance between the RWR and NWR tasks. The RWR scores from this present study fell in between the scores obtained by Sundström et al. and Vance et al. (see Table 6.3 below); however, the Swedish and English-speaking children showed significantly lower scores on the NWR task, while the Arabic-speaking children obtained highly similar scores on both RWR and NWR tasks .

Table 6.3: Scores from 4 year old children on real word repetition RWR and non-word repetition NWR from three studies.

Language of participants	Scoring method	Number of item	RWR	NWR
Arabic (present study)	Proportion of WWA	30	0.85	0.84
Swedish (Sundström et al, 2014)	PPC	25	91%	81%
English (Vance et al., 2005)	Percentage of WWA	60	79%	71%
Note: Percentage of phonemes correct =PCC. Whole-word accuracy=WWA				

For children between the ages of 5;0-5;11, there were no differences between processing demands, where the overall scores on the three tasks reached ceiling (SSR 0.97> RWR .096 > NWR 0.95). The findings did not meet the predictions and did not replicate cross-linguistic speech processing studies such as Sundström, et al. (2014) on Swedish-speaking children or Budd et al. (2012) and Vance et al. (2005) on English-speaking children. The Arabic-speaking 5 year olds showed RWR scores that were very similar to the 5-year old participants in Budd et al.'s and Sundström et al.'s studies, where accuracy scores also reached ceiling. However, the 5 year old children in Budd et al.'s and Sundström et al.'s studies showed significantly lower NWR scores compared to RWR scores; this was not observed with the Arabic-speaking participants (see Table 6.4 bellow). In contrast, the results were in line with Williams and Stackhouse (2000) study, were no significant difference between tasks at this age were observed.

Table 6.4: Scores from 5 year old children on real word repetition RWR and non-word repetition NWR, from four studies.

Language of participants	Scoring method	Number of items	RWR	NWR
Arabic (present study)	Proportion of WWA	30	0.96	0.95
English (Budd et al, 2011)	Proportion of WWA	56	0.96	0.84
Swiss (Sundström et al, 2014)	PCC	25	95%	87%
English (Vance et al., 2005)	Percentage of WWA	60	87%	77%
Note: Percentage of phonemes correct =PCC. Whole-word accuracy=WWA				

Findings from the present study did however replicate, to some degree, early findings from Williams and Stackhouse's (2000) study, where children did not show different profiles of performance. In their study, children of all age group (3, 4 and 5) performed equally well on single repetition of RWs, NWs, and SSs. The possible difference between Williams and Stackhouse's study and the studies conducted by Budd et al. (2012), Sundström, et al. (2014) and Vance et al. (2005) could be due to the smaller item list, which included only 10 bisyllabic items and 6 trisyllabic items compared to other studies which include a list of mon, bi and trisyllabic items, each with more than 15 stimuli.

6.2.1.2 *Developmental Sensitivity - Group Comparisons*

Unexpectedly, the overall task sensitivity to developmental change did not meet the predictions, as children's general performance on the three repetition tasks did not improve between age bands; significant effects of age on accuracy of repetition were only evident between the 3 year old group and the 5 year old group. The results were not consistent with other speech processing studies (e.g. Chiat and Roy, 2007; Sundström, et al., 2014; Vance et al, 2005; Williams & Stackhouse, 2000). For example, in Williams and Stackhouse's (2000) study, age effects were significant between the 3 and 4 year olds but not between 4 and 5 year olds; this was observed on repetition tasks of RWs, NW's and SSs. Similarly, in Vance et al.'s (2005) study, age effects on RWR and NWR tasks were observed between 3- and 4 year olds and between 4 and 5 year-olds. Furthermore, many other study have found age effects in children as young as 2 years of age such as Roy and Chiat (2005) and Chiat and Roy (2007). Interestingly, these studies used whole word accuracy (Vance et al, 2005; Williams & Stackhouse), as with the present study, so the scoring method could not be argued to be a reason for the discrepancy between the results.

As mentioned in the onset of this section, the selection and design of stimuli for the tasks could have contributed greatly to children's performance. The investigating children's sensitivity to tasks and age relate changes could be observed when different stimuli lengths are examined.

6.3 Results 2: Effects of Stimuli Length

In the previous section the effects of test condition RWR, NWR and SSR and age on children's performance were presented and analyzed. In this section, the effects of item length of each task are presented and analyzed. As stated in the methods chapter (Chapter 3, Section 4.3.5), the items for each task increased in the number of syllables; where each task included 10 monosyllabic items, 10 bisyllabic items and 10 trisyllabic items. This section answers the questions related to the accuracy of single repetitions when increasing stimuli length and the developmental sensitivity of each length, the questions were:

Research questions addressed:

Question 2: How do typically developing Arabic-speaking children perform on the repetition tasks (real words, nonwords and syllable sequences) when stimuli length increases?

- **a)** Does the stimuli length (monosyllabic vs bisyllablic vs multisyllabic stimuli) affect children's repetition accuracy on the repetition tasks?
- **b)** Are there age-related (developmental) changes in children's accuracy on the tasks when stimuli length increases?

Descriptive statistics of A1 scores on different stimuli length in each test condition is shown in Table 6.6. General inspection of the mean show that bisyllables were the most accurate, with ceiling effects seen in the 5-year old age group. Analysis was carried out to determine whether length of the different tasks (RWR, NWR and SS) affected children's A1 scores. Mauchly's test indicated that the assumption of sphericity has been violated for the main effects of stimuli length, $p < .05$, therefore the degrees of freedom was corrected to meet sphericity using the Greenhouse-Geisser estimate of sphericity ($\varepsilon > .5$).

Real Word Repetition There was a significant main effect of length F(1.4, 149) = 24.7, $p < .001$, and a significant main effect of age F (2,105) =21.8 p< .001, and no significant interaction between length and age F(2.9,149) = 1.72, p =.17. The main effect of word length on accuracy was further analysed using Bonferroni post-hoc pairwise comparisons. Children were significantly more accurate when repeating bi-syllabic words compared to mono and tri-syllabic words (mean difference =.064 and

.081 respectively, p's <.05), there was no significant difference between the mono and trisyllabic words in accuracy of one repetition. For the main effect of age, the Games-Howell post-hoc comparisons revealed that there was only a significant change in accuracy performance between the 3 and 5 year olds (mean difference =0.2, $p<.05$). There was nonsignificant developmental changes between the 3 and 4-year olds (mean difference = .08) and between the 4 and 5-year olds (mean difference = .1) all p's > .05 (Figure 6.2 illustrated the difference in in length and age group).

Table 6.5: Means, standard deviations (s.d) and ranges of (A1) scores by different stimulus syllable lengths (mono, bi and tri syllables) in the three test conditions (real-words non-words and syllable sequences).

TASK		Age group		
REAL WORD RW		3year old	4year old	5year old
mono-syllables	Mean (s.d)	**.74 (.2)**	**.85 (.16)**	**.96 (.08)**
	Range	.40-1	.70-1	.80-1
bi-syllables	Mean (s.d)	**.83 (.22)**	**.91 (.1)**	**.99 (.001)**
	Range	.40-1	.70-1	.99-1
tri-syllables	Mean (s.d)	**.73(.19)**	**.81(.16)**	**.95(.15)**
	Range	.30-.90	.40-1	.70-1
NONWORD NW				
NWR mono-syllables	Mean (s.d)	**.77 (.23)**	**.86 (.14)**	**.95 (.08)**
	Range	.84-1	.60-1	.80-1
NWR bi-syllables	Mean (s.d)	**.76 (.2)**	**.90 (.1)**	**.99(.001)**
	Range	.40-1	.70-1	.98-1
NWR tri-syllables	Mean (s.d)	**.61 (.18)**	**.80(.17)**	**.92(.1)**
	Range	.30-.90	.40-1	.70-1
SYLLABE SEQUENCE SS				
SS mono-syllables	Mean (s.d)	**.93 (.09)**	**.95 (.05)**	**.98(.03)**
	Range	.80-1	.90-1	.90-1
SS bi-syllables	Mean (s.d)	**.85 (.1)**	**.91 (.09)**	**.99(.00)**
	Range	.50-1	.70-1	.98-1
SS tri-syllables	Mean (s.d)	**.67(.22)**	**.83(.17)**	**.94. (.1)**
	Range	.44-1	.40-1	.70-1

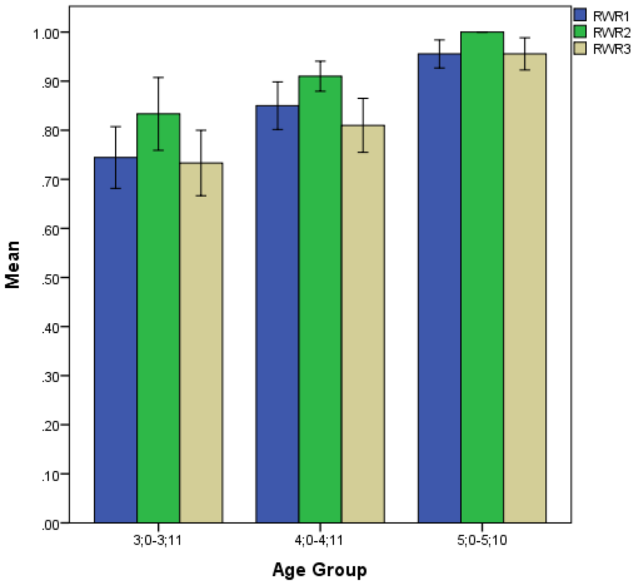

Figure 6.2: Mean proportion scores on A1 by age group and length of real-words. RWR1=monosyllables, RWR2=bisyllables, RWR3=trisyllables. Error bars represent 95% Confidence interval

Non-Word Repetition There was a significant main effect of length F(1.34, 135) = 27, p< .001, and a significant main effect of age F (2,101) =27.3 *p*< .001, and no significant interaction between length and age F(2.6,135) = 2.7, *p* =.074. The main effect of non-word length on accuracy was further analysed using Bonferroni post-hoc pairwise comparisons. Bisyllabic non-words were significantly more accurate than trisyllables (mean difference = .11, *p*<.05) but not significantly more accurate than mono-syllables (mean difference =.03, *p* >.05). Although mean scores of monosyllable were greater than the trisyllables, the difference failed to reach significance (mean difference= .08, p>.05). For the main effect of age, the Games-Howell post-hoc comparisons revealed similar results to the RWR, were the only significant difference was between the 3 and 5-year olds (mean difference =.24, p<.01). There was no significant difference between the 4 and 5-year olds (mean difference =.1, p>.05) or the 4 and 3-year olds (mean difference =.12, p>.05). The differences between the groups and syllable lengths are illustrated in Figure 6.3.

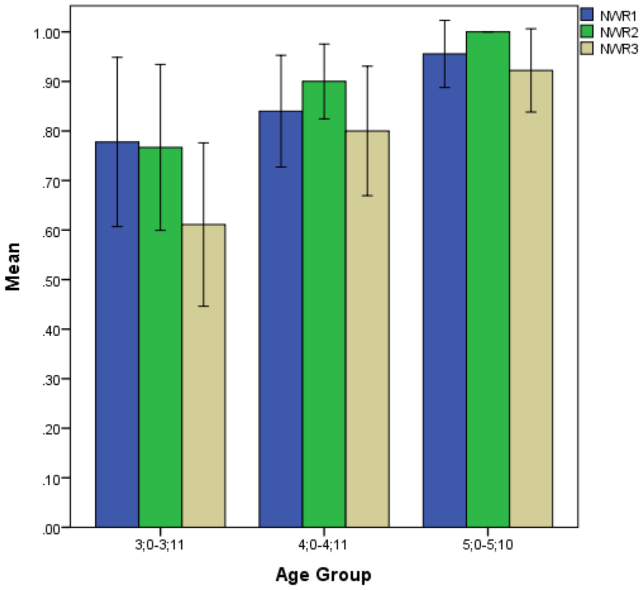

Figure 6.3: Mean proportion scores on A1 by age group and length of non-words. NWR1=monosyllables, NWR2=bisyllables, NWR3=trisyllables. Error bars represent 95% Confidence interval.

Syllable Sequence Repetition There was a significant main effect of length F (1.39, 140) = 587, p< .001, and a significant main effect of age F (2,101) =20 *p*< .001, and a significant interaction between length and age F(2.8,140) = 9.3, p < .001. For the main effects of length, Bonferroni comparisons showed that although bisyllables had greater mean scores, they were not significantly different than monosyllables (mean difference = .01, p>.1), the significant difference was noted between the trisyllables and both the mono and bisyllables (mean difference =.14 and .10, *p*'s <.01. To break down the interaction between length and age, Games-Howell post-Hoc comparison revealed that for the monosyllables, there was no significant difference between the groups on accuracy of monosyllables (mean difference between 3 and 4 =.01, mean difference between 4 and 5yera olds =.03) all *p*'s>.05. For the bi-syllables, there was only marginal difference between the 3 and 5-year olds (.1, p=.056), while there was no difference between the 3 and 4-year olds (mean difference =.04, p>.1) and no significant difference between the 4 and 5-year olds (mean difference= .09, p>.05). Trisyllabic items also failed to reach significance between the group bands, although mean difference was observed. There was only significant difference was between the 3 and 5-year olds (mean difference =.26, p<.05). From Figure 6.4 and the Table 6.6, which illustrates the syllable sequence data, the mean differences between the groups are different, however the differences failed to reach statistical significance. This could be due to the large variation in performance and large CI's which overlapped, resulting in insignificance and therefore no developmental change from a statistical point of view. The same observation is seen on the NWR task, were different lengths of stimuli failed to differ between groups. Again, this could be due to large variability and CI overlapping.

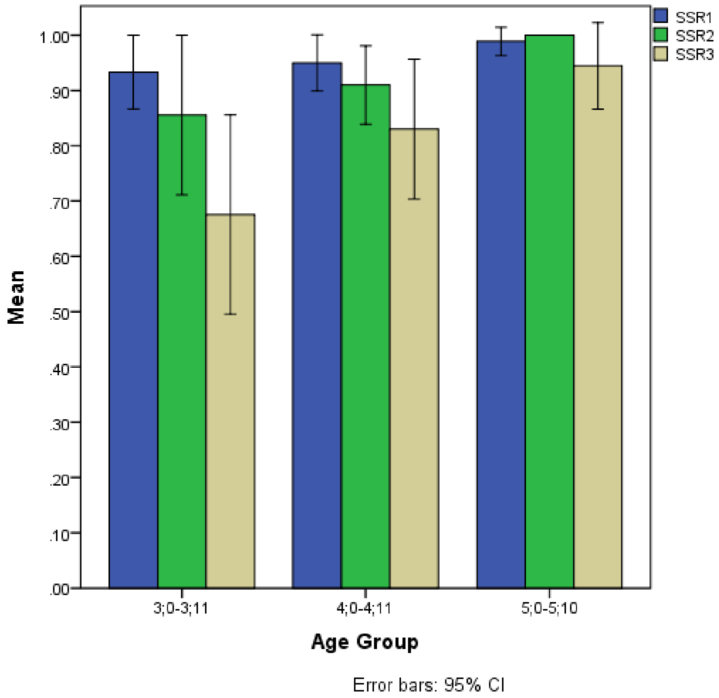

Figure 6.4: Mean proportion scores on A1 by age group and length of syllable-sequences. SSR1=monosyllables, SSR2=bisyllables, SSR3=trisyllables Error bars represent 95% Confidence interval.

Summary of Results:

a) Speech processing	
RWR	*Sig mono = trisyllables > bi in 5yo no sig.
NWR	= mono = bi > tri
SSR	Mono > bi > tri

b) Developmental	
RWR mono	3 = 4 =5
RWR Bi	
RWR tri	
NWR mono	3< 4 =5
NWR bi	
NWR tri	
SSR mono	3=4=5
SSR bi	
SSR tri	Onley sig 3<5

6.4 Discussion of Research Question 2

The results presented above on the effects of item length on children's performance on single repetition of tasks are discussed within the speech processing model of Stackhouse and Wells (1997) and the literature reviewed in chapters 2 and 3. The discussion will focus on the effects of length and developmental change on each task.

Within the speech processing model of Stackhouse and Wells (1997), the effects of increased performance load on different levels of processing (stored phonological knowledge and motor programs, motor programming and execution) were explored in Arabic-speaking children by presenting items of increasing length: including monosyllabic, bisyllabic and trisyllabic items.

6.4.1 *Length Effects and Developmental Progress*

The predictions were partially supported, where as predicted, repetition accuracy was sensitive to different stimuli lengths. However, the effects of stimuli length on repetition accuracy generally did not replicate findings on English-speaking children, rather, language specific factors influenced children's performance on the tasks (Chiat & Ray, 2007; Vance et al, 2005; Williams & Stackhouse, 2000). To be more precise, the children in this study showed greater facility at repeating bisyllabic lexical items than monosyllabic and trisyllabic items. Furthermore, surprisingly a clear profile of developmental change was only evident between the youngest and oldest age groups i.e. 3 and 5-year olds. Generally, this was not in line with many reported findings on English-speaking children who were at least younger than 5-years old (e.g. Roy & Chiat, 2004; Vance et al., 2000; Williams & Stackhouse, 2000).

6.4.1.1 *Real word Repetition RWR*

The 3 and 4 year old Arabic-speaking children were more accurate at single repetitions of bisyllabic RW items than on monosyllabic and trisyllabic items, while unexpectedly, the children did not show any statistically significant difference between mono and tri-syllabic items. The following trend was observed: **mono = trisyllables (both) < bisyllables.** At first glance, these findings are in contrast to results found by Chiat & Ray (2007) and Vance et al, (2005) on RWR accuracy in English speaking children. These studies found effects of stimuli length in children ages 3 and 4 years old, where accuracy of repetition was found to be highest for monosyllabic words followed by bisyllables then trisyllables. On the other hand, the present result supports the suggestion on the influence of language-specific factors on repetition accuracy, where similar to Dipsldro et al.'s (2013) study, Italian-speaking children, different effects of length were observed in these children, where no differences were found between bisyllabic and trisyllabic RWs. Furthermore, similar to Vance et al.'s findings, the present study found that children aged 5 years, showed ceiling effects and no differences between different lengths were observed, as children were highly accurate on different word lengths, especially bisyllabic items (mean proportion score was 0.99).

These finding were not entirely surprising, given that in Arabic, bisyllables on the whole constitute a large portion of children's early words which are influenced by the rich biometric word shapes of the Arabic language (Khattab & Al-Tamimi, 2013; Watson, 2002). Recall from the literature review (Chapter 3) that early phonological representations are argued to be language–specific templates. The different distribution of word shapes in different languages provides insight into how children's early words vary across languages. For example, data on the early words of English show a dominance of monosyllables with a concentration of CVC shape (see Elsen, 1996; Vihman & Keren-Portnoy, 2013; Vihman & Velleman 1998). In contrast, as reviewed in previously, bisyllables in Arabic are more common than monosyllables, with bisyllables or trisyllables occupying the majority of nouns (and verbs) (Watson, 2002). This dominance is evident in children's early templates, where Arabic acquisition studies found that bisyllables were targeted most in children's productions (Abdoh, 2011; Khattab & Al-Tamimi, 2013); constituting 60.9% of children's productions while monosyllables were at 38.2%, flowed by trisyllables at 0.9% (Abdoh, 2011). Furthermore, a close look at the present study's list of bisyllabic words and their syllable structure (refer to Appendix ..), shows that 9 out of the 10 bisyllabic words had a CVCV(C) syllable structure (for example: CVCV /kæ sæː/ 'cup' and CVCVC /kiˈtæ:b/ 'book'). This structure is one of the most commonly occurring shapes in the Arabic language -for both nouns and verbs- and this syllable shape CVCV was produced with the highest percentage (at 29%) in the early words of Arabic-speaking children followed by CVC structure then a CVC:V (Abdoh, 2011).

A particularly interesting finding that should be closely construed was the insignificant difference between mono and trisyllabic words. These findings do not replicate any of the studies on English speaking children; however, they are closely similar to findings in Italian speaking children (Dispaldro et al., 2013), where there were no significant difference between short bisyllabic and trisyllabic words. The present results could be interpreted to be a result of both methodological and language-specific factors (this will be discussed below along with the non-words).

6.4.1.2 *Non-word repetition NWR*

When the Arabic-speaking children were asked to repeat NWs, the following trend was observed: **mono = bisyllables (both)** > **trisyllables,** where children were significantly better at repeating bisyllables and compared to trisyllables. The findings replicate previous studies on English-speaking children ages 3, 4 and 5 (Chiat & Ray, 2007; Roy & Chiat, 2004; Vance, Stackhouse, & Wells, 2005). Within a speech processing model, the ability to create new motor programs are assessed through NWR.

Like RWR, monosyllables of NWR were repeated more accurately than trisyllables, however this effect was only found in the 3-year old age group. It seems that trisyllabic NW's were more challenging and more evident at this young age, therefore more sensitive to individual variation.

An interesting observation was that monosyllabic RW's were still less accurate than bisyllables in all age groups, even when a close inspection of the items shows that both sets of stimuli included a distribution of early acquired consonants, while bisyllabic items included later acquired consonants that were not present in the monosyllabic list. Table below illustrates consonant position in RWs, where the consonants were matched to non-words.

CVC (C) word	**SI-I**	**SI-F**	**SF-I**	**SF-F**
b æ b	b	b		
d ʊ b	d	b		
ʃ æ ms	ʃ	s		
t æ mr	t	r		
f i: l	f	l		
g ɪ rd	g	d		
k ɛ lb	k	b		
f æ: r	f	r		
m ɔ z	m	z		
j ɛd	j	d		
CVCV word				

k e: k æ	k		k	
ˈkæ sæ:	k		s	
g i ˈtˤa: r	g		tˤ	r
ʕ i ˈ g æ l	ʕ		g	l
ˈdʒ æ m æ l	dʒ		m	l
ˈʕinæb(ʕunæb)	ʕ		n	b
K i' t æ: b	k		t	b
g æ m æ ɾ	g		m	r
g æ l æ m	g		l	m
l æ m.b æ	l	m	b	
Tri-syllabic word				
b u r t u ˈq æ: l	b	r	q	l
q u: ˈb.b æ ʕ æ	q		b	b
t i ˈl i f o n	t		f	n
mæ xæ ˈd.da	m		d	
n æ ˈðˤ . ɑ r æ	n		r	
k æ n æ b æ	k		b	
θ æ ˈl .læ dʒ æ	θ	l	dʒ	
t u ˈf .fæ ħ æ	t	f	ħ	
s æ m æ ˈk ə	s		k	
m ɛ l. ˈʕ æ g ə	m	l	g	

6.4.1.3 *Motor execution level – SSR task*

Turning to the SSR task, the following pattern of performance was observed in all age groups: trisyllables < (both) monosyllables =bisyllables. In Chapter 2 (Section 2.1.3) that within the processing model of Stackhouse and Wells, SSR is a lower level output task that is hypothesized to involve the level of motor execution i.e. the vocal tract. It assesses a child's ability to imitate sounds in isolation and in sequences of sounds without the imposing linguistic demands of stress and vowel changes i.e. without confirming to stored knowledge of their language. The design of the monosyllabic items included a simple CV structure pattern, for example / fæ/, which would assess children's ability to imitate speech sounds in isolation. The design of the disyllabic items included a CVCV structure, for example /kæ tæ/, to assess children's ability to produce two sequences of sounds, and finally trisyllables with a structure of CVCVCV, for example / bæ tæ qæ/, where sequences of sounds with more articulatory adjustment were assessed. The participants did not find single repetitions of the mono and bisyllabic items challenging enough to capture increasing processing demands of length and articulatory placement adjustments. On the other hand, items including three syllables were more challenging to the children and accuracy scores reduced significantly. The articulatory adjustments required when repeating the CVCVCV structure of the trisyllables posed a higher processing challenge to the children, and accuracy scores could have reduced due to the different positions of articulations. It is possible that typical phonological processes and articulatory errors observed in children such as deletion and substitution were more likely to occur with trisyllables containing different consonantal sounds than they would when imitating sounds in isolation or only two sound sequences. Furthermore, while it is argued that syllable sequences with equal stress and no vowel changes are non-meaningful and are likely to measure speech motor execution that is less contaminated by linguistic factors, it is possible that a simple CV or CVCV structure is readily stored within word templates in lexical representations.

Furthermore, there was a significant effect of syllable length. Meaning that syllable length significantly affected the results (rating) performance of the participants. However the results of syllable length was not affected by the age group of the child. Meaning that all children were significantly affected by syllable length.

6.5 General Discussion of Chapter 6: Single Repetitions

In general, repetition of items at a single-word level showed interesting findings. As a whole, the present study for the speech processing tasks were not sensitive to processing demands for 5 year old children.

Single repetitions on the tasks were therefore, not a sensitive measure to developmental change with the stimuli set used in this study. The unexpected findings raise issues of linguistic and methodological influences.

6.5.1 *Phonological Working Memory Account*

The speech processing model described by Stackhouse and Wells (1997) does not account for phonological short-term memory in speech processing task performance, as it was argued that comparison between performances on the tasks is considered to be informative as a task itself, however, as mentioned previously, the model does not disregard the memory load. Furthermore, it is not the intention of this present study to investigate the role of phonological short-term memory on children's performance.

However, NWR has been extensively studied in the literature as a measure of PSTM (see Coady & Evans, 2008; Gathercole, 2006) and length effects. It would therefore be interesting to consider the assumed role of phonological memory capacity in auditory repetition, particularly NWR.

