Arabic Fluency Assessment: Procedures for Assessing Stuttering in Arabic Preschool Children

Roa'a Alsulaiman

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

Speech Research and Experimental Intervention Lab
Division of Psychology and Language Sciences
Faculty of Brain Sciences
University College London
2022

Declaration of Authorship

I, Roa'a Alsulaiman, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Acknowledgment

I remember wanting to focus my PhD on language development. Like good prospective PhDs would do, I contacted professors "around the globe" asking if they were taking students that fall. Searching UCL, I came across Professor Peter Howell, and thought I would give his area a try. He replied and the story of the best educational journey a PhD student can ask for began since then. I will forever be indebted to you, Pete, as a mentor and a role model. Indeed, you instilled my interest in understanding speech processes and how it can go wrong!

I would like to thank my secondary supervisor, Professor John Harris, for introducing me to the fascinating world of phonetics and phonology, and for his valuable comments on my testing materials.

A big thank you to my examiners, Professors Steven Frisson and Andrew Nevins. I truly enjoyed discussing my thesis with you, and I appreciate your shared motivation in producing the best thesis possible.

To my cousin, Aseel Alkhadhi, thanks for keeping your promise to always be available when I need ventilation; and for giving me tips at different milestones during the PhD

My old and gold friend, Shatha AlNafe'a, for always letting me decide where we eat at the Brunswick; and for entertaining me about your experiments in oral medicine and oncology, for real!

A special thanks to the stuttering support groups in Saudi Arabia who helped tremendously in recruiting many of the participants: and to the two schools, in Riyadh and London, in which pre-schoolers testing took place. I very much appreciate you being generous with time and space to help make this research possible.

Finally, to my front-line team, my siblings, Osama, Ziad, Doa'a, Lamees and Iyad. My words can't express how much I appreciate you being an endless source of love, support and encouragement, at all times. I love you آل منصور

Conference Presentations

1. **Alsulaiman**, R., Howell, P. (2019). Assessment of speech fluency in preschool-aged children: scale structure, design and development. Poster presented at UCL Psychology and Language Science peer conference, Cumberland, United Kingdom.
2. **Alsulaiman**, R., Howell, P. (2019). Indications of speech fluency elicited from Arabic Grade-school children. Poster presented at the Child Language Symposium in Sheffield, United Kingdom
3. **Alsulaiman**, R. Harris, J. & Howell, P. (2020). Screening English and Arabic children for speech dysfluency. Poster presented at the 2nd workshop on Speech perception and production, London United Kingdom
4. **Alsulaiman**, R., Rachmand, A. & Howell, P. (2020). Speech fluency assessment for adults and children: self-report scales design and development. Poster presented at the 2nd workshop on Speech perception and production, London United Kingdom
5. **Alsulaiman**, R., Sharif, H. & Howell, P. (2020). Comorbid Child-onset Fluency Disorder and ADHD. Poster presented at the European Congress of Psychiatry.
6. **Alsulaiman**, R. Harris, J. & Howell, P. (2020). Treating Word-finding difficulty in Arabic Children with English as Additional Language. Poster presented at MPAL, Max Plank Institute for Psycholinguistics. Nijmegen, The Netherlands.
7. **Alsulaiman** & Howell, P. (2021). Children fluency assessment: development and validation of a self-report instrument for pre-schoolers. Talk presented at Oxford Dysfluency Conference. Oxford, United Kingdom
8. **Alsulaiman**, R. Harris, J. & Howell, P. (2021). Screening Arabic and English children for stuttering and treating word-finding difficulty. Poster presented at International Association for the study of Child Language, Philadelphia, United Stated

بِسۡمِ ٱللَّهِ ٱلرَّحۡمَٰنِ ٱلرَّحِيمِ

وَقُل
رَّبِّ زِدۡنِي عِلۡمٗا

/O my Lord advance me in knowledge/

Dedication With a deep feeling of gratitude, to my warm mother, for protecting me with her lovely prayers, until the last minute before my PhD exam; and to my kind father, the first and best scientist I've known, for exemplifying what a scientist should do to accomplish. "O my Lord bestow on them your mercy as they cherished me when I was small" (Holy Quran, Isra 17:23-24).

Abstract

The primary aim of this thesis was to screen school-aged (4+) children for two separate types of fluency issues and to distinguish both groups from fluent children. The two fluency issues are Word-Finding Difficulty (WFD) and other speech disfluencies (primarily stuttering). The cohort examined consisted of children who spoke Arabic and English. We first designed a phonological assessment procedure that can equitably test Arabic and English children, called the Arabic English non-word repetition task (AEN_NWR). Riley's Stuttering Severity Instrument (SSI) is the standard way of assessing fluency for speakers of English. There is no standardized version of SSI for Arabic speakers. Hence, we designed a scheme to measure disfluency symptoms in Arabic speech (Arabic fluency assessment). The scheme recognizes that Arabic and English differ at all language levels (lexically, phonologically and syntactically).

After the children with WFD had been separated from those with stuttering, our second aim was to develop and deliver appropriate interventions for the different cohorts. Specifically, we aimed to develop treatments for the children with WFD using short procedures that are suitable for conducting in schools. Children who stutter are referred to SLTs to receive the appropriate type of intervention. To treat WFD, another set of non-word materials was designed to include phonemic patterns not used in the speaker's native language that are required if that speaker uses another targeted language (e.g. phonemic patterns that occur in English, but not Arabic). The goal was to use

these materials in an intervention to train phonemic sequences that are not used in the child's additional language such as the phonemic patterns that occur in English, but not Arabic. The hypothesis is that a native Arabic speaker learning English would be expected to struggle on those phonotactic patterns not used in Arabic that are required for English.

In addition to the screening and intervention protocols designed, self-report procedures are desirable to assess speech fluency when time for testing is limited. To that end, the last chapter discussed the importance of designing a fluency questionnaire that can assess fluency in the entire population of speakers. Together with the AEN_NWR, the brief self-report instrument forms a package of assessment procedures that facilitate screening of speech disfluencies in Arabic children (aged 4+) when they first enter school. The seven chapters, described in more detail below, together constitute a package that achieves the aims of identifying speech problems in children using Arabic and/or English and offering intervention to treat WFD.

Impact Statement

Speech production difficulties such as stuttering are very common and should be identified at an early age. Stuttering should also be distinguished from other speech disfluencies such as word-finding difficulties (WFD). This is particularly important in the context of the impact WFD can have on everyday communication and education. Children with English as an additional language (EAL), including Arabic pre-schoolers with EAL are at an additional challenge; they may have difficulties locating English vocabulary items during speech articulation. Consequently, they may experience WFD on some words which interrupts the forward flow of their speech. This type of disfluency can be treated in-school by trained staff or school teachers using practical procedures (efficient and not time-consuming).

Overall, the work in this thesis was motivated by the paucity of standardized instruments that can equally assess Arabic and/or English children for speech disfluencies. The thesis presents a package of practical instruments that can be used to screen preschool children for stuttering when they first enter school. Such instruments are crucial knowing that the number of children learning English as an additional language has increased and continues to do. Equally important, having these instruments available aids in identification of: (1) stuttering in which case children are referred to speech and language therapy to receive intervention; and (2) word-finding difficulty which can be enhanced by training procedures such as phonological training which we test in this thesis. When the findings of our research are used in schools, this will therefore lead to significant improvements in the quality of life of Arabic children with WFDs. Furthermore, the findings in this thesis can be further developed to establish assessment guidelines that can eventually form a national screening program for pre-school children, similar to what is available for hearing and eyesight.

The research in this thesis also presents novel findings with respect to phonological factors in Arabic that attract stuttering. For the first time in the literature, a detailed analysis of the linguistic factors associated with predicting stuttering in Arabic in children and adults is provided in this thesis. The findings from the linguistic analysis (Chapter 3) opens new avenues of research exploring severity of stuttering in Arabic taking into consideration the unique morphological and phonological characteristics of the language. Further work on these findings can give a better understanding of how and why there is a link between learning EAL and first occurrence of speech disfluencies. Our findings open avenues for research on the application of machine learning techniques to automatically score speech utterances such as the non-word repetition task that was presented as part of the screening instruments.

In conclusion, the Arabic language is in its infancy as far as stuttering research is concerned. There is still a lot to learn about speech disfluencies in Arabic, and in Arabic children with EAL. Nevertheless, with continued research and the work in this thesis, we hope that we will continue to expand on the current data and findings so that researchers and clinicians will have clearer guidelines for the assessment and treatment of Arabic speakers who stutter.

Table of Contents

List of Abbreviations

AIPC – Arabic Index of Phonological Complexity

AWNS – Adults who do not stutter

AWS – Adults who stutter

CWNS – Children who do not stutter

CWS – Children who stutter

DSM - American Psychiatric Association Diagnostic Manual

EAL – English as an additional Language

EO – English Only

FS – Fluent Speakers

IPC – Index of Phonological Complexity

NWR – non word repetition

PC – Physical concomitant

RT (Reaction Time in milliseconds)

SLT – Speech and Language Therapist

SSI – Stuttering Severity Instrument

WFD – Word Finding Difficulty

WWR – Whole word repetition

List of Tables

List of Figures

Summary of Chapter Contents

Chapter 1: General Introduction

In the introductory chapter, a selection of theoretical approaches to stuttering are examined; and the general features of stuttering are also reviewed. The nature of WFD is then briefly outlined, mainly to highlight differences in symptoms between the two communication disorders.

Chapter 2: Identifying types of disfluencies: The non-word repetition task as a potential marker of speech disfluency

A test was required that can distinguish children with WFD from those with other fluency problems such as stuttering. To this end, we devised a non-word repetition test that should be sensitive to fluency difficulty but not WFD (Howell et al., 2017). We first investigated the phonological structures of Arabic and English to identify the features that they share. This allowed the AEN_NWR to be designed so that it provided a fair test that is applicable to children who speak Arabic, English or both languages. The AEN_NWR uses non-words that are generated based on shared phonotactic properties of Arabic and English. The AEN_NWR can be validated for English speakers using Riley's stuttering severity instrument (SSI_3; 1994). The SSI allows estimation of stuttering severity based on analysis of spontaneous speech sample. The instrument uses the percentage of stuttered syllables, the average duration of the three longest stutters and a measure of physical concomitant as an index of fluency difficulty. However, there is no standardized form of SSI_3 that can be used for Arabic speakers. Thus, an Arabic fluency scheme (A_FS) was designed as part of the screening package to measure speech disfluency when children first enter school. Details on how the Arabic fluency scheme was designed and tested are provided in Chapter 3.

Chapter 3: Phonological factors that attract stuttering in Arabic: Towards formulation of an Arabic fluency scheme

The purpose of this study was to study investigate factors that affect stuttering in Arabic and use them as a basis to develop a procedure that can better assess stuttering in Arabic. Speech samples were collected from adults who stutter (AWS) and children who stutter (CWS) in Arabic and the samples were first analyzed using a direct translation of SSI from English to Arabic. At the same time, issues not addressed or not relevant to Arabic by the English form of SSI were noted. Next, the samples were re-analysed taking into consideration the parameters of the Arabic index of phonetic complexity (AIPC; Al-Tamimi, Khamaiseh & Howell, 2013). The AIPC was developed for assessment of the idiosyncratic phonological, phonotactic and morphological properties of Arabic relative to the English form of IPC (Jakielski, 1998). The findings were integrated into a two-step algorithm that(1) counted the number of syllables and (2) counted the number of disfluencies in extracts of Arabic speech incorporating features noted during the SSI analysis. The two algorithms together represent components of the proposed Arabic fluency scheme (A_FS) that parallels the syllable and stutter counts done in SSI.

Chapter 4 Towards Establishing Instruments to assess fluency in Arabic: Validation Study

The purpose of this chapter was to validate the AEN_NWR. In order to provide support that the AEN_NWR is a reliable measure of phonological skills and speech fluency, the work in this chapter tested for a relation between AEN_NWR scores and the percentage of stuttered syllables (%SS). Performance on A_FS was also compared with the %SS. Hence, the second goal was establishing the new A_FS as a framework for the characterization of syllabification and syllabic structure of words in spontaneous Arabic speech.

Chapter 5: Does phonological priming improve WFD in Arabic children with English as additional language (EAL), English only (EO) children and children with EAL other than Arabic?

This chapter presents an innovative form of phonological intervention for Arabic children with EAL who experience WFD. The technique involved providing phonological training with non-words, which focused on the difficult structures of the English language for Arabic speakers. The hypothesis was that WFD can result when children who are native speakers of one language (e.g., Arabic) have to use unfamiliar phonological structures when they use an additional language. This predicts that non-word repetition material designed to train Arabic children with EAL on phonemic sequences needed for English that are absent in Arabic should improve word-finding. Whilst the AEN_NWR in Chapter 2 employed phonological material that is equally easy for children speaking English, Arabic or both languages to use, it also excluded NWR material that would pose articulatory challenges for native speakers of one of these languages that are acquiring the other. In this study, we used the excluded material to train Arabic children with aspects of English phonology that they lack in order to improve word-finding abilities. A phonological training intervention protocol was designed in which the NWR stimuli were used as training material. Two groups of children were tested; Arabic children with EAL in Saudi Arabia and English children who are studying Arabic as additional language in a Saturday school in the UK. Comparisons were made between the groups. We tested the hypothesis that phonological training should prime a child's use of unfamiliar phonological networks and improve lexical access. A picture-naming task was conducted as part of the procedure to measure the children's speech production performance. Pictures in this task carried phonological structures that correspond to the non-word primes. It was expected that priming effects will arise when the primes activate the phonological representations associated with the pictures and thus reaction time will reduce following the intervention. Several other language measures were obtained to evaluate the effectiveness of the intervention in improving (1) phonological ability as measured by the AEN_NWR; (2) speech comprehension; and (3) overall measures of disfluency (i.e. %Whole Word repetition as an indicator of WFD, %Stuttered syllables). These measures were taken prior to the intervention (i.e., pre-test), immediately following the intervention and around two-weeks later in a follow-up session.

Chapter 6: Self–report measures of speech disfluency

We outline the importance of designing a fluency questionnaire that (1) meets the new DSM_V criteria for assessing disfluency by including measures of social anxiety etc.; and (2) considers fluency on a continuum rather being an either/or condition (fluent/disfluent).

Chapter 7: General Discussion

The final chapter provides an overall summary and explores the theoretical significance of the results. It is focused on the approaches that have been introduced in the thesis and how the findings can differentiate these theories. The problems encountered in the design and research practice are highlighted and future directions for this line of investigation are suggested.

Chapter 1 Introduction to Disfluencies in Spontaneous Speech

" By speech alone the eye and ear, nay the feelings of all senses, are united in one, and centre in commanding thought, to which the hands and other members are only obedient instruments…"*Jonathan Rée", I see a voice*

1.1 Overview of the introduction

Taken together, the theoretical work and the focused empirical investigations in this thesis contribute to the main goal of assessing speech disfluencies in young children. The theoretical work in the current chapter addresses the general question of distinguishing between different disfluency symptoms; particularly when the speakers use (1) Arabic as their first language and English as an additional one; (2) English as a first language and Arabic as an additional one; and (3) a mixture of both languages. An overview of the existing procedures that assess children speech disfluency is then presented.