Findings from the present study showed that overall, with immediate single repetitions, the Arabic-speaking children repeated NW's as accurately as RW's. The overall finding is not supported by the account of phonological memory capacity (Gathercole et al., 19994). Within the working memory system, NWs are thought to rely on PSTM more than RWs which relies more on stored lexical knowledge (e.g., Dispaldro et al, 2009). However, the insignificant results between RWR and NWR could be due to the reduced storage demands required for the repetition of NWs, where it is possible that storage demands were minimized with immediate single repetition. This suggestion is supported by Kamhi and Catts (1986) study, who used NWR tasks so that children would not rely on stored lexicon knowledge; however, they disregarded short term memory for their findings, arguing that immediate repetition reduces demands on memory and is not accountable for results. As for SSR task, it is suggested with high certainty that this task would also require phonological working memory (when viewed within the concept of working memory), as the task investigates the lower execution to assess sound production skills of a child and their ability to sequence sounds accurately and requires auditory input, to repeat sound and sequences of sounds. However, it is suggested that the memory demands are reduced compared to the other tasks, due to the design of the stimuli, where no stress or vowel change occurred. Therefore, the storage demands are reduced compared to other tasks and the task does not require holding the string of sound and sound sequences in memory for long. Therefore, memory could partially explain results, where SSR task scores were significantly higher than other tasks.

However, the present design does not allow for conclusions to be drawn about the exact mechanism of this account. It is emphasised that the interpretation of the present results within the phonological working memory account should be approached with caution. The present findings suggest that, indeed, scores fell lower as length of items increased from bisyllables to trisyllables; however, both RW and NW monosyllabic and trisyllabic accuracy scores were insignificantly different, possibly due to linguistic factors specific to Arabic. It is reasonable to suggest that the present findings supports the proposition of Baddeley's (2003), who argues that the phonological loop although a separate construct does not function totally independent of stored language knowledge, and Snowling et al., (1991) who argue that NWR is influenced by prior linguistic knowledge and is not a pure measure of phonological working memory. However, on a sturdier note, the findings could be in accordance with MacDonald and Christiansen's (2002) notion; that there is no distinction between working memory and language knowledge; a distinct phonological working memory does not exist. Rather, children's performance on a task of phonological working memory such a NWR task can be explained as a reflection of language experience and cognitive constrains, such as worldliness and precision of underlying phonological representations.

6.5.2 *Language-specific Factors*

One factor that is suggested be a driving force wich influences children's performance on the tasks, especially RWR and NWR, are the properties of the Arabic language, which includes the functions and properties of consonant and vowel within the Arabic language.

Within the consonant vowel - CV - hypothesis (Nespor, Pene & Mehle, 2003) it is speculated that consonants and vowels have different functional roles in processing speech and language. Consonants are thought to carry lexical information while vowels carry grammatical and prosodic information. In Arabic, the ratio of consonants and vowels are remarkably different, with a highly rich consonantal inventory, including emphatics and pharyngealised consonants, and a small vocalic inventory (refer to chapter 3 Arabic language). Words consists of a root-and-template/pattern (non-concatenative morphology) (Watson, 2002 and McCarthy, 1982; McCarthy & Prince 1990), where the root carries the basic semantic information and vowels are inserted in between to create different relations of nouns and verbs.

It is suggested that with the consonants carrying the rich lexical cues, the effect of vowel change in the list of NW items was not strong enough to reduce accuracy scores compared to its counterpart RW, since the consonantal root was maintained; as NWs were derived from RWs by only changing the vowels and maintaining the consonants, and only consonant accuracy was scored. For example, the root (k t b) is related to writing and when used with the interpolated vowels /i and æ/ creates the noun /kitæb/ 'book', while the vowel change created the non-word /kutib/. Furthermore, since the consonantal root carries the basic information, it is also speculated that high accuracy scores observed on the SSR were also influenced by the simple root-pattern; especially since the SSR also maintained the consonant root of the RW; for example the bisyllabic sound sequence /kæ tæ/, although non meaningful, contains the basic consonantal root.

Moreover, studies that compared segment type repetition performance, i.e. comparing consonants versus vowels, suggest that vowels are easier to repeat than consonants (Santos, Bueno & Gathercole, 2006; Sundström, Samuelsson & Lyxell, 2014; Yuzawa & Saito, 2006). Generally, vowels are acoustically more prominent due to their duration and amplitude, therefore having higher acoustic energy and are thought to be easier to perceive compared to consonants (Ladefoged & Disner, 2012). Sundström et al., (2014) suggested that consonants might be harder to perceive than vowels when there is less support from long-term memory representations, as with the case of NWR. Therefore, with the NWR task where consonants from the RW were maintained and only vowels were changed, the consonants even though harder to perceive, would have the advantage of stored knowledge of the consonant root, which would have supported repetition accuracy.

6.5.3 *Lexical and Sublexical Influences*

In light of the findings, it seems particularly interesting to consider lexical and sublexical factors that could have affected performance on real word and non-word repetition. It should be emphasised, however, the design of the items in this present study does not allow for conclusions to be made about the effects of lexical and sublexical factors on word and non-word repetition accuracy, such as word frequency, word-likeness or phonotactic probability. The stimuli design controlled for item length but not for other lexical and sublexical factors. Nevertheless it is without doubt that these factors played a role in children’s performance, and could present as confounding variables. Frequency effects of word and structural frequency could have both played a role in children’s performance. The suggestion is supported by young children’s performance on overall RW and NWs and by children’s higher scores on bisyllables compared to monosyllables. These factors will be discussed in more detail in the general discussion. However, for this section, general observations about children’s overall performance could suggest that lexical factors were influential in the youngest age groups performance; more precisely the 3-year old children. The design of this study included a picture naming task, which was designed as a pre-request task to ensure that a child was familiar with the presented stimuli words, and therefore, the word was stored within their lexical representations i.e. had stored motor programs. However, even if the child was familiar with a word, this does not mean that a word was frequent in the child's mental lexicon. Lexical factors such as word frequency (high frequency vs low frequency words), was a factor that was not controlled for in the present study due to lack of research in this area. The role of frequency (word and type frequency) is well establishes in the literature as an important aspect of language processing (Bybee, 2000; Sosa & Bybee, 2011; Vihman, 1992). With children as young as 3-year old, language experience varies from child to child depending on environmental factors, adults input and induvial variations; it is a vital age of rapid language acquisition. Therefore, 3-year old children showed huge variability in their responses to the picture naming task. For example, during the picture naming task, some 3-year old children were able to spontaneously and immediately name less frequently used words such as / ʕinæb / "grapes", /ʕigæl/ "men's head wear", /kæˈnæbæ/ "couch/sofa" while other children needed some time to think of a response or needed verbal prompting, such as a description of the item or producing the initial syllable of the word.

On the other hand, some children were not able to name a picture, and needed to repeat after the examiner, and then were later asked to name the picture to ensure the child had stored the word. Therefore, the word stimuli had two major confounds that could have resulted in the large individual variation in performance (as seen by their wide standard deviations); first, the frequency by which a word occurs and is stored within the child's lexicon is a factors that could have affected children performance; especially during this vital developmental age where language acquisition. Second, the list of stimuli words did not account for word frequency or age of acquisition.

Within the speech processing model of Stackhouse and Wells, it was assumed that when children repeat RWs they utilize the pre-existing lexical rout and benefit from top-bottom processing. This will be more likely with high-frequency words were existing motor programs are already stored within lexical representations. However, the child does not have to use existing motor programs and can treat the word as an unfamiliar item and create a new motor program to repeat it. It is possible that words that were low-frequency or occurred less frequently in the child's environment were less likely to be repeated using an existing motor program and were therefore treated as a NW. This could explain the 3-year olds large variation in performance of both RWs and NWs.

The insignificant developmental sensitivity could have been the result of the items also, as was mentioned early in this section, complex syllable structures, including geminate structures (as in the trisyllables), final consonant clusters (such as in the monosyllables) are acquired very early in language development are present in the templates of Arabic-speaking children form the age 3 years old (Amayreh, 2003; Ghada Khattab & Al-Tamimi, 2013).

On striking finding from this study was that monosyllabic items of RWs and NWs were found to be as challenging as their trisyllabic items. This finding was not consistent with what has been found in the literature on English-speaking children and other languages (e.g. Chiat & Roy, 2007; Roy & Chiat, 2004; Vance et al, 2005). As reported earlier, they could have been a result of methodological or linguistic differences.

In terms of the speech processing model; the three repetition tasks require a speech output acoustic signal, therefore all tasks share two levels of processing. They involve access to a) motor planning where the gestural targets are assembled in correct sequences and then b) the plan is accurately executed at the motor execution level which gives rise to the acoustic signal. Increasing syllable number and the inclusion of consents clusters both presumably require a more elaborate articulatory plan. The trisyllabic items have more complex articulatory sequences where the child has to perform rapid articulatory adjustments, on the other hand the monosyllables included four items with consonant clusters, therefore, also increasing processing load, nevertheless, there were six other items with a simple CVC structure (/b æ b/, /d ʊ b/, /f i: l/, /z ɪ r/, /m ɔ z/, /j ɛd/). Explaining, the results is not straightforward, and could be due to many factors. It could be that children performed well on the trisyllabic items and monosyllabic items with consonant clusters, as Arabic children acquire a wide range of complex syllables structures from an early age and would consonant clusters are suggested to be mastered by 4 years old (W. Ammar, 1999). On the other hand, it could be suggested that children showed speech/phonological errors on both trisyllabic items and monosyllabic consonant cluster items and items with later acquired phonemes (such as /s/ and /z/ in /z ɪ r/, /m ɔ z/), which are acquired around 4 and 5 years old (M. Amayreh & A. Dyson, 2000)

CHAPTER 7 RESULTS AND DISCUSSION II: RAPID CONSECUTIVE PRODUCTIONS – SPEECH MOTOR PERFORMANCE

7.1 Results 3: Performance on Rapid (Multiple) Consecutive Productions

Children's overall performance on rapid consecutive productions – speech motor tasks are presented and analyzed in this chapter. To answer the research question, this section is organised based on the behavioural measures used in this study for rapid repeated productions, namely, accuracy and consistency of performance. The research question was:

Research questions addressed:

Question 3: How do typically developing Arabic-speaking children perform on the – rapid consecutive repetition/speech motor - tasks (real-words vs non-words vs syllable sequence) using measures of accuracy and consistency?

- **a)** Are accuracy and consistency of responses affected by the test conditions (real word vs non-word vs syllable sequence)?
- **b)** Are there age-related (developmental) changes in children's accuracy, consistency performance on the DDK tasks (real word, non-word and syllable sequence repetition)? i.e., are the tasks developmentally sensitive when accuracy and consistency of repetitions are measured?

7.1.1 *Accuracy of Consecutive Repetitions (Ar): Performance across Test Conditions (RWs, NWs, SS) and Developmental Change*

This section tests weather children's accuracy of rapid consecutive productions differs between tasks, and whether developmental progression is evident between the ages of 3 to 5 years. Descriptive statistics of the total average proportion of correct (Ar) responses on the three tasks (RW, NW, SS) in each age group are presented in Table 7.1 and illustrated in Figure 7.1.

Table 7.1: Means, standard deviation (s.d) and range on Accuracy of consecutive repetitions (Ar) for the three test conditions in each age group.

		Age group		
Tasks		**3 year olds**	**4 year olds**	**5 year olds**
Real words	Mean (s.d)	**.67 (.2)**	**.83(.1)**	**.95(.05)**
	Min-Max (Range)	.29-.89	.79-1	.86-1
Non-words	Mean (s.d)	**.60 (.2)**	**.79 (.1)**	**.90(.05)**
	Min-Max (Range)	.27-.88	.74-1	.84-1
Syllable sequences	Mean (s.d)	**.77(.14)**	**.87(.1)**	**.97(.04)**
	Min-Max (Range)	.55-1	.82-1	.89-1

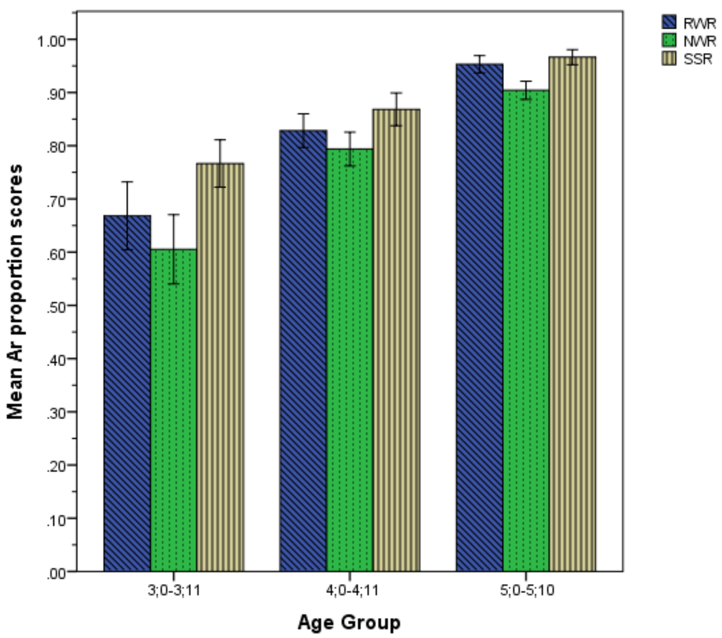

Figure 7.1 Mean proportion scores of Accuracy of consecutive repetitions (Ar) by age group and test condition. Error bars represent 95% Confidence interval.

Visual inspection of the means show that the children in the three age groups were overall more accurate when repeating stimuli consecutively on the SSR task and were least accurate on the NWR task, with RWR scores falling between SSR and NWR tasks. From a developmental perspective, the observed Ar means scores seem to increase with age and 5 year old children showed overall the highest Ar scores, with scores reaching ceiling, especially on the SSR task. Further analysis was carried out to investigate the significance of the effects of test condition and age on children's accuracy of consecutive repetition scores using a mixed ANOVA; 3 (test condition: RWR, NWR, SSR) × 3 (age group: 3, 4, 5). Mauchly's test indicated that the assumption of sphericity was met for the main effects of test condition, χ^2 (2) =2.3, *p*= .312. Results showed that there was a significant main effect of test condition F (2, 202) =169, *p*< .001 and age F (2,101) = 45.8, *p< .001*, also, there was a significant interaction between test condition and age F (4, 202) = 16.3, *p*< .001. This indicates that children's Ar performance on the test conditions differed with age.

To examine this interaction, the effects of test condition in each age group were investigated using Bonferroni's post-hoc comparison. Results showed that in the 3 and 4-year old age groups, children were significantly more accurate when consecutively repeating SS compared to both RW and NW (4-year olds showed marginal significance *p =.052)*. Interestingly, for both the groups there was only marginal significant difference between the RWR and NWR tasks (all *p's* = .05). Although there were marginal significant differences, the following trend was observed for test conditions (SSR> RWR > NWR). On the other hand, 5-year old child showed no significant difference in Ar scores on the SSR and RWR tasks (*p* > .05); however their performance on both RWR and SSR were significantly better than the NWR task (*p*< .01). Mean differences between the tasks and *p* values are summarised in Table 7 below.

When it comes to age-related changes on Ar performance, the Games-Howell post-hoc comparison showed that for the three tasks, RWR, NWR and SSR; there was a significant difference between the age groups, were 3 year old children showed significantly lower Ar scores than 4 year olds; and 4 year olds were significantly less accurate at consecutive repetitions than 5 year olds (all *p's* < .01). The following developmental trend was observed (3 <4 <5) A summary of mean differences and *p* values are provided in Tables 7.3 and 7.4.

7.1.2 *Consistency of Consecutive Repetitions (C) measure: Performance Across Test Conditions (RWs, NWs, SSs) and Developmental change*

A closer inspection of the data provided in Table 7.2 and Figure 7.2, suggests that children's consecutive repetitions in general were highly consistent; and that the children in the three age groups performed better on the SSR task, where their scores reached ceiling, especially the 5-year olds, and were least consistent on the NWR task. Effects of test condition and age were analysed with a 3 (Test condition) × 3 (Age group) ANOVA. Results revealed that there was a significant main effect of test condition, *F* (1.88 , 190)

= 146.4, p < .05. This indicated that overall, children's consistency was affected by different test conditions. Furthermore, there was a significant main effect of Age F (2, 101) = 54.7, p< .001, indicating that there was a developmental change in consistency of repetitions.

Table 7.2: Means, standard deviation (s.d) and range of Consistency scores (C) for the three Test Conditions in each Age group.

		Age group		
Tasks		**3 year olds**	**4 year olds**	**5 year olds**
Real words	Mean (s.d)	.87 (.08)	.95 (.033)	.98 (.02)
	Min-Max (Range)	.75-.96	.88-.99	.93-1
Non-words	Mean (s.d)	.80 (.1)	.92 (.04)	.94 (.04)
	Min-Max (Range)	.64-.93	.87-.98	.88-1
Syllable sequences	Mean (s.d)	.93 (.06)	.96 (.04)	.99 (.01)
	Min-Max (Range)	.86-1	.88-1	.98-1

There was also a significant interaction effect between the age of the participants and the test condition F (3.7, 109) = 19, p <.001. First, the effects of test condition was investigated. Bonferroni post-hoc comparisons revealed that although children were slightly more consistent on the SSR than the RWR task, the difference was not significant (p's >.05). However, although RWR was more consistent than NWR, only 4 and 5-year old age groups showed a significant difference between the tasks, while the 3-year old children did not show significant differences between the tasks. The insignificant difference between the RWR and NWR in the 3-year old age group was unexpected. Although the means evidently are different, the insignificance between the RWR and NWR on consistency measure could be due to children's larger standard deviations compared to RW and SS. Furthermore, the large deviations from the mean on the NWR could have resulted in the confidence interval (CI) of the NWR and RWR to overlap. The CI of NWR slightly overlapped with the mean of the RWR (see Figure 4), the mean difference between RWR and NWR = .067; CI (95%) - .004 to + .138), thus CI included a zero indicating no difference between the means and an insignificant result.

Analysis of effects of age and test condition was carried out using Games-Howell post-hoc comparisons. Results showed that there were no significant difference between the 5 and 4-year old children on all the three tasks, RWR, NWR and SSR (all p's >.05). However, 3 year old children were significantly less consistent than the 4 and 5-year olds on the RWR and NWR tasks (p*'s* < .05), while on the SSR task, the 3-year olds were only significantly less consistent than the 5-year olds, but not the 4-year olds. For post-hoc comparisons of mean differences between the tasks see Table 7.3 and between age groups see Table 7.4.

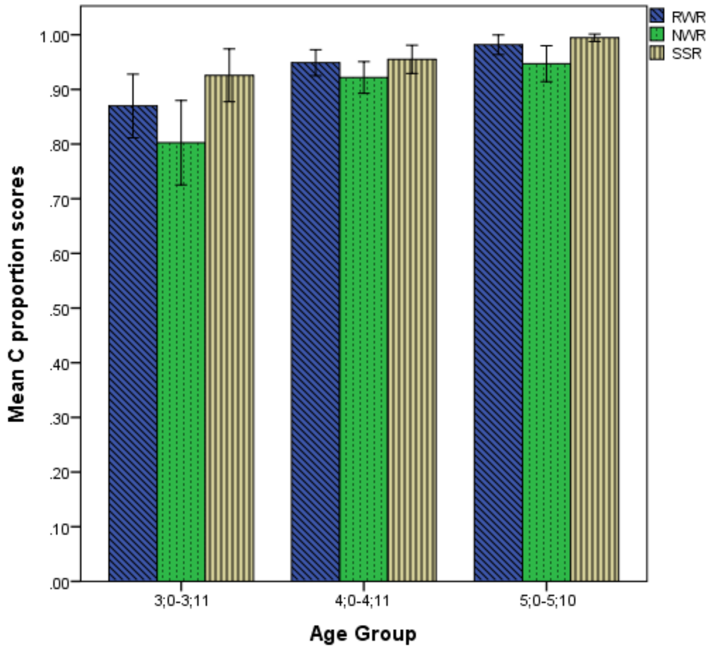

Figure 7.2: Mean proportion scores of Consistent repetitions (C) by test condition and age group. Error bars represent 95% Confidence interval.

Table 7.3 Bonferroni post-hoc comparisons of within-group differences on test conditions - (speech processing within each age group) for accuracy and consistency scores.

		Tasks					
		SSR > RWR		**RWR > NWR**		**SSR > NWR**	
3-year olds	**Accuracy Ar**	.10	**p*=.016	.063	^*p*=.055	.161	**p*=.003
	Consistency	.056	*p*=.10	.067	^*p*=.064	.123	**p*=.001
4-year olds	**Accuracy Ar**	.040	^ *p*=.052	.034	^*p*=.055	.075	***p*=.000
	Consistency	.01	*p*=.33	.03	* *p*=.016	.04	*p*=.085
5-year olds	**Accuracy Ar**	.02	*p* =.26	.049	* *p*=.004	.068	**p*=.001
	Consistency	.013	*p*=.23	.035	* *p*=.036	.047	**p*=.017

P value: ^ = marginal significant difference, * = significant difference at p< .05, ** significant at p<.001

Table 7.4 Between-group post-hoc comparisons for accuracy and consistency scores on each of the test conditions; i.e. developmental progression on each task.
P value: ^ = marginal significant difference, * = significant difference at p< .05

		Tasks					
		3 < 4		**4 < 5**		**3 < 5**	
RWR	Accuracy Ar	.16	**p*=.039	.12	* *p*=.04	.28	* *p*=.000
	Consistency	.08	**p*=.005	.03	*p*=.4	.11	* *p*=.000
NWR	Accuracy Ar	.18	*p*=.002	.11	**p*=.02	.29	* *p*=.000
	Consistency	.12	*p*=.00	.02	*p*=.1	.14	**p*=.000
SSR	Accuracy Ar	.11	*p*=.01	.1	* *p*=.02	.2	**p*=.01
	Consistency	.03	*p*=.4	.03	*p*=.1	.055	**p*=.006

Summary of Results:

Speech output Processing skills:

Age group	Accuracy Ar	Consistency C
3 years	SSR > RWR = NWR	SSR = RWR = NWR
4 years	SSR $>^{\text{marginal } p=.052}$ RWR $>^{\text{marginal } p=.050}$ NWR	SSR = RWR > NWR
5 years	SSR = RWR > NWR	SSR = RWR > NWR

Developmental progression on accuracy scores:

3 < 4 <5

Developmental progress on consistency scores:

RWR and NWR → 3< 4 =5

SSR → only sig between 3 < 5

7.2 Discussion of Research Question 3

The results presented above focused on the overall accuracy and consistency scores of rapid consecutive repetition i.e., speech motor tasks. It was attempted to examine the effects of speech processing demands and age on speech motor performance as measured by accuracy and consistency. Overall, several key findings emerged from the results, replicating cross-linguistic evidence and supporting predictions. As with the presentation of the results, accuracy of consecutive repetitions will be discussed first and subsequently followed by consistency of productions.

7.2.1 *Accuracy of Rapid Productions*

First, accuracy of rapid consecutive repetitions was affected by test condition i.e. speech processing demand, where producing a stimulus multiple times consecutively and at speed resulted in different profiles of performance on speech tasks compared to single repetitions. This replicated findings from Williams and Stackhouse’s (2000) study, where task effects were evident with rapid consecutive repetitions, but not with single repetitions. However, the findings from the Arabic-speaking children did not replicate the pattern of performance seen in the English-speaking children in Williams and Stackhouse's study.

Williams and Stackhouse (2000) found differences between NWR and RWR only in 3 and 4-year olds, whereas 5-year olds were at ceiling with no task effects. On the other hand, the present study found that both 3 and 4-year olds performance on rapid multiple productions of RWs and NWs were similar to their performance on single repetitions, where NWR scores fell only marginally below RWR scores (p=.05). The findings suggest that even with increasing processing demands on the motor levels of planning and execution, children at this age are arguably using an interaction of both top-bottom and bottom-up processing i.e., presumably using both stored motor programs and creating new ones. However, when there was no reference to linguistic knowledge, and linguistic influences were at its minimal, SSR task was the least demanding in terms of processing load at this age group. Furthermore, interestingly, the 3 year olds were significantly more accurate at the SSR task compared to the more linguistically demining RW and NW repetition tasks. This suggests that at the age of 3, constrains of linguistic processing demands could have interfered with their performance.

A slightly different profile appeared with the older age group (4;0-4;11), NWR scores fell marginally below RWR (p=.05), while the difference between RWR and SSR was shrinking, resulting in a marginal significant difference (p=.052). This suggests that around the age of 4, motor processing demands of the tasks start to differentiate and top-bottom vs bottom-up processing rout begin to operate. Children rely on their stored motor programs - where the assembled specifications of articulatory gestures are stored - to support accurate pronunciations of the RW, and trigger that program rapidly and repeatedly. In contrast, NWR becomes an increasingly challenging task, as there is no reference to stored motor programs and children had to devise a new motor program and trigger it repeatedly. This processing rout was also seen in the 5-year old children, even though there scores were at ceiling, the difference between the RWR and NWR reached significance.

7.2.2 *Consistency of Rapid Productions*

Regarding consistency of rapid consecutive productions, the findings showed that the youngest group of children (3;0-3;11) did indeed show inconsistency, however, this remained relatively low. The children were able to produce items consistently with an average proportion of 0.87 on RWs, 0.80 for NWs, and 0.93 for SSs. By 4 years of age and older, children were highly consistent with ceiling proportion scores at > 0.94 on all tasks. Interestingly, the level of intraword consistency found in the current study highly confirmed to that found in Holm et al.'s (2007) study, in their study the majority of children's productions were consistent, where children aged between 3;0-3;11 were approximately 87% consistent and 4 year old children and older were > 95% consistent (Holm et al., 2007, p.478). Remarkably, despite the linguistic and methodological differences of the two studies, the general proximity of consistency levels is striking. In Holm et al.'s study, words were elicited by asking the child to name pictures or repeat words on three separate occasions during a single session, while the present study was a speech motor (DDK) task with high levels of motor planning and execution demands, where rapid consecutive productions were elicited after repetition. In the more methodologically comparable study of Williams and Stackhouse (2000), where rapid consecutive productions were scored for consistency, direct comparison to the current study was not straightforward, as scores on the RWR, NWR and SSR tasks were reported as the number of items produced consistently rather than percentages. However, generally as with the Arabic-speaking children from the current study, the English-speaking 3, 4 and 5 year old participants in Williams and Stackhouse's study were very consistent in their productions. In their study, children's consistency scores were averaged on all items of the three tasks (RWs, NWs and SSs), and consistency levels of 84.5% were found for 3 year olds, this was close to the 3 year olds in the current study. On the other hand consistency was 89% for the 4 year olds and 91.6% for the 5 year olds, these overall scores were close, although lower, than scores from the Arabic-speaking children in the current study (scores on each task were >0.94). This small discrepancy could be due to the linguistic stimuli and item length selection; the Arabic items contained monosyllabic items along with bi and trisyllabic items while the items included in Williams and Stackhouse's study contained 10 bisyllables, 6 trisyllables and did not include monosyllables.

Furthermore, an important aim of the present study was to investigate the effects of different processing demands on consistency of productions The effects of processing demands on production consistency were interesting, showing that: first, there were no differences between RWR and SSR on consistency of productions in all age groups, suggesting that within the course of normal development, the increased speech motor processing load of rapid productions, did not affect production behaviour (similar to immediate single repetitions, RWR and SSR were not significantly different). Within the speech processing model of Stackhouse and Wells (1997), children were therefore able to first access their stored motor programs, plan and execute that program in real time to produce rapid articulatory adjustments of RWs as consistency as they would with planning and executing linguistic sounds and syllable sequences. Second, with regards to RWR and NWR, the 3 year old age group showed similar profiles of performance on consistency of productions with only marginal (p=.064) differences; thus, both accuracy and consistency scores of rapid speech output at this age did not differentiate processing demand of RWs and NWs. On the other hand, although the 4 and 5 year old children showed celling scores on consistency of rapid productions, the distinction in consistency of productions of RWs and NWs were significant. This difference was not evident with single immediate repetition (discussed in Chapter 6 Section 6.2). The findings were similar to 4 year olds found in Williams and Stackhouse's (2000) study, were scores on rapid consecutive productions of RWs were greater than NWs. On the grounds of the speech processing model, NWR and RWR tasks – which were matched on consonant, stress and structure – share the processing skills of input and output motor skills of planning and execution. However, the skills required to consecutively produce a NW was to create a motor program and trigger that program repeatedly at speed, while the processing requirement for the RWs were minimised due to existing motor programs. Therefore, it could be argued that with the stimuli used in the present study, 3 year old children's underspecified motor programs affected children's performance on the RW repetition task and therefore, the children could have used the rout of motor programmes to create RWs.

Another goal of the study was to document the effects of age on children's speech motor performance (as measured by accuracy and consistency) during the repetition tasks that included stored linguistic stimuli with prior experience producing a word, novel phonological sequences and sounds and sound sequences. With accuracy scores, developmental progression was evident between groups on all tasks 3< 4< 5; with 5 year old children reaching ceiling (proportion scores on all tasks > 0.90). On the other hand, children were highly consistent compared to accuracy scores and developmental progression was evident between 3 and 4 year olds on RW and NW tasks; scores reached ceiling by the age of 4. These findings were comparable to earlier studies on the effects of age on accuracy and consistency of productions and met the predictions (Holm et al., 2007; Macrae & Sosa, 2015; Williams & Stackhouse, 2000).

The literature reviewed in Chapter 2 Section 2.3 highlighted the potential sources of inconsistency found in young children, which included linguistic and motoric-based sources and the overlap between them. Although the present study does not directly address the sources of speech inconsistency, the tasks presented to children in the current study are speech motor - DDK – tasks in nature, that require oromotor/articulatory control, and therefore both linguistic and speech motor systems in young children should be discussed as potential sources of their reduced consistency.

The present findings of reduced consistency scores in the youngest age group (3;0-3;11) compared to the older groups, and the developmental increase in consistency scores on the speech motor tasks, support existing motor behaviour theories and kinematic studies. Generally, young children demonstrate movement variability and reduced control and coordination, which is considered an important and normal aspect of motor control development (Green &Nip, 2009; Kent, 1992; Thelen, 1991; Thelen, 1995). More specifically, speech motor movement variability is greater in young children compared to adults during repeated production of linguistic units, and is characterised by increasing accuracy and speech movement consistency of the articulators which continue to increase up to adolescence (Kent, 1992; Smith & Goffman, 1998).

Kent (1992) suggests that inconstancy of repeated productions of linguistic units is rooted partly, if not largely, to speech motor control immaturity; based on this view, the present findings could possibly suggest that constrains in oral motor and articulatory control skills of the 3-year olds compared to older children limited their ability to maintain consistent rapid productions. On the other hand, recall from earlier literature review that although constrains in oro motor control skills could be one interpretation of reduced segmental consistency, emerging cognitive and linguistic skills act as catalysts to speech motor skills (Green & Nip, 2009) and linguistic processing demands "top-down" processing influences motor execution movement (see Smith & Goffman, 2004). The evidence from the present study reinforces this proposition based on the following observations.

First, 3 year old children were significantly less consistent on RWs and NWs compared to the older groups on speech motor tasks. It is possible that at this stage of development, where cognitive and linguistic skills are rapidly accelerating, these oromotor and articulatory control and movement coordination constrains could have restricted accurate and consistent productions. Evidence from the literature shows that a dip in in lip and jaw movement stability was observed in children at the age of 2 (Green, Moor & Reilly, 2002), and research has also shown that peaks in lexical inconsistency (as measured by whole-word variation) are observed in typically developing children at 2-year of age i.e. during the vital period of rapid language acquisition, phonology and vocabulary growth (Sosa & Stoel-Gammon, 2006). Figure 7.3 below shows a schematic presentation of the developmental time course of oromotor movement viability and lexical consistency based on empirical evidence. It is noted that the dip in children's oromotor control and lexical consistency at 2-year of age, which gradually but significantly increases. Therefore, after the dip, the gradual and steady decrease in movement variability are observed along with gradual increase in lexical consistency is assumed to continue up to the age of 3.

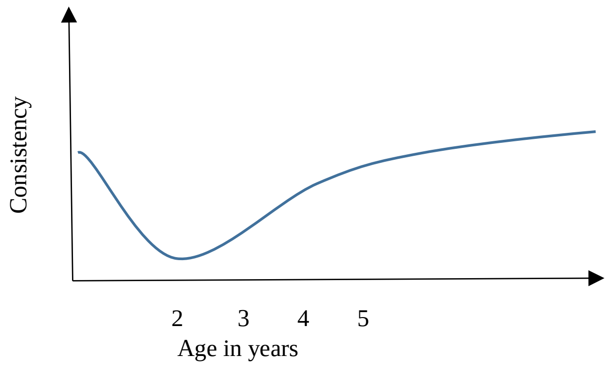

Figure 7.3: schematic presentation of the development of oromotor movement stability and lexical production consistency.

A second more compelling observation was the SSR task scores; there was no developmental progression between 3 and 4-year old nor between 4 and 5-year old children on the SSR task on consistency measure (nor accuracy). Therefore, lower execution level "bottom-up" processing did not affect children's performance, rather, that influences of linguistic and cognitive "top-down" processing on speech motor behaviour were most influential and sensitive to developmental change.

Caution should be made when drawing conclusions regarding speech motor control system immaturity influence of the task. The presence of inconsistency in rapid repeated productions should not be taken as evidence of motor immaturity (or motor disorder). There is evidence that the relationship between motor control output and word/segmental accuracy and variability is complex, and they do not directly or automatically coincide. Earlier work by Goffman and Smith (1999) showed that high spatial and temporal movement variability in young children (compared to older children and adults) as recorded using instrumental kinematic methods was present even when children produce an accurate linguistic target; and that word inconsistency during repeated productions does not necessary imply variability in motor movement implementation (Goffman, Gerken & Lucchesi, 2007). Models of speech processing and speech production should account for complex and bidirectional interactions across motor and phonological levels that, though systemic, are multidimensional, as demonstrated by the different roles of representation and execution levels of processing.

Furthermore, the current findings of reduced accuracy and consistency in 4 and 5 year olds on the NWRs compared to RWRs could arguably be interpreted from a kinematic stance of dynamic speech motor learning. In the present study both real words and non-words were matched on consonants, syllables and stress, and only the vowels of the real words were changed to create the non-words, therefore, it is suggested that the items are matched in terms of their kinematics. Within the concept of transfer of learning or generalization of dynamic learning of motor control (introduced in literature review, Chapter 2 Section 2.3.22), speech motor control is highly specific and learning speech fails to generalise, speech motor learning is local and specific. Even with the matched stimuli, the subtle changes in articulatory movement of the novel words could have resulted in reduced accuracy and consistency of speech motor control. This suggests that there was no transfer in speech motor learning or a generalisation of dynamics learning even when items have similar articulatory movements, therefore supporting, at least partially, the evidence that speech motor control for a new utterance/word is highly local and specific and that board generalising of dynamic representations do not apply to speech motor control (Tremblay, Houle, & Ostry, 2008).

In summary, the current findings do not explain or directly address sources of token to token inconsistency during the period of language and phonological acquisition which have been attributed to factors such as unstable or incomplete underlying phonological representations and lack of segmental detail, and immaturity of speech motor control. However, the findings do suggest that token to token inaccuracy and inconsistency identified using broad phonemic transcription is a common feature during rapid motor productions (kinematic tasks) in typically developing Arabic-speaking children. It does not necessary indicate underlying linguistic-phonological or motor planning deficits, rather, it may reflect the underlying processes and motor skills maturation during typical speech development, such as the gradual acquisition of holistic representations or production strategies used by children when attempting to produce complex phonological sequences.

7.3 Results 4: Effect of Stimuli length on Rapid Consecutive Productions

The section will present the results on the effects of increased item length of each test condition RWR, NWR and SSR and age on children's accuracy and consistency performance. This section answers the questions related to the accuracy of single repetitions when increasing stimuli length and the developmental sensitivity of each length, the questions were:

Research questions addressed:

Question 4: how do children perform on accuracy and consistency of rapid consecutive repetition- tasks when stimuli length increase?

- **a)** Does stimuli length of the test conditions (monosyllabic vs bisyllablic vs multisyllabic stimuli) affect accuracy and consistency of responses?
- **b)** Are there age-related (developmental) changes in children's accuracy and consistency performance when length of stimuli increases? i.e., is stimuli length developmentally sensitive when measured by accuracy and consistency?

7.3.1 *Accuracy of Consecutive Repetition (Ar): Performance across Test Conditions Stimuli Lengths (mono vs bi vs trisyllables) and Developmental Change*

Descriptive statistics for each test condition based on syllable length are provided in Table 7.5. Visual inspection of the data show that generally, children performed better on the bi-syllabic items more than the mono and their least accurate repetitions were on the tri-syllabic items. In addition, 5-year old children showed higher scores on all syllable lengths of the three tasks, followed by the 4 year olds, with 5-year old children reaching ceiling on all syllable lengths. Further analysis was carried out to investigate the effect of syllable length in each test conditions (RWR vs NWR vs SSR) and the possible effects of developmental change i.e. age (3 vs 4 vs 5-year olds).

Real Word Repetition there was a significant effect of length $F(1.7, 180) = 23.79$, $p < .01$, and of age $F(2, 105) = 48.8$, $p<.001$, but there was no significant interaction between them $F(3.43,180) = 1.16$, $p = .33$. The main effects of length was further examined using

Bonferroni post-hoc pairwise comparison. Overall, for the three age groups, bi-syllabic items were significantly more accurate with consecutive repetitions than both mono (mean difference=.08, $p < .01$) and trisyllabic RW's (mean difference=.09, $p < .05$). However, although children performed slightly better on the monosyllables and showed an overall higher Ar ranges (see Table 7.5 and see Figure 7.3 below.); the difference was not significant between the mono and trisyllables (mean difference =.01, $p = .61$). As for the simple effects of age, Games-Howell post-hoc comparisons revealed that there was a developmental change, only between the 3 and 4 year olds (man difference = .15, p =.05). Whereas there was no significant difference between the 4 and 5-year olds on all syllable lengths (mean difference = .12, p =.13) on all word lengths.

Table 7.5: Means, standard deviations (s.d) and ranges of Accuracy scores of consecutive repetitions (Ar) by different stimulus syllable lengths (mono, bi and tri syllables) in the three test conditions (real-words non-words and syllable sequences).

TASK Ar		Age group		
REAL WORD RW		3year old	4year old	5year old
RWR mono-syllables	Mean (s.d)	**.65 (.2)**	**.81 (.1)**	**.93 (.2)**
	Range	.30-.96	.70-1	.80-1
RWR bi-syllables	Mean (s.d)	**.76 (.22)**	**.88(.13)**	**.99 (.02)**
	Range	.36-1	.66-1	.90-1
RWR tri-syllables	Mean (s.d)	**.63(.23)**	**.80 (.14)**	**.94(.1)**
	Range	.20-.88	.64-.94	.68-1
NON-WORD NWR				
NWR mono-syllables	Mean (s.d)	**.62 (.2)**	**.80 (.13)**	**.87 (.12)**
	Range	.24-.94	.64-.98	.68-1
NWR bi-syllables	Mean (s.d)	**.67 (.2)**	**.88(.1)**	**.97(.04)**
	Range	.30-.90	.80-.98	.90-1
NWR tri-syllables	Mean (s.d)	**.55 (.1)**	**.76(.17)**	**.86(.12)**
	Range	.26-.86	.62-1	.60-1
SYLLABE SEQUENCE SS				
SS mono-syllables	Mean (s.d)	**.92 (.1)**	**.95 (.04)**	**.98 (.03)**
	Range	.80-1	.90-1	.90-1
SS bi-syllables	Mean (s.d)	**.83 (.17)**	**.92 (.1)**	**.99(.01)**
	Range	.50-1	.90-1	.97-1
SS tri-syllables	Mean (s.d)	**.55(.2)**	**.73(.15)**	**.93(.11)**
	Range	.34-1	.66-1	.70-1

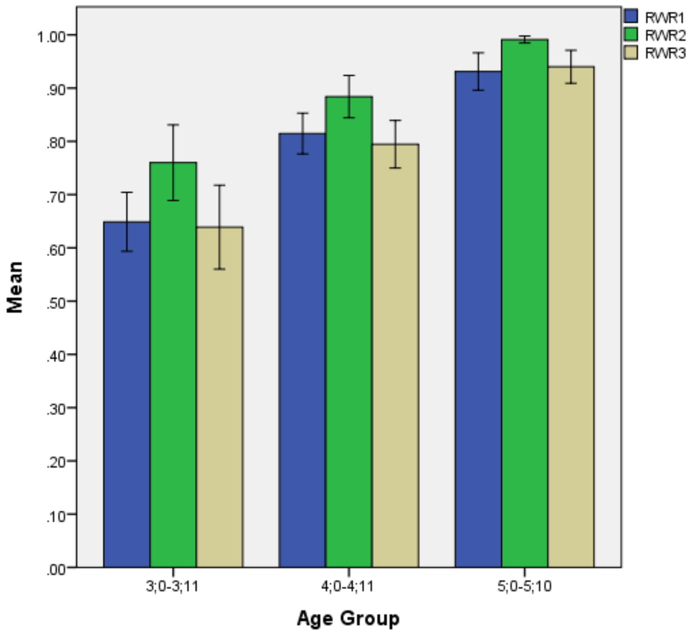

Figure 7.4: Mean proportion scores on Ar by age group and length of real-words. RWR1=monosyllables, RWR2=bisyllables, RWR3=trisyllables Error bars represent 95% Confidence interval.

Non-Word Repetition as with the RWR, there were both effects of length F (1.6, 170) = 39, $p < .01$ and effects of age F (2, 109) = 50, $p < .001$, with no significant interaction between them F (3.1,170) = 2, $p = .20$. Further analysis using post-hoc pairwise comparison revealed similar effects of length and age as with the RWR (see Figure 7.4 for a visual illustration). As for syllable length, bisyllabic items were significantly more accurate at consecutive repetitions than both mono and trisyllabic real words (mean difference = .075 and .12 respectively, all *p*'s < .01) and no significant difference between the mono and trisyllables (mean difference = .05, *p = .60).* Games-Howell post-hoc comparisons revealed that developmental change was significant only between the 3 and 4-year old children (mean difference = .19, $p < .05$) with no significant difference between the 4 and 5–year olds (mean difference = .11, *p* =.24) on all word lengths.

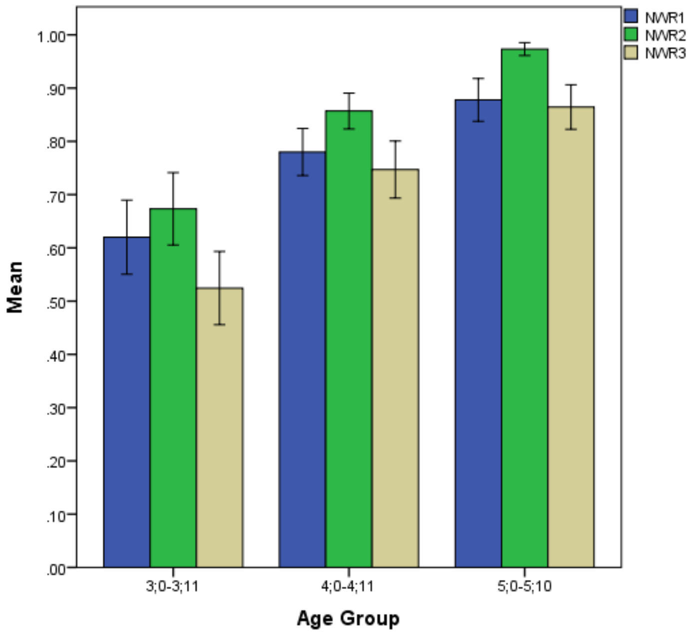

Figure 7.5: Mean proportion scores on Ar by age group and length of non-words. NWR1=monosyllables, NWR2=bisyllables, NWR3=trisyllables Error bars represent 95% Confidence interval.

Syllable Sequence Repetition there was a significant effect of length F (1.55, 156.8) = 174.8, and age F (2, 101) = 34.8. However, unlike the RWR and NWR, the length by age interaction was significant F (3.1, 156) = 26.3, (all *p's* > .01). That means that children's accuracy on the SS differed within each age group (see Figure 10 for a visual illustration). For the 3 and 4-year olds, pairwise comparisons revealed no significant differences between the mono and bisyllables (mean difference for the 3-year olds = .01 and for the 4-year olds= .03, *p's*> .05), while there were significant differences between the trisyllables and the mono and bisyllables, (for the 3-year olds: mean differences = .37 and .27 respectively. For the 4-year olds: mean difference =.22 and .19 respectively; all *p*'s <.01). There was no significant differences on the length of stimuli for the 5-year old children (all *p*'s>.07).

To break down the interaction of effects, post-hoc comparisons using Game-Howell correction revealed that there was no significant developmental difference between the groups on monosyllabic items (mean difference between the 3 and 4 year olds = .02, and between the 4 and 5-year olds =.04, all *p's* >.05). There was also no significant developmental change on the bisyllables, the only significant difference was between the 3 and 5-year olds (mean difference =.17, *p*<.05). As for the trisyllable sequences, there was a significant developmental difference between the groups, where 4-year olds were significantly more accurate than 3-year olds (mean difference= .18, $p < .05$), and 5-year olds were significantly better than 4-year olds (mean difference = .2, p=.48).

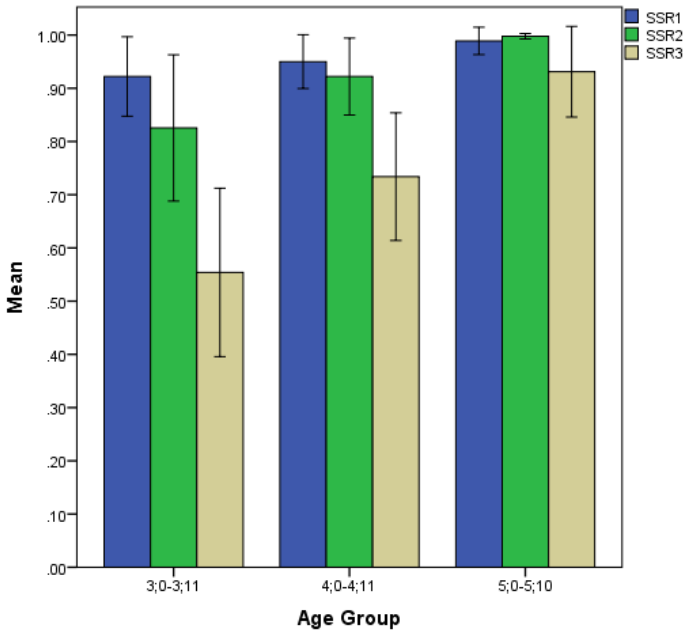

Figure 7.6: Mean proportion scores on Ar by age group and length of syllable sequences. SSR1=monosyllables, SSR2=bisyllables, SSR3=trisyllables. Error bars represent 95% Confidence interval.

7.3.2 *Consistency of Consecutive Repetitions (C) measure: Performance across Test Conditions Stimuli Lengths (mono vs bi vs trisyllables) and Developmental Change*

Analysis was carried out to determine whether there was an effect of stimuli length of each test condition using mixed design ANOVA with 3 (Age group) × 3 (stimuli length of test condition). Descriptive statistics are provided in Table 7.6 which shows that children were generally highly consistent, especially on the SSR task.

Table 7.6: Means, standard deviations (s.d) and ranges of Consistency of consecutive repetitions (C) by different stimulus lengths (mono, bi and tri syllables) in the three test conditions (real-words non-words and syllable sequences).

TASK (C)		Age group		
REAL WORD RWR		3year old	4year old	5year old
RWR mono-syllables	**Mean (s.d)**	**.86 (.06)**	**.94 (.04)**	**.97 (.04)**
	Range	.78-.94	.84-.98	.90-1
RWR bi-syllables	**Mean (s.d)**	**.90 (.07)**	**.97 (.03)**	**.99 (.02)**
	Range	.80-1	.88-1	.94-1
RWR tri-syllables	**Mean (s.d)**	**.83(.13)**	**.93(.05)**	**.98 (.02)**
	Range	.62-.98	.84-1	.94-1
NON-WORD NWR				
NWR mono-syllables	**Mean (s.d)**	**.76(.1)**	**.89(.05)**	**.90(.08)**
	Range	.65-.94	.78-.98	.84-1
NWR bi-syllables	**Mean (s.d)**	**.86(.11)**	**.94(.05)**	**.98(.02)**
	Range	.60-1	.86-1	.94-1
NWR tri-syllables	**Mean (s.d)**	**.78(.11)**	**.92(.07)**	**.96(.04)**
	Range	.58-.98	.82-1	.90-1
SYLLABE SEQUENCE SS				
SS mono-syllables	**Mean (s.d)**	**.97 (.07)**	**.1 (.00)**	**.1 (.00)**
	Range	.8-1	1	1
SS bi-syllables	**Mean (s.d)**	**.96(.02)**	**.99 (.01)**	**.99(.006)**
	Range	.88-1	.98-1	.98-1
SS tri-syllables	**Mean (s.d)**	**.83 (.14)**	**.87(.1)**	**.98(.03)**
	Range	.65-1	.76-1	.94-1

Real Word Repetition: There was significant main effect of syllable length F (1.8, 191) = 20.8, p < .001, and a significant main effect of age F (2, 109) = 62.6, p <.001, and no significant interaction between the length of the word and age F (3.6, 191) = 1.2, p <.001. Bonferroni post-hoc comparisons revealed that overall, children were significantly more consistent when repeating bi-syllabic RW items than mono-syllabic and tri-syllabic items (mean difference = .032, and .041 respectively both p's< .05). There was no significant difference in consistency between the mono-syllabic items and tri-syllabic items (mean difference= .01, p > .1). From a developmental perspective, post-hoc investigation showed that, as with Ar proportion scores of RWR, there were no significant differences between the 4 and 5-year olds on consistency scores (mean difference = .032, p >.1), and the only significant developmental change was seen between 3 and 4-year olds (mean difference= .08, p< .01). See Figure 7.6 for visual illustration.

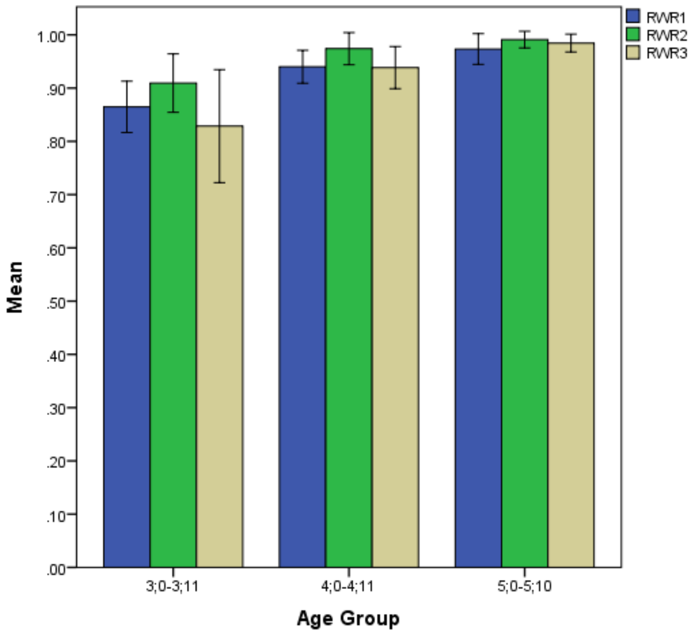

Figure 7.7: Mean proportion scores on C by age group and length of real-words. RWR1=monosyllables, RWR2=bisyllables, RWR3=trisyllables Error bars represent 95% Confidence interval.