The purpose of this chapter is to set the stage for the importance of assessing speech fluency in preschool children. In order to familiarize the reader with some of the issues addressed in this thesis, this chapter briefly discusses definitions and general features of stuttering. Differences between stuttering and other speech disruptions that tend to occur frequently in preschool children who speak English only (EO) or English as an additional language (EAL) are highlighted. The specific research pertaining to models of fluency control and the clinical pattern of WFD are reviewed.

1.2 What is Stuttering?

Humans are able to produce fluent speech rapidly. Speech can easily be produced at the rate of two to three words per second; these are selected from thousands of words in the mental lexicon. This is done effortlessly (Levelt, 1983). However, the production of spontaneous speech is not always without errors. In fact, speech production is an extremely complex process that depends on the precise coordination of laryngeal, orofacial, and respiratory muscles (Simonyan et al., 2016). The flow of speech may be disrupted if the speaker detects an error in any of these aspects. These disruptions create speech disorders such as stuttering.

Although stuttering has been researched for many years, there is still no consensus on its definition. Most definitions are descriptions of a list of behaviours that are common within people who stutter. Wingate (1964) gave a well-respected definition of stuttering based on a careful analysis of observable elements identifiable in disfluent speakers. It emphasized three basic areas that are considered to define stuttering. The main emphasis was on (1) speech characteristics; and the other two were (2) accessory features and (3) other associated features. The three areas are described in more details below.

Speech characteristics. Stuttering is characterized by interruptions to the forward flow of speech (Sheehan, 1953). The type of disruption in stuttering is different from the normal disfluencies experienced by fluent speakers. There are three particular disfluencies in stuttered speech; sound and syllable repetition, prolongation of voiced or voiceless sounds and a failure to produce a sound, known as “breaks”. The location of stuttered events (sounds or syllables) is not random within a stream of connected speech. There is a tendency for more stuttering on word-initial syllables than on word-final syllables (Hubbard, 1998; Richels et al., 2010). This suggests that initiation of speech sounds might pose a challenge to speakers who stutter.

Accessory features. Disruptions in speech are sometimes accompanied by accessory activities involving movement of body structures (Wingate, 1964); these may be voluntary or involuntary. Voluntary movements include tapping of the hand or feet during a speech block (Riley, 1972). Involuntary movements take the form of abnormal muscle tension particularly in face muscles such as rapid eye blinks that accompany stuttering moments.

Associated features involve the speaker's personal reactions, feelings or attitudes toward stuttering; commonly referred to as affective components. For a long time, it was not clear whether anxiety occurred as a result of stuttering or whether speakers were anxious because of stuttering. However, recent studies have showed it is an association between social anxiety and stuttering. It is likely that stuttering moments are accompanied by negative feelings which may increase susceptibility to social difficulties (Iverach & Rape, 2014). The relationship between fluency and social anxiety has been investigated extensively (e.g., Mahr & Torosian, 1999, Messenger et al., 2004). These studies used self-report instruments such as fear of negative evaluation (FNE; Watson & Friend, 1969) to examine social anxiety in fluent and dysfluent speakers. The association between social anxiety and type of speaker highlights the clinical importance of measuring both dimensions to determine how they interact. Information about the specific type of anxiety is needed by speech and language pathologists (SLPs) so that they can address clients' concerns when these negative aspects are present (Howell, 2010a).

Data regarding stuttering incidence, age of onset and recovery rate differ between studies; although estimates are generally quite consistent (Alm, 2005). The incidence of stuttering was given in pioneering work by Andrew and Harris (1964) who reported that the incidence of stuttering was 4.9%. The researchers followed more than 1000 children from birth until the age of 15 years. Yairi and Ambrose (2005) indicated that approximately 5% of pre-school age children exhibit episodes of stuttering; and that the mean onset age for those children is 33 months. The mean age of onset was also confirmed by Mansson (2000). Around 80% of these children spontaneously recover by puberty (Howell, 2007) and the highest rate of recovery is three years post onset (Yari & Ambrose, 2005). Stuttering affects more boys than girls; a finding that has been consistently reported. Yairi and Ambrose (2005) reported that four times as many boys continue to stutter into adulthood than girls.

1.3 What is Word Finding Difficulty?

Not all disruptions of speech fluency indicate stuttering. During spontaneous speech, the flow of speech might be interrupted while a word is being retrieved. This is commonly referred to as word-finding difficulty (WFD). The clinical pattern of WFD is different than that of disfluent speech. WFD is another form of a communicative disorder that occurs when the speaker has difficulty in retrieving a word during discourse (German, 1991). This retrieval difficulty is probably because of deficiency in vocabulary and likely to be a particular problem for children with EAL. Alternatively, the target word may actually be in the speaker's lexical system but one reason for the slow retrieval is that the target word has a low frequency of usage. According to the definition given, we all may experience WFD at certain times. Children, in particular, can also experience WFD during spontaneous speech or when naming objects due to their small lexicon. However, in some cases it may develop into a chronic word-finding problem which has implications on the child's self-esteem and overall well-being (Best, 2005). Therefore, a clear understanding of the problem and the nature of causes of WFD in important.

Recently, there has been a growing body of research studies investigating the neurobiological processes underpinning word finding in brain damaged populations. The neurological basis of word retrieval has also been widely discussed in language-disorder types such as aphasia. For example, the brain regions that impact naming in participants with aphasia have been studied to identify the core regions associated with word finding (Migliaccio et al., 2016). It was found that the left temporal pole and left posterior temporal cortices impact semantic and lexical naming mechanisms. Brain imaging techniques are not as easily available for young children as for adults to allow the neural basis of WFD in childhood to be studied. Although indirect, perhaps the analysis of WFD errors provide the best evidence of the nature of deficiencies in word retrieval (McGregor, 1997). Another type of speech disruption is the "tip of the tongue (TOT)" phenomenon. TOT states occur on a regular basis in normal speakers; and they are characterized by the ability to only retrieve part of the information about a target word such as the first phoneme, number of syllables or stress location (Gollan &

Silverberg, 2001). TOT is associated with the word-frequency effect, which in turn is partly an effect of age-acquisition (Levelt,1999). Thus, it is likely that young children will exhibit this phenomenon as they are trying to find words; particularly ones with low frequency. This will eventually lead to non-fluent speech with possibly hesitations, interjections and repetition of syllables (Telser, 1971).

In the next section, we provide a brief overview of various models of fluent speech control. Differences in the characteristics of disfluent speech will be highlighted (i.e., hesitancy in speech vs stuttering). This contributes to understanding the importance of identifying the right disfluency symptoms that distinguish stuttering from other disfluencies that can occur in fluent speech.

Models of speech production and fluency failure

The issue of what can lead to producing disfluent speech has been approached from different theoretical views. We begin this section by outlining part of Levelt's (1989) model of lexical access. While it does not specifically address stuttering, it is a prominent psycholinguistic model that influenced many authors who attributed stuttering to linguistic factors. For example, the covert repair hypothesis (CRH), which was proposed by Postma and Kolk (1993) is a language-based theory of stuttering that has been influenced by Levelt's model. Both Levelt's model and the CRH attribute many disfluencies in speech to the speaker's attempt to repair a detected linguistic error before producing an utterance. Work by Howell (2010) showed that while many stuttering theorists attributed stuttering to language factors, there are other models that attribute it to motor factors. Four areas have been interpreted as motor factors involved in stuttering: (1) altered auditory feedback, (2) local and global speech rate changes, (3) articulatory variability, (4) articulatory coordination. At the other end of the spectrum, there are theories that attribute stuttering to the interaction between motor and language factors like EXPLAN (Howell & Au-Yeung, 2002) theory, which will be described below. It is a theory of fluent speech control that applies to fluent speakers and speakers who stutter. One important contribution of this theory lies in how it makes a distinction between speech disfluencies that are more characteristics of stuttering and other disruptions that can be exhibited by all speakers. Implications on how this affects the work we do on screening children for fluency will be shown.

Lexical access in speech production

Lexical retrieval involves a number of cognitive components. To better study lexical retrieval difficulties in children with WFD, it is essential to understand the nature of the underlying impaired mechanisms that might impact the processes involved in retrieval and naming. The current conceptualizations of WFD in children are largely motivated by adult processing models. In adult models of lexical retrieval, there is a consensus that producing a word involves multiple aspects of stored knowledge about words including word meaning, syntactic properties and sound structure. According to Levelt's (1989) model, there are two stages in the lexical access process (1) lexical selection and (2) phonological encoding. The first stage of lexical access involves a semantically driven search for an appropriate word from a set of lexical candidates or lemmas and a selection of a target lemma. The lemmas are unspecified for phonological forms, but their semantic and syntactic properties are specified

(Levelt, 1992). Syntactic specification involves word class, subcategorization information for verbs, and grammatical gender and countability (i.e., mass, count) for nouns. Semantic specification involves conceptual organization in which all perceptual, emotional and usually other additional information about a target word must be fulfilled. For example, to initiate lexical selection for the word "Tiger", the speaker must focus on lexical concepts including that it is an animal, dangerous and has stripes. Activation of other related concepts in the speaker's mental lexicon is also possible (Levelt, 2001). The "lemma" refers to the syntactic properties of the lexical item. For example, the lemma for Tiger specifies that it is a noun and it is countable. As soon as the target lemma is selected, this indicates that lexical selection is complete, and the second phase of lexical access starts (phonological encoding). The phonological encoding stage involves the activation and selection of phonological forms (lexemes) resulting in speech production. To explain that, the lemma information is transferred to the formulator; in which a phonetic plan is being created to prepare for articulation. The formulator gives the lemma a linguistic representation and transforms it into a sequence of sounds that are grammatically organized. Finally, the phonetic plan is converted into muscle movements for the articulators (e.g. tongue, lips, larynx, etc) and the speaker pronounces the sounds of a word. Figure 1.1 gives a schematic representation of the two-step approach to lexical access in Levelt's model of speech production.

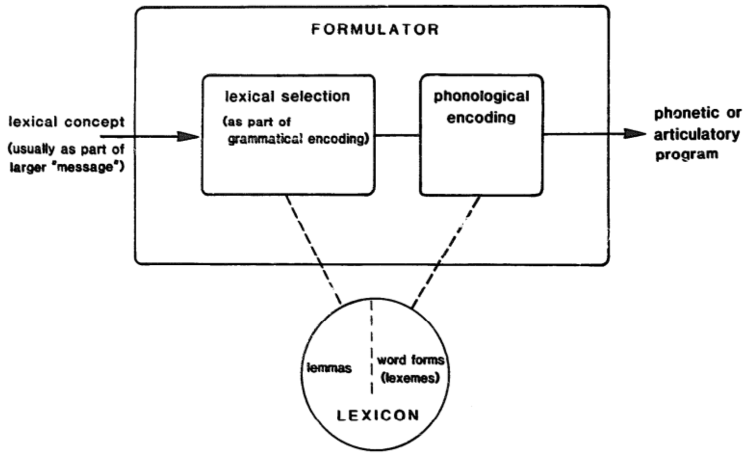

Figure 1.1: A schematic representation of the two steps of lexical access in speech production (adapted from Levelt, 1992).

Levelt's model also incorporates a perceptual monitor system in which the speakers can get feedback about their utterances (Levelt, 1992). Through the perceptual system, the speaker is able to monitor the utterance and correct any errors. Errors can occur at any of the two stages as described above (lemma or lexeme levels). Figure 1.2 below shows the architecture of the Levelt's model including the formulator system that was described above. The bottom part of the model shows the two ways that the speaker can receive feedback about utterance: internal feedback and overt speech. In internal feedback, the speaker can receive feedback about their utterance before it is spoken. In this case, the errors are corrected internally (the listener would not hear the error). There are cases, however, in which the listener would hear the error in the utterance (overt speech). In this case, the speaker uses the external monitoring system to monitor the utterance. The internal loop allows a speech repair to take place before producing speech with an error (Howell, 2010). The resultant utterance will be a speech with some hesitancy, and repetition to some parts that are not necessarily essential. The speaker then carries on and these hesitations and possibly repetition of some of the words should not interrupt the forward flow of the speech.

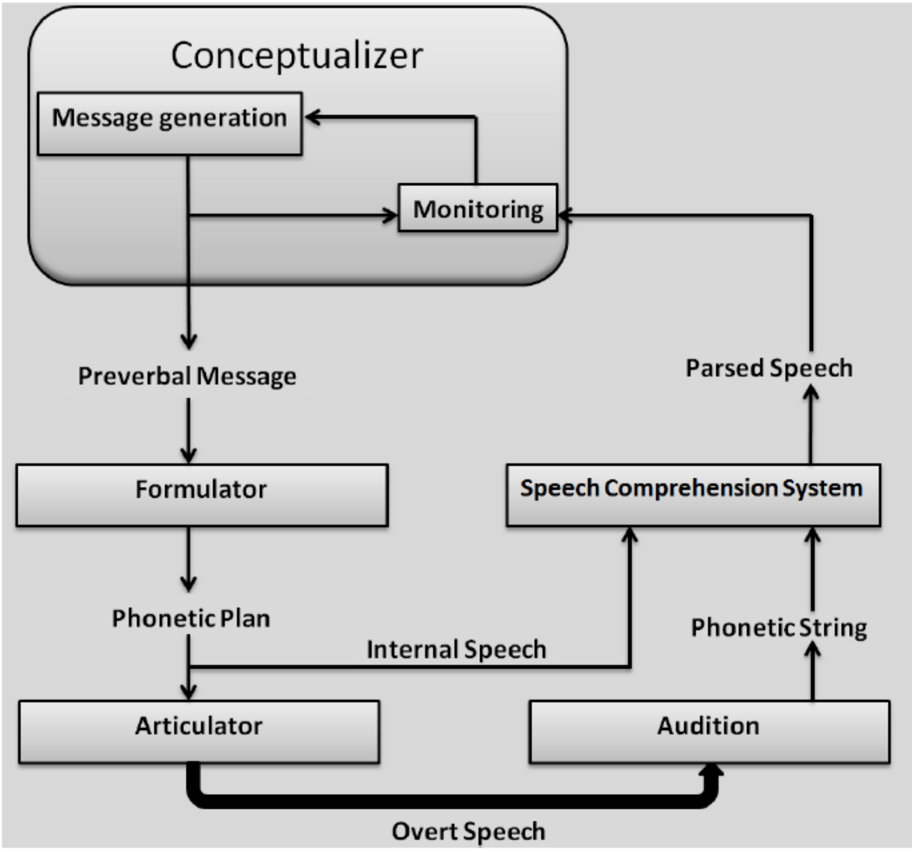

Figure 1.2: Levelt's blueprint showing how production and perception of sounds are involved in speech output control.

This section has summarized the main elements in Levelt's model of speech production. Levelt's model is comprehensive, and we have only selected parts of it to (1) provide a brief introduction on the topic of lexical access; and (2) to demonstrate the main focus of this model on speech errors. And, as a consequence, we must consider not only the error type of interest, but also, the potential contribution of co-occurring features of performance.

In the next section, two possible causes of WFD will be explored, including the role of semantic and phonological representations. The identification of specific causes of WFD may lead to the development of effective intervention protocols.