Non-word Repetition: visual inspection of the data (Table 7.6 and Figure 7.7) indicate that the mean scores of the mono and trisyllabic items were very similar, although the younger group showed greater variation. Analyses of variance showed that there was a significant main effect of syllable length F (2, 202) = 34.3, p < .001, a significant main effect of age F (2, 101) = 47.8, p <.001. There was a significant interaction between the length of the non-word and age F (2, 202) = , p < .05. For the main effects of length, Bonferroni post-hoc comparisons revealed that overall, children were significantly more consistent when repeating bi-syllabic items compared to mono-syllabic (mean difference = .08, p < .05). Whereas, there was no significant difference between the bi-syllabic and tri-syllabic items (mean difference= .043, p >.05), nor was there a difference between the mono and trisyllabic items (mean difference= .036 p >.05). For the age and interaction effects, post-hoc Games-Howell comparisons revealed similar developmental trends as with the consistency scores on RWR, where that 3-year olds children were significantly less consistent at repetition compared to 4 and 5-year old children (mean difference = .12 and .14 respectively, all p's < .001). There was no significant developmental difference between the 4 and 5-year olds (mean difference = .025, p >.05), whereas 3-year old children were significantly less consistent (mean difference =.14 p >.05).

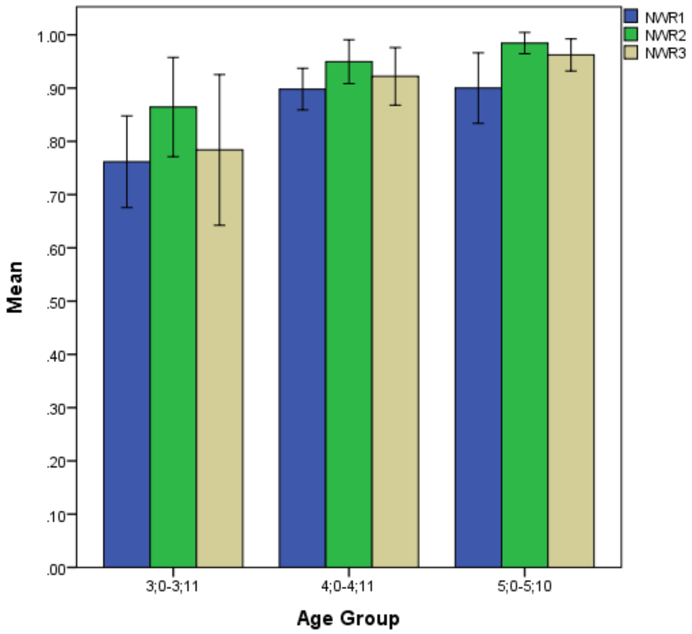

Figure 7.8: Mean proportion scores on C by age group and length of non-words. NWR1=monosyllables, NWR2=bisyllables, NWR3=trisyllables Error bars represent 95% Confidence interval.

Syllable sequence: a close inspection of the means (Table 7.6 Figure 7.8) show that all age groups were highly consistent on the SR task, with ceiling effects on the monosyllabic items while variation in consistency scores was seen on the tri-syllabic items. For the analyses of variance, Mauchly's test indicated that the assumption of sphericity has been violated χ^2 (2) =213, P< .05, therefore the Greenhouse-Geisser estimates were used (ε = .532). There was a significant main effect of syllable length F (4.06, 107) = 85.2, p < .001 and a significant main effect of age F (2, 101) = 23.1, p <.001. There was a significant interaction between the length of the syllable sequence and age F (2.1, 107) = 15, p < .01.

For the main effects of length, length of the stimuli of the SSR task did not affect consistency performance in the 5-year old children, as children performed at ceiling and there was no significant difference (all p's >.1). As for the 4 and 3-year old children, trisyllables were significantly less consistent than bisyllables (mean difference = .12 and .14 respectively, p's <.05). For the interaction effects, 5 and 4-year old children both performed at ceiling with no inconsistencies while reaping the monosyllable items, (mean difference = .00, p= 1), and they were both nnot signifcantley different than the 3-year olds (both mean differences =.02, p>.01). Similar result were seen for the bi-syllabic items where the children performed at ceiling and were not significantly different than each other (all *p's*>.1). On trisyllabic items, there was no significant difference on consistency scores between 3 and 4-year olds (mean difference =.04,p>.1. However, although the 4-year old groups were highly consistent, they were not as consistent as the 5-yaer olds which performed at ceiling, and marginal significance was observed (mean difference= .12, p=.069). The 3 and 5-year olds were significantly different (mean difference = .047, p<.05)

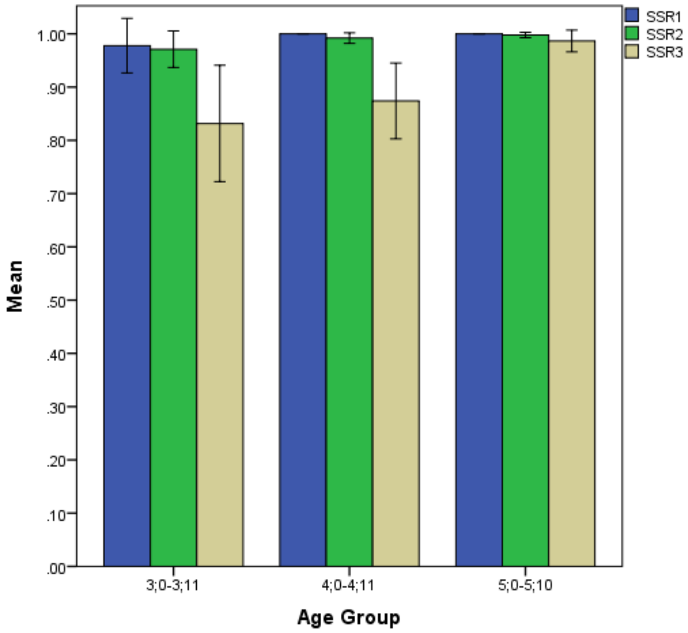

Figure 7.9: Mean proportion scores on C by age group and length of syllable sequences. SSR1=monosyllables, SSR2=bisyllables, SSR3=trisyllables Error bars represent 95% Confidence interval.

7.4 Discussion of Research Question 4

The Arabic-speaking children of this study showed similar profiles of performance seen in English-speaking children. The current data quantifies children's production accuracy and consistency, and describes developmental change with different item lengths. The children were able to produce a highly structured and constrained stream of acoustic energy, by rapidly changing the shape and position of articulators within the vocal tract.

An interesting observations, that although not the primary focus of the current study is worth highlighting, was children's performance on the monosyllabic word /ʃæms/ "sun" and its non-word /ʃims/ during rapid consecutive repetitions. The children in all age groups (3 to 5 year olds) found this item extremely challenging and both the accuracy and consistency of productions reduced. This word/non-word is a short monosyllable starting with a fricative and ending with a consonant cluster containing a voiced nasal occlusive and a voiceless coronal sibilant; therefore, to produce rapid consecutive productions, a child had to rapidly join a final consonant cluster fricative with another fricative. Within the psycholinguistic framework of Stackhouse and wells (1997) the ability to join words (particularly around word boundaries) in a cohesive utterance is a supra-lexical intonation feature called segmental juncture features. in English-speaking children it starts to emerge during the second year of life with the emerging two-word utterance, and continues to develop up to the age of 8 years old (Corrin, Tarplee, & Wells, 2001)

Although some authors have reported equivocal findings when comparing accuracy of non-words with difficult articulatory motor targets such as fricative and clusters, liquids to less demanding motor targets (Edwards & Lahey, 1998).

Furthermore, it was observed that some children adjusted their productions of the item to include a filler neutral vowel or CV syllable at the beginning of the item to reduce the challenge of continuous fricatives at the junctions (e.g., / ʃæms/ is produced with five consecutive repetitions as → [ʃæms.ɛʃæms.ɛʃæms.ɛʃæms.ɛʃæms]). Interestingly, filler vowel and syllables as an initiation strategy is not uncommon in children (Ghada Khattab & Al-Tamimi, 2013; Peters, 2001). Khattab and Al-Tamimi (2013) reported that young Lebanese Arabic-speaking children used initial vowel and CV (glottal stop followed by a vowe) syllable fillers and provided an example of the word /ʃæms/. Khattab and Tamimi suggest that these fillers could be used by young children springboard to initiate articulation or they could possibly be used as dummy syllables based on the definite article /ʔæl/ “the” that frequently occurs in Arabic language which is usually assimilated [ʔæʃʃæms] “the sun”. These suggestions could also be an alternative explanation of the filler behavior observed in children’s productions in the current study. In cases were CV was added, this would have resulted in a bisyllabic item, which is, as discussed in previous sections, the most frequent shape of the Arabic language and therefore highly prevalent in children’s productions form an early age. Children could have used this strategy to increase their accuracy and consistency of productions, as it would more likely to match their rich bisyllabic underlying phonological templates. This behavior has also been reported in Kattab and Al-Tammimi’s study, were the children in their study often adapted mono and multisyllabic words to bisyllabic shapes.

.

CHAPTER 8 GENERAL DISCUSSION & CONCLUSION

The present thesis was interested in the evaluation of speech output processing skills in 3 to 5 year old typically developing Saudi Arabic-speaking children, which was motivated by theoretical and empirical evidence. The main aims of this study was to address central questions of **a)** how speech processing demands of three tasks affects children's speech output performance, **b)** does the speech processing demands of the tasks capture developmental change in speech processing behaviour; and **c)** how does increased processing demands of length affect performance on the tasks. The aims were addressed by implementing the theoretical speech processing model of Stackhouse and Wells (1997) to design tasks for a cross-sectional study that includes children ages 3, 4 and 5 years.

The three output tasks used in the present study included real word repetition, non-word repetition and syllable sequences repetition. In chapter 6 the results examined whether speech processing demands affect typically developing Arabic-speaking children's performance as measured by whole word accuracy.

In general, many unpredicted findings emerged from this study, many of which were influenced by methodological and linguistic factors. In contrast, many findings were in line with cross-linguistic empirical evidence.

This final chapter will bring together chapters 6 and 7 by highlighting and discussing the main findings. This chapter will also include the studies limitations, directions for future research and theoretical and clinical implications.

8.1 Speech Output Processing Skills and Developmental Sensitivity

Within the psycholinguistic framework of Stackhouse and Wells (1997), tasks are used to illustrate the processing demands of the different levels. The current study devised a set of tasks and stimuli to tap different levels of speech output processing, the tasks were real words, non-words and syllable sequences; each task included 30 items. Within the speech processing model of Stackhouse and Wells, all three tasks share the skills that tap input levels, motor planning levels and execution. However, the central levels for repeating items of real words (familiar words), generally, would assess a child’s ability to access stored phonological knowledge and motor programs. Non-word repetition assess a child’s ability to create new motor programs (i.e. motor programming) and syllable sequence repetition assess a child’s ability to produce sounds and sound sequences at the motor execution level.

In this study, the output tasks did not generally differentiate between different levels of processing with the stimuli set that was used. In particular, the immediate repetition of real words and non-words were not sensitive to processing skills emerging between the ages 3 to 5 years. However, children aged 5 years old were at ceiling on their performance on the repetition tasks, which supports Arabic acquisition studies that suggest that most of the phonemes of the Arabic language are typically acquired by the time a child reaches this age (Amayreh, 2003). Children aged 4 years also reached high scores on the tasks, showing no discernible difference between tasks. Conversely, the 3 year olds—the youngest group in the study—showed large standard deviations (ranges) on the tasks; this was not observed with the older groups. Furthermore, developmental progress on the tasks was only evident between 3 and 5 year olds, which indicates that the immediate single repetitions children made in the tasks were not a sensitive measure to assess their developmental change according to the given stimuli.

The ceiling effects found in the typically developing children, particularly, the 5 year olds, may have been due to the simplicity of each task. This is perhaps due to the fact that each item was designed within the range and vocabulary of a 3 year old, and non-words matched the consonantal root of the real word. Therefore, ceiling effects were not entirely a surprising result among the older children, who also had the advantage of language experience. Further, since the tasks were easy for older children who had the advantage of greater language experience, it is not clear whether this factor sufficiently closes the gap between 5 year olds with speech difficulties and those without.

This study also investigated the effects of increased speech motor control demands, by using a diadochokinetic task (DDK) using measures of accuracy and consistency. The present study found that, by the age of 4, children performed similarly on the accuracy and consistency measures for both the real word and syllable sequences tasks. Therefore, it is clear that by that age, children were able to plan an accurate and consistent repetition, both when using an existing level of stored information (motor programs) and when no reference to lexical influence at the level of motor execution was provided. However, children younger than 4 (i.e. the 3;0-3;11) appear to have different speech processing profiles. Further, the results show that both real word and non-word repetition tasks were a developmentally sensitive measure of the DDK task at least between the ages of 3 and 4. Conversely, the syllable sequence repetition task was not a developmentally sensitive measure as age differences only occurred between 3 and 5 years old. The Arabic-speaking children of this study showed similar profiles of performance seen in English-speaking children. The current data quantifies children’s production accuracy and consistency, and describes developmental change; however, it does not explain it. Generally, all three tasks of real word, non-word and syllable sequences involved input and output skills; children’s overall performance on the tasks did not suggest that planning and executing a real word with complex CV structures and syllable stress had less of an advantage compared to syllable sequences that have simple-least linguistically loaded stimuli and less consonants. This could suggest that complexity and syllable stress did not necessary have higher processing loads on the motor planning and execution levels, rather this suggests an advantage of stored phonological/motor programs in a child’s productions.

Furthermore, maximum performance tasks, specifically the DDK task, are used to investigate speech motor skills in children, however, they are commonly used with school-aged children, and are not routinely used with preschool children in clinical settings. Consequently, there are relatively few studies that investigate the performance of young children on DDK tasks (Robbins & Klee, 1987; Rvachew, Ohberg, & Savage, 2006; Williams & Stackhouse, 2000). Rvachew and colleagues suggested a number of reasons for the apparent limitation of normative studies, such as that preschool-aged children generally lack the motivation to complete the rapid repetition tasks and have difficulty understanding the tasks and/or their responses are not reliable showing variability in responses from trial to trial (Rvachew & Brosseau-Lapré, 2012; Susan Rvachew, Hodge, & Ohberg, 2005). These factors may have influenced how the 3 year olds performed on the rapid tasks in this study, which might explain why the results from this age group associated with each task showed high variabilities (ranges). Rapid repetition tasks were highly demanding and challenging for the young developing child, on both higher language levels and lower motor execution levels. Tellingly, in response to the speech motor task, one 4 year old boy when instructed to rapidly and consecutively repeat an item five times said: 'I'm four years old, so I should only repeat four times'. His comment prompts one to consider whether consecutive productions should actually be tailored to a child's developmental stage of language acquisition.

For example, according to Brown's Stages of Syntactic and Morphological Development, Stage IV English-speaking children aged between 2;11 and 3;4 years would have a mean length utterance of 3.0–3.75, and children at Stage V, aged approximately 3;5–3;9 years, would produce a mean length of 3.75–4.50 utterances (McLaughlin, 1998; McLeod & Bleile, 2003). It is possible that a child who just turned 3 years should repeat items that equate to their age. This same method could also be considered when assessing a child with speech and language difficulties. Calculating the mean length of an utterance in accordance to a child's age could help to determine the number of repetitions they are expected to produce. However, conversational speech is entirely different from rapid consecutive repetitions and the elicitation of such responses in young children is fraught with difficulty. In fact, there is evidence to suggest that the DDK task may not reflect a child's speaking skills. For example, Haselager, Slis and Rietveld (1991) reported a low correlation between children's DDK rates and their articulation rate in conversation. Furthermore, Yaruss and Logan (2002) reported that typically developing children aged between 3 and 7 years produced numerous speech errors during the DDK task, which included consonant deletion, articulatory placement errors and voicing errors, affecting approximately 15 percent of the consonants. Overall, participants produced more errors during their DDK productions than is typically expected in young children's conversational speech. Their study suggests that the frequency of speech errors during rapid productions in a DDK task does not reflect a child's conversational speech, and a DDK task may be more difficult for young children than conversational speech. Hence, eliciting rapid productions in young children, particularly in those aged 3, presents more challenges for the child's developing oro-motor skills than previously presumed. Asking a young child to rapidly repeat items five times is not only challenging for their limited motor skills, it also tests their cognitive level and their attention span. For example, in the present study, eliciting five repetitions from children who were younger than 4 years old required more practice, feedback and reinforcements, especially that at a young age children are not skilled with counting the number repetitions using their fingers or by memory. On the contrary, testing the older age groups (especially the 5 year olds) was less challenging, as they stopped when prompted to, and who in instances, repeated items more than five times—in this case, only the first five consecutive repetitions were calculated. However, Rvachew, Ohberg and Savage (2006) reported that typically developing preschool children in their study (which included 20 children aged between 4 and 6 years) were able to complete all tasks of maximum performance—including prolongation and the DDK task. Rvachew et al. (2006) indicated that these tasks could be applied to preschool children; however, it should be noted that, unlike the present study, theirs did not include children younger than 4 years old.

Furthermore, with regards to the syllable sequence stimuli design, there was the possibility of a slight bias in the stimuli presented to the children. The monosyllables had less consonants compared to the other tasks (real words and non-words), therefore, this reduced the possibility of inaccuracies to occur at the execution level of processing. Moreover, the syllable sound and sound sequences in the task was a simple CV structure, this sequence did not challenge children’s speech perception, recognition, memory or even their planning levels. This simple sequence could have resulted in the ease of processing load on motor planning when rapid consecutive repetitions were required. This concept supports *the ease of articulation construct* (e.g., Locke, 1972) where it is suggested that children only had to plan movements for the same consonant place and manner (homorganic) when repeating the same simple CV sequence.

Increasing processing demands or what is referred to as performance load is considered informative. Arabic-speaking children produce many bisyllabic word shapes and acquire a range of complex syllable structures from an early age due to their frequency in adult input (Khattab & Al-Tamimi, 2013). Therefore, the frequency by which bisyllables occur in Arabic and their dominance in the mental lexicon of the Arabic-speaking child from an early age, results in this form being accessed more accurately. Although one could argue that a repetition task increases the accuracy of productions, and from a speech processing perspective, children could use bottom-up processing to produce the words, the fact that bisyllabic words were significantly more accurate in all age groups supports the proposition that phonological structure of words are language-specific phonotactic templates (Vihman & Keren-Portony, 2013). These findings give insight into the construction of templates of Arabic-speaking children and offer a window into their phonological representations, and supports other studies on Arabic-speaking children (e.g., Shaalan, 2010) who conducted a study on a small sample of 11 typically developing Qatari Arabic-speaking children (ages 5;0-6;9) and found that syllable length and consonant clusters significantly affected children's performance.

With regard to the rapid sound and syllable sequences production task, trisyllabic syllable sequences (e.g., / bæ tæ qæ/) showed sensitivity to length as they showed reduced accuracy and consistency scores and were sensitive to developmental change (3 < 4 < 5). The syllable sequence repetition task was not sensitive to increasing processing demands of item length between one and two syllables. The simple CV and CVCV structure was not challenging enough for children and no significant difference on accuracy scores were noted with single repetitions. The three syllabic items however, captured processing behaviour changes, this could have been a result of motor planning challenges. This was confirmed by children's performance on the rapid consecutive repetitions when measured by accuracy and consistency of repetitions.

One striking observation was that Arabic-speaking children generally showed greater facility at repeating bisyllabic real words and non-words both at a single repetition level and when produced multiple times. This was observed for both measures of accuracy and consistency. While mono and tri-syllabic items were equally challenging to the children. In contrast to single repetitions, rapid multiple repetitions i.e. DDK tasks were sensitive to developmental change.

The children's early phonological representations are assumed to be holistic and reduce as the child gains segmental detail (this comes with growing vocabularies), thus consistency increases. With non-word repetition, studies show that by 5 years of age the child has gained sufficient vocabulary size and the child can use lexical and sub-lexical information to facilitate non-word repetition accuracy performance. This could mean that the child has sufficient segmental detail and thus have sufficient detail in their phonological representations to extract and assemble the new novel string and plan a new program. So presumably, by this age there should be no difference in consistency between non-word repetition and real word repetition. These findings from the different typically developing age groups are considered to be of central importance to understanding developmental perspectives on output motor processing (i.e., programming and planning).

With regard to the phonological short-term memory account, the evidence suggests that with immediate repetition, there was little or no demand on phonological short-term memory. However, this demand could have increased and was necessary with immediate consecutive repetitions, as evident with reduced accuracy scores with consecutive productions. Furthermore, the findings emphasis the role of frequency in speech processing as a wells established phenomena in children ages 3, 4 and 5 years old– although not directly investigated. Specifically, the frequency by which bisyllables occur in Arabic and its dominance in the mental lexicon of the Arabic-speaking child from an early age, could have resulted in this form to be accessed more accurately (i.e., the children in this study were more accurate at repeating bisyllabic items). The Arabic-speaking children produce many disyllabic word shapes and acquire a range of complex syllable structures from an early age due to their frequency in adult input (Khattab & Al-Tamimi, 2013). Therefore, although word frequency effects were not measured in the present thesis; the results of increased bisyllabic item accuracy may reflect the internal mental lexicon of the Arabic-speaking child.

The study of consistency also allows for the investigation of Arabic-speaking children's underlying representations. By comparing different levels of processing (real words, non-words and syllable sequences) on consistency scores children's processing skills could be explained more thoroughly compared to only including one task. It was revealed in the literature review that studies found that consistency is affected by factors such as age and vocabulary knowledge. Although this study does not directly address the effects of vocabulary on performance, some assumptions could be made. It is possible that children's vocabulary knowledge affected their performance on the consistency scores, consequently, this could have affected how they treated complex words (effects of lexical and sub-lexica factors). It is generally expected that as children get older their knowledge and vocabulary (both expressive and receptive) expand, therefore, this could have influenced children's consistency performance on the tasks. The youngest age group showed high variability in their performance on the tasks, which is in line with other studies with the same age group. It is possible that at this age, the children have varying vocabulary knowledge and frequency of occurrence; which could have affected their performance.

8.2 Study Evaluation and Future Direction

8.2.1 *Strengths, Limitations and Future Direction*

This section will consider the strengths and limitations of the present study. In particular it will cover aspects of the design, the participants, the stimuli used and scoring methods.

8.2.1.1 *Recruitment, Participants Issues and Study Design*

The present study included a wide range of age groups; with ages ranging from 3;0 to 5;10 (years; month), they have all attended preschool, either nursery or kindergarten, therefore they were largely homogenous in terms of their experience with formal schooling. However, not all children within this age range have had the opportunity to be within the preschool educational system. In fact, during data collection for the present study, it was noted that the number of nursery class children (children ages 3;0-3;11) had far less students in their classrooms compared to older kindergarten classes (hence the fewer number of consent forms returned during data collection).

Compulsory education in Saudi Arabia is from year 1 primary school which is around 6 years of age. It is a challenging assignment to figure the differences between homebased children and children in nursery, as their experiences many vary greatly.

Furthermore, the present study showed high ranges in scores (standard deviations) among the youngest age group (3 year old children), with standard deviations being sometimes half of the mean value. Large standard deviations is commonly reported/observed in children within this age range (see Section 8.1 above); however, high variation could be due to the smaller sample size of this age group (29 participants), compared to the other groups (4 year olds: 50 participants and 5 year olds:50 participants). This could be at least partially responsible for the insignificant difference between 3 and 4 year olds on single repetition tasks. Furthermore, this high variation in scores could also be due to the wide age range of the 3 year old sample, ranging from the ages of 3;0 to 3;11. It is not unreasonable to assume that children's performance on the repetition tasks of real word and non-word and any difference between them is dependent on the age of the child. First, performance on tasks of real word and non-word repetition is dependent on lexical and sublexical factors, and 3;0 year old's experience with language (both receptively and expressively) is by no means similar to a 3;11 year old child. With the younger age group showing high levels of variations, on both real word and non-word tasks, this age groups lower vocabulary levels and their school curriculum is fare from the higher KG-2 and KG-3 grads for 4 and 5-year olds. Therefore, the wide age range of the 3 year old age group could have at least been partially responsible for the insignificant difference between the real word and non-word repetition tasks on both single repetitions and rapid multiple repetitions. Due to the 3 year olds high variability on their performance, their reliability was questionable.

Second, young children have short attention spans and could be easily distracted during testing; although efforts were made to reduce fatigue and distractions during testing, these factors could have been confounding variables affecting children's performance; especially with the many items/tasks presented to the children, their smaller attention span could reduce and distracted (as experienced during testing). With the non-word task, the young child had only one/two opportunities to hear each novel phonological string and repeat it, therefore, if they were distracted or they misperceived part of a non-word, repetition accuracy will be affected.