Potential origins of WFD

Semantic representations

The association between semantic representations and WFD have been investigated through the study of semantic errors, semantic priming, and semantic fluency. In the case of naming errors, research documents the assumption that semantic errors are likely to be the result of incomplete semantic representation, while phonological errors are likely to be the result of incomplete phonological representation (Messer & Dockrell, 2006). Research by Rubin and Liberman (1983) explored the nature of word retrieval errors in children using a picture naming task. They found that semantic substitution is the most frequent error type (e.g., "cat" for "dog"). McGregor (1997) conducted a study to examine the clinical pattern of WFD errors made by preschool children. Twelve pre-schoolers with WFD and a matched control group of twelve typically developing children participated. All children performed three tasks involving word-finding: the noun-naming and verb-naming subtests of the Test of Word Finding Difficulties (TWF; German, 1991) and a story retelling task. While the general error profiles of the two groups were similar, semantic errors were more common than phonological errors on all three tasks. Children with WFD made more errors in connected speech (story retelling) compared to the control group. The types of semantic error were not the same in the three tasks. In the two groups, most semantic errors on the nouns and verbs involved thematic relations between error and the target such as semantic circumlocution (e.g. "something you tell the time with" when the target is "watch"). It was suggested that this is an indicator that disrupted semantic representations are a cause of WFDs. However, because children with WFDs are likely to have a less well-developed language system compared to controls, this may result in an overall higher rate of errors.

Dockrell et al. (2001) examined the nature of naming errors in 31 children with a formal diagnosis of WFD according to German's Test of Word Finding TWF (1991). Three control groups also took part in the study and they were matched on chronological age, naming age and level of receptive grammar. All three groups had 31 children aged between 6;4 and 7;10 year. The groups were assessed on comprehension of words, accuracy of naming and latency to name. Children with WFD performed significantly worse than their age-matched controls (less accurate and longer latency). When compared to their language-matched control group, the performance of children with WFD on the naming task was equivalent, and the error patterns were also similar. Additional analysis showed that children with WFD and the language-control groups performed significantly worse than the age-match control groups. These findings suggest that semantic errors are common across different groups of children; and that the WFD are not necessarily a result of poor semantic representation. When interpreting the nature of speech errors, one must consider not only the type or error, but also the context in which the error occurs in. In a study that examined speech production errors in patients with aphasia, Caramazza and Hillis (1990) showed that aphasics made more semantic than phonological errors. According to Harley and MacAndrew (2001), speech errors arise because the target words are aberrant in terms of properties such as word frequency and length. In semantic errors, there is a semantic relationship between the intended word and the produced one. The form of relationships could be: (1) antonym (e.g. hot for cold); (2) coordinates from the same category (e.g. truck for canoe; or (3) associates (e.g. price for wage). These mixed types of errors due to semantic relatedness might make semantic errors occur during speech production. On the other hand, there is no semantic basis for phonological errors; rather it is due to phonological similarity between the target word and the produced one. This can be in the form of sharing the same initial phoneme or a later segment overlap (Harley & MacAndrew, 2001).

Phonological representation

While a number of studies have pointed out that deficiencies in semantic representation contribute to WFD, other studies showed evidence for impoverished phonological representation as the cause of WFD. For example, Faust et al. (2003) examined the source of WFD in children with dyslexia and a matched control group of typically developing children. A picture naming task was used to evaluate performance. It was shown that children with dyslexia showed fewer correct responses and spontaneous recalls. The differences between the two groups in phonological substitutions were significant, but no differences were found on semantic information. Children with dyslexia did not differ from the control group in the partial semantic information, but they gave invalid partial phonological information. Thus, it seems that children with dyslexia have difficulty in accessing the phonological codes of pictures of known names. These findings suggest that the naming problems of children with dyslexia arise because of their difficulty in retrieving the phonological word forms (i.e. lexeme) after the corresponding abstract lexical representation has been successfully accessed. In that case, the target word may have been selected appropriately but the phonological code of the word fails to be recognized.

Additionally, German (2002) conducted a single-subject, multiple-baseline- design study to evaluate the effectiveness of a phonological treatment protocol in improving WFD. Two 8-year-old, monolingual English-speaking boys participated. They were assessed using the Test of Word Finding, Second Edition (TWF-2; German, 2000). It was found that naming errors were reduced when the participants were provided with metalinguistic reinforcement and phonological cues. The results may suggest that deficiencies in phonological storage and phonological representation are the cause of WFD. Thus, a WFD intervention that is based on phonological training may aid in improving word-finding.

In summary, this section of the review has shown that WFD may originate due to errors at the lemma or lexeme levels during the word retrieval process. Thus, phonological and semantic components could be incorporated into future interventions for WFD.

EXPLAN (2002) theory

While the focus in Levelt's work was on errors, EXPLAN theory views fluent speech control differently. The focus in this theory is on timing control as described next. The argument is that neither language, nor motor factors alone can account for disfluent speech (Howell, 2004). EXPLAN embodies two processes: language processing (PLAN) and motor programming (EX). Both processes can take different amount of time. The planning process supplies a symbolic linguistic representation that the execution system can use. The time the planning process takes depends on linguistic factors such as phonological complexity of the speech segments to be produced. The execution process implements the speech plan based on what has been passed over from the language system. The idea is that an utterance must be planned, and the plan needs to be complete before the speaker attempts the motor process. Importantly, the two processes have to be timed together to overcome fluency failure. In fluent speech, PLAN and EX are two independent processes that operate in parallel as a chaining process (Howell & Au-Yeung 2001). This allows planning to continue without stopping the execution of a previous word. Figure 1.3 below gives a diagrammatic representation of the EXPLAN model showing the temporal relationship between planning and execution. Time is represented along the abscissa (X-axis). The top row represents planning, and the second row represents execution. The figure shows that execution of the word (n) begins after its plan is ready, and that there is enough time to plan the next word (n+1) while (n) is being executed. Thus, fluent speech is produced.

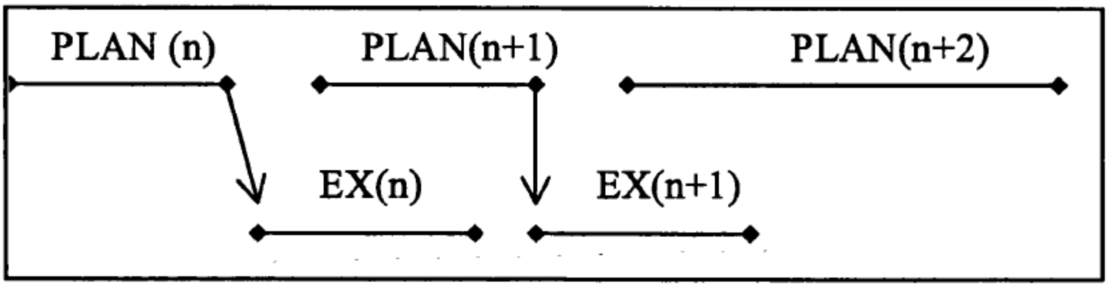

Figure 1.3: A schematic representation of the synchronization between linguistic planning (PLAN) and speech execution (EX) during fluent speech production.

EXPLAN account for two ways fluency failures can arise in which the execution of a word finishes before the plan for the next word (n) has been completed. The speaker then has two options to deal with the incomplete language plan: (1) to repeat the word that has just been executed "stalling"; or (2) to produce only the first part of a word “advancing”. EXPLAN also associates stalling and advancing with different word types (easy vs difficult words). Easy words are often function words; this is a closed set of words that have a functional role (e.g., the conjunction “then”). They are considered to be phonologically simple as they have few consonant clusters, are usually monosyllabic words that tend to occur frequently. On the other hand, difficult words are often content words, which is an open class that include nouns, verbs and adjectives and they are considered phonologically complex.

The stalling class includes whole function word repetition and hesitation which are symptoms that can be shown by fluent speakers (Howell, 2004). In this thesis, we will follow work by Howell (2004) and refer to these symptoms as disfluencies. Stalling on a function word before a content word involves repeating a function word that has already been planned and produced fluently until the complete plan for the following content word is ready. The speaker repeats a function word before a content word to gain extra time to prepare the plan of the content word until it is ready for output (hence the name stalling). Figure 1.4 below shows a schematic of the temporal relationship between planning and execution for speech produced with stalling. For the first word (n), the planning and execution are synchronized appropriately. After n is produced, the plan for the next word (n+1) is not ready. In this case, the speaker will repeat the word (n) again. By the time the word(n) has been executed, the plan for the next word (n+1) is ready and the word can be produced fluently. Stalling can be in the form of repeating a function word for a few times, and it can also be in the form of pausing. This type of disfluency in speech can be particularly frequent when children use EAL as will be shown in the next section.

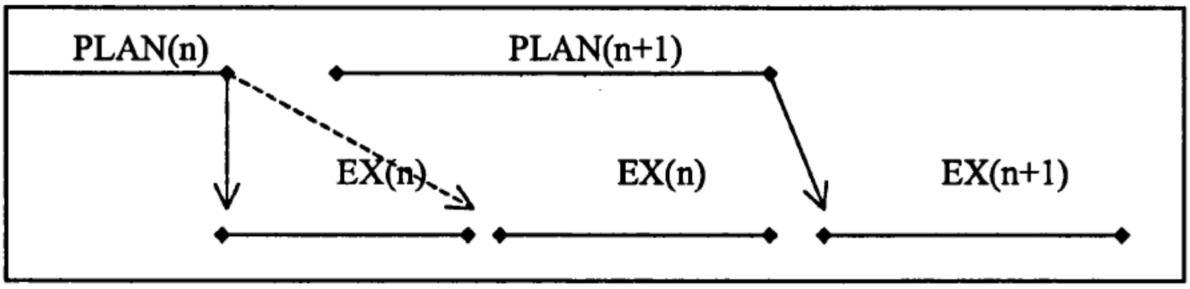

Figure 1.4: A schematic representation of the synchronization between linguistic planning and speech execution during speech produced with stalling (a whole word is being repeated)

The advancing class of speech disfluency, however, includes part-word repetitions, broken words and prolongations, which are characteristics of stuttered speech according to Riley (2009). These stuttering symptoms occur usually on the first part of content words and they show that the speech is fragmented (Howell, 2007). What happens in advancing is that the speaker completes executing a function word, and before the plan for the next word is ready, the speaker commences that word and the plan runs out of time. Thus, because the content word is not fully prepared, the speaker will have to repeat, break or prolong part of that word until the linguistic planning is complete.

Unlike function words, content words may pose particular challenges due to their phonological complexity. Words in this category may contain more than one syllable, have certain later emerging consonants that pose articulatory challenges and/or include consonant clusters. Word onsets are usually the point where most speakers find it difficult to transition from there to the remaining part of the word which could lead to the word being fragmented (Howell & Au-Yeung, 2002).

Another aspect of EXPLAN theory that is also relevant to work in this thesis is the account for developmental changes in the stuttering symptoms and word types. As speakers who stutter get older, they show more advancing symptoms than stalling (Howell, 2008). As implied above, EXPLAN maintains an association between word type (function vs content words) and the type of disfluency. Thus, more problems will occur on content words in the form of fragmentation in speakers whose stuttering problems will persist. In that case, word-initial consonants may be repeated (e.g., "s.s.step"), prolonged (e.g., "ssstep"), or a break within a word may occur (e.g., "s-tep). On the other hand, a lot of stalling on function words would be expected in children who will later recover. Children may stall by repeating a function word before a content word until the linguistic plan for that word is available. This pattern of fluency failure symptom is not problematic, it is like a normal disfluency and as children get older, they have a high chance of recovery (Howell, 2008). A child stalls on a function word in order to avoid having to commence a word in which only part of its plan is completed. Hence, the child is trying to avoid producing a fragmented word which indicates again that this is a case of a WFD.

The implication drawn here is that whole word repetitions (WWR) should be excluded from symptom counts when screening preschool children for fluency (Mirawdeli, 2015). This is particularly important when screening children with EAL. Repeating whole function words alone by such children may be an indication that they have WFD and they should not be classified as children who stutter. In fact, only when a child shows high rates of WWR and any of the speech fragmentation symptoms in the advancing class, a referral for treatment should be made. It is not always practical to obtain speech samples from all children when they first enter their preschool year. The AEN_NWR that will be introduced later in this chapter was designed to fill the need for a rapid tool that can assess speech disfluency.

To summarize this section, unlike Levelt's theory, EXPLAN rejects the notion that errors are responsible for how fluent speech is controlled. What is important, however, to produce fluent spontaneous speech is an appropriate synchrony between the language and motor factors. Also, EXPLAN shows that adult speakers who continue to stutter stop stalling on function words and start showing advancing on content words. On the other hand, children who will recover show a lot of WWR on function words.

1.4 WFD characterization and EAL

An opinion dating back to at least 1937 was that speaking two languages is a risk factor for stuttering. The claim was that having to select a word from two competing languages could slow down speech production and increase chances of stuttering (De Houwer, 2011). This was later criticized in many studies and empirical evidence was provided. For example, Howell et al. (2009) examined the effect of bilingualism on stuttering during late childhood (i.e., 8 to 12 years). A total of 317 Children were examined; they were classified into three groups; (1) CWS who speaks of an alternative language exclusively; (2) bilingual CWS; or (3) bilingual fluent control. The association between stuttering and bilingualism was significant by $\chi 2$ test. This calls into question the idea that bilingualism causes stuttering. One potential explanation of the stuttering classification for bilingual children is that they show high rates of WWR in their speech. And when including WWR to symptoms of stuttering, many children would be classified as CWS when they are fluent. To explain, there tends to be an overlap between symptoms shown by fluent and disfluent speaker, and this unsurprisingly has led to a long debate on what is the right set of symptoms to diagnose stuttering (Howell, 2010). Another reported explanation is the TOT phenomenon, which occurs in children as they trying to retrieve words, and this process disrupts the flow of their speech (Telser, 1971).

One of the earliest attempts at characterizing stuttering was that of Johnson and Associates (1959). They proposed that the following eight symptoms are commonly observed in the stuttered speech: 1) incomplete phrases; 2) Revisions; 3) Interjections; 4) Phrase repetitions; 5) Whole-word repetitions (WWR); 6) Part-word repetitions; 7) Prolongations; and 8) Broken words. There seems to be a general agreement that the last three symptoms are characteristics of stuttering, including in Riley's (2009) stuttering severity instrument (SSI), which is used to assess severity of stuttering. There has been, however, mixed opinions on whether WWR is a typical symptom to characterize stuttered speech. Identifying the correct set of symptoms is important to minimize misdiagnoses of fluent children as stuttering when they do not stutter. As pointed out above in Howell's work on EXPLAN, WWR and hesitancy are not indicators of stuttering but rather an indication of WFD. They constitute disfluency symptoms that fall in the stalling class, which resembles normal disfluencies. Additionally, empirical evidence has been supplied from neuroimaging work that evaluated the neural basis to symptoms classification by Jiang et al. (2012). The results showed that different brain activity patterns were associated with symptoms that are more typical of stuttering (e.g., prolongation and broken words) and the ones that are less typical of stuttering (e.g. revisions and interjections). The results also showed that WWR were more similar to fluent speech than to non-fluent speech. Although WFD occur in the speech of monolingual children as they develop their language, they are more frequent in the speech of young children who use EAL (Howell et al., 2016). EAL children might have a smaller vocabulary to select from when speaking spontaneously. Thus, they repeat whole words or have silent delays to delay the time at which the subsequent word is being retrieved. The difficulty in retrieving a target word may be because of its low frequency (i.e., the number of times a word occurs in the language) which slows down the speed of access and production.