8.2.1.2 *Lexical and Sublexical Factors*

Researchers have become increasingly interested in the relationship between phonological and lexical development and studies have found a strong relationships between them (Stoel-Gammon, 2011). Throughout this thesis it has been noted that studies have documented the effects of lexical and phonological sublexical factors of word frequency, phonological neighbourhood density, age of acquisition and phonotactic probability on the accuracy and consistency of word and non-word production in children from infancy to 5 year of age (Coady & Aslin, 2004; Jamie L Metsala, 1999; Munson, Edwards, & Beckman, 2005; Anna Vogel Sosa & Stoel-Gammon, 2006; Anna V Sosa & Stoel-Gammon, 2012; Zamuner, 2009; Zamuner, Gerken, & Hammond, 2004). Notably, this study, like many other studies of intra word consistency/inconsistency, found that word consistency/inconsistency was word specific and child specific. Some words were produced with complete consistency while other words showed variability across productions; suggesting that consistency/ inconsistency is influenced by the child's experience with the words and the sound combinations that create the words, i.e. lexical and sublexical factors. Although these factors are not the primary focus of the current study, as mentioned previously, their effects could have contributed to the results of this study and presented as confounds or as possible explanations for some of the speech processing output skills, therefore, raising immediate questions and implications for the design of future studies. In this section, strengths/weaknesses of the study stimuli in terms of these factors will be addressed along with suggestions for future research design and directions.

First, one strength of this study was the inclusion of the picture naming task as a prerequisite task. This speech production activity aimed to ensure that target words were actually stored within a child's lexical representations and therefore, not treated as a novel word. This task is not usually included in many speech-processing studies that compare real-word and non-word repetition (e.g. Chiat and Roy, 2004; Roy and Chiat, 2007; Vance et al., 2005). However, there were limitations regarding the words used, in particular, when it came to the youngest 3 year old age group. As mentioned previously, if a child did not know a word the tester would provide a cue or a verbal prompt to facilitate picture naming; consequently, familiarity and/or frequency of a word within a child's lexical representations presented as a confounding variable. Lexical variables of word frequency, neighbourhood density and age of acquisition have facilitative effects on children's speech output processing; where it influences accuracy of word repetition and influence word consistency/variability across productions (see Stoel-Gammon, 2011 for a review). Word frequency (the number of times a word occurs in a language's spoken corpus) influences speech processing output accuracy and consistency (Anna V Sosa & Stoel-Gammon, 2012; Stackhouse, Vance, Pascoe, et al., 2007; A Tyler & Edwards, 1993) where high frequency words are more likely to be accurately and consistently produced compared to low frequency words, as high frequency words have stored phonological representations/motor programs and are accessed repeatedly which strengthens the pathway to production. Also, for age of acquisition (the age when a word is stored in a child's mental lexicon) the general assumption is that early acquired words are produced with greater facility than later acquired words (Anderson, 2008). When designing the stimuli for the present investigation, the age group of the children tested and the vocabulary knowledge at that age were taken into consideration; however, due to the lack of word frequency and database of age of acquisition, they were not controlled with further detail (See methods, Chapter 5 Section 5.4.1). Future research should control for such variables and investigate the effects they have on Arabic-speaking children's speech output skills. To do so with limited resources available, detailed cross-sectional or longitudinal data from individual children need to be analysed. For example, future target words could be selected based on an individual child by using parental reports of the child's productive vocabulary; and whenever possible similar words used by different children could be selected as target words. This method would create a more naturalistic set of target words; however, it would still lack control of phonetic characteristics. Future research could study the average age of acquisition of the consonants of the target words (see Sosa & Stole-Gammon, 2012). Data from Amayreh (2003) and Amayreh and Dyson (2000) on Arabic-speaking children's consonant inventories could be the base for scoring consonants of the words on age of acquisition, where a consonant acquisition is defined as the age in which 75% mastery occurs. Word familiarity effects on accuracy and consistency of productions could also be the primary investigation since both word frequency factor and age of acquisition are suggested to contribute to the familiarity of a word (Garlock et al., 2001). However, the challenges ahead not only lie with creating the stimuli to investigate lexical effects, but also in the fact that these factors are highly correlated, where earlier acquired words were higher in frequency and high neighbourhood density words rather than words with low frequency of occurrence and low neighbourhood density (Storkel, 2004).

Second, studies confirm that preschool children are sensitive to phonotactic probability i.e., sublexical factor (phonological factors), even children as young as 2 years olds old (Coady & Aslin, 2004; Edwards, Beckman, & Munson, 2004; Munson et al., 2005; Torrington Eaton et al., 2015; Zamuner et al., 2004). Studies show that high phonotactic probability facilitates accurate repetition of non-words more than low phonotactic probability, however, studies are not consistent in their findings when the effects of frequency of phonotactic probability is applied to real word repetition (Sosa & Stoel-Gammon, 2012; Torrington Eaton et al., 2015). The present study designed non-words that confirm to phonotactic constraints of the Arabic language by creating words that match the real words. Nevertheless, phonotactic probability could have influenced results. It was discussed previously (See Chapter 6 Section 6.4) that the unexpected nonsignificant difference between monosyllabic items of real words and non-words compared to trisyllabic items, could be explained, among other explanations, by the low phonotactic probability of monosyllables, particularly non-word monosyllables. Shalaan (2010) notes that it is difficult to create monosyllabic non-words without possibly violating the phonotactic-rules of Arabic, as they could be very low in phonotactic probability due to the nature of the root-and-pattern of the Arabic language. Therefore, the challenge of monosyllabic items could suggest that monosyllables are not necessarily valid in Arabic when it comes to differentiating performance between different lengths, particularly for a task of non-word repetition.

Third, one of the strength of the non-words stimuli was that there were very limited lexicality effects of syllables within the bisyllabic and trisyllabic non-words (i.e there were limited syllables that were actual words)[14]; therefore, reducing influence of stored words within lexical representations. This confounding factor was controlled in the present study since non-words containing embedded real words increase accuracy of repetition in children (Dollaghan, Biber, & Campbell, 1993; Dollaghan, Biber, & Campbell, 1995). On the other hand, the effects of word-likeness of the non-words were controlled to a satisfactory level in the present study but not in significant detail during the design and revision of the non-word stimuli. Recall that during the pilot study, the researcher and the 6 year old participants subjectively judged some non-words as highly word-like (see Chapter 3, Section 3.4.4.2), with some non-words placed as sounding dialectically similar to other Arabic accents. Subsequently, those non-words were revised by exchanging the vowels with another selection of Arabic vowels to reduce the factor of word-likeness. The word-likeness of the list of non-word stimuli was not part of the main investigation and could have in part affected the overall results of non-significant difference on the real word and non-word repetition tasks, since studies confirm that the degree to which non-words are word-like affects repetition accuracy in preschool children (Gathercole, 1995; Gathercole et al., 1991). Future research could control for this confound in more detail by creating non-words that differ in subjective rating by adult listeners, non-words with the lowest word-likeness rating could be used for investigating repetition accuracy or future research could further investigate the effects of different word-likeness ratings of non-words on repetition accuracy. For example, Sundström et al. (2014) developed a list of 131 non-words complying with Swedish phonotcatic rules, and asked eight adults to score each non-word from 0 to 8 based on the degree the item sounded like a real-word. The study included the non-words that were least word-like, resulting in a total of 25 non-word items.

Nevertheless, creating Arabic non-words should be approached with careful consideration to the properties of the Arabic language. To be more specific, due to the nature of the present thesis, where task comparisons and matched task stimuli were emphasised, non-words were derived from the real words by maintaining the consonants and changing the vowels, this method was used in many studies (e.g. Vance et al., 2005; Williams & Stackhouse, 2000). However, it was speculated that the stimuli list of the

[14] The syllables that resemble real words were usually in the vocabulary of adults (refer to methods Chapter 4 Section 4.2.2.2).

non-words had an influence on children's performance that resulted in the non-significant difference between the real words and non-words on immediate repetitions. That is, as discussed previously, since Arabic is a root-and-pattern language, with the consonant root acting as the lexical unit that holds semantic meaning; it was therefore possible that changing the vowel patterns while maintaining the trilateral roots created non-words that were highly word-like. Therefore, this reduced the possibility of different processing levels to be distinguished or of creating demands on phonological short-term memory (See discussion in Chapter 6 Section 6.5). Consequently, it is recommended that future speech processing research should develop non-words that do not match a consonant root patterns of a real word. Alternatively, non-words that match real words could be develop by maintaining consonant-vowel syllable pair and placing them in new combinations; this procedure has been used in many studies (e.g. Hoff et al., 2008; Roy & Chiat, 2004; Torrington Eaton et al., 2015). For examples, Torrington Eaton et al. moved syllables between non-words (for example: real words ‘‘cookie’’ and ‘‘puppy’’; non-words ‘‘pookie’’ and ‘‘kuppy’’), while Roy and Chiat altered syllables within nonwords (for example: real word "ladder" to non-word "daller" and real word ‘‘magazine’’ to non-word ‘‘gazameen’’). It is therefore recommend that children's speech processing performance on these patterns of non-words could be further investigated by comparing performance against non-words created by maintaining the consonantal root and only changing the vowel pattern. Adult rating of word-likeness of the two methods of non-word pattern combinations would lead to more detail on the different stimuli non-words and could provide valuable information on processing skills in Arabic-speaking participants.

To sum up, with regards to non-word repetition, studies have found that children can more accurately repeat non-words that reflect the properties of the lexicon. Therefore, controlling for these factors may provide a stronger insight into children's speech processing skills and underlying representations. the findings from the present study strongly recommends future research that investigates the effects of lexical and sublexical factors on Arabic-speaking children's single repetition accuracy and accuracy and consistency of both real words and non-words.

One point to consider regarding the present study was the lack of information on the lexicon (vocabulary) size of the children tested. To be more specific, in the literature review it was pointed that many studies found a strong correlation/relationship between non-word repetition accuracy and receptive vocabulary (Edwards et al., 2004; Munson et al., 2005) and between expressive vocabulary and consistency of multiple productions (Macrae & Sosa, 2015; Sosa & Stoel-Gammon, 2012), where smaller vocabularies predicted lower non-word accuracy and consistency scores. In fact, some studies have found that vocabulary knowledge, not age, predicted performance task performance on consistency measure (Sosa & Stoel-Gammon, 2012). It is fair to say that that there is indirect evidence on the link between non-word repetition performance in in the young children and their vocabulary knowledge (see Chapter 6 Section 6.5.3). However, future research should consider studies that identify predictors and correlates of Arabic-speaking children's performance on accuracy and consistency of repetitions that would increase our knowledge about Arabic children's speech processing and supported empirical evidence. Although using predictors/collations with vocabulary knowledge or expressive/receptive language skills is important, there still remains the key challenge of impoverished standardised test of vocabulary and language.

8.2.1.3 *Scoring Responses*

Regarding accuracy of productions, the present investigation applied a whole-word scoring method, where children's responses on items were evaluated as a whole and scored as either entirely correct or as incorrect, without regard to the number of phonological errors. However, this method of scoring has its limitations, particularly, for example, when measuring the severity level of a child's speech sound difficulty, when measuring intelligibility or when measuring progress (for a review see Newbold et al., 2013; Rvachew et al., 2012, p. 334 & 475). However, the whole-word method was favoured over the scoring methods of percentage of phonemes correct or Percentage of consonant correct on the grounds that whole word scoring has a practical advantage, as it is easy and quick to calculate and would be very appealing to clinicians working with full caseloads as it is less time consuming. Under this scoring method, clinicians would not be expected to have experienced knowledge of phonetic transcription or developmental phonology. Furthermore, this method has been used by many speech processing studies (e.g. Chiat & Roy, 2007; Dispaldro et al., 2013; Vance et al., 2005; Williams & Stackhouse, 2000). Nevertheless, it could be argued that the scoring method applied in the present study may have contributed, at least to some extent, to the nonsignificant difference between the real word and non-word repetition tasks on single repetitions along with the lack of significant developmental progression. If the younger children had even a few more phonological/speech sound errors compared to older children when repeating items, the whole-word scoring method could penalize them disproportionately compared with a more comprehensive scoring method such as the percentage of consonants correctly produced. Although this could potentially be a source of the nonsignificant differences, it was noted previously that, several studies using different speech stimuli and different scoring methods have shown that overall accuracy of real words was higher than non-words in typically developing children, and that developmental progression was evident (e.g. Budd et al, 2011; Chiat & Roy, 2007; Pascoe et al, 2016; Roy & Chiat, 2004; Sundström et al., 2014; Vance et al, 2005). Furthermore, the whole word scoring method has been used in many nonword repetition tests, such as the Preschool Repetition test [PSRep] (Chiat & Roy, 2007) and the Children's Test of Nonword Repetition test [CNRep] (Gathercole et al., 1994)[15]. Moreover, Dispaldro et al.'s 2013 study showed that using different scoring methods did not find different results in typically developing children. In their study, 17 typically developing Italian-speaking children's repetition accuracy on real words vs non-words (ages 4;1 to 5;7) were scored using two different scoring methods; the first scoring method was the percentage of phonemes correctly produced and the second was percentage of whole word correctly produced; the two scoring methods did not yield different results. Furthermore, when performance of typically developing children was compared against the performance of children with language difficulties, the whole-word scoring method showed a greater magnitude of differentiation between the groups (however, it is recognised that this may not necessarily apply with children to speech sound difficulty). Generally, in order to determine whether different scoring methods could have an impact on the overall results of typically developing Arabic-speaking children, further investigation is needed. Future research should examine whether scoring methods using whole-word or percentage of phonemes correct or percentage of consonant correct (Lawrence D Shriberg, Austin, Lewis, McSweeny, & Wilson, 1997) produces differences between a) real words, non-words and syllable sequences b) differences between age groups.

When considering how best to score children's responses on rapid consecutive repetitions, the present study showed two strengths. The first is the use of headphones, which were used to listen to and transcribe children's recorded speech productions to score their accuracy and consistency; the second, and perhaps most important strength, was the use of Praat, an acoustic program that provided a high audio quality level during

[15] These tests are designed to diagnostically differentiate children with language impairment from typical development.

transcription and analyses. The program was especially helpful when transcribing rapid productions, as it was necessary to listen to individual productions for detailed and accurate transcription. Although this procedure has proven helpful and necessary for the scrupulous nature demanded of this research, it may not be an appealing future method for practitioners dealing with caseloads that involve high numbers of children with speech and language difficulties. Within a clinical setting, clinicians might find this method of transcription and scoring time consuming and may abstain from doing so. Nonetheless, an applicable alternative to this method would involve scoring children's responses in real time—that is, during the administration of their tasks. This method of online transcription is not uncommon; it is found in children's speech assessment kits such as the Nuffield Apraxia Program-third edition (NDP3) (Williams, Stephens, Williams, McLeod, & McCauley, 2004), which allows for scoring accuracy and consistency based on live transcription of rapid consecutive productions. The Diagnostic Evaluation of Articulation and Phonology assessment kit (DEAP) (Dodd et al., 2002), also allows practitioners to transcribe and score the consistency of children's naming and repetition of words live. However, it is important to note here that the DEAP scores the inconsistency of words produced three times separately throughout a single session and not consecutively as in a diadochokinetic task. Therefore, the underlying reasons of inconsistencies found in children's productions are different for both tasks. However, from a scoring perspective, there still remains a dilemma with direct transcription of children's responses, especially with rapid consecutive productions—which could have an effect on accuracy and consistency scores. For example, Sosa (2015) used similar procedures to those described in Home et al.'s (2007) study to investigate intra-word inconsistency/consistency in typically developing children. However, Sosa found that intra-wad inconsistency in children ages 3;6-3;11 was greatly higher than that found in Home et al.'s study. According to Sosa, the large discrepancy between the two studies was likely to be due to the transcription method used, which was the major difference between the two studies. Holm et al. used online transcription during the assessment, while Sosa's study transcribed children responses from audio and video recordings; no online transcriptions were made. Thus, according to Sosa, it is possible that transcribing responses from recorded samples could lead to detailed and scrutinized transcription of a child's productions and the listener would be less likely to filter out rather minor differences across productions. In a typical clinical setting, it is less likely that a clinician would be able to detect subtle differences and errors in a child's productions during live transcription and especially, as in the case of the present study, during rapid productions. It is therefore recommended that future research focuses on different transcription methods used to transcribe and detect differences across children's productions and whether they yield different results. Interestingly, as Sosa (2015) argued, it is possible that using phonetic transcription could be an unreliable method in detecting (or quantifying) inconsistency and consistency, especially for children with speech difficulties. Rather, using more advanced acoustic and/or kinematic methods to investigate production inconsistency and consistency would provide more reliable and informative results.

Furthermore, one of the limitations of this study was the omission of vowels when scoring items on accuracy and consistency, although, this procedure was not without precedent (Ferguson & Farwell, 1975; Sosa & Stoel-Gammon, 2006; Sosa & Stoel-Gammon, 2012). The inclusion of vowel calculations could lead to a more sensitive measure of word accuracy and consistency. Their inclusion would indeed be interesting, since some children have vowel difficulty, and children suspected or diagnosed with developmental verbal dyspraxia (DVD) are characterised by having "inconsistent errors on vowels and consonants" (ASHA, 2007). For example, when faced with a non-word repetition task, a child could change its vowels to mirror a lexicalised item, which could indicate a weak motor programming skillset that is necessary to create and learn new words. Further analysis could help quantify the types of speech errors that are produced during the rapid multiple productions task, both for typically developing children and those with speech difficulties Two basic speech categories of analysing error patterns are identified: 1) segmental features, which include sequencing errors and vowel change errors; and 2) supra-segmental features, which include dysfluencies and stress errors (see Figure 8.1).

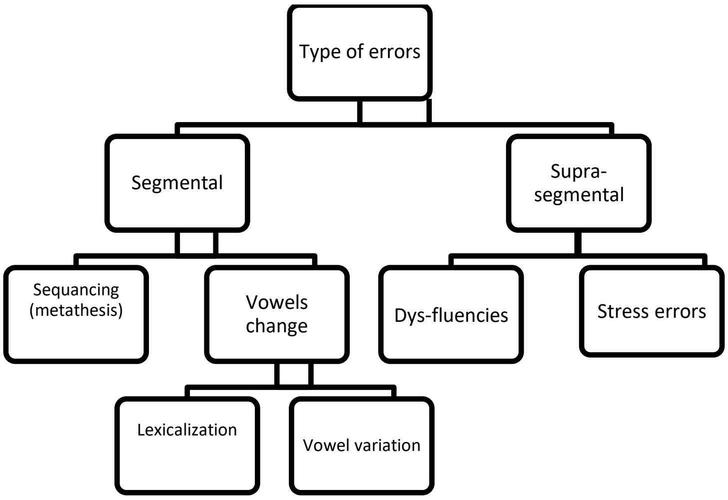

Figure 8.1: Graph representing the types of speech errors for analysis.

Another point to consider when using repetition tasks to measure a child's production accuracy is how it runs the risk of benefiting from the adult model. Studies have documented the influence of elicitation method (spontaneous versus imitation tasks) on the accuracy of speech productions in preschool English and Spanish-speaking children aged 3 to 5, with and without speech difficulties, on single-word and sentence level productions (Goldstein, Fabiano, & Iglesias, 2004; Shea & Blodgett, 1994; Summers & Larson, 1992, November ; Weston, 1997). The studies found that words/sentences were produced more accurately in an adult-modelled immediate imitation task compared to self-generated – spontaneous- productions. However, when Goldstein et al. (2004) investigated the difference between spontaneous and imitated single-word productions in twelve Spanish-speaking children with speech difficulty between the ages of 3 and 4, they found that the majority of words (62 percent) were produced identically in both tasks. However, they also found that children were more likely (20 percent) to show more adult-like productions with imitation than on spontaneous productions. Within the model of Stackhouse and Wells (1997), the distinctions between imitative and spontaneous productions are differentiated and are outlined with important clinical implications; the underlying nature of a child's difficulty and an individual child's responses to both tasks would provide valuable information on their processing abilities and their emerging phonological and lexical development. Furthermore, such repetition tasks are diagnostically informative, particularly with children diagnosed with speech motor disorders such as DVA, who show more speech errors on imitation tasks than spontaneous productions (ASHA, 2007; Velleman, 2006). The present study used different repetition tasks for comparison as per the essences of the theoretical model of Stackhouse and Wells. However, a thorough evaluation of speech-processing skills necessitates that a distinction between typically developing children's performance on spontaneous and imitation tasks be made to provide a reliable basis of comparison against children with speech difficulties.

Furthermore, rate measurements were excluded from the present thesis. The two key factors informing this decision were the challenges of administering a motor task of rapid consecutive repetitions to young children and the difficulty of scoring responses when the accuracy of repetitions was affected. In fact, one of the most observed challenges when measuring DDK's concerned how best to record the live tasks and score and calculate the time for children to repeat the sequences (Gadesmann & Miller, 2009). The consequences of such challenges usually results in variability in children's performances across different trials and across young children. To reduce such challenges, Thoonen et al. (1996, 1999) and Rvachew, Hodge and Ohberg (2005) recommended using standardised protocols when administering DDK tasks. Theses authors suggested that clinicians should record responses and measure the rate of repetitions using an acoustic waveform editor. Rvachew et al., (2005) described a standard procedure for obtaining DDKs using a software they developed, which can be applied to preschool children; this procedure was based on Thoonen et al's. (1996, 1999) procedure to abate the challenges of measuring maximum performance tasks in children. Primarily, the software helps to facilitate the recording process of children's repetitions, measuring and retrieving children's responses. Rvachew et al. (2005) also recommend that clinicians use any acoustic sound file editor when recording responses because it can accurately measure the duration of children's repetition rates.

It is noted in this thesis (see the Methods in Chapter 5, Section 5.4) that the Praat software was used to visualise the recorded responses of children as a spectrogram and waveform. This provided a procedure to precisely score the accuracy of each individual production of the consecutive repetitions (See Chapter 5, Figure 5.2 for a display of the rapid consecutive repetition of a word). The children's quick repetitions of monosyllabic, bisyllabic and trisyllabic items was displayed on the Pratt editor and the boundaries between each repetition were determined by visual inspection of the image and supported by the auditory recordings. Here, if rate calculations were to be conducted and analysed, they would be determined by the onset of the syllable (located at the beginning of the burst of the consonant) and the end of the syllable of the fifth repetition. Repetition rate would be calculated by rate of repetition per second. However, accurate repetitions in young children seldom occur; speech sound errors and phonological processes are very common in young children. In fact, young children are more likely to produce inaccurate repetitions in DDK tasks than in conversation, which could ultimately affect the subsequent rate of repetitions (Canning & Rose, 1974; Yaruss & Logan, 2002). Further, other challenges in measuring rate such as inhalations/exhalation or dysfluencies between repetitions usually occur with young children. To overcome this issue, researchers such as Thoonen et al. (1996) exclude inaccurate productions, such as a flat syllables, which are assigned as a missing value. However, using acoustic visualisations and the lengthy procedure of scoring rate when excluding less accurate responses from analysis raises the question of the utility and significance of such a measure in clinical practice. This segmentation procedure is not practical when considering additional clinical factors, such as the environment and the caseload of a speech-language therapist; while research facilities generally employ acoustic analysis to measure rate of repetitions, a stopwatch is typically used in clinic settings to calculate the rate of repetitions. Further, as debated in the literature, it is difficult to interpret timed repetitions, which places its reliability into question (Gadesmann & Miller, 2009). Gadesmann and Miller (2009) suggest that caution should be taken when measuring DDK rates in clinical practice because it involves diagnostic outcomes. Their study showed lower inter- and intra-rater reliability, which is not acceptable for clinical diagnosis.

Although measuring rate has its challenges, it also has its advantageous. First, its inclusion provides an opportunity to draw comparisons between typically developing Arabic-speaking children and English-speaking children on the basis of their performance on DDK rates. They can also gauge whether rate measure captures developmental changes between different age bands in preschool Arabic-speaking children. Further, when incorporating the Praat software (or any other acoustic software), rate performance could still be investigated when articulatory errors occur during children’s word production. Future research might benefit from using such software to facilitate the recording and measurement processes when investigating young Arabic-speaking children’s rate performance. It could also help facilitate how clinicians introduce instructions to children during trials and help them systemically administer DDK tasks to increase child cooperation and further manage time.

Finally, one potential direction that future research might take concerns gemination, or the elongated pronunciation of spoken consonants. This would be particularly interesting, and would test how children, both typically developing and those with speech difficulties, perform on individual items involving bi-syllabic and tri-syllabic words and non-words with a geminate structure The interest derives from the notion that young Arabic-speaking children's early words are characterised by many complex structures and produce geminates words from an early age due to their early linguistic exposure (Khattab & Al-Tamimi, 2013). Therefore, children's speech-processing profiles could be compared using items with and without geminates.

8.2.2 *Theoretical Implications*

The present study has contributed greatly to our understanding of speech output processing skills in Arabic-speaking children by using the theoretical psycholinguistic approach of Stackhouse and Wells (1997). Furthermore, this study has supported, at least not directly, theoretical accounts of whole-word phonology or template-based approach to phonological representations (Vihman & Keren-Portnoy, 2013).

First, the Stackhouse and Wells (1997) framework has been developed after years of psycholinguistic and neuropsychological research. The approach profiles a child's speech difficulty based on systemic hypothesis testing of speech processing abilities through a series of activities, whether the child has a known causal factor or an unknown origin. However, Stackhouse and Wells stress the need to refer to group studies of normal control data when working with children, if the nature of a child's difficulty is to be understood and compared to typical development. The present thesis used Stackhouse and Wels's model to investigate the hypothesised levels of output processing of repetition tasks (real words, non-words and syllable sequences). Children's performance on the tasks did not capture different hypothesised levels of processing, as children's performance did not differ on the tasks with immediate repetitions. However, this does not mean that the theoretical concept of the approach was not supported, rather it could suggests that the items used to capture the levels of processing were not sensitive enough (as discussed in earlier sections). The design of the linguistic items, show how both linguistic and psycholinguistics are dependent on one another, and that carful control of variables within a set of stimuli is required to test hypothesized levels. For example, it was hypothesised that non-words would not require access to stored phonological representations/motor programs, rather it requires online motor programming that creates non-words from a selection of phonological unites at a sub-lexical level. The available selection would expand as the child expands his/her lexical knowledge. Therefore, to capture a child's ability to use motor programming skills for non-word repetition against levels of stored phonological/ motor programs, the stimuli items must consider factors, such as sub-lexical factors of phonotactoc probability, and reduce influence from other processing levels, if conclusions are to drawn from dissociated hypothesised levels of processing.