To conclude this section, it is important to distinguish between two patterns of fluency failures that can disrupt children's speech (stalling and advancing). According to EXPLAN, there are differences between disfluencies that are shown by fluent speakers (i.e., WWR on function words) and speech fragmentations that are usually shown by speakers who stutter. EXPLAN attributed the former to the stalling class, and the latter to the advancing class. WWR indicates WFD, which is a vocabulary problem that can be improved through phonological and/or semantic training.

A phonological training procedure for WFD has been designed and is discussed (see Chapter 5 for design requirement details). The choice of phonological treatment to treat WFD was motivated on several bases. First, stuttering also needs to be identified at an early age so that an effective and appropriate intervention by SLT can take place before the problem becomes acute. In the next section, we give details on available procedures to separate children with WFD from those who stutter.

Procedures for Identifying children with speech disfluency.

It is widely agreed that identification of language difficulties in the early school years is vital so that the appropriate intervention can be conducted (Snow, 2014; Bercow, 2008). Delaying intervention until after the disorder has begun may result in multiple risks that lead to lower educational achievement and behavioural and social problems in the child's later life (Bercow, 2008). A school-based screening procedure has been developed recently to identify children with Speech Language Communication Needs (SLCN) in reception class children and to separate those children with SLCN who have speech difficulty from those who have WFD (Howell et al., 2017). The procedure employs a spontaneous speech sample obtained from each child to detect symptoms of speech disfluency and WFD. Whilst WFD is indicated by WWR and pauses on words prior to words failing to be retrieved, stuttering is characterized by the last three symptoms from Johnson's list which are also the symptoms used by Riley (2009) (i.e. prolongation, part word repetition and broken words). In addition to the three stuttering symptoms, children who stutter may also exhibit WWR, but they have to exhibit both (WWR and stuttering symptoms) to be diagnosed as stuttering (Howell et al., 2020). The analysis of disfluent symptoms for screening is based on Riley's (2009) Stuttering Severity Instrument (SSI). SSI uses three components to measure fluency: the percentage of stuttered spoken syllables (%SS), the average duration of the three longest stutters and a measure of physical concomitants (PC). Examples of PC include tics, distracting sounds and facial movements. When used in screening, the percentage occurrence of the symptoms out of all symptoms (%SS) is calculated for every child and if the score is above a threshold %SS then he/she is classified as having fluency difficulty. Due to time-constraints when testing children in schools, a speech-based procedure might not be practical, and an alternate form of testing is needed. Howell et al. (2017) developed a novel NWR test, the 'Universal' Non-word Repetition task (UNWR) to assess NWR ability in children with diverse language backgrounds. The UNWR is an efficient procedure that applies to 20 different languages[1] and can differentiate between PWS and people who do not have a speech deficit. To validate the newly designed UNWR, its performance was evaluated against the SSI since the test has already been standardized for English. A significant correlation was found between participants' scores on SSI and UNWR, which supports the use of UNWR as a sensitive measure of speech fluency.

Arabic was not one of the twenty languages that the UNWR applies to. To bridge this gap, an Arabic – English NWR (AEN-NWR) that can equitably test English and Arabic speakers was developed as part of this thesis. In Chapter 2, I explain the design requirements of an Arabic – English NWR (AEN-NWR) that can equitably test English and Arabic speakers, including those who speak both languages to different extents.

Similar to how the UNWR was validated, our intention was to develop and standardize an Arabic version of SSI to validate AEN_NWR. The rationale behind using SSI is that it is the standard way of assessing fluency for speakers of English. SSI was employed as a basis for assessing stuttering in English and Arabic speakers; this is because its English form has been assessed for reliability. Additionally, there is a clear and easy-to-follow manual. SSI has many flexible features that allow assessments to be made in different situations (Howell, Soukup-Ascencao, Davis & Rusbridge, 2011). As the central concern of this thesis is assessment of preschool children's speech, SSI is important as it

[1] The 20 languages are English, Polish, Romanian, European Portuguese,Bulgarian, Serbo-Croat-Bosnian,Czech,Dutch, French,German, Hungarian, Slovene, Swedish, Danish, Norwegian, Russian, Latvian, Ukrainian, Urdu-Hindi and Bengali

has a format for assessing pre-schoolers who cannot read. It provides an important benchmark for assessment of stuttering and is considered one of the best attempts at setting objective criteria for the clinical assessment of stuttering (Conture, Caruso, Rustin, Purser & Rowley, 1987).

The data used to evaluate SSI statistically were collected from English speakers who stuttered, and it used English passages that had to be read as well as others that were spoken spontaneously. It has been translated for use in other language including Persian (Bakhtiar, Seifpanahi, Hossein, Ghanadzade & Packman, 2010), Jordanian Arabic (Alqhazo & Al-Dennawi, 2018) and Egyptian Arabic (Hassan, Sallam, Khodeir & Mahmoud, 2017). Translated forms need re-evaluating statistically as the English norms would not apply. Thus, the norms do not apply to the Arabic version as the standardization has not been conducted. Moreover, the test is not appropriate for assessment of Arabic in terms of the procedures used in SSI for counting the number of production units (i.e., syllables for English) as well as the specifications of stuttered events.

Translating and adapting a standardized test is more efficient than creating a new one because the resulting test carries some of the scientific integrity and the psychometric history of the test in the original language. Also, it is desirable to have versions of the same test so that the needs of participants who speak different languages can be served (Kristjansson, Desrochers & Zumbo, 2003). Translation is an easy option for languages where there is no objective diagnostic tool for assessing persons with stuttering. However, the translated forms need scrutinizing, and possibly adapting, to ensure that they are appropriate for assessing the language. Karimi et al. (2011) explained the steps taken to translate

SSI-3 to Persian based on the World Health Organization (WHO)'s guidelines. The procedure involves forward translation by a translator who should be knowledgeable of the English-speaking culture and who speaks the language of the target culture as a mother tongue. The next step is back translating the test by a translator who speaks English as a mother tongue. The instrument then needs pretesting before a final version can be released. In Chapter 3, I propose a speech-based scheme that takes into account the phonological and morphological features of Arabic. An empirical study is reported in which the instrument was administered to Arabic speaking adults and children.

Chapter 2 Identification of Fluency and Word-finding Difficulty using a Phonological Performance Measure

"The delicate organs of speech, must be considered as the rudder of a reason, and speech as the heavenly spark, that gradually kindles our thoughts and senses to a flam" *Jonathan Rée", I see a voice*

2.1 Introduction

This chapter presents a thorough description of the design of an Arabic and English non-word repetition task (AEN_NWR) that can be used to equitably assess Arabic and English children, including those who speak both languages to different extents. It is intended for use in identifying stuttering and fluency difficulty. The chapter begins by a brief review of the literature on NWR as a potential behavioural clinical marker for differentiating between WFD and speech disfluency. This literature has focused on nonword repetition and aspects of phonological performance in children who stutter as well as children with other language disorders such as Specific language impairment (SLI) and dyslexia. NWR can potentially inform work in these areas pertaining to speech disfluency. Following this review, we turn to a description of the Arabic language phonology specifying areas that aided in the generation of Arabic and English non-words.

2.2 Non-word repetition as an assessment tool

The ability to repeat a novel phonological sequence is a basic language skill that humans possess. Infants under the age of 12 months attend to speech sounds, especially when the sounds are spoken to them by adults (Zangl & Mills, 2007). Infants can spontaneously imitate the words of others and by the time they are 2 years old they repeat non-word when requested (Gathercole, 2006). Children also repeat "non-words" spontaneously when they mimic real words spoken to them by adults (Coady & Evans, 2008). Although non-word repetition (NWR) tasks appear simple, they rely on the following cognitive processes. First, the person must process the acoustic signal, extract phonemes and match the signal with phonological representations in memory. Then, the person must plan the articulatory movements for achieving production of the non-word and execute this plan as their response (Munson, Kurtz & Windsor, 2005). Correlations between phonological working memory and non-word repetition was first examined by Gathercole and Baddeley (1989). Performance on non-word tasks has been examined in children with several language disorders such as dyslexia (Snowling,1981) and specific language impairment (Archibald & Gathercole, 2007). Performance on NWR tasks presents a challenge for children with these language disorders. For instance, Snowling's (1981) study that examined NWR non-word abilities in children with dyslexia tested two groups of children: 22 typical readers and 20 dyslexic children whose age ranged between 7 and 17 years old. The children repeated 30 non-words that were two, three or four syllables long. non-words posed more difficulty for dyslexic children than for the control group with dyslexic children making more repetition errors. Significant differences between the groups were found when four syllable non-words were repeated. The finding was interpreted as indicating a phonological deficit in dyslexic children. Subsequent studies (e.g., Broom & Doctor, 1995) agree that dyslexia should be considered a phonological deficit indicating language weakness rather than impaired low-level auditory difficulties (Snowling, 2001).

Before turning to studies on NWR and stuttering, details are given about phonological memory to illustrate its role in repetition of non-words or unfamiliar words. The phonological loop is part of working memory (WM), which is a cognitive system that temporarily holds and manipulates information whilst people perform tasks such as comprehension and learning (Baddeley, 1986). The WM-model of Baddeley and Hitch (1974) proposed three major components: (1) the central executive system, which is the supervisory controlling system that is aided by the other two components; (2) The visuospatial sketchpad, which is concerned with the visuo-spatial memory; that is, it stores and processes information in a visual or spatial form; and (3) the articulatory loop (now referred to as the phonological loop) that is responsible for rehearsing and storing speech-based verbal information. It transforms the verbal stimuli into phonological codes which have the associated acoustic and temporal properties of the stimuli. Matches between the phonological codes and codes that exist in the long-term memory system (i.e., phonemes and words) are sought. The phonological loop can be further divided into two sub-components: the phonological short-term store and the subvocal rehearsal component (Baddeley, 1996). The phonological short-term store is like an inner ear which holds speech-based information for up to about two seconds. Speech-based information can be maintained by the subvocal articulatory rehearsal component; a process that can be used to enter information into the phonological store. The subvocal rehearsal component is like an inner voice which allows information to be rehearsed, for example, object names that are articulated either overtly or sub-vocally (Baddeley, 1996). Rehearsal allows a person to remember a telephone number by circulating the phone digits to oneself. Figure 2.1 provides a graphical representation of the phonological loop model based on Baddeley (1996) adapted from Gathercole (2008).

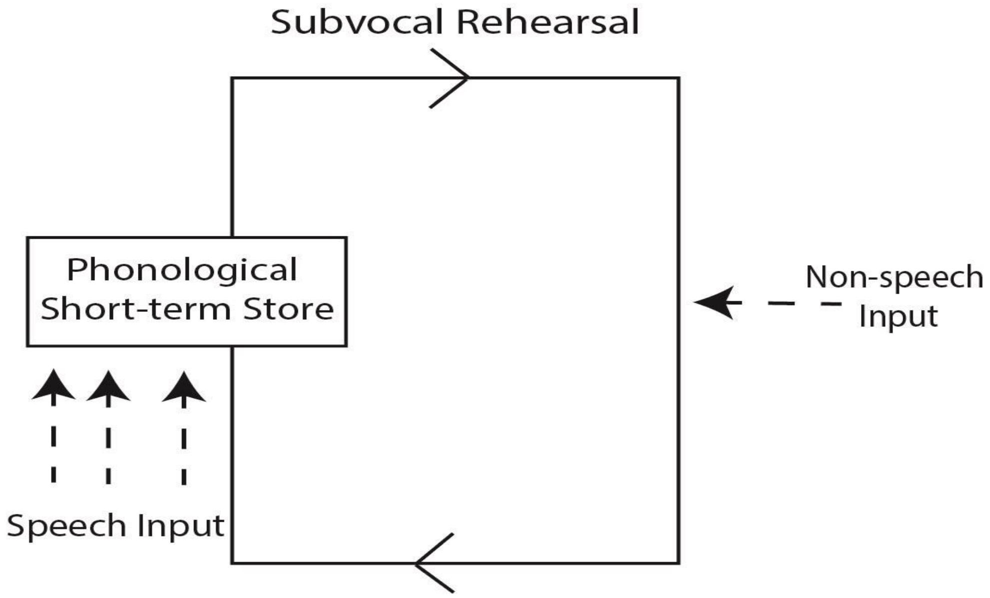

Figure 2.1: Graphical representation of the phonological loop model based on Baddeley (1996) adapted from Gathercole (2008), See text for further detail

We now provide a brief overview of the way the phonological loop operates when repeating non-words. As implied earlier, it has been established that there is a strong relationship between NWR performance and the phonological loop component of WM (Gathercole et al., 1994). Repeating non-words requires temporary storage of unfamiliar phonological sequence in the phonological loop; and it is assumed that success when holding the sequence depends on the short-term memory capacity of the phonological loop. The rehearsal component of the phonological loop serially reactivates the unfamiliar phonological sequence stored in the phonological store, where this process does not necessarily involve movements of speech articulators. As long as rehearsal is maintained, the phonological store can hold on to the speech information. Indeed, the process of rehearsal is time-limited; the longer a phonological sequence is, the longer it takes to reactivate the sequence leading to fewer rehearsals in a given time (Gathercole, 2008).

Several studies have reported how NWR performance is affected in CWS. Hakim and Ratner (2004) investigated the performance of eight CWS and eight children who do not stutter (CWNS) on non-word repetition where the children's ages ranged between 4 and 8 years. Children attempted to repeat 40 non-words from Gathercole et al.'s (1994) Children's Test of Non-word Repetition (CNRep). The task consisted of 40 non-words of 2-, 3-, 4-, and 5-syllables in length. Results showed that CWS were less accurate at repeating non-words at all syllable-lengths, although statistical differences between participant groups only occurred for the 3-syllable non-words but not the longer 4- and 5-syllable words. Anderson et al. (2006) replicated and extended the findings from Hakim and Ratners's (2004)'s work in a sample of younger children (aged 3 to 5 years). The authors argued that examining non-word repetition performance at this young age could provide an opportunity to assess phonological memory during a time of critical language development, relative to school aged children. Anderson et al. (2006) administered the CNRep to 12 CWS and a matched control group of 12 CWNS. CWS were significantly less accurate in repeating non-words of two and three syllables. However, no significant differences were found between the groups in their accuracy of repeating the longer non-words (four and five syllable non-words). The findings of this study is partially consistent with results from Hakim and Ratner (2003). The authors suggested that the significant differences between the groups even at the two syllable non-words were probably because this is a younger group of children; thus, their performance was not impacted by ceiling effects. On the other hand, he lack of significance effects for the longer non-words was due to the impact of floor effects in both groups. Subsequently, Anderson and Wagovich (2010) examined the relationships between measures of linguistic processing speed and between two aspects of cognition: phonological working memory and attention. Nine CWS and a matched control group of 14 CWNS, aged 3 to 5 years, participated. Gathercole's (1994) CNRep test was again used in this study. Children were asked to repeat 40 non-words, where there were 10 each of 2-, 3-, 4-, and 5-syllable non-words. There were significant differences between the two groups in their accuracy of repeating non-words of two and three syllables. However, the differences between the two groups were not significant when repeating non-words of four and five syllables, although CWS performed worse (i.e., were less accurate in their repetition) with them than were controls. These findings are consistent with Anderson et al. (2006) study and the lack of significant differences on the longer nonwords could be attributed to floor effects in both groups.