Stackhouse and Wells model accounts for the central cognitive-linguistic processes involved in speech production and they could be a source of speech difficulty in children. The central level in Stackhouse and Wells model is the lexical representations that include phonological representations. This representation is thought to be holistic in nature and less segmented (See Chapter2 Section 2.11) at the early stages of language development and become increasingly segmented as the child increase his/her vocabulary. Within the whole-word or template-based approach to phonological representations (Vihman & Keren-Portnoy, 2013) this concept is supported by children's performance across word productions, where they produce the same word differently each time. The present study's use of the behavioural measure of consistency has important theoretical implications, since reduced consistency is considered to be a sign of the maturing phonological representations. The developmental improvement observed in this study on consistency scores of rapid consecutive repetitions between the ages 3 to 4 years supports this theoretical concept. Further support for the template-based approach to phonological representations comes from children's performance on different stimuli length. The Arabic speaking children in this study showed higher accuracy and consistency on bisyllabic structures of words and non-words compared to monosyllables. This gives insight into the underlying representations and structure of the Arabic children and supports the whole-word phonology, since the concept assumes that children's templates are influenced by the adults input and templates emerge out of the shapes that are frequent in the adult's language, and these shapes are applied to new words. In Arabic, the rich bisyllabic shapes in the Arabic language are exhibited in children's early word productions, where they are dominant in the output of the Arabic-speaking child compared to other shapes (Khattab & A-Tamimi, 2013).

Furthermore, the influences of higher top-down linguistic-cognitive processes to bottom-up lower levels of execution was explored in the present study through the theoretical psycholinguistic model. Children performance on a kinematics task i.e. rapid repeated productions or speech motor control was investigated using different linguistic tasks that tap different levels of motor processing, responses was scored by behavioural measures of accuracy and consistency. The present study has shown that nonwords were less consistent than real words in 4 and 5 year olds, although their scores were at ceiling. This study therefore has supported research using experimental kinematic approaches to investigate the influence of linguistic demands on oromotor movement variability (Kent, 1992; Krishnan et al., 2013; Sadagopan & Smith, 2008; Smith & Goffman, 2004), and supported, with the behavioural measure used, the concept that language skills are linked to motor abilities. Since both real words and nonwords were highly matched in terms of their kinematics, the lower consistency scores on the nonwords supported the idea that speech motor skills are highly complex and local and learning dynamic representations does not transfer.

8.2.3 *Clinical Implications*

As noted in Chapter 1 of the literature review, psycholinguistic model-based approaches could be appealing to speech-language therapist in Saudi Arabia that deal with the heterogenic nature of children with speech difficulties, and who are challenged with the diminished resources for assessing an Arabic-speaking child and labelling their speech difficulty. This is because the psycholinguistic approach does not burden the therapist with a diagnostic label. Rather, the model is an inclusive diagnostic system that can be applied to any child regardless of their condition—whether known or unknown cause of speech difficulty. The speech therapist can continue with their usual assessment process of a medical (if necessary) perspective to understanding the cause of a child’s speech difficulty (such as Cerebral Palsy or Cleft lip /palate) and a descriptive-linguistic analyses alongside complimenting their assessment with the psycholinguistic approach. In fact, as mentioned previously, the approaches complement each other, and have to be administered side by side.

Conversely, speech and language therapists in Saudi Arabia may have an initial anxiety when introduced to the approach. A British study (Joffe & Pring, 2008) reported that a clinical practice respondent mentioned that she and many other speech-language therapist were “terrified” by psycholinguistic models. This response to the psycholinguistic model is not surprising and would be expected from clinicians in Saudi Arabia. The psycholinguistic model might seem rather complicated with its “boxes-and-arrows” and its complex use of terminology, especially that clinicians in Saudi Arabia are accustomed to long-established assessment systems that identify medical causes of difficulty and further describes linguistic behaviour in specific ways—although, speech-language therapist also draw somewhat on psycholinguistic assessments in some way. However, this does not encourage progress, to assess children in a comprehensive way and intervene and monitor their speech skills, a speech and language therapist must move from the long-held view that speech difficulties are assessed at an overt segmental (sound) or phonological level, and should develop children's speech skills beyond that level of difficulty (Stackhouse & Wells, 1997).

With regards to the findings from the present study, the norms and material used have noteworthy potential clinical implications, which could be seen in several aspects when assessing Arabic-speaking children with speech difficulties. First, the normative data presented in this thesis offers a baseline of performance for typically developing children which can thus serve as a useful point of reference for clinicians assessing children's speech. Second, there were implications relating to the tasks and material used in this study. Whilst the DDK task is a tool that is already being utilised by clinicians in clinical practice, it typically only involves the repetition of syllable sequences to measure a child's peripheral speech motor skills. However, clinicians could supplement this task with the stimuli and data from this present study. Specifically, they could compare a child's performance on the syllable sequence repetition task to test the child's motor planning and execution skills, then compare their performance to the repetition of real words and nonwords to measure a child's abilities to plan and execute stored motor programs and to create new motor programs.

Further, the study's findings relating to the consistency measure is of central importance. The typically developing children from this study ages 3, 4, and 5 years old were generally consistent on all the repetition tasks. However, the 3 year old children were significantly less consistent than both the 4 and 5 year old children. By the age of 4, children were generally consistent and no developmental changes were observed compared to 5 year olds. This finding could help clinicians determine the nature of a child's speech difficulty; a child would be expected to be consistent on the tasks by age 4. Therefore, if a child with speech difficulties shows inconsistencies, this could suggest underlying difficulties that are most likely, but not necessarily, motoric in nature (such as a speech motor planning difficulty). It should be noted that the normative data on consistency provided in the present study are of particular importance, given the apparent lack of published research on typically developing children's consistency of repeated productions.

Moreover, the present study focused on tracking speech processing and developmental changes when increasing performance load by increasing stimuli length—or number of syllables of an item. The speech processing literature reviewed previously (Chapter 2, section 2.2.2), has documented the effects of item length in many languages (including English) where multisyllabic items of words, non-words and syllable sequences capture the inaccuracies of children's productions more than monosyllabic and bisyllabic items (e.g., Bernhardt & Major, 2005; Chiat & Roy, 2007; James et al, 2008; Metsala & Chisholm, 2010; Roy & Chiat, 2004; Vance et al., 2005;). However, the present study has reported that bisyllabic words and non-words were more accurately produced than both mono and trisyllabic words and non-words. These findings build on previous literature and supported acquisition studies such as Khttab and Al-Tamimi (2013), suggesting that typically developing Arabic-speaking children produce many bisyllabic word shapes compared to monosyllabic shapes. Consequently, it can be recommended that clinicians assessing Arabic-speaking children make a clear distinction between the child's ability to accurately produce bisyllabic items compared to other item lengths. Ultimately, a child aged 3, 4 and 5 years old would be expected to produce bisyllabic items more accurately than other forms. Therefore, clinicians could introduce different stimuli lengths as part of the assessment battery, and can use the stimuli set developed in this study for its comprehensive qualities that can be used for children as young as three. Alternatively, clinicians can select stimuli of varying lengths and include them when assessing a child's articulation skills.

However, one of the challenges when using multisyllabic items in DDK tests is the different nature of speech difficulties in children. Tasks that are otherwise appropriate for children within a certain age group and level of speech difficulty could become inappropriate for diagnostic and assessment purposes within other age ranges and severity levels. For example, children suspected of DVD (or a speech motor disorder) could have significant difficulty producing multisyllabic items, likewise, these items could be difficult for very young children as their productions may be limited to mono or bisyllabic words. Furthermore, speech and language difficulties tend to co-occur in children, this further complicates the use of items with longer syllable lengths to test children, who might otherwise be unable to produce them.

Generally, two issues can arise for clinicians using these assessment material on young children. First, the stimuli items developed for the present study are long, and require a child to repeat the items multiple times and at speed. Consequently, a young child might consider the list of stimuli long and they are likely to be distracted and disinterested in the task, likewise, clinicians might find the long list time consuming. However, clinicians could reduce the number of items and resort to selecting a collection of items of different lengths. Similarly, they could also incorporate the stimuli items and tasks with other activities. For example, an advantage of the stimuli used for the present study is that many of the items are commonly found in many articulation tests used with young children, therefore, a clinician can incorporate articulation test activities with the rapid repetition test. The second issue that clinicians might encounter during these assessments concern the rapid repetitions tasks. Young children might be challenged in understanding the instructions, to perform and complete the task and to grapple against a waning attention span; all which could lead to difficulties administering such tasks to very young children. In fact, these very observations were noted in the present study when testing the 3 year old children, which forced their completion of the tasks over multiple shorter sessions. However, multiple short sessions might not be applicable in a real-life clinical setting where a clinician has large caseloads of children.

There still remains some key issues to consider upon the potential future application of Stackhouse and Wells' framework in everyday clinical practice. First, there needs to be an assessment method that evaluates all the levels and components of the model, not just the current study's focus on output repetition tasks, but also the hypothesised levels of input and pure knowledge of children's lexical store. These assessments should be established for reliability and validly according to different typically developing age groups. Second, an increased understanding of underlying difficulties in a child will be effective in therapeutic planning, however, using a psycholinguistic model to assessment does not function alone. Rather, they should be applied as a complementary method to the medical and linguistic perspectives. Finally, future research should also consider whether the psycholinguistic assessment could capture subgroups of children with speech difficulty who have similar diagnostic labels. Ultimately, if the measures identified in this study were to be used in clinical settings and considered useful for diagnostic purposes with children with speech difficulties, future research should examine and compare whether children with speech difficulties show more errors on accuracy and consistency measures than typically developing children.

8.3 Conclusion

To our knowledge, this study is the first investigation of preschool Arabic-speaking children's speech output skills using a box-and-arrow model of speech processing. The speech processing model of Stackhouse and Wells (1997) has been adopted in this study to typically developing 3, 4 and 5 year old children, therefore enhancing our understanding of the normal developmental processes in children speaking Arabic within this age band.

The findings presented in this study contribute to the existing cross-linguistic psycholinguistic literature. The present study showed that the Arabic speaking children presented different patterns of speech processing behaviour, with the stimuli set and tasks presented. With single repetitions of the stimuli presented, children's performance on the tasks were not differentiated, nor was there a developmental progression between age bands. Therefore, the stimuli used to test the children did not capture different levels of motor programming for non-word repetition and stored phonological/motor programs for real word repetition in all age groups, while only the productions of syllable sequences at the execution level was highly more accurate than other tasks. On the other hand, rapid productions of the tasks, differentiated processing demands and captured developmental progress, as measured by accuracy and consistency of productions, at least between 3 and 4 year olds, suggesting the possible increase in demands on output levels of motor planning and execution. Furthermore, increasing processing load on the assumed levels of possessing of each task by increasing the length of items were not straightforward. Children generally performed better on bisyllabic items than monosyllabic items, and generally, no significant difference between mono and trisyllables, suggesting and supporting a language-specific approach to understanding increased processing loads and demands in children.

This study presented results that provide a privileged insight into Arabic-speaking children's underlying phonological representations and motor performance using behavioural measures of accuracy and consistency.

This study has contributed to the **theoretical** knowledge of speech output skills in Arabic-speaking children:

- It has shown that the Arabic language has its unique properties and therefore its own psycholinguistics. The Arabic children's performance on the tasks has shown that the properties of a language determines the utility of a theoretical model or approach in differentiating processing demands or determining its sensitivity to processing demands.

- This study attempted to increase processing load by increasing the number of syllables of words and non-words, however, this study has shown that speech processing load is defined by a child's ambient language that influences a child's underlying template representations. To break this down, Arabic-speaking children's phonological templates are influenced by the rich bisyllabic word shapes that occur frequently in their input by the adult's model i.e. their ambient Arabic language, and these word patterns are prevalent in children's productions. The present study has supports this view of template behaviour in children, by providing evidence that bisyllabic words were produced more accurately than short monosyllabic words. Therefore, this study has added to the evidence that supports the theoretical template-based approach to phonological representations. Further support of language-specific templates of phonological representations comes from two findings, first, bisyllabic non-words were produced more accurately than monosyllabic non-words, supporting template-based concept that prominent shapes of phonological templates are applied to new words. Second, bisyllabic words and non-words were produced more consistently across multiple productions, which also supports the prominence of particular patterns of a language in shaping children's developing internal phonological structure.

- The reduced consistency of productions in the youngest 3 year old children compared to older children could reflect the developing and incomplete underlying lexical representations of the Arabic-speaking child at that age. This could be one source of reduced consistency scores in young children; that supports the concept of "holist" phonological representations and lack of segmental detail of phonological representations during critical stages of language acquisition.

- The findings from the present study supports the dynamic system theoretical account of development, where reduced consistency of rapid consecutive productions i.e. speech motor tasks was shown to be a normal behaviour in typically developing children and that increase in consistency of productions with age is a sign of developmental change.

The present study has also **methodological** implications:

- It has shown that careful consideration to the special properties of the Arabic language (internal nonconcatenative morphology of the language), where its consonantal root of underlying lexical unit and the vowel and consonant phonetic systems should be taken into account when devising stimuli, particularly non-word stimuli.
- This study points to possible lexical and sublexical factors such as word frequency, phonotactic probability and word-likeness effects that could have affected children's performance on the speech output tasks. Therefore, caution should be made when drawing conclusions about children's performance on processing tasks and carful stimuli design and control should be taken into consideration for future research designs.
- It has shown that designing stimuli of increasing complexity and processing load to assess children's processing performance should be comprehensively designed taken into account the child's ambient language.

Finally, the study has **practical** implications

- The assessment of diadochokinetic tasks is a technique already utilised by clinicians in clinical practice. Therefore, clinicians could carry out their usual protocol for a diadochokinetic task and add the procedures and stimuli presented in this study to their measures.
- The test battery designed in this study included a comprehensive set of assessment stimuli that were matched phonetically; this comprehensive set has an advantage clinically. It can be found within a clinics articulation test battery and can be used with very young children. Therefore, a clinician could use both the stimuli for the regular articulation test along the diadochokinetic tasks.
- Increasing processing load by increasing the number of syllables of items could provide valuable information on children with speech difficulties. Typically developing Arabic-speaking children have been found to use, and more accurately produce, bisyllabic items, compared to other item lengths. This implies that clinicians would benefit from using speech assessment materials that contain

different stimuli lengths when assessing a child with speech difficulty. The stimuli developed for the current study can be used for such purposes, as it contains ag-appropriate material with different stimuli lengths.

- The normative data from this study is helpful for comparative purposes with children with speech difficulties. The data suggests that speech motor tasks of rapid consecutive repetitions are developmentally sensitive and can be used for comparative purposes for children who are not developing accordingly. Where consistency scores were developmentally sensitive between 3 and 4 years old, while accuracy was sensitive developmentally between 3, 4 and 5 years.
- Speech inconsistency is prevalent in the speech of typically developing children ages of 3 years old up to the age of 4 years. Therefore, one should be cautious when interpreting any intra-word inconsistency as indicative of specific subtypes of a speech disorder. An important note to consider when interpreting the findings from studies measuring accuracy and consistency is that 3-year old children showed more variability in their performance on processing tasks compared to older groups. Ultimately, this rendered the reliability of the tasks with these age groups questionable.
- The data from this study also implies that more challenging material is required for children aged 5 years old. Generally, this age band performed at ceiling, which suggests that the tasks did not capture their true range of performance; thus, any comparisons drawn to a child with speech difficulty could be fraught.

Certainly, as with any study of this nature demonstrates, the conclusions reached in this study can only be determined by its unique stimuli used to test the children. Therefore, generalised conclusions cannot be drawn and applied equally to understanding all 3-5 Arabic speaking children’s speech processing skills. To do so, subsequent research must be designed to evaluate the validity of this study’s conclusions and to develop more comprehensive accounts of children’s performance.

REFERENCES

Abdoh, E. M. A. (2011). *A study of the phonological structure and representation of first words in Arabic.* (Doctoral thesis), University of Leicester. Retrieved from https://lra.le.ac.uk/handle/2381/10221

Abou-Elsaad, T., Baz, H., Afsah, O., & Mansy, A. (2015). The nature of articulation errors in Egyptian Arabic-speaking children with velopharyngeal insufficiency due to cleft palate. *International Journal of Pediatric Otorhinolaryngology, 79*(9), 1527-1532. doi:10.1016/j.ijporl.2015.07.003

Abou-Elsaad, T., Baz, H., & El-Banna, M. (2009). Developing an articulation test for Arabic-speaking school-age children. *Folia Phoniatrica et Logopaedica, 61*(5), 275-282. doi:10.1159/000235650

Al-Ghamdi, Y. S., Omer, M. I., Khalil, M. K., Ali, S. A., Barmada, R. A., & Abdelgader, M. H. (2002). Clinical evaluation of disabled children in Al-Qassim region, Saudi Arabia. *Neurosciences, 7*(4), 272-277.

Al-Sulaiman, A. A., Bademosi, O. F., Ismail, H. M., Al-Quliti, K. W., Al-Shammary, S. F., Abumadini, M. S., . . . Magbool, G. M. (2003). Cerebral palsy in Saudi children. *Neurosciences, 8*(1), 26-29.

Alawaji, N. N. (2014). *Speech Production in Arabic Speaking Children with Operated Cleft Palate* (Doctoral thesis), University of Sheffield, Sheffield, United Kingdom. Retrieved from http://etheses.whiterose.ac.uk/id/eprint/7520

Albustanji, Y. M., Albustanji, M. M., Hegazi, M. M., & Amayreh, M. M. (2014). Prevalence and types of articulation errors in Saudi Arabic-speaking children with repaired cleft lip and palate. *International Journal of Pediatric Otorhinolaryngology, 78*(10), 1707-1715. doi:10.1016/j.ijporl.2014.07.025

Aldabas, R. A. (2015). Special Education in Saudi Arabia: History and Areas for Reform. *Creative Education, 6*, 1158-1167.

Alt, M. (2011). Phonological working memory impairments in children with specific language impairment: Where does the problem lie? *Journal of communication disorders, 44*(2), 173-185. doi:10.1016/j.jcomdis.2010.09.003

Amayreh, M. (2003). Completion of the Consonant Inventory of Arabic. *Journal of Speech, Language, and Hearing Research, 46*, 517-529. doi:10.1044/1092-4388(2003/042)

Amayreh, M., & Dyson, A. (2000). Phonetic inventories of young Arabic-speaking children. *Clinical Linguistics & Phonetics, 14*(3), 193-215. doi:10.1080/026992000298823

Amayreh, M. M., & Dyson, A. T. (2000). Phonetic inventories of young Arabic-speaking children. *Clinical Linguistics & Phonetics, 14*(3), 193-215. doi:org/10.1080/026992000298823

Ammar. (2002). Acquisition of syllabic structure in Egyption Colloquial Arabic. In L. M. Windsor & N. Hewlerr (Eds.), *Investigations in clinical phonetics and lingustics* (pp. 153-160). Mahwah, NJ and London: Lawrence Erlbaum.

Ammar, W. (1999). The acquisition of consonant custers in Egyption children from two to four years. *Language Science, 2*(3), 10-37.

Anderson, J. D. (2008). Age of acquisition and repetition priming effects on picture naming of children who do and do not stutter. *Journal of fluency disorders, 33*(2), 135-155. doi:10.1016/j.jfludis.2008.04.001

Arabia, G. A. f. S. i. t. K. o. S. (2016). Population by Single Age , Nationality (Saudi/Non-Saudi) and Gender. Retrieved from https://www.stats.gov.sa/en/43

ASHA. (2007). Childhood Apraxia of Speech [position Statment and Technical Report] Retrieved from www.asha.org/policywww.ashsa.org/policy

Awaad, E. N. (2008). اضطرابات النطق لدى تلاميذ المرحلة الابتدائية بمدينة جدة [Speech difficulties in primary school children in Jeddah city] http://www.kau.edu.sa/Show_Res.aspx?Site_ID=372&Lng=AR&RN=855 Retrieved from http://www.kau.edu.sa/Show_Res.aspx?Site_ID=372&Lng=AR&RN=855

Ayyad, H. S., Bernhardt, B., & Stemberger, J. P. (2016). Kuwaiti Arabic: acquisition of singleton consonants. *International Journal of Language & Communication Disorders, 51*(5), 531-545. doi:10.1111/1460-6984.12229

Aziz, A. A., Shaheen, E. A., Osman, D. M., & El Sabagh, A. (2012). Arabic psycholinguistic screening tool: A preliminary study. *The Egyptian Journal of Otolaryngology, 28*(1), 64. doi:0.7123/01.EJO.0000411080.90712.1c

Azza, A., Gad Alla, M., Shoubary, A., & Abdel-Wahab. (2007). *Standardization and modification of illinois test of psycholinguistic abilities on Egyptian children.* (16)

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annual review of psychology, 63*, 1-29. doi:10.1146/annurev-psych-120710-100422

Bakalla, M. H. (1984). *Arabic Culture through its Language and literature*. London ; Boston Kegan Paul International.

Baker, E., Croot, K., McLeod, S., & Paul, R. (2001). Psycholinguistic models of speech development and their application to clinical practice. *Journal of Speech, Language, and Hearing Research, 44*(3), 685-702. doi:10.1044/1092-4388(2001/055)

Balladares, J., Marshall, C., & Griffiths, Y. (2016). Socio-economic status affects sentence repetition, but not non-word repetition, in Chilean preschoolers. *First Language, 36*(3), 338-351. doi:10.1177/0142723715626067.

Bernhardt, B., & Major, E. (2005). Speech, language and literacy skills 3 years later: a follow-up study of early phonological and metaphonological intervention. *International Journal of Language & Communication Disorders, 40*(1), 1-27. doi:10.1080/13682820410001686004.

Boersma, P. & Weenink, D. (2013): Praat: doing phonetics by computer [Computer program]. Retrieved 2014 from http://www.praat.org/.

Bowen, C. (2014). *Children's speech sound disorders*: John Wiley & Sons.

Broomfield, J., & Dodd, B. (2004a). Children with speech and language disability: Caseload characteristics. *International Journal of Language and Communication Disorders, 39*(3), 303-324. doi:10.1080/13682820310001625589

Broomfield, J., & Dodd, B. (2004b). The nature of referred subtypes of primary speech disability. *Child Language Teaching and Therapy, 20*(2), 135-151. doi:10. I 191/0265659004ct267oa

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*(3), 311-335. doi:10.1080/136828299247432

Bybee, J. (2000). The phonology of the lexicon: Evidence from lexical diffusion. In M. Barlow & S. Kemmer (Eds.), *Usage-based models of language* (pp. 65-85). Standford, CA: CSLI Publications.

Chiat, S., & Polišenská, K. (2016). A framework for crosslinguistic nonword repetition tests: Effects of bilingualism and socioeconomic status on children’s performance. *Journal of Speech, Language, and Hearing Research, 59*(5), 1179-1189. doi:10.1044/2016_jslhr-l-15-0293

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*(2), 429-443. doi:10.1044/1092-4388(2007/030)

Coady, J. A., & Aslin, R. N. (2004). Young children's sensitivity to probabilistic phonotactics in the developing lexicon. *Journal of Experimental Child Psychology, 89*(3), 183-213. doi:10.1016/j.jecp.2004.07.004

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

Cohen, W., Waters, D., & Hewlett, N. (1998). DDK rates in the paediatric clinic: a methodological minefield. *International Journal of Language & Communication Disorders, 33*(S1), 428-433. doi:10.3109/13682829809179463

Corrin, J., Tarplee, C., & Wells, B. (2001). Interactional linguistics and language development. In M. Selting & E. Couper-Kuhlen (Eds.), *Studies in interactional linguistics* (pp. 199). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Crary, M. A. (1993). *Developmental Motor Speech Disorders* San Diego, Califotnia: Singular Publishing Group.

Dispaldro, M., Deevy, P., Altoé, G., Benelli, B., & Leonard, L. B. (2011). A cross-linguistic study of real-word and non-word repetition as predictors of grammatical competence in children with typical language development. *International Journal of Language & Communication Disorders, 46*(5), 564-578. doi:10.1111/j.1460-6984.2011.00008.x

Dispaldro, M., Leonard, L. B., & Deevy, P. (2013). Real-word and nonword repetition in Italian-speaking children with specific language impairment: A study of diagnostic accuracy. *Journal of Speech, Language, and Hearing Research*, *56*(1), 323-336. doi:10.1044/1092-4388(2012/11-0304)

Dodd, B. (2005a). Children with speech disorder: defining the problem. In B. Dodd (Ed.), *Differential diagnosis and treatment of children with speech disorder* (2nd ed., pp. 10-13). London: Whurr Publishers.

Dodd, B. (2005b). *Differential diagnosis and treatment of children with speech disorder*. London: Whurr Publishers.

Dodd, B., Russell, T., & Oerlemans, M. (1993). Does a past history of speech disorder predict literacy difficulties? In R. M. Joshi & C. K. Leong (Eds.), *Reading disabilities: Diagnosis and component processes* (pp. 199-212). Dordercht: Kluwer Academic Publishers.

Dodd, B., Zhu, H., Crosbie, S., Holm, A., & Ozanne, A. (2002). *Diagnostic evaluation of articulation and phonology (DEAP)*: Psychology Corporation.