Sasisekaran and Byrd (2013) investigated NWR accuracy in 14 CWS and a matched control group of CWNS aged between 8 and 15 years of age. Participants repeated a set of 36 non-words consisting of 12 non-words at each syllable lengths (2-, 3-, 4-, and 7-syllables). CWS were less accurate when producing two-syllable non-words compared with the CWNS. However, differences between CWS and CWNS on accuracy at each syllable length was not reported. The non-words at four-syllables posed most difficulty for children in both groups (i.e. had the lowest percent of correct repetitions). Based on this result, the authors suggested that their findings are consistent with previous studies; confirming that CWS show a trend to perform poorly on NWR tasks. Whilst the studies above focused on English-speaking children, Sugathan and Maruthy (2020) explored NWR performance in Kannada-speaking school-aged children. Seventeen CWS and a matched group of CWNS were tested. The non-words consisted of 2-, 3- and 4-syllables, and for each syllable length there were 12 non-words. These were language specific in that they conformed with the phonotactic constraints of the Kannada language. CWS were less accurate in producing the non-words compared to the CWNS at all syllable lengths. Significant differences between the two groups were reported for the mean number of correct non-words; however, whether the differences at each syllable length were significant was not reported. In summary, NWR has a strong potential as a screening instrument that can separate CWS from CWNS based on their accuracy in repeating nonwords.

Howell et al. (2017) investigated NWR ability in a group of ninety-six 4-5-year old monolingual English children and children with EAL, who came from diverse language backgrounds, using the UNWR. The goal was to evaluate the effectiveness of the UNWR in distinguishing between CWS and children with WFD, irrespective of which language they speak. Children with

EAL often have to produce phonological sequences in English; a language they are not familiar with. This is similar to what happens in NWR tasks where children are required to repeat a novel sequence of phonemes that does not exist as a word in their first language; nor in any additional language they speak. Howell et al.'s (2017) study faced the methodological problem, common to many NWR tasks, that materials tend to be biased towards the language for which the test was developed. Howell et al.'s (2017) 'universal' NWR task was developed to apply to various languages spoken in UK schools hence the name 'universal' and avoids confounds between language ability in the test language and presence of stuttering symptoms. If a person has WFD but no stuttering symptoms, this should not be evident when repeating non-words, whereas CWS are expected to struggle performing the task (Howell et al., 2017). UNWR scores were predicted by %SS, but not by WFD (as measured by the percentage of whole word repetition (%WWR)). This relationship between UNWR and %SS provided empirical evidence that UNWR provides a sensitive measure of stuttering. The authors also attributed this relationship to the fact that both measures (UNWR and %SS) reflect phonological planning, whilst WFD is more of a vocabulary problem rather than an articulation one. The results also showed that monolingual English children and children with EAL did not differ in their performance on UNWR. Thus, accuracy in repeating non-words on the UNWR was not affected across language groups who showed different levels of %WWR. This again highlights that the test eliminated the problem in other NWR tasks that favour the language for which the test was designed. The UNWR has a strong potential as a screening instrument for language-diverse samples that can separate CWS from CWNS based on their accuracy in repeating non-words. The task applies to various languages spoken in UK schools hence the name 'universal'. Thus, this test eliminates the problem in other NWR tasks that favour the language for which the test was designed. Arabic language, as will be shown below, has different phonological and morphological structures to English and the other Indo-European languages that UNWR applies to. Consequently, a language-specific NWR was needed that tests Arabic- and English-speaking children in an unbiased way. This was achieved by generating non-words that meet phonotactic constraints that apply to both languages. The steps involved in generating "easy" AEN_NWR stimuli that incorporate phonotactic constraints shared by English and Arabic to equitably test speakers of English and Arabic are described in detail in the next section.

2.3 The Arabic language

Arabic is the most widely spoken Semitic language in the world and is spoken by more than 250 million people as their first language in the Middle East. Arabic is considered a Diglossic language (Ferguson, 1959), There is one standard form, modern standard Arabic (MSA) and a large number of regional colloquial Arabic (CA) forms. MSA is used in most Arab countries in different situations and places including on street signs, in newspapers, Television news, university, schools and books. When spontaneous speech samples are collected from participants, speakers use their own CA dialect. However, when a read passage is used, participants would use MSA because CA forms are not usually used in written texts (although CA is used in, for example, text messaging). Figure 2.2 summarizes the situations where MSA and CA are typically used.

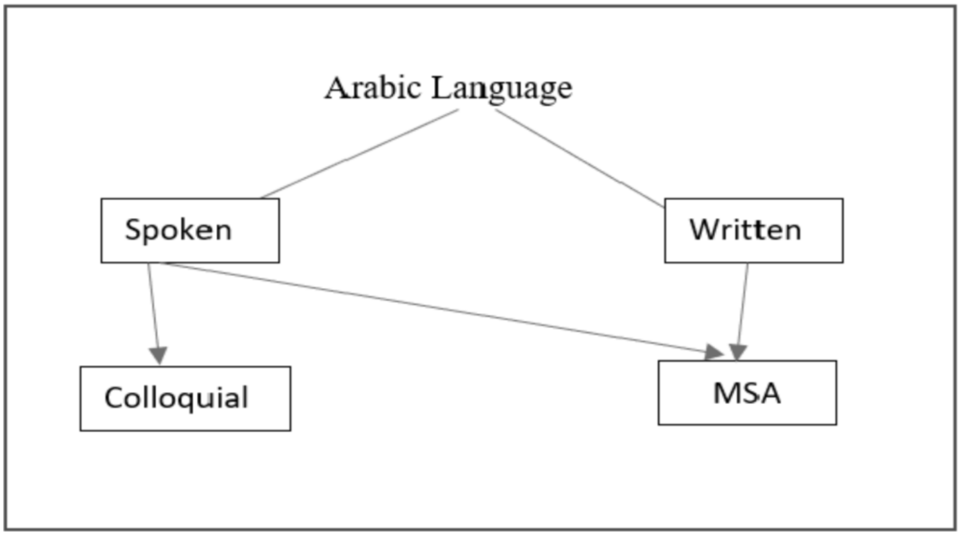

Figure 2.2: Usage of MSA and CA

Arabic has a rich and complex morphology (Kardi & Nie, 2006). It exhibits a discontinuous morphology that is based on the combination of two abstract phonemes: the root and the word pattern (Boudelaa & Marslen-Wilson, 2005). The root is exclusively made up of a sequence of consonants (usually three) that carry the core semantic information. The word pattern specifies the phonological structure and the morphosyntactic properties of the vowels, prefixes and suffixes that are then attached to the root to derive lexical meanings. In Arabic and in other Semitic languages, roots and word patterns are intertwined to form words across different lexical categories (e.g., nouns, verbs, and adjectives). KTB is an example of a root that is commonly reported in the literature, and it implicates the meaning of writing (Tucker, 2011). Table 2.1 shows 10 examples to distinct word forms derived from the root. Those examples are from different lexical categories and are supplied with different vowels and prefixes.

Table 2.1: Words derived from the /ktb/ root

Word	**Meaning**
Kataba	He wrote
Katabat	She wrote
Kitaab	Book
Maktaba	Library; bookstore
Kaatib	Writer
Kutayyib	Booklet
Maktuub	Written

2.3.1 The Arabic syllable

Arabic and English differ in several respects including syllable structure. Syllables in Arabic also vary in features like stress, length and intonation which depend on the CA.

The following are the syllable types that occur in MSA and CA.

Syllable template   Example

1) CV   /walad/ 'boy'

2) CVC   /fam/ 'mouth '

3) CVV   /laa/ 'No'

4) CVVC   /na:r/ 'fire /fatra:t / periods

5) CVCC   /bint/ 'girl'/jalast/ 'I sat down'

Most of these templates occur in monosyllabic words. They also occur in multisyllabic words, but no word can include more than one heavy syllable (Abdoh, 2010). Syllables in Arabic can be light or heavy. Light syllables have the pattern CV; a consonant and a short vowel. Heavy syllables in most Arabic dialects can either have a consonant followed by a long vowel CVV, or a consonant followed by a short vowel and another consonant in the coda position CVC. A third syllable type, the super heavy syllables, can have two consonants word finally CVCC or a long vowel followed by a consonant CVVC. The occurrence of this type is restricted to word-final positions in many Arabic dialects (Watson, 2002). The syllable always starts with a single consonant as its onset. To emphasize, no word can start with a vowel in Arabic. The maximum number of consonants in the onset position is one, in MSA and most CA dialects. Word-final positions can take up to two consonants. A sequence of more than two consonants does not occur anywhere and a vowel is obligatory within a syllable.

In discussing the Arabic syllable, it is worth introducing the sonority sequencing phonotactic principle (SSP), which outlines the syllable structure in terms to sonority Sievers (1881). To explain more, it is a cross-linguistic generalisation that the most well-formed syllables are characterised by a sonority rise throughout the onset to the nucleus, and a fall from the nucleus throughout the coda (Daland et al., 2011). For example, in the English word "trend", the first consonant *t* in the syllable onset is a stop, the lowest on the sonority scale; next is r, a liquid that is more sonorous, then we have the vowel *e* – the sonority peak; finally, in the syllable coda, is *n*, a nasal, and last is another stop, *d*. In the next section, we describe how the SSP was adhered to when creating the AEN_NWR items.

2.4 Considerations for designing Arabic English NWR stimuli

This section explains the steps taken to generate the AEN_NWR stimuli

Step 1: Overlapping phonotactic constraints across English and Arabic were identified to create the non-word candidates. This entailed the following considerations. For each syllable template, all possible phone sequences were created according to the following constraints.

1. **Syllable patterns**: Two syllable templates were selected as they are permitted in both languages.
    1. CV - A short open syllable (Consonant – vowel)
    2. CVC - A medium closed open syllable (Consonant – vowel – consonant)

A final dull syllable was also employed word finally only to generate word-final C.C clusters. Since the second segment of the dull syllable occupies an onset and a syllable onset must be supported by a nucleus, that onset then must be followed by an empty nucleus, hence the name dull syllable (Harris & Gussmann, 2002). We now take a step back and provide further details on the concept of "dull" syllable. Traditionally, the conventional 'Western perspective' of syllabification views any consonant following the last vowel of a word as a syllable coda. Harris and Gussmann (2002) questioned this assumption drawing on a substantial body of evidence relating to language typology, words stress, vowel length and cluster phonotactics. The authors compared the Western view with the Eastern prospect that maintains that a word-final consonant occupies the onset of a final 'dull' syllable, one with an empty nucleus. The evidence provides further support to the conclusion that a final consonant is not a coda; but rather it forms the onset of a syllable containing a silent with no audible vowel (Harris & Gussmann, 2002).

In Arabic, the syllable is always initiated with a single consonant, which requires an obligatory onset. To emphasize, no word can start with a vowel in Arabic. The maximum number of consonants allowed in the onset position is one, in MSA and in most CA dialects. In English, the 'C' in the onset is optional (an example where the 'C' is absent in the word 'eye'). However, there is a strong preference in English for a syllable to begin with a consonant as zero onset syllables (Ø) are rare (Harris, 1994). Thus, all syllables generated are well-formed with one C in the onset.

The selected templates were strung together to form polysyllabic non-words that systematically increase in phonological difficulty. Testing progressed through successive syllable lengths since testing children with 2-syllable non-words is expected to be less challenging than when testing with 3- or 4-syllable non-words. For example, a 4-syllable non-word (e.g., CV.CV.CVC.CVC) has sequences that are phonologically more complex than non-words with 2- or 3-syllables. Also, starting tests using non-words with a lower number of syllables primarily eases children into the experiment and possibly allows them to progress to longer and more complex syllables that they are capable of producing.

Word-initial clusters were not allowed (they are permitted in English but not Arabic dialects). For example, Hijazi Arabic, which is spoken in the west region of Saudi Arabia precludes word-onset clusters. The non-words were generated to be appropriate for a range of dialects of Arabic, so they are suitable for use with participants from diverse geographical and socioeconomic backgrounds. Therefore, care was taken that the phonotactic constraints were not specific to one Arabic dialect because speakers of any dialect are potential participants. Sequences of more than two consonants do not occur in any syllable position in Arabic and a vowel is obligatory within a syllable except for a final dull syllable. Because of the conflicting points of view about the existence of a coda in Arabic, and about the phonotactic restrictions on the consonant clusters that appear word finally (whether they constitute a coda, or they are just adjacent consonants that appear word finally), only the phonotactic constraints that the cluster should adhere to are discussed.

All possible phone sequences were created for each syllable template, with the following additional constraints:

1. **Consonants and vowels selection**: To be included in the AEN-NWR, a consonant or vowel had to occur as a phoneme or an allophone in both languages. The specific consonants and vowels that were selected were as follows:

Consonants. Consonant phonemes that exist in English and Arabic are the following: /f/ , /b/, /d/, /m/, /n/, /s/, /z/, /k/, /g/, /ʃ/, /θ/ , /t/ /r/ and /l/ . The remaining consonant phonemes that exist in English or Arabic were not selected when the consonant phoneme does not have a corresponding phoneme in the other language. There are no corresponding phoneme for /tˤ/, /dˤ/, /ðˤ/, /sˤ/, /x/ , /ɣ/ , /ħ/ in English. Also, there are no corresponding consonant phoneme for /p/, /v/ and /tʃ /or /j/, in

Arabic. The consonant phoneme /t/ was selected although it is realized differently in the two languages. In Arabic, /t/ is dental involving simultaneous contact with the upper front teeth and the tongue tip and stopping the oral passage of air. In English, however, it is alveolar alone except when it occurs before dental fricatives (e.g., in words like ‘eighth’ or between words like ‘at that’) when it is dental. These are sub-phonemic differences, and the phoneme category is used in the NWR test, as it is one of the phonemes that has an early age of acquisition as well as a high frequency of occurrence in both languages as will be shown next.

A) Frequency of occurrence. Table 2.2 shows that the selected phonemes have high frequency of usage and frequency of usage is comparable in the target languages. The occurrence of phoneme frequency is well documented in English (e.g. Mines, Hanson & Shoup, 1978). However, it is still unknown for MSA and Arabic dialects, despite several attempts in spoken Arabic (e.g. Alqattan, 2015). In Table 2.2 below, we gave phoneme frequencies for both languages.