Dollaghan, C., Biber, M., & Campbell, T. (1993). Constituent syllable effects in a nonsense-word repetition task. *Journal of Speech, Language, and Hearing Research, 36*(5), 1051-1054. doi:10.1044/jshr.3605.1051

Dollaghan, C., & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41*(5), 1136-1146. doi:10.1044/jslhr.4105.1136

Dollaghan, C. A., Biber, M. E., & Campbell, T. F. (1995). Lexical influences on nonword repetition. *Applied Psycholinguistics, 16*(2), 211-222. doi:10.1017/S0142716400007098

Dyson, A. T., & & Amayreh, M. M. (2007). Jordanian Arabic speech acquisition. In S. McLeod (Ed.), *International Guide to Speech Acquisitio* (pp. 288–299). Clifton Park, NY: Thomson Delmar Learning Association.

Edwards, J., Beckman, M. E., & Munson, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research, 47*(2), 421-436. doi:10.1044/1092-4388(2004/034)

Edwards, J., & Lahey, M. (1998). Nonword repetitions of children with specific language impairment: Exploration of some explanations for their inaccuracies. *Applied Psycholinguistics, 19*(02), 279-309. doi:10.1017/S0142716400010079

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*(4), 865-878. doi:10.1044/jslhr.4304.865

Ferguson, C. A., & Farwell, C. B. (1975). Words and sounds in early language acquisition: English initial consonants in the first fifty words. *Language, 51*, 419-439.

Field, A. (2009). Discovering statistics using SPSS:(and sex and drugs and rock 'n'roll). Introducing statistical methods: London: Sage.

Fletcher, S. G. (1978). *The fletcher time-by-count test of diadochokinetic syllable rate*: CC Publications.

Fox, A. V. (2004). *Kindliche Aussprachestörungen-Phonologischer Erwerb, Differenzialdiagnostik, Therapie*: Schulz-Kirchner Verlag GmbH.

Fox, A. V., Dodd, B., & Howard, D. (2002). Risk factors for speech disorders in children. *International Journal of Language & Communication Disorders, 37*(2), 117-131. doi:10.1080/13682820110116776

Gadesmann, M., & Miller, N. (2008). Reliability of speech diadochokinetic test measurement. *International Journal of Language & Communication Disorders, 43*(1), 41-54. doi:10.1080/13682820701234444

Garlock, V. M., Walley, A. C., & Metsala, J. L. (2001). Age-of-acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults. *Journal of Memory and language, 45*(3), 468-492. doi:10.1006/jmla.2000.2784

Gathercole, S. E. (1995). Is nonword repetition a test of phonological memory or long-term knowledge? It all depends on the nonwords. *Memory & Cognition, 23*(1), 83-94. doi:10.3758/BF03210559

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics, 27*(04), 513-543. doi:10.1017/S0142716406060383

Gathercole, S. E., & Baddeley, A. D. (1990). The role of phonological memory in vocabulary acquisition: A study of young children learning new names. *British Journal of Psychology, 81*(4), 439-454. doi:0.1111/j.2044-8295.1990.tb02371.x

Gathercole, S. E., Willis, C., Emslie, H., & Baddeley, A. D. (1991). The influences of number of syllables and wordlikeness on children's repetition of nonwords. *Applied Psycholinguistics, 12*(03), 349-367. doi:10.1017/S0142716400009267

Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory, 2*(2), 103-127. doi:10.1080/09658219408258940

Goldstein, B., Fabiano, L., & Iglesias, A. (2004). Spontaneous and Imitated Productions in Spanish-Speaking Children with Phonological Disorders. *Language, Speech, and Hearing Services in Schools, 35*(1), 5-15. doi:10.1044/0161-1461(2004/002)

Gray, S. (2003). Diagnostic accuracy and test–retest reliability of nonword repetition and digit span tasks administered to preschool children with specific language impairment. *Journal of communication disorders, 36*(2), 129-151. doi:10.1016/S0021-9924(03)00003-0.

Hayden, D., & Square, P. (1999). Verbal Motor Production Assessment for Children. San Antonio, TX: The Psychological Corporation.

Henry, C. E. (1990). The development of oral diadochokinesia and non-linguistic rhythmic skills in normal and speech-disordered Young children. *Clinical Linguistics and Phonetics, 4*(2), 121-137. doi:10.3109/02699209008985476.

Hewlett, N. (1990). Processes of development and production. In P. Grunwell (Ed.), *Developmental speech disorders* (pp. 15-38). London: Whurr Publishers.

Hickman, L. (1997). Apraxia Profile. San Antonio, TX: The Psychological Corporation.

Hof, Y. V. Z.-o. t., Wijnen, F., & Dejonckere, P. (2009). A test of speech motor control on word level productions: The SPA test (Dutch: Screening Pittige Articulatie). *International Journal of Speech-Language Pathology, 11*(1), 26-33. doi:10.1080/17549500802617689

Hoff, E., Core, C., & Bridges, K. (2008). Non-word repetition assesses phonological memory and is related to vocabulary development in 20- to 24-month-olds. *Journal of Child Language, 35*(4), 903-916. doi:10.1017/S0305000908008751

Holes, C. (1995). *Modern Arabic: Structures, Functions and varieties*. London: Longman.

Holm, A., Crosbie, S., & Dodd, B. (2007). Differentiating normal variability from inconsistency in children's speech: normative data. *International Journal of Language & Communication Disorders, 42*(4), 467-486. doi:10.1080/13682820600988967

Ingham, B. (1994). *Najdi Arabic: central Arabian*. Amsterdam: John Benjamin Publication.

Iuzzini, J., & Forrest, K. (2011). *Lexical and segmental inconsistency in disordered and typical developing speech.* Paper presented at the annual convention of the American Speech-Language-Hearing Association, San Diego, CA.

Jakobson, R. (1968). *Child language, aphasia and phonological universals, trans. A.R. Keiler.* The Hague: Mouton. (Originally Published as Kindersprache, Aphasie und allgemeine Lautgesetze. Uppsala: Almqvist & Wiksell, 1941).

James, D. G. (2006). *Hippopotamus is so hard to say: Children’s acquisition of polysyllabic words.* (Doctoral Thesis), The University of Sydney.

James, D. G., Van Doorn, J., & McLeod, S. (2008). The contribution of polysyllabic words in clinical decision making about children's speech. *Clinical Linguistics & Phonetics, 22*(4-5), 345-353. doi:10.1080/02699200801919240

Kent, R. D. (1992). The biology of phonological development. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 65-90). Timonium,MD: York Pess.

Kent, R. D. (2000). Research on speech motor control and its disorders: A review and prospective. *Journal of communication disorders, 33*(5), 391-428. doi:10.1016/S0021-9924(00)00023-X

Khattab, G. (2007). Lebanese Arabic speech acquisition. In S. McLeod (Ed.), *International Guide to Speech Acquisition* (pp. 300–312). Clifton Park, NY: Thomson Delmar Learning Assoc.

Khattab, G., & Al-Tamimi, J. (2013). Influence of geminate structure on early Arabic templatic patterns. In M. M. Vihman & T. Keren-Portnoy (Eds.), *The emergence of phonology: Whole-word approaches and cross-linguistic Evidence* (pp. 374). New York: Cambridge University Press.

Krishnan, S., Alcock, K. J., Mercure, E., Leech, R., Barker, E., Karmiloff-Smith, A., & Dick, F. (2013). Articulating novel words: Children's oromotor skills predict nonword repetition abilities. *Journal of Speech, Language, and Hearing Research, 56*(6), 1800-1812. doi:10.1044/1092-4388(2013/12-0206)

Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (2000). Prevalence and natural history of primary speech and language delay: Findings from a systematic review of the literature. *International Journal of Language and Communication Disorders, 35*(2), 165-188.

Lee, H. J., Kim, Y. T., & Yim, D. (2013). Non-word repetition performance in Korean-English bilingual children. *International Journal of Speech-Language Pathology, 15*(4), 375-382. doi:10.3109/17549507.2012.752866.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A Theory of Lexical Access in Speech Production. *Behavioral and Brain Sciences*, 22, 1-75.

Levelt, W. J. M. (2013). *A History of Psycholinguistics: the pre-Chomskyan era*. Oxford: Oxford University Press.

Lewis, B. A., Freebairn, L. A., & Taylor, H. G. (2000). Follow-Up of Children with Early Expressive Phonology Disorders. *Journal of Learning Disabilities, 33*(5), 433-444. doi:10.1177/002221940003300504

Mace, J. (1998). *Arabic Grammar: a reference guide*. Edinburgh: Edinburgh University Press.

Macrae, T. (2013). Lexical and child-related factors in word variability and accuracy in infants. *Clinical Linguistics & Phonetics, 27*(6-7), 497-507. doi:10.3109/02699206.2012.752867

Macrae, T., & Sosa, A. V. (2015). Predictors of token-to-token inconsistency in preschool children with typical speech-language development. *Clinical Linguistics & Phonetics, 29*(12), 922-937. doi:10.3109/02699206.2015.1063085

Majerus, S., Poncelet, M., Elsen, B., & Van der Linden, M. (2006). Exploring the relationship between new word learning and short-term memory for serial order recall, item recall, and item recognition. *European Journal of Cognitive Psychology, 18*(6), 848-873. doi:10.1080/09541440500446476

Mancini, F., & Voghera, M. (1994). Lunghezza, tipi di sillabe e accento in italiano [How Italians speak]. In T. De Mauro (Ed.), *Come parlano gli Italiani* (pp. 211- 241). Firenze, Italy: La Nuova Italia.

McLaughlin, S. (1998). *Introduction to language development*. San Diego, CA: Singular.

McLeod, S., & Bleile, K. (2003). *Neurological and developmental foundations of speech acquisition*. Paper presented at the American Speech-Language-Hearing Association Convention, Chicago.

McLeod, S., & Hewett, S. R. (2008). Variability in the production of words containing consonant clusters by typical 2-and 3-year-old children. *Folia Phoniatrica et Logopaedica, 60*(4), 163-172. doi:10.1159/000127835

Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology, 91*(1), 3. doi:10.1037/0022-0663.91.1.3

Metsala, J. L., & Chisholm, G. M. (2010). The influence of lexical status and neighborhood density on children's nonword repetition. *Applied Psycholinguistics, 31*(03), 489-506. doi:10.1017/S0142716410000081

Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations:

Precursors to phonemic awareness and early reading ability In J. L. Metsal & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89-120). Mahwah, NJ: Erlbaum.

Mullen, R., & Schooling, T. (2010). The national outcomes measurement system for pediatric speech-language pathology. *Language, Speech, and Hearing Services in Schools, 41*(1), 44-60. doi:10.1044/0161-1461(2009/08-0051)

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*(1), 61-78. doi:10.1044/1092-4388(2005/006)

Newbold, E. J., Stackhouse, J., & Wells, B. (2013). Tracking change in children with severe and persisting speech difficulties. *Clinical Linguistics & Phonetics, 27*(6-7), 521-539. doi:10.3109/02699206.2013.790479

Nip, I. S., Green, J. R., & Marx, D. B. (2009). Early speech motor development: Cognitive and linguistic considerations. *Journal of communication disorders, 42*(4), 286-298. doi:10.1016/j.jcomdis.2009.03.008

Pascoe, M., Rossouw, K., Fish, L., Jansen, C., Manley, N., Powell, M., & Rosen, L. (2016). Speech processing and production in two-year-old children acquiring isiXhosa: A tale of two children. *South African Journal of Communication Disorders, 63*(2), 1-15. doi:10.4102/sajcd.v63i2.134

Peters, A. M. (2001). Filler syllables: what is their status in emerging grammar? *Journal of Child Language, 28*(01), 229-242.

Pollock, K. E. (1991). The identification of vowel errors using traditional articulation or phonological process test stimuli. *Language, Speech, and Hearing Services in Schools, 22*(2), 39-50. doi:10.1044/0161-1461.2202.39.

Preston, J. L., & Edwards, M. L. (2009). Speed and accuracy of rapid speech output by adolescents with residual speech sound errors including rhotics. *Clinical Linguistics & Phonetics, 23*(4), 301-318. doi:10.1080/02699200802680833

RCSLT. (2011). *Royal College of Speech and Language Therapists Policy Statment: Developmental Verbal Dyspraxia*. Retrieved from London: www.rcslt.org

Robbins, J., & Klee, T. (1987). Clinical assessment of oropharyngeal motor development in young children. *Journal of Speech and Hearing Disorders, 52*(3), 271-277. doi:10.1044/jshd.5203.271

Roy, P., & Chiat, S. (2004). A Prosodically Controlled Word and Nonword Repetition Task for 2-to 4-Year-OldsEvidence From Typically Developing Children. *Journal of Speech, Language, and Hearing Research, 47*(1), 223-234. doi:10.1044/1092-4388(2004/019)

Rvachew, S., & Brosseau-Lapré, F. (2012). *Developmental Phonological Disorders: Foundations of clinical practice*. San Diego: Plural Publishing.

Rvachew, S., Hodge, M., & Ohberg, A. (2005). Obtaining and interpreting maximum performance tasks from children: A tutorial. *JOURNAL OF SPEECH LANGUAGE PATHOLOGY AND AUDIOLOGY, 29*(4), 146.

Rvachew, S., Ohberg, A., & Savage, R. (2006). Young children's responses to maximum performance tasks: preliminary data and recommendations. *Journal Of Speech Language Pathology and Audiology, 30*(1), 6.

Sadagopan, N., & Smith, A. (2008). Developmental changes in the effects of utterance length and complexity on speech movement variability. *Journal of Speech, Language, and Hearing Research, 51*(5), 1138-1151. doi:10.1044/1092-4388(2008/06-0222)

Salem, H. (2000). *Study of the acquisition of the syllable structure in sentance perspective in the speech of normal egyption children.* (Doctoral thesis), University of Alexandria.

Schaefer, B., Fricke, S., Szczerbinski, M., Fox-Boyer, A. V., Stackhouse, J., & Wells, B. (2009). Development of a test battery for assessing phonological awareness in German-speaking children. *Clinical Linguistics & Phonetics, 23*(6), 404-430. doi:10.1080/02699200902770187

Shaalan, S. (2010). *Investigating grammatical complexity in Gulf Arabic speaking children with specific language impairment (SLI).* (Doctoral thesis), University College London, London, United Kingdom.

Shahin, K. (2006). Remarks on the speech of Arabic-speaking children with cleft palate: three case studies. *Journal of Multilingual Communication Disorders, 4*(2), 71-77.

Shea, T., & Blodgett, E. (1994). *Spontaneous and imitated productions in preschoolers with phonological disorders.* Paper presented at the Annual Convention of the American Speech–Language–Hearing Association, New Orleans, LA, USA.

Shriberg, L. D., Austin, D., Lewis, B. A., McSweeny, J. L., & Wilson, D. L. (1997). The Percentage of Consonants Correct (PCC) MetricExtensions and Reliability Data. *Journal of Speech, Language, and Hearing Research, 40*(4), 708-722. doi:10.1044/jslhr.4004.708

Shriberg, L. D., & Kwiatkowski, J. (1994). Developmental phonological disorders I: A clinical profile. *Journal of Speech and Hearing Research, 37*(5), 1100-1126. Retrieved from

Smith, A., & Goffman, L. (2004). Interaction of motor and language factors in the development of speech production. In B. Maasen, R. D. Kent, H. Peters, P. H. Van Lieshout, & W. Hulstijn (Eds.), *Speech motor control in normal and disordered speech* (pp. 227-252): Oxford University Press.

Snowling, M., Chiat, S., & Hulme, C. (1991). Words, nonwords, and phonological processes: Some comments on Gathercole, Willis, Emslie, and Baddeley. *Applied Psycholinguistics, 12*(03), 369-373. doi:10.1017/S0142716400009279

Sosa, A. V. (2015). Intraword variability in typical speech development. *American Journal of Speech-Language Pathology, 24*(1), 24-35. doi:10.1044/2014_AJSLP-13-0148

Sosa, A. V., & Stoel-Gammon, C. (2006). Patterns of intra-word phonological variability during the second year of life. *Journal of Child Language, 33*(01), 31-50. doi:10.1017/S0305000905007166

Sosa, A. V., & Stoel-Gammon, C. (2012). Lexical and phonological effects in early word production. *Journal of Speech, Language, and Hearing Research, 55*(2), 596-608. doi:10.1044/1092-4388(2011/10-0113)

Speech disorders. (2011). *the Jeddah Institute for Speech and Hearing (JISH).* Retrieved from http://www.jish.org/SiteNav/SpeechLanguageServices/SpeechDisorders.aspx

St Louis, K. O., & Ruscello, D. M. (2000). *Oral Speech Mechanism Screening Examination (OSMSE)* (third ed.). austin, TX: Pro-ED.

Stackhouse, J., Vance, M., Michelle, P., & Bill, W. (2007). *Compendium of auditory and speech tasks: childrens Speech and Litracy Difficulties*: John Wiley & Sons.

Stackhouse, J., Vance, M., Pascoe, M., & Wells, B. (2007). *Compendium of auditory and speech tasks: Children's speech and literacy difficulties 4* John Wiley & Sons.

Stackhouse, J., & Wells, B. (1997). *Children's speech and literacy difficulties: A psycholinguistic framework*. London: Whurr Publisher.

Stoel-Gammon, C. (2011). Relationships between lexical and phonological development in young children. *Journal of Child Language, 38*(01), 1-34. doi:10.1017/S0305000910000425

Storkel, H. L. (2004). Do children acquire dense neighborhoods? An investigation of similarity neighborhoods in lexical acquisition. *Applied Psycholinguistics, 25*(02), 201-221. doi:10.1017/S0142716404001109

Summers, P., & Larson, G. (1992, November). *Patterns of articulatory shift under two testing conditions.* Paper presented at the the annual convention of the American Speech-Language Hearing Association, San Antonio, TX.

Sundström, S., Samuelsson, C., & Lyxell, B. (2014). Repetition of words and non-words in typically developing children: The role of prosody. *First Language*, 0142723714550213. doi:10.1177/0142723714550213

Thelen, E., & Bates, E. (2003). Connectionism and dynamic systems: Are they really different? *Developmental Science, 6*(4), 378-391. doi:10.1111/1467-7687.00294

Torrington Eaton, C., Newman, R. S., Ratner, N. B., & Rowe, M. L. (2015). Non-word repetition in 2-year-olds: Replication of an adapted paradigm and a useful methodological extension. *Clinical Linguistics & Phonetics, 29*(7), 523-535. doi:10.3109/02699206.2015.1029594

Tremblay, S., Houle, G., & Ostry, D. J. (2008). Specificity of speech motor learning. *Journal of Neuroscience, 28*(10), 2426-2434. doi:10.1523/JNEUROSCI.4196-07.2008

Tyler, A. (2010). Subgroups, comorbidity, and treatment implications. In R. Paul & P. Flipsen (Eds.), *Speech sound disorders in children: In honor of Lawrence D. Shriberg* (pp. 71-92). San Diego, CA: Plural.

Tyler, A., & Edwards, M. L. (1993). Lexical acquisition and acquisition of initial voiceless stops. *Journal of Child Language, 20*(02), 253-273. doi:10.1017/S0305000900008278

Vance, M. (1996). Portuguese translation of "assessing speech processing skills in children: a task analysis". In M. Snowling & J. Stackhouse (Eds.), *Dislexia, Fala e Linguagem: um Manual do Profissional.* So Pablo: artmet.
Vance, M., Stackhouse, J., & Wells, B. (2005). Speech-production skills in children aged 3-7 years. *Int J Lang Commun Disord, 40*(1), 29-48. doi:10.1080/13682820410001716172
Velleman, S. (2006). *Childhood apraxia of speech: Assessment/treatment for the school-aged child.* Paper presented at the Annual Convention of the American Speech-Language-Hearing Association, Miami, FL.
Versteegh, K. (1997). *The Arabic Language*. Edinburgh: Edinburgh University Press.
Vihman, M. M., & Keren-Portnoy, T. (2013). *The emergence of phonology: Whole-word approaches and cross-linguistic Evidence*. New York: Cambridge University Press.
Waring, R., & Knight, R. (2013). How should children with speech sound disorders be classified? A review and critical evaluation of current classification systems. *International Journal of Language & Communication Disorders, 48*(1), 25-40. doi:10.1111/j.1460-6984.2012.00195.x
Wells, B., Stackhouse, J., & Vance, M. (1999). La conscience phonologique dans le cadre d'une evaluation psycholinguistique de l'enfant. *Rééducation orthophonique, 197,* 3-12.
Weston, A. D. (1997). The influence of sentence elicitation variables on children's speech production. *Journal of Speech, Language, and Hearing Research, 40*(5), 975-989. doi:10.1044/jslhr.4005.975
Wilcox, K., Morris, S., Speaker, K., & Catts, H. (1996). The effect of syllable shape on articulatory rate and stability. In T. W. Powell (Ed.), *Pathologies of Speech and Language: Contributions of clinical phonetics and linguistics*. New Orleans: I.C.P.L.A.
Williams, P., & Stackhouse, J. (1998). Diadochokinetic Skills: Normal and Atypical Performance in Children Aged 3–5 Years. *International Journal of Language & Communication Disorders, 33*(sup1), 481-486. doi:10.3109/13682829809179472
Williams, P., & Stackhouse, J. (2000). Rate, accuracy and consistency: diadochokinetic performance of young, normally developing children. *Clinical Linguistics & Phonetics, 14*(4), 267-293. doi:10.1080/02699200050023985
Williams, P., Stephens, H., Williams, A., McLeod, S., & McCauley, R. (2004). Nuffield Centre Dyspraxia Programme. *Windsor: Miracle Factory*.
Zamuner, T. S. (2009). The structure and nature of phonological neighbourhoods in children's early lexicons. *Journal of Child Language, 36*(01), 3-21. doi:10.1017/S0305000908008829
Zamuner, T. S., Gerken, L., & Hammond, M. (2004). Phonotactic probabilities in young children's speech production. *Journal of Child Language, 31*(03), 515-536. doi:10.1017/S0305000904006233.
Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (1992). PLS-3: Preschool Language Scale-3. San Antonio, TX : The Psychological Corporation.