Table 2.2: Phoneme's frequency of occurrence in conversational Arabic and English

Phoneme	Token Frequency in Arabic Alqattan (2015)	Frequency of Phonemes Occurrence in conversational English.
/f/	3.01	1.55
/b/	6.97	1.90
/d/	3.93	3.33
/m/	5.81	2.99
/n/	7.89	6.72
/s/	4.26	4.61
/ʃ/	2.36	.56
/z/	1.07	2.75
/k/ and /kʰ/ [2]	3.84	3.10
/g/	2.46	1.18
/θ/	0.32	.70
/ ð/	0.6	3.14
/t/ and /tʰ/ [3]	6.29	5.78
/r/	6.18	3.87

Consonant Acquisition. The age of acquisition of Arabic phonemes is generally similar to the corresponding ones in English, for common consonants although some allophonic variations between the two have been noted (Amayreh & Dyson, 1998). The small differences between the age of acquisition in Arabic and English is probably due to subtle differences in the mastery criteria in the studies. An Arabic-speaking child acquires /b/, /d/, /k/, /f/, /m/, /n/, /l/, /w/ in early childhood (<2:0 to 3:10) whilst /s/, /h/ and /ʃ/ are acquired later in age (4:0 to 6:4) (Amayreh & Dyson, 1998). The acquisition criteria for English in the present study were adapted from Sander (1972). According to Sander, children acquire /m/, /n/ and /w/ before the

2 The English voiceless stop /k/ is aspirated word initially

3 The English voiceless stop /t/ is aspirated word initially

age of 3. The consonant phonemes /s/ and /ʃ/ are acquired later at around the age of 4.5. Table 2.3 shows the approximate age of acquisition for each of the selected consonant phonemes in both Arabic and English as reported in the studies that were mentioned above. It is important to note that these are average age estimates and the upper age limit in Sander (1972) stops at an age level at which 90% of children are customarily producing the consonant phoneme. Information on this table was considered when designing non-words where consonants that are acquired earlier constituted most of the stimuli created making it possible for young children to produce them.

Table 2.3: Age of Acquisition of the selected consonants in Arabic and English

Consonant Phoneme	**Approximate age of consonant acquisition in Arabic** (**Amayreh & Dyson**, **1998**)	**Approximate age of consonant acquisition in English Sander** (**1972**).
/f/	(2;0 to 3;10)	(2;5 to 4;0)
/b/	(2;0 to 3;10)	(2;0 to 4;0)
/d/	(2;0 to 3;10)	(2;0 to 4;0)
/m/	(2;0 to 3;10)	(2;0 to 3;0)
/n/	(2;0 to 3;10)	(2;0 to 3;0)
/s/	(4;0 to 6;4)	(3;0 to 8;0)
/z/	(After 6;4)	(3;5 to 8;0)
/k/	(2;0 to 3;10)	(2;0 to 4;0)
/θ/	(After 6;4)	(4;5 to 7;0)
/ð/	(After 6;4)	(5;0 to 8;0)
/t/	(2;0 to 3;10)	(2;0 to 6;0)

B) **Vowels**. Three short vowels were selected [i,a,u] because they exist in both languages where each has an equivalent or a near equivalent that can be mapped across languages as shown in Table 2.4 below. Thus, a speaker of any of the two languages is not expected to have difficulty perceiving and then producing those vowels.

Table 2.4: Short vowels mapping from Arabic to an equivalent or near equivalent in English

Arabic short vowels	**English short vowels**
/ɪ/ front high unrounded short	/ɪ/ Front high unrounded lax
/ʊ/ back high unrounded short	/ʊ/ back high unrounded lax
/a/ central low unrounded short	/æ/ front low unrounded

Long vowels, however, were excluded because of differences between the two languages that may result in the speaker having difficulty perceiving contrasting forms because of the absence of some vowels from the participants' first language. Vowels in Arabic are less variable among speakers of different Arabic dialects which employ the three short vowels that were listed above and another three long vowels. There are cases where long vowels might be analyzed as a sequence of two nuclei rather than a single branching nucleus (Kaye, 2013). English has a larger number of vowels that do not have an equivalent in Arabic and thus may be problematic for an Arabic speaker to identify and produce. The quality of English long vowels also varies considerably between English accents. Alshangiti (2015) investigated how Saudi Arabic learners of English perceive and produce English vowels. She examined the problematic phonemic contrasts for learners of British English. The results showed that vowels that do not have counterparts in Arabic were more challenging for Arabic listeners to recognize. Also, Shafiro, Levy, Khamis-Dakwar and Kharkhurin (2013) investigated the perception of American English vowels and consonants by native Arab speakers and Arab-English bilinguals. The results showed that vowel perception was less accurate than consonant perception in both groups. The authors attributed low accuracy in perceiving vowels to the bilingual participants' mapping of the larger Arabic English inventories to the smaller inventories of Arabic vowels. Moreover, a speaker might substitute one phoneme with another according to the speaker's first language. For example, in an Arabic word like /ka:n/ (where the vowel is front low unrounded and long), English speakers would be expected to assimilate to the nearest vowel, which is the back low unrounded vowel [ɑ] such as in the word /calm/ . This was reported in a study by Huthaily (2003) where he examined the difficulties in producing vowels for English learners of Arabic. Huthaliy presented participants with the Arabic word [ra:tib] where the long vowel was front unrounded and long. All participants substituted the long vowel with [ɑ] and used [ɑ] to substitute for [a:] whenever it occurred. Since long vowels are commonly substituted, long vowels were excluded. Short vowels were included because they only differ in narrow phonetic details that should not be problematic for the child to recognize when listening to them.

2. **Phonotactic constraints**: To screen Arabic and English children for speech disfluency using non-words, the following phonotactic constraints in both languages were implemented to ensure that the non-words that were created abide by the constraints for both languages

A) **Coda-onset clusters**. 'Coda' refers to a word-internal consonants, not a word-final consonant. This applies when using a syllable template with a coda like CVC in a creating a non-word with two syllables and more, for example: CVC.CV , CVC.CVC. Word-internal codas are restricted by the following six conditions: (1) A coda must be a sonorant /m/,/n/,/w/, /r/,/l/ or an obstruent /k/ or /s/, /f/ ; (2) A post-coda onset must be a plosive /t/,/k/,/d/,/g/,/b/; (3) No geminates are allowed as they are not allowed in English; (4) A nasal must be homorganic with a following onset: [mb, nt, mb, nd]; (5) [s] can only appear before a plosive voiceless onset: [st, sk, sf]; and (6) [r] can appear word finally in General American English but not in non-rhotic accents like British English hence its usage will be restricted to onset positions alone. All of the clusters allowed by these restrictions are shared by Arabic and English.

A) **Word Final clusters**. As mentioned above, the phonotactics of the word-final consonants in English are similar to those of the internal coda. In other words, a word final cluster behaves like an internal C.C because it is also a coda onset cluster. Consequently, because of the parallelism between the two domains, the phonotactic constraints need to be stated once only (Harris & Gussmann, 2002). Generally, the two phonemes in the word final consonant cluster in English must conform to the sonority sequencing phonotactic principle (SSP). Thus, a coda onset C.C cluster should decrease in sonority. There are exceptions when word-final pairs of consonants violate the SSP. These occur in a sequence of two stops that are not homorganic (e.g. act), and a sequence of a stop + /s/. Moreover, there are final sequences that show no evidence of phonotactic constraints which are usually generated by suffixation. In Arabic, however, the SSP is not a reliable predictor of the sequence of two consonant phonemes that occur word-finally. There are many Arab words, in MSA and in several other Arabic dialects that violate the SSP. Al Tamimi and Al Shboul (2013) conducted a study of MSA coda cluster phonotactics to assess applicability of SSP. They assumed that any consonant that follows the last vowel of a word belongs to a coda. The authors used a sample of around 500 CVCC lexical items that were collected from The Hans Wehr Dictionary of Modern Written Arabic. It is important to note here that looking at the words that were selected from the dictionary, many of them are nouns of high frequency and they are usually pronounced the same way in MSA and Arabic dialects as in /ʕaks/ 'reverse' , /s'ubh/ 'morning', /wadJh/ 'face'. The results showed that contrary to what is widely reported in the literature on the compliance of the CC coda with SSP, this was only true for 42% of the cases. The remaining 58% of the cases showed a violation to the SSP sonority hierarchy in the form of: (1) reversal (49% of the C.C coda clusters showed a rise in sonority); or (2) plateau in which the C.C cluster consonants are of almost equal sonority. The results challenge the fact that the phonotactics of the Arabic C.C coda are sonority based because SSP was violated in more than 50% of cases. This means that the last two consonants in the C.C cluster adhered to the sonority hierarchy only in less than 50% of the words; whilst it was violated in the remaining 58% of the words examined. Violation surfaced in one of two forms: (1) sonority reversals (i.e. the second C is more sonorous than the first); and (2) sonority plateau (i.e. the two Cs are identical in sonority). The authors raised the idea of re-considering a more theoretical model outside the scope of SSP that has been long thought of as to govern complex coda syllables. It is important to remember that the results here were based on a dictionary data; thus, different results might be found when examining spoken Arabic.

2. **Step 2**: Non-word candidates were selected for each syllable length as follows: First, all permitted syllable combinations which could serve as templates were created for each syllable length; 100 of these templates per syllable length were then selected at random. Next, consonant and vowel phones were selected at random, apart from those barred by the constraints listed above (i.e., consonants and vowel selection). The selected phones were entered into the template for each syllable length. Finally, individual syllables were combined. Given all the constraints above, the NWR test should meet the goals of including segments that a child speaking Arabic or English can pronounce and do so in comparable ways across these languages when children are equally fluent. In the test, even when phonologically complex materials for the two languages are tested, testing starts with 'easy' materials with

minimal use of clusters. Then, complexity is increased systematically in terms of the number of segments and syllable structures. Appendix 1 shows the orthographic transcription of the final two, three and four syllable AEN_NWR stimuli.

3. **Step 3**: Stress, as a linguistic phenomenon, occurs in both languages. The two languages are similar in terms of the association between stress and heavy syllables; Arabic and English are both quantity-sensitive languages where heavy syllables attract stress (Watson, 2011; Archibald, 1998). In order to meet the phonological requirement of testing across the two languages, language-specific stress contrasts were avoided by producing all syllables with equal stress, except for the CVCC heavy syllables which mark the end of stimuli utterance. Additionally, another set of identical non-word unstressed syllables was developed where short vowels were reduced to schwa; hence the vowels were homogenised. This was done because there are differences in vowel quality and stress between Arabic and English. When evaluating a child's repetition of an AEN_NWR stimulus, any differences in produced stress patterns or vowel quality are ignored and only consonants were considered in scoring. The reason for ignoring vowels during scoring is because of the subtle differences within the Arabic dialects. Additionally, previous research in English showed that vowels are less important demonstrating that reducing vowels to schwa in synthetic sentences has minor effects on speech intelligibility (Whiteside, 1996).

Step 4: Checks were made to ensure none of the phone sequences are words in either of the languages using Aralex (Boudelaa & Marslen-Wilson, 2010) which is a lexical database for MSA that provides token frequencies of roots and word patterns. Aralex integrates information from two sources: 1) a 40-million-word corpus derived from different newspapers covering various topics such as politics, sport and culture and 2) a dictionary compromised of 37,494 entries that provides information and token and type frequencies of Arabic words and morphemes. Although Aralex was created using MSA rather than a spoken dialect, it still meets the requirements for a lexical database that can be checked for lexicality effects for the following reasons. 1) Despite the existence of many dialects across different Arabic countries, the speakers of the language still have one single inventory of phonemes. 2) MSA and spoken Arabic present with different phonological, syntactic and lexical systems where each fulfills distinct sociolinguistic functions. The database was built from MSA using text from newspapers but when dialectal words were found they were retained in the corpus but flagged as such (Boudelaa & Marslen-Wilson, 2010). It is worth noting here that Arab dialects are rarely written, and they are mostly used for speech communication alone. The Aralex database (2015) has a user-friendly interface that consists of 12 boxes where filled boxes can be used to input a search string. The user has the option of displaying the results in English or Arabic Unicode. For example, the orthographic form window takes as input an Arabic or English script and displays the selected results either in Arabic or English.

In this chapter, the Arabic_English non-word repetition task has been outlined and presented. The test has a high potential for identifying preschool children with speech disfluency. The proposed set of stimuli conforms to accepted standards for NWR tasks including the following language-specific phonotactic constraints (Arabic and English), avoiding later-developing consonants, and minimizing potential resemblance between real words and nonwords. Furthermore, the test does not require knowledge of lexical semantics for either of the two languages; a factor that might influence performance. Taking all factors together, the AEN_NWR is potentially a useful tool for identifying WFD from speech disfluency in preschool children. Testing using those new stimuli with large groups of typically developing Arabic and English children as well as Arabic and English CWS is vital to determine if the test can distinguish between the two groups. In Chapter 4, the AEN_NWR undergoes validation as an indicator of speech disfluency by administering it to a sample of Arabic speakers who stutter. Then, in Chapter 5, the test is used in two groups of Arabic and English typically developing children to measure their phonological performance as part of an array of tests conducted to evaluate a WFD intervention. In the next chapter, an Arabic fluency scheme (A_FS) was designed as part of the screening package to measure speech disfluency with details on disfluency symptoms in Arabic connected speech samples. The development of A_FS is deemed essential to (1) have a scheme that can identify distinct types of speech symptoms; and (2) for use in validating phonological performance measures such as the AEN_NWR.

Chapter 3 The Need for a Redesigned Arabic Fluency Assessment

"It is the grammatical fact which explains, more satisfactorily than physics or metaphysics, why the objects of hearing elusive and insubstantial wraiths compared with those of vision" *Jonathan Rée", I see a voice*

3.1 Introduction

As indicated in Chapter 1, SSI-4 is an established measure that is used widely when assessing the fluency of English-speaking children (Riley, 2009). However, it does not take into account the structural features of Arabic. To date, no tool is available for assessing fluency in Arabic. One of the main aims of this PhD was therefore to establish a speech measurement scheme that can be used in a package to screen Arabic children when they first enter school. An Arabic fluency assessment scheme was created for this purpose that addresses syllable and disfluency count issues which were discussed in the first section of Chapter 1 (syllable and disfluency counts are measured when obtaining SSI-4 in English). To start, samples of stuttered speech in Arabic were analyzed strictly according to the SSI guidelines with respect to stuttering symptoms and syllables count. These revealed issues associated with making SSI-4 type counts in Arabic. Any issues not addressed or not relevant to Arabic when SSI-4 was applied were noted and used as a basis for designing the A_FS reported in the current chapter. This chapter then reports development of two algorithms used in A_FS (one for counting syllables, and one for counting disfluencies in Arabic). The theory behind the algorithms is presented first followed by an experiment that reports on the accuracy of the methods (A_FS and SSI-4) for counting syllables and disfluencies when assessing stuttering in Arabic.

3.2 Preliminary Considerations for Algorithm 1: Syllable count

In order to achieve reliable and accurate assessment of stuttering, it is important to have a clear and explicit definition of what to count and how counts are made. According to Roberts (2011), counting stuttering events in languages other than English can be problematic or even impossible. Languages vary in their syllable structure; monosyllabic words can play different roles to multisyllabic words depending on the language (Selkirk,1995). Moreover, words that belong to function words display phonological properties that are significantly different from words that belong to content word categories. For example, in English, monosyllabic function words appear in either stressless or stressed strong form depending on where they appear in sentences, whereas content words always appear in a stressed unreduced form (Selkirk,1995).

Bloodstein (1970) pointed out that procedures that express the frequency of stuttering as numbers or percentages of affected words or syllables has been extensively used in research that documents the variability of stuttering. However, Howell and Rusbridge (2011), whilst acknowledging that was true for English, reported on the lack of standardization when assessing stuttering in different languages and emphasized the need for clear definitions of symptoms and assessment when studying stuttering in other languages.

Irrespective of the length or type of speech sample, the frequency of stuttering can be expressed as the percentage of words or syllables affected by disfluency and significant differences in scores between PWS and typical speakers occur (Conture el., 1987). The decision about whether to measure word or syllable units depends on several factors including the purpose of the assessment (e.g., whether it is a clinical evaluation or for research purposes). Also, the variation in structure of syllable forms across languages plays a role. For instance, there are languages that do and do not allow high numbers of syllables per word. Calculating the percentage of affected units using either syllables or words can lead to big discrepancies across language types because using either number of syllables or words in the denominator of an expression like units affected/all units leads to different results (Yairi, 1997). This leads to a related issue when comparisons across languages are made with either unit. Thus, if the number of syllables per word differs across two languages, the opportunity for discrepancies between the two counting methods increases. This discrepancy may be moderated by converting word counts to the average number of syllables per word in each language – e.g., approximately 1.5 for English (Yaruss, 2000). Note, that a different divisor would be required for Arabic. Moreover, research is needed to establish the conversion factor (average number of syllables per word for Arabic) and this is not a straightforward process. The high usage of inflection in Arabic can lead to inconsistent syllable counts and possibly poor agreement between observers. In Arabic, the syllable must start with a consonant. Therefore, all Arabic words also begin with a consonant. When a word starts with a definite article /ʔal/, the glottal stop /ʔ/ may be dropped when the word is pronounced in combination with a subsequent word. The definite article then takes the form of a proclitic prefixed to nouns and adjectives (Winder & Ziyadeh, 2019). The glottal stop, a sound produced by obstructing the flow of air in the vocal tract, is pronounced in some variations of English as an allophone of /t/ like in 'witness' or 'Scotland'. An example where it is dropped in Arabic is when /kitaab/ 'book' is used along with /ʔalwalad/ 'the boy' which is pronounced /ki.taa.bal.wa.lad/ for 'the boy's book'. In this case, the first sound from the second word starts the additional syllables of the compound word and /ʔ/ becomes/al/ (as in the example), /el/ or /ul/. A short vowel suffix is added to the end of the word /kitab/ to serve as a liaison between the two parts of the compound word. Thus, depending on the context and position within a sentence, Arabic words may be produced with a basic canonical form where the syllables are fully articulated, or they may be re-syllabified. The following examples illustrate this using two and three syllable words.

a. A two-syllable word : /ki-taab/ 'book'
    i. /ki-taab/ Pause form.
    ii. /ki-ta:-bu/ Full form (e.g., when the word is the first part of a genitive construction, it is added to a definite noun).
    iii. /ki-taa-bun/ Full form ("nunation[4] a final /n/ sound used as a suffix in Arabic" is added to indicate an indefinite article)

[4] In Modern Standard Arabic, nunation refers to the process of making nouns or adjectives indefinite. Nunation is rare in spoken Arabic.

b. A three-syllable word: / da- ʤa:- ʤa. / ‘chicken’

i. /da- ʤa:- ʤa/ Singular pause form

ii. /da- ʤa:- ʤa-tun/ Plural full form (“nunation“ is added to indicate an indefinite article)

iii. /da- ʤa:- ʤa:t/ Plural pause form.

iv. /da- ʤa:- ʤa-tu-ha:/ Plural full form with a suffix.

The division of English words to function/content word types was introduced in Chapter 1 to provide a context for the analysis of disfluency types. Howell and Rusbridge (2011) described a scheme where English words can be grammatically classified into two types. Classification depends on whether stuttering occurs on function or content words. Function words are a closed set consisting of pronouns, articles, prepositions, conjunctions, and auxiliary verbs. Content words, on the other hand change “dynamically” and include nouns, main verbs, adverbs, and adjectives. The classification of word-types in English to function or content words is helpful in determining different disfluency patterns (Howell, 2004). Additionally, it aids in demonstrating how there is a developmental pattern in how disfluency loci change from childhood to adulthood. Work by Howell, Au-Yeung and Sackin (1999) was one of the earliest attempts at studying the mechanism of the association between word types and stuttering. They showed that children who stutter have a tendency to repeat whole function words. As they age, this type of disfluency decreases, disfluency on content word increases. Similar results were obtained for Spanish speakers who stutter (Au-

Yeung, Gomez & Howell, 2003). The authors evaluated whether there is a developmental pattern in the loci of disfluency from function words to content words. Participants were divided into five groups in the age range from 3 to 5 (the youngest group) to 20 to 68 (the oldest group). It was found that disfluency rate of function words dropped off and content word disfluency rate increased across age groups. The classification into function and content words might not be appropriate for analysis of Arabic speech samples. In Arabic, the syllable plays a functional role, even when the syllable is short. For example, the Arabic /wa/ وِ functions as a proclitic that is attached before a word to refer to coordination (e.g., the pen and the paper) or to swear. Similarly, the syllable /fa/ is another proclitic that is a one-syllable function word used to refer to ordinality (e.g., the first then the second). The same holds true for other one-syllable Arabic proclitics such as /li/ 'for' , /bi/ 'by' , /ka/ 'like', /li/ 'causation' and /ʔa/ 'questioning' . This shows that both the content and the function words in Arabic can be combined into one single word. In these cases, the addition of function word to content words results in function-content words (Al-Tamimi et al., 2013). Table 3.1 below describes the three lexical types of word in Arabic, along with an example on each.

Table 3.1 : A description of word types in Arabic grammatically classified into three categories.

Lexical Category	Description	Example
Function word	Pronouns, articles, prepositions, and conjunctions	/fii/ (‘in’)
Content word	nouns, adverbs and adjectives	/kabiir/ (‘big’)
Function-content word	content words with one or more bound function words. A possible example is a noun with bound possessive, inflected verbs and definite articles with a noun	/ʔalkaatib/ (‘masculine writer’)

Given this difference in lexical categories between Arabic and English, it might not be appropriate to convert between word and syllable count when measuring stuttering in Arabic. Estimates of stuttering frequency when the measurement unit was based on word or syllables can be substantially different because of the extra word type. Additionally, due to the differences in the syllabic structure of Arabic compared to English, reliability in syllable counts may be compromised. Thus, it is important to have clear guidelines on what should be counted as syllables by investigating the sub-constituents of individual Arabic syllables. Equally important is to examine issues that arise in conversion between word and syllable counts when assessing stuttering in Arabic.

3.3 Preliminary Considerations for Algorithm 2: Disfluency count

The number of disfluencies and their specific forms are regarded as a vital indicator of stuttering Ambrose & Yairi, 1999). The previous section looked briefly at issues in counting the number of syllables and words in Arabic speech samples to obtain valid measures for that language. In this section, the features of the disfluencies identified in Arabic are examined to determine how to classify disfluencies in Arabic.SSI-4 was standardized for English speakers. Therefore, it does not account for many unique phonological, morphological and syntactic features of Arabic. Because currently there is no equivalent of SSI that has been developed and standardized for Arabic speakers, the guidelines in the English SSI-4 manual were used. Thus, a word is denoted as disfluent if the speaker: 1) repeats 2) prolongs a sound or a syllable or 3) breaks part of the word before completing it. Repetition of whole words or phrases do not count as stuttering instances (Riley, 2009) and these events were excluded when the norms for English speakers were established (Riley 1994). There are currently no indications about what to count as disfluency in Arabic, which, if such disfluencies differ from those in English, would require re-standardization of instruments like Arabic forms of SSI-4. In other words, there is no scheme for analysis of stuttered speech in Arabic that shows which types of disfluencies should be counted. Therefore, we relied on the Arabic index of phonetic complexity as a basis (AIPC) to inform parts of our analysis to develop the new disfluency algorithm. AIPC was devised by Al-Tamimi et al. (2013). The AIPC was developed as a framework that provides detailed assessment of which Arabic words are phonologically complex. This was achieved by attributing difficulty of words to the phonetic factors that a word may possess, which might make it more susceptible to stuttering. The AIPC gives an aggregated complexity score based on summing up the number of phonological factors within each word. It was adapted from the Index of phonetic complexity (IPC) metric that was developed by Jakielski (1998) for assessing early phonetic development of US English. The AIPC was developed to account for the unique phonological, phonotactic and morphological features of Arabic. It took into consideration the complexity in articulating certain consonant phonemes that are identified as difficult sounds due to either late age mastery or complex articulatory movements that they involve. Additionally, a commendable feature of AIPC is that it accounts for geminated consonants through a new category "consonants by length". The rationale behind adopting it was to minimize arbitrary decision making when assessing spontaneous speech samples. The nine complete categories identified in the AIPC are summarized in Table 3.2, which allows the metric to assess words with respect to the nine distinct factors listed. The way AIPC is counted is that when any of the phonetic attributes occurs in a word, it is given a score of one point. The scores are then added up and an overall AIPC score is obtained. The factors that score a point and those that do not for each of the nine categories are shown in Table 3.2.

Table 3.2: Nine factors in the Arabic index of phonetic complexity (AIPC) adapted from Al-Tamimi et al. (2013)

Category	Points assigned for	No points assigned for	One point for each
1	Consonants place of articulation	Labials, coronals, glottals	Dorsal, radical
2	Consonants by manner of articulation	Plain stops, nasals, glides	Emphatic sounds, fricatives, affricates, liquids
3	Vowels by class	monophthongs	rhotic
4	Word shape	Words that end with a vowel	Words that end with a consonant
5	Word length in syllables	Monosyllabic or disyllabic words	Words with three or more syllables
6	Singleton consonants by place variegation	Words with place reduplicated singletons	Words with place variegated singletons
7	Contiguous consonants	No clusters	Consonant cluster
8	Cluster by type	Homorganic cluster	Hetrorganic cluster
9	Consonant by length	Singleton consonants	Geminate consonants

Geminated consonants in Arabic play a key role in formulating the syllable template words can take (Ibrahim, 2016). Al-Ani (1970) defined gemination as “the prolongation of the continuants and a longer closure of Stops”. Gemination in Arabic can occur word-initially, word-medially, and word-finally and geminated consonants are considered consonant clusters. Word-final geminated consonants do not have a distinct morphological function in Arabic (Ibrahim, 2016). This means that whether an Arabic word is articulated with or without a geminate consonant, this would never affect its overall semantic meaning. For example, a word such as /fann/ "art" (terminated with a geminate consonant) and a word like /fan/ "art" (terminated with a non-geminate consonant) are semantically identical.

With respect to the syllable boundary, the first member of the geminated cluster occurs as a coda of the preceding syllable, while the second member is the onset of the following syllable (Ibrahim, 2016). The following two examples illustrate the syllable boundary in geminated consonants.

1. /bal.lal/ CVC-CVC /he made wet/
2. /sam.mam/ CVC-CVC /he poisoned/

According to Al-Ani (1970), every consonant cluster involves a transition, which means that the first consonant in the cluster is not released until the second member is uttered. The geminated clusters contrast with their corresponding single consonants as illustrated in the following example.

1. /kataba/ 'he wrote'
2. /kattaba/ 'he caused to write'

Transitioning from one consonant to another may be problematic for Arabic speakers. It is predicted that geminated consonants are a factor that attracts more disfluency in the form of repetition (e.g. /kas.kas.sart/), prolongation (e.g. /bbbitʃak.karik/) or breaks (e.g.,/ʔit-tisˤaal/ where the '-' indicates the break location before the second segment is produced). In the three forms, disfluency might occur on the geminated cluster itself or prior to it, as the speaker plans to produce the gemination. A geminated consonant in Arabic is written once only and a small diacritic called 'shadda' is placed over the consonant (Ibrahim, 2016). However, the occurrence of 'shadda' is not usually required as its existence can usually be inferred from the contexts. Nevertheless, presenting it is useful, especially for children or in books for early readers of Arabic to avoid confusion between single and geminate consonants.

Predictions resulting from considerations about syllable and disfluency count algorithms

The investigation that follows is an empirical examination about what should be counted when measuring stuttering and what and how disfluencies should be counted. Furthermore, the study also aimed to identify which of the AIPC factors were most important in attracting stuttering. The goal was then to use results from this study to establish guidelines for accurate measurement of stuttering in Arabic spontaneous speech samples. Comparisons were made between measures of stuttering frequency based on word counts and measures based on syllable counts. Additionally, tests were conducted to investigate whether the three-word types (function, content and function-content) differ in their mean AIPC scores and to look at associations between individual AIPC factors and stuttering. For instance, are stuttered/non-stuttered words related to words that include gemination/words that do not include gemination. Tests were also conducted to assess the order of importance of AIPC categories.

With respect to syllable count, a hypothesis was developed concerning establishing guidelines on how syllables are counted when measuring fluency in Arabic spontaneous speech samples. In English, a conversion factor of 1.5 syllables per word has been suggested as a reliable convertor with sufficient precision (Yaruss, 2000). Because of the complex morphological structure of Arabic, the conversion factor cannot simply be applied. Thus, it seemed sensible to first examine the possibility of converting between word and syllable counts. We tested the hypothesis that a significant difference should occur in the percentage of stuttering depending on the measurement unit (i.e., syllable count vs word count) (Hypothesis one).

With respect to disfluency count, the following three hypotheses were tested.

1) There will be a significant difference in phonological complexity of different word types. Content-function words are expected to be more phonologically complex than content words, which are in turn are expected to be more phonologically complex than function words. AIPC was used as the metric to measure phonological complexity. Thus, AIPC scores of content words would be expected to be higher than the AIPC score of function words (Hypothesis two). Research on English and German using the IPC as a metric has shown that content words were phonologically more complex than function words (Dworzynski & Howell, 2004). It would be of interest to examine which AIPC factors predict stuttering and the strength of these factors. When word types are analysed separately, stuttered words (where a stuttering instance is a repetition, prolongation or a break) will have high AIPC scores compared to their non-stuttered counterparts (Hypothesis three).
2) On the basis of the above description about how gemination attracts more disfluencies, it was predicted that there should be an association between geminated consonants and stuttering. Geminated consonants pose difficulties for PWS (Al-Tamimi et al., 2013) and therefore they are more likely to be stuttered than their counterpart non-geminated consonants (Hypothesis four).

To ensure precise quantification to stuttering symptoms, it is essential to know which of the AIPC factors has the most impact on stuttering and to assess the order of importance of those factors. Thus, a research question concerning the order of factors leading to stuttering was examined.