APPENDIXES

Appendix 1: The Initial Stimuli List of Words and their Syllable Structure

	Meaning in English	Structure of the word	Arabic words
Mono-syllables	Door	CVC	b æ b
	House		b e t
	Berar		d ʊ b
	Elephant		f i: l
	Button		z ɪ r
	Banana		m ɔ z
	hand		j ɛd
	Sun	CVCC	ʃ æ ms
	Date		t æ mr
	Monkey		g ɪ rd
	Dog		k ɛ lb
Bi-syllables	Cake	CVCV	k e: k æ/ ə
	Cup		ˈkæ sæ:
	Men’s head wear	CVCVC	ʕ i ˈ g æ l
	Camel		ˈdʒ æ m æ l
	Graps		ʕinæb (ʕunæb)
	Book		k iˈ t æ: b
	Moon		ˈg æ m æ ɾ
	Pen		ˈg æ l æ m
	Train		g i ˈtˤa: r
	Light bulb	CVC.CV	l æ ˈm.b æ
	Cockroach	CVC.CVC	sˤur. ˈsˤ æ:r
	Pepper		fil.fil
	Watermelon		ħ æb .ħ æb
Multisyllables	orange	CVC.CVCVC	b u r. t u ˈq æ: l

	Hat	CVC:VCV	q u: ˈb.b æ ʕ æ
	Eyeglasses		n æ ˈðˤ.ðˤ ɑ r æ
	Fridge		θ æ ˈl.læ dʒ æ
	Apple		t u ˈf .fæ ħ æ
	Couch	CVCVCV	k æ n æ b ə
	Fish		s æ m æ ˈk ə
	Telephone	CVCVCVC	t i ˈl i f o n
	Pillow	CVCVCCV	mæ xæ ˈd.da
	Spoon	CVC.CVCV	m ɛ l. ˈʕ æ g ə
	Balcony	CVCVCVCV	b æ l æ k ɔ n æ

Appendix 2: The Vowel Changes Leading To Creating the Non-words

Meaning	Real words (Arabic) - monosyllables	Vowel change	Non-word
door	b æ b	Near-open front long vowel /æ:/ →to front high long vowel [i:]; [bib], this non-word cold be an English word, or a sound describing a car horning, thus the long back vowel was used [o]	b o b
house	b e t	Changes to any Arabic vowel will always result in an English word – this item was excluded	Excluded/ not replacement
Bear	d ʊ b	Changing the short back high vowel /ʊ/ to a short low back vowel [ɑ] results in an illegal non-word. The vowel /ɑ / occurs in Arabic after an emphatic consonant. Thus, the short mid-high /ɔ/ was selected instead.	d ɔ b
Sun	ʃ æ ms	Front high long vowel /i:/ → front low long vowel [æ:]	ʃ i ms
Date(food)	t æ mr	Front high long vowel /i:/ → front low long vowel [æ:]	t i mr
Elephant	f i: l	Front high long vowel /i:/ → front low long vowel [æ:]	f æ: l
monkey	g ɪ rd	Mid-Front high short vowel /ɪ / → front low short vowel [a]	g a rd
dog	k ɛ lb	Mid-open mid-front short vowel /ɛ/→ near close mid-front short vowel /ɪ /	k ɪ lb
button	z ɪ r	Changing the short front vowel to the front low vowel [a] will result in an Arabic word "visit", and changing the vowel to any other front short vowels will still maintain the real-word. Therefore, the vowel was substituted to the short mid-back vowel [ɔ]	z ɔ r
banana	m ɔ z	Mid-back short mid high vowel to /ɔ/ → mid-back shorthigh vowel [ʊ]	m ʊ z
hand	j ɛd	Changing the vowel /ɛ/ to any other front vowels still maintains the real-word and dose not result in a non-word. Therefore, the back [ʊ]	j ʊ d
Meaning	**Real word - bisyllables**		**Non-word**
Cake	k e ˈk æ æ/ə	Front long vowels → substituted from high to low vowels and vice versa.	k æ k i
Cup	ˈkæ sæ:	substituting the front low long vowel /æ:/ to the long front high vowel [/i:] creates a real-word "my bag", thus, the back long high vowel was used instead [u]	ku su
Train	g i ˈtˤa: r	Front long vowels → substituted from high to low vowels and vice versa.	g æ: tˤi: r
(Men's head wear)	ʕ i ˈ g æ l	Front long vowels → substituted from high to low vowels and vice versa.	ʕ æ ˈ g i l
Camel	ˈdʒ æ m æ l	Front long vowels → substituted from high to low vowels and vice versa.	ˈdʒ i m i l
Grapes	ˈʕi næb (ʕunæb)	Front long vowels → substituted from high to low vowels and vice versa.	ˈʕ æ n i b
Book	k iˈ t æ: b	When changing the front close vowel /i/ to near-open front vowel [æ:] and the /æ:/ in the second syllable to [i] a real-word was created "writer" thus the vowel /i/ in the first syllable was changed to a back close long vowel [u].	k uˈ t i: b
Moon	ˈg æ m æ ɾ	Front low long vowels /æ:/ → front high long vowel [i:]	ˈg i m i r

Pen	ˈg æ l æ m	Front low long vowels /æ:/ → front high long vowel [i:]	ˈg i l i m
Light Bulb	ˈl æ m.b æ	Front low long vowels /æ:/ → front high long vowel [i:]	ˈl i m.b i
Cockroach	sˤur sˤ æ:r	Back long high vowel /u/ → to back low long vowel [ɑ:] vowels and vice versa. Front low long vowel [æ:] → front high long vowel /i:/	sˤ ɑ: r sˤ i r
Pepper	fil fil	Front high long vowels /i:/ → front low long vowel [æ:]	fæl fæl
Watermelon	ħ æb ħ æb	Front low long vowels /æ:/ → front high long vowel [i:]	ħ ib ħ ib

Appendix 3: Full ***Preliminary*** Stimuli List

Mono-syllables	**Meaning**	**Real word (Arabic)**	**Non-word**	**Syllable Sequence**
1	door	b æ b	b o b	b æ
2	Bear	d ʊ b	d ɔ b	d æ
3	Sun	ʃ æ ms	ʃ i ms	ʃ æ
4	Date(food)	t æ mr	t i mr	t æ
5	Elephant	f i: l	f æ: l	f æ
6	monkey	g ɪ rd	g a rd	g æ
7	Dog	k ɛ lb	k ɪ lb	k æ
8	Button	z ɪ r	z ɔ r	zæ
9	Banana	m ɔ z	m ʊ z	m æ
10	Hand	j ɛd	j ʊ d	j æ
Bi-syllables	**Meaning**	**Real word**	**Non-word**	**Syllable**
1	Cake	k e: ˈ k æ/ə	k æ ˈ k i	k æ k æ
2	Cup	ˈkæ sæ	ˈ ku su	ki si
3	Train	g i ˈtˤɑ: r	g æ ˈ tˤi r	g æ tˤ æ
4	(Men's head wear)	ʕ i ˈ g æ l	ʕ æ ˈ g i l	ʕ æ gæ
5	Camel	ˈdʒ æ m æ l	ˈ dʒ i m i l	dʒ æ m æ
6	Grapes	ˈʕinæb (ʕunæb)	ˈ ʕ æ n i b	ʕ æ næ
7	Book	k iˈ t æ: b	k u t i: b	k æ t æ
8	Moon	ˈg æ m æ ɾ	g i m i r	g æ m æ
9	Pen	ˈg æ l æ m	g u l u m	g æ l æ
10	Light Bulb	ˈl æ m.b æ	l i mb i	l æ mæ
11	Cockroach	sˤur sˤ æ:r	sˤ ɑ: r sˤ i r	sˤ æ sˤ æ
12	Pepper	fil fil	fæl fæl	f æ f æ
13	Watermelon	ħ æb ħ æb	ħ ib ħ ib	ħ æ ħ æ
Tri-syllables	**Meaning**	**Real word**	**Non-word**	**Syllable**
1	Orange	b u r. tu ˈq æ :l	b æ r. t æ q i: l	b æ t æ q æ
2	Hat	q u ˈb.b æ ʕ æ	q a ˈb.b u ʕ u	q æ b æ ʕ æ
3	Telephone	t i ˈl i f o n	t u l u f æ n	t æ l æ f æ
4	Pillow	mæ xæ ˈd.da	mixi ˈd.di	mu xu d u
5	Glasses	n æ ˈðˤ . ðˤ ɑ r æ	n i ˈðˤ. ðˤ i r i	n u ðˤ u r u
6	Couch	k æ n æ b ə	k i ˈn i b i	k u nu bu
7	Fridge	θ æ ˈl.læ dʒ æ	θ i ˈl.l u dʒ u	θ i l i dʒ i
8	Apple	t u ˈf .fæ ħ æ	t æ f.f u: ħ u	t æ f æ ħ æ
9	Fish	s æ m æ ˈk ə	s u m u ˈ k u	s i m i k i
10	Spoon	m ɛ l. ˈʕ æ g ə	m æ l. ˈʕ u g u	m æ ʕ æ g æ
11	Balcony	b æ l æ k ɔ n æ	b i l i k u n i	bæ læ kæ næ

Appendix 4: Ethical Approval for the Pilot Study

ETHICS REVIEWER'S COMMENTS FORM

This form is for use when ethically reviewing a research ethics application form.

1. Name of Ethics Reviewer:	Richard Body Catherine Tattersall Ray Wilkinson

2. Research Project Title:	Speech accuracy, rated and consistency in Arabic speaking children
3. Principal Investigator (or Supervisor):	Prof Joy Stackhouse
4. Name of Student (if applicable):	Noaf Al-Kheraiji
5. Academic Department / School:	Human Communication Sciences
6. I confirm that I do not have a conflict of interest with the project application	

7. I confirm that, in my judgment, the application should:

Be approved:	Be approved with *suggested* amendments in '8' below:	and/or	Be approved providing *requirements* specified in '9' below are met:	NOT be approved for the reason(s) given in '10' below:
	X			

8. Approved with the following suggested, optional amendments (i.e. it is left to the discretion of the applicant whether or not to accept the amendments and, if accepted, the ethics reviewers do not need to see the amendments):

- It might be nice to have a sentence or two in the information sheets about why a new speech assessment for Arabic speaking children is necessary.
- On the parent information sheet under 'does my child have to take part' you should add that they can withdraw without any negative consequences.
- Information Sheet: 'What will happen . .? ':*sgined* -> signed
- On your consent form there should be a space for the child's name as well as a space for the parent to sign otherwise you are not gaining consent for a specific participant.
- Appendix 1 , questionnaire: on this it would be better to use a code (linked to the consent form). That way you do not need to have the child's name on the questionnaire itself. Then when storing the data you can easily separate the identifiable consent form from the coded questionnaire.
- Clarify with supervisors what you will do if you identify a child that potentially has speech and language or other difficulties.
- Appendix 3: Information sheet for parents and head teacher: Ddepartment -> Department
- Appendix 5 (last para): porject -> project.

9. Approved providing the following, compulsory requirements are met (i.e. the ethics reviewers need to see the required changes):

10. Not approved for the following reason(s):

11. Date of Ethics Review: 30 November 2012

RA Body

Appendix 5: Illustration of the final set of pictures used for the "picture naming task" and the practice items.

	mono-syllables	**Bi-syllables**	**Tri-syllables**	**Practice Items**
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				

Note: All pictures were purchased from 123RF limited. The pictures are being used with permission under license (www.123rf.com) and are copyrighted property of their contribution or licensed partner.

Appendix 6: The Final Revised Stimuli list

Mono-syllables	**Meaning**	**Real word (Arabic)**	**Non-word**	**Syllable Sequence**
1	door	b æ b	b o b	b æ
2	Bear	d ʊ b	d æ b	d æ
3	Sun	ʃ æ ms	ʃ i ms	ʃ æ
4	Date(food)	t æ mr	t i mr	t æ
5	Elephant	f i: l	f æ: l	f æ
6	monkey	g ɪ rd	g a rd	g æ
7	Dog	k ɛ lb	k u l b	k æ
8	Button	z ɪ r	z ɔ r	zæ
9	Banana	m ɔ z	m ɛ z	m æ
10	Hand	j ɛd	j ʊ d	j æ
Bi-syllables	**Meaning**	**Real word**	**Non-word**	**Syllable**
1	Cake	k e: ˈ k æ/ə	k æ ˈ k i	k æ k æ
2	Cup	ˈkæ sæ	ˈ ku su	ki si
3	Train	g i ˈtˤɑ: r	g æ ˈ tˤi r	g æ tˤ æ
4	(Men's head wear)	ʕ i ˈ g æ l	ʕ æ ˈ g i l	ʕ æ gæ
5	Camel	ˈdʒ æ m æ l	ˈ dʒ i m i l	dʒ æ m æ
6	Grapes	ˈʕinæb (ʕunæb)	ˈ ʕ æ n i b	ʕ æ næ
7	Book	k i ˈ t æ: b	k u t i: b	k æ t æ
8	Moon	ˈg æ m æ ɾ	g u m u r	g æ m æ
9	Pen	ˈg æ l æ m	g u l u m	g æ l æ
10	Light Bulb	ˈl æ m.b æ	l u mb u	l æ mæ
Tri-syllables	**Meaning**	**Real word**	**Non-word**	**Syllable**
1	Orange	b u r. t u ˈq æ :l	b æ r. t æ q i: l	bæ tæ qæ
2	Hat	q u ˈb.b æ ʕ æ	q a ˈb.b u ʕ u	qæ bæ ʕæ
3	Telephone	t i ˈl i f o n	t u l u f æ n	t æ læ fæ
4	Pillow	mæ xæ ˈd.da	mu xu ˈd.du	mi xi di
5	Glasses	n æ ˈðˤ. ðˤ ɑ r æ	n i ˈðˤ. ðˤ i r i	nu ðˤ u r u
6	Couch	k æ ˈnæ bə	k u ˈnu bu	k i n i b i
7	Fridge	θ æ ˈl.læ dʒ æ	θ i ˈl.l u dʒ u	θ i l i dʒ i
8	Apple	t u ˈf .fæ ħ æ	t æ f.f u: ħ u	t æ f æ ħ æ
9	Fish	s æ m æ ˈk ə	s u m u ˈ k u	s i m i k i
10	Spoon	m ɛ l. ˈʕ æ g ə	m æ l. ˈʕ u g u	m æ ʕ æ g æ

Appendix 7: Examples of Questionnaire provided to parents/caregivers

This questionnaire is to help us analyse the data we collect for this project. It is not for any other purpose and will not affect any other services available for you or your child in any way. If you would prefer not to complete any of the questions, please leave blank and move onto the next one. All responses will be kept confidential.

Childs name:

Date of birth:

1- **Hearing**

Have you ever had concerns about your child's hearing?	YES/ NO
If yes, what caused this concern?	
Was help/treatment given?	YES/ NO
If yes, what help/treatment was given?	
Do you still have concerns?	YES/ NO
If yes, what concerns do you have?	

2- **Vision**

Have you ever had concerns about your child's vision?	YES/ NO
If yes, what caused this concern?	
Was help/treatment given?	YES/ NO
If yes, what help/treatment was given?	
Do you still have concerns?	YES/ NO
If yes, what concerns do you have?	

Continue of Appendix 7

3- Speech and Language

Approximately how old was your child when s/he began to say words?	
Approximately how old was your child when s/he began to say sentences?	
Have you ever had concerns about your child's speech or language development?	YES/ NO
If yes, what caused this concern?	
Was help/treatment given?	YES/ NO
Do you still have concerns?	YES/ NO
If yes, what concerns do you have?	
Has your child ever seen a speech and language therapist?	YES/ NO
If yes, Can you tell us what your childs speech therapist said about your child's speech /language difficulty?	
How many treatment sessions has your child attended?	

Appendix 8: Ethical Approval for the Main Study

ETHICS REVIEWER'S COMMENTS FORM

This form is for use when ethically reviewing a research ethics application form.

1. Name of Ethics Reviewer:	Richard Body Catherine Tattersall Ray Wilkinson

2. Research Project Title:	Speech Accuracy, Rate and Consistency in Arabic Speaking Children
3. Principal Investigator (or Supervisor):	Prof. Joy Stackhouse
4. Name of Student (if applicable):	Noaf Al-Kheraiji
5. Academic Department / School:	Department of Human Communication Sciences

6. I confirm that I do not have a conflict of interest with the project application

7. I confirm that, in my judgment, the application should:

Be approved:	**Be approved with *suggested* amendments in '8' below:**	and/or	**Be approved providing *requirements* specified in '9' below are met:**	**<u>NOT</u> be approved for the reason(s) given in '10' below:**
	X			

8. Approved with the following suggested, optional amendments (i.e. it is left to the discretion of the applicant whether or not to accept the amendments and, if accepted, the ethics reviewers <u>do not need</u> to see the amendments):

- Application Form: the corhorts of children with speech difficulties have different age ranges (3-5 in Saudi Arabia; 3-7 in Sheffield). The reason for this is not apparent from the application. Researcher should consider if this difference is warranted.
- Consent Forms, point 5: The use of the phrase 'tested again ... prior to his/her initial test' is confusing. Reword.

9. Approved providing the following, compulsory requirements are met (i.e. the ethics reviewers need to see the required changes):

10. Not approved for the following reason(s):

11. Date of Ethics Review: 03.05.13

RABody

Appendix 9: Example of children's scoring sheet
only the behavioural measures presented in the current study are included in this sheet.

Appendix 9a : Front scoring sheet for score calculations

Child's Code:__________________

SCORE CALCULATION

#RP=Number of Repetitions (1=5 Repetitions, 0= 4 Repetitions)
PN=Picture naming Accuracy
A1=Whole word Accuracy of one repetition
Ar= proportion of Accurate productions of consecutive repetitions
C= Consistency of consecutive repetitions

			Segmental						
			Consonants						
TASK		**# RP**	**PN**	**A1**	**Ar**	**C**	**SE**	**SEQ**	**LEX**
REAL WORD	**RW**								
RWR mono-syllables	WR1								
RWR bi-syllables	WR2								
RWR tri-syllables	WR3								
TOTAL SCORE	WR Total								
NON-WORD	**NWR**								
NWR mono-syllables	NWR1								
NWR bi-syllables	NWR2								
NWR tri-syllables	NWR3								
TOTAL SCORE	NWR Total								
SYLLABE SEQUENCE	**SS**								
SS mono-syllables	SS1								
SS bi-syllables	SS2								
SS tri-syllables	SS3								
TOTAL SCORE	SS Total								

School: ______________________

Class: ________________________

Child's Code: _________________

Gender: ______________________

Audio	Folder	
	Track	

	Day	Month	Year	Age
Date of birth (Gorgonian)				
Date of birth (Hijri)				
Date of T-1				
Date of T-2 (If applicable)				

Comments/Observations:

____

____

____

____

____

____

ADMINISTERED TESTS

TESTS	**MARK (X) FOR TASKS COMPLETED BY CH**			**REFUSED**
PLS-3 Receptive				
PLS-3 Expressive				
Picture Naming				
Real-word Repetition (RWR)	Mono-syllabic Items	Bi-syllabic Items	Tri-syllabic Items	
Non-word Repetition (NWR)	Mono-syllabic Items	Bi-syllabic Items	Tri-syllabic Items	
Syllable Sequence Repetition	Mono-syllabic Items	Bi-syllabic Items	Tri-syllabic Items	

TEST SCORES

TASK	**Raw Score**	**Standard Score**	**Age Equivalent**
PLS-3 Receptive			
PLS-3 Expressive			
PLS-3 Total language			

Appendix 9 b: An example of the transcription sheet, where scores are calculated per item. This sheet is for the real word repetition task.

			Single Repetition	Multiple Rapid Production – DDK -					Scores					
#	Adult Model	Picture naming	Child 1st attempt	Repetition 1	Repetition 2	Repetition 3	Repetition 4	Repetition 5		PN A	A1	Ar	C	C5 SL
Monosyllables														
1	b æ b													
2	d ʊ b													
3	ʃ æ ms													
4	t æ mr													
5	f i: l													
6	g ɪ rd													
7	k ɛ lb													
8	z ɪ r													
9	m ɔ z													
10	j ɛd													
Bisyllables														
11	=k e: k æ													
12	ˈkæ sæ:													

13	g i ˈtˤɑː r													
#	Adult Model	Picture naming	Child Model 1st attempt	Repetition 1	Repetition 2	Repetition 3	Repetition 4	Repetition 5		PN A	A1	Ar	C	C5 SL
14	=ʕ i ˈg æ l													
15	ˈdʒ æ m æ l													
16	ʕinæb/ʕunæb													
17	k i' t æ: b													
18	ˈg æ m æ ɾ													
19	ˈg æ l æ m													
20	ˈl æ m.b æ													
Tri-syllables														
21	bur.t uˈq æ: l													
22	q u:ˈb.bæʕ æ													
23	t i ˈl i f o n													
24	mæ xæ ˈd.da													
25	næˈðˤ.ðˤɑ ræ													
26	k æ n æ b æ													
27	θ æˈl.lædʒ æ													
28	t u ˈf .fæ ħ æ													
29	s æ m æ ˈk ə													
30	mɛ l.ˈʕ æ g ə													

Appendix 10: Values of Shapiro-Wilk's test (*significant at alpha level > .05; non-normal distribution), Skewness and Kurtosis (Z-value should be within -+ 1.96, negative values= negative skeweness; missing values = ceilling score). A1= Accuracy of one repetition; Ar= Accuracy of consecuitve repetitions, C=consistancy

Test Measure-Condition	Syllable length	Age group	Shapiro-Wilk's test; p significnat at > .05	Skewness			Kurtosis		
				Value	SE	Z-Value	Value	SE	Z-Value
Accuracy1 - Real words	monosyllabic items	3 YO	.353	-0.730	.687	--1.063	- 0.233	1.334	-0.175
		4YO	.006	-0.565	0.717	-0.788	-1.816	1.400	-1.297
		5YO	.000	-1.620	0.717	-2.259	0.735	1.400	0.525
	Bisyllabic items	3 YO		-1.478	0.687	-2.151	0.889	1.334	0.666
		4YO	.002	-1.600	0.717	-2.232	3.194	1.400	2.281
		5YO	.006	-	-	-	-	-	-
	Trisyllabic items	3 YO	.009*	-1.597	0.687	-2.325	2.058	1.334	1.543
		4YO	.001*	-2.330	0.717	-3.249	6.183	1.400	4.416
		5YO	.001*	-2.506	0.717	-3.495	6.337	1.400	4.525
Accuracy 1- Non-words	monosyllabic items	3 YO	.02	-0.897	0.717	-1.251	-0.994	1.400	-0.71
		4YO	.113	-0.947	0.717	-1.320	-0.018	1.400	-0.013
		5YO	.00	-1.620	0.717	-1.259	0.735	1.400	0.525
	Bisyllabic items	3 YO	.08	-0.760	0.717	-1.0599	-1.088	1.400	-0.777
		4YO	.04	-1.094	0.717	-1.526	0.611	1.400	0.436
		5YO		-	-	-	-	-	-
	Trisyllabic items	3 YO	.44	-0.851	0.717	-1.187	0.414	1.400	0.296
		4YO	.013	-1.777	0.717	-2.478	4.315	1.400	3.082
		5YO	.00	-1.289	0.717	-1.798	0.770	1.400	0.55
Accuracy1- Syllable sequences	monosyllabic items	3 YO	.00	-1.192	0.717	-1.662	-0.446	1.400	-0.319
		4YO	.00	-0.857	0.717	-1.195	-1.714	1.400	-1.224
		5YO	.00	-3.000	0.717	-4.184	9.000	1.400	6.429
	Bisyllabic items	3 YO	.02	-1.203	0.717	-1.678	0.165	1.400	0.118
		4YO	.006	-1.600	0.717	-2.232	3.194	1.400	2.281
		5YO	.00	-	-	-	-	-	-
	Trisyllabic items	3 YO	.14	0.033	0.717	0.046	-2.010	1.400	-1.436
		4YO	.003	-2.160	0.717	-3.013	5.460	1.400	3.9
		5YO	.00	-2.121	0.717	-2.958	4.647	1.400	3.319

Continue Appendix 10

Test Measure-Condition	Syllable length	Age group	Shapiro-Wilk's test	Skewness			Kurtosis		
				Value	SE	Z-Value	Value	SE	Z-Value
Ar- Real words	monosyllabic items	3 YO	.90	-0.662	0.687	-0.964	0.674	1.334	0.505
		4YO	.40	-0.236	0.717	-0.329	-1.236	1.400	-0.883
		5YO	.02	-1.635	0.717	-2.280	1.669	1.400	1.192
	Bisyllabic items	3 YO	.151	-0.961	0.687	-1.399	0.034	1.334	0.025
		4YO	.045	-1.347	0.717	-1.879	1.867	1.400	1.334
		5YO	.00	-2.506	0.717	-3.495	6.337	1.400	4.526
	Trisyllabic items	3 YO	.42	-0.675	0.687	-0.983	-0.629	1.334	-0.472
		4YO	.25	-1.503	0.717	-2.096	2.355	1.400	1.682
		5YO	.00	-2.387	0.717	-3.329	6.076	1.400	4.34
Ar- Non-words	monosyllabic items	3 YO	.54	-0.559	0.717	-0.7796	-0.113	1.400	-0.081
		4YO	.69	-0.193	0.717	-0.269	-0.643	1.400	-0.459
		5YO	.062	-0.889	0.717	-1.2399	-0.441	1.400	-0.315
	Bisyllabic items	3 YO	.17	-0.884	0.717	-1.233	-0.444	1.400	-0.317
		4YO	.36	-1.140	0.717	-1.5899	1.358	1.400	0.97
		5YO	.01	-1.145	0.717	-1.597	0.064	1.400	0.0457
	Trisyllabic items	3 YO	.84	-0.132	0.717	-0.184	-0.642	1.400	-0.459
		4YO	.69	-0.924	0.717	-1.289	1.125	1.400	0.804
		5YO	.030	-1.428	0.717	-1.991	1.269	1.400	0.906
Ar- Syllable sequences	monosyllabic items	3 YO	.002	-0.549	0.717	-0.766	-2.011	1.400	-1.436
		4YO	.00	-0.857	0.717	-1.195	-1.174	1.400	-0.839
		5YO	.00	-3.000	0.717	-4.184	9.000	1.400	6.429
	Bisyllabic items	3 YO	.07	-0.916	0.717	-1.278	-0.479	1.400	-0.342
		4YO	.00	-1.935	0.717	-2.699	4.148	1.400	2.963
		5YO	.00	-3.000	0.717	-4.184	9.000	1.400	6.429
	Trisyllabic items	3 YO	.68	0.844	0.717	1.177	0.665	1.400	0.475
		4YO	.29	-1.317	0.717	-1.837	2.501	1.400	1.786
		5YO	.00	-1.400	0.717	-1.953	0.953	1.400	0.681

Continue Appendix 10

Test Measure-Condition	Syllable length	Age group	Shapiro-Wilk's test	Skewness			Kurtosis		
				Value	SE	Z-Value	Value	SE	Z-Value
C- Real words	monosyllabic items	3 YO	.06	-0.467	0.717	-0.651	-1.355	1.400	-0.968
		4YO	.01	-1.595	0.717	-2.225	2.862	1.400	2.044
		5YO	.001	-1.440	0.717	-2.008	0.764	1.400	0.546
	Bisyllabic items	3 YO	.05	-0.658	0.717	-0,918	-1.027	1.400	-0.734
		4YO	.0	-1.925	0.717	-2.685	3.380	1.400	2.414
		5YO	.0	-2.506	0.717	-3.495	6.337	1.400	4.526
	Trisyllabic items	3 YO	.095	-0.781	0.717	-1.089	-1.403	1.400	-1.002
		4YO	.06	-0.910	0.717	-1.269	1.115	1.400	0.796
		5YO	.06	-1.289	0.717	-1.798	0.770	1.400	0.55
C - Non-words	monosyllabic items	3 YO	.023	0.014	0.717	0.0915	-1.983	1.400	-1.416
		4YO	.00	0.697	0.717	0.972	-1.344	1.400	-0.96
		5YO	.00	-0.242	0.717	-0.338	-1.338	1.400	-0.956
	Bisyllabic items	3 YO	.32	-1.587	0.717	-2.213	2.503	1.400	1.788
		4YO	.031	-1.105	0.717	-1.541	-0.213	1.400	-0.152
		5YO	.02	-1.394	0.717	-1.944	0.184	1.400	0.131
	Trisyllabic items	3 YO	.2	-1.279	0.717	-1.784	2.362	1.400	1.687
		4YO	.06	-0.454	0.717	-0.633	-1.583	1.400	-1.131
		5YO	.03	-0.578	0.717	-0.806	-1.466	1.400	-1.047
C- Syllable sequences	monosyllabic items	3 YO	.001	-3.000	0.717	-4.184	9.000	1.400	6.429
		4YO	-	-	-	-	-	-	-
		5YO	-	-	-	-	-	-	-
	Bisyllabic items	3 YO	.009	-2.076	0.717	-2.895	4.399	1.400	3.099
		4YO	.00	-3.000	0.717	-4.184	9.000	1.400	6.429
		5YO	.00	-3.000	0.717	-4.184	9.000	1.400	6.429
	Trisyllabic items	3 YO	.007	-0.867	0.717	-1.209	1.099	1.400	0.785
		4YO	.004	-0.959	0.717	-1.338	1.133	1.400	0.809
		5YO	.00	-0.467	0.717	-0.651	-1.355	1.400	-0.968