3.4 Methods.

3.4.1 Participants

The researcher collected speech samples from as many children and adult PWS as possible to obtain a representative sample of Arabic speakers with developmental stuttering. Speech therapy clinics and nurseries as well as administrators of stuttering self-help groups were contacted. It was difficult to find candidates because of certain cultural sensitivities. Specifically, only five female candidates provided reading samples; thus, their data were excluded[5]. This is because spontaneous speech has a higher external validity with respect to the experience of stuttering since it is a natural reflection of the participants' language abilities (Howell & Rusbridge, 2011). Despite these logistic and cultural difficulties, the current experiment included 10 monolingual Arabic speakers. They were ten males aged between 6 and 28 years old. In terms of Arabic dialects, 10 participants spoke Saudi Dialects (4 Najdi and 2 Hijazi) and one each spoke Syrian, Yemini, and Bahraini and Omani dialects. Seven participants were diagnosed as stuttering by a Speech language therapist (SLT); and three self-reported

[5] Due to cultural reasons, all female participants agreed to only provide reading samples and were hesitant to provide spontaneous speech. This was especially the case when they were informed that speech samples will be recorded

that they exhibited stuttered disfluencies but never attended speech therapy due to lack of services in the areas where they live. Table 3.3 contains information on participants' gender, age, spoken dialect, and the number of stuttered syllables and overall syllables from the transcriptions. Participants were divided into two groups according to their ages. Group 1; N=3 (6 to 14), Group 2; N=7 (15 to 28). The cut-off age of 14 was applied because it approximates the teen age of 12 when change in speech pattern occurs (Howell, 2004).

Table 3.3: Participants demographic Information, words and syllable counts and stuttering frequency

ID	Group	Gender	Age	Stuttering Diagnosis	Arabic Dialect	Syllables Count	Word Count	Stuttered Syllables	%SS	Number Stuttered words
1	1	M	6	Yes	Najdi	200	96	18	9	14
2	1	M	6	Yes	Najdi	176	65	39	11.9	35
3	1	M	7	No	Omani	200	99	31	15.5	31
4	2	M	14	Yes	Najdi	200	96	29	14.5	26
5	2	M	20	Yes	Syrian	200	97	17	8.5	13
6	2	M	21	Yes	Najd	200	88	3	1.5	2
7	2	M	23	No	Yamani	200	93	21	10.5	17
8	2	M	24	No	Hijazi	200	88	7	3.5	4
9	2	M	28	Yes	Hijazi	200	105	2	1	2
10	2	M	28	Yes	Bahraini	200	124	18	7.5	14

3.4.2 Procedure

Transcription

Spontaneous speech samples were orthographically transcribed verbatim by the researcher who is a native Arabic speaker and has phonetic training. The ten transcripts (from three children and seven adults) were used for analysis. All samples had accompanying source audio files, available as MP3 files that can be played on audacity. This allowed double-checking of transcripts if queries arose about the reliability of transcripts of particular words. Syllable and disfluency counts were made on monosyllabic, disyllabic, tri-syllabic and tetra-syllabic function, content and content-function words separately and then words were classified as fluent or disfluent for analysis. All stuttering characteristics were marked on the transcription and then transferred to tables along with an English translation. Interjections (e.g., /ah/, /um/), revision (e.g. I like / I want chips), abandoned utterances (e.g. he said/ we all went) and phrase repetition (e.g. we ate and then we ate) were not counted as disfluencies in this analysis as these are less typical of stuttering. Yairi and Ambrose (1999) indicated that, based on their long clinical experience, these utterances tend to occur frequently and sound fluent. The separate analyses on words of different lexical types (i.e., content, function and content-function) were conducted for each age group. Individual words were extracted from the transcripts; listed in columns in an excel sheet and coded for word type, fluency, AIPC parameters and participant number.

3.4.3 Reliability Measures

The researcher was trained in transcription skills. In this respect, English speech samples (used in later chapters) were transcribed and cross-checked by another researcher until the given sections could be transcribed without error. The researcher also took an intermediate course in Phonetics and phonology course in University College London, Linguistics Department. Due to the unavailability of experienced Arabic transcribers at the time this study was conducted, rather than using inter-judge reliability values, a measure of consistency rather than reliability was taken. Even though this procedure has its own limitations, work by Ingham and Cordes (1997) suggests that judges who show a high intra-rater agreement also have the tendency to have the highest inter-rater agreement. For the re-transcribed and re-coded samples alpha values ranged from 0.86 to 0.92, indicating a high level of consistency.

3.5 Results

Hypothesis one concerned differences in counting the number of syllables depending on the measurement unit (i.e., syllables or words). A Pearson correlation coefficient was computed to assess the relationship between the measure of stuttering based on word counts and measure of stuttering based on syllable count. There was a strong positive correlation between the two variables, which was statistically significant ($r=.80$, $n = 10$, $p = .005$). We also wanted to evaluate if a conversion factor can be used to convert between syllables and words. To obtain this factor, for each participant, the percentage of stuttered words was divided by the percentage of stuttered syllables. The mean value was then calculated, which was 1.96 syllables per word and the SD was .93.

Hypothesis two concerned the relationship between phonological complexity using AIPC as an index and between word types across the two age groups. Related t-tests corrected for multiple comparisons by Bonferroni revealed the following: (1) The overall AIPC scores of content and of function-content words were significantly higher than those of function words for the two age groups (Content words versus function words: Group 1: $n = 3$, $t = 2.5$, $P<.0$; Group 2: $n = 7$, $t = 8.7$; $p < .001$); (Function-content words versus function words: Group 1: $n = 3$, $t = 6.15$, $P<.001$; Group 2: $n = 7$, $t = 16.42$: $p < .001$). (2) The overall AIPC scores of function-content words were significantly higher than those of content words for the two age groups (Group 1: $n = 3$, $t=4.179$, $p < .001$; Group

2: n = 7, t = 8.76; $p < .001$). Figure 3.1 below displays the mean AIPC score for each word type separately for both groups.

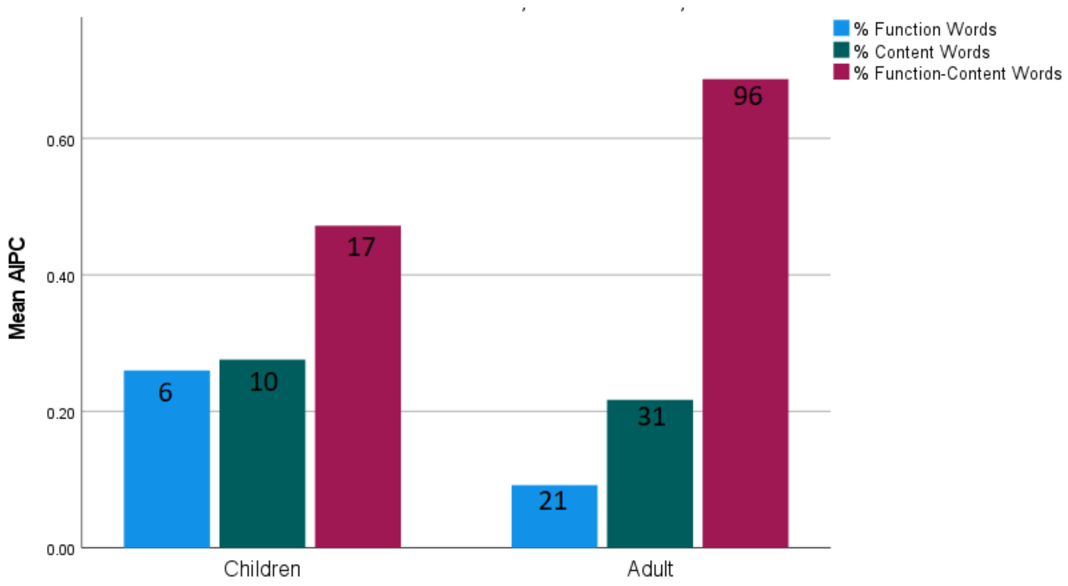

Figure 3.1: Mean AIPC scores for the three-word types for Children and adults

T-tests were used to compare AIPC scores of stuttered words with their non-stuttered counterparts (Hypothesis three). The stuttered words were significantly higher than the non-stuttered counterparts on their AIPC for the function-content words and content words based on these results (Table 3.4). Function words were not included in this analysis due to the low number of function words with high AIPC scores.

Table 3.4: Differences between stuttered and non-stuttered function-content and content words on the AIPC

	Stuttered words	**Non-stuttered words**	**T test result**
Group 1	function-content	function-content	t = 4.91 $p < .001$
Children	Content	Content	t = 3.75 $p < .001$
Group 2	function-content	Content Function	t =.3.3, $p < .001$
Adults	Content	Content	t =1, p<.3

A chi-square test was conducted to determine whether there was an association between stuttering and gemination (Hypothesis four). All words with more than three syllables were included in this analysis. Words with three or fewer syllables were excluded. The rationale behind this analysis relates to the environment of gemination in Arabic syllable structure. Geminated consonants usually occur in word-medial position, and they rarely occur in word-initial positions (Ghalib, 1984). Thus, they seemed to be more likely to occur with longer words. Results of the test showed that the relationship between the two variables was significant X^2 (1, 368) = 4.38, p = .03. Thus, we reject the null hypothesis and conclude that there is a relationship between stuttering and geminated consonants. Logistic regression was used to determine which AIPC factors predicted stuttering and the strength of these factors. Logistic regression was selected as a statistical analysis method because it allows prediction of a discrete dependent variable (stuttering), from a set of independent variables (the AIPC factors). The Wald statistics was used as an index to indicate relative importance of factors. The Wald statistics uses a chi square distribution and indicates whether the independent variables are significantly different from zero. The test was then employed for all AIPC parameters to determine whether each factor was related to stuttering across content and function-content words. Due to the small number of content and content-function word types, they were grouped together during this stage of analysis for both age groups. This can be justified because some Arabic words have content and function aspects. Cases like this include when a letter (not pronoun) precedes a verb to indicate the present tense. Tables 3.5 and Table 3.6 below provides the results of the chi square tests for children and adults, in order of importance.

Table 3.5: Important AIPC factors in adult's' speech

Adult	**Content Function and Content Words**	
Affricates	Wald = 19.50	p<.000
Gemination	Wald = 4.28	p=.03
Fricatives	Wald =4.13	p=.04
Emphatic Sounds	Wald = 1.34	p=.2
Word Length	Wald = .83	0=.3
Word Shape	Wald = .229	p=.63
Radicals	Wald = .041	p=.8

Table 3.6: Important AIPC factors in children's speech

Children	Content Function and Content Words	
Fricatives	Wald = 11.32	P<.000
Gemination	Wald = 7.9	p =.005
Word Length	Wald = 3.79	p=.05
Radicals	Wald = .95	p = .3
Affricates	Wald = .08	p=.7
Emphatic Sounds	Wald = .06	p=.8
Word Shape	Wald = .04	p=.9

3.6 Discussion

The initial part of this discussion is organized in terms of hypotheses to enable understanding of the syllabic and phonetic parameters in Arabic stuttered speech and the relevance of the theoretical part to the empirical stuttered data.

Hypothesis 1: Frequency of stuttering on the basis of measurement unit

The strong positive correlation between the number of syllables counts and word count may suggest that it is possible to use either unit as a metric to obtain measure of stuttering frequency. However, when we attempted to obtain a conversion factor, the range of values for the conversion ratio was wide (between 1.30 and 4.5). This was expected for two reasons: 1) the small number of participants, with one of the participants providing only 175 syllables; and 2) the analysis combined data from adults and children where we expected marked differences in the quality of the utterances as children develop their language.

Hypothesis 2: The relationship between AIPC and word types

The results showed that for both children and adults, the content and the function-content words were phonologically more complex than the function words. Additionally, the function-content were phonologically more complex than the content words. This is consistent with the results of Al-Tamimi et al. (2013). This is also in accordance with results by Dworzynski and Howell (2004) who showed that content words were phonologically more complex than content words in groups of German adults and children.

Hypothesis 3: **Phonological complexity of stuttered vs non-stuttered words**

The AIPC scores of stuttered function-content words were significantly higher than their non-stuttered counterpart for both groups. This validates previous findings by Al-Tamimi et al. (2013). For the content words, the difference between AIPC of stuttered and non-stuttered words was significant only for the children group. Although no statistical tests were conducted due to the small sample size, a general observation when examining the stuttered content words (See Table 3.8 below) showed that the symptoms tend to occur in early position of utterance. This is in agreement with Wingate (2002), who stated that onsets of words are frequently reported to be problematic locations for speakers who stutter. At this stage, word onsets were not included as a factor in AIPC, but it would be of interest to examine this effect as a factor in future studies.

Hypothesis four: **The relationship between stuttering and gemination**

The results showed that stuttering and gemination are associated. The analyses showed that all speakers exhibited stuttering symptoms occurring in words where there was a geminate consonant. In those instances, either the geminated consonant phoneme or one of the consonant phonemes that preceded the geminated consonant in the same word was repeated. It is possible that the speaker repeats a consonant phoneme to prepare for the process of transitioning from one phoneme to another. Alternatively, the speaker might lengthen the vowel that precede the geminate and shortens the one that follows. This pattern has been reported to occur in other languages that have geminated consonants such as Kannada, Italian, Finnish and Benghali (Homma & Yamada, 2014). Table 3.7 gives examples of stuttering events that occurred on geminated consonants. In relation to this, the /l/ sound of the Arabic definite article 'ʔal-' in a phrase like /ʔal-ʃams/ ' the sun' is assimilated to the subsequent single consonant. Consequently, the assimilated form is alternatively pronounced /ʔaʃʃams/.

Table 3.7 Geminated consonants example words and disfluency types.

Word Transcription	Translation	Stuttered Phoneme	Stuttering Type	Stuttering location
ku.kul.li.yat	All of them	/k/	Repetition	First phoneme
ʔal-ʃa.χaṣ	The person	-	Break	Before the geminated consonant
Mmmmmma.ħam.mad	Name	/m/	Prolongation (8 seconds)	First phoneme. Note that the word also included the phoneme /ħ/, a complex one that is produced at the back of the oral cavity

A research question addressed the order of importance of AIPC factors. We analysed which specific factors have the most impact on stuttering in both groups. This provided a way of obtaining a numerical value across AIPC factors to characterize words' difficulty. This is a property that will be included in the A_FS scheme that is outlined in the next section.

Based on Wald statistics results, it was established for children that the factors that made words more difficult were in order of importance 1) consonants by manner of articulation (words that contain fricatives), and 2) gemination. The two factors that made words difficult for adults were in order of importance 1) consonant by manner of articulation (words that contained affricates); 2) gemination and 3) consonant by manner class (fricatives). The late age of acquisition of many Arabic complex sounds (i.e. around the age of 6 and above) is expected to have a role in changing the pattern of how phonetic properties affect speech control in children who stutter which explains why fricatives might have posed difficulty for children, Other consonant by manner of articulation groups did not yield significant results such as the emphatic sounds possibly because those phonemes have low frequency of occurrence in the language. For a child aged 6 or 7 with a small-sized lexicon, chances are low that the child will produce words with those rare phonemes. Thus, a possible explanation for the insignificant results is possibly because of low power due, again, to the small sample size. The differences in the importance of consonant by manner of class (fricatives) between the two groups can be attributed to the late age of mastery for these sounds in this class during phonological acquisition (Bernhardt et al., 2015). Articulating fricatives requires precise control of airflow to force air through a narrow channel. This has been shown in a number of studies that investigated the anatomical and articulatory differences between children and adults and how this may affect fricative production. For example, one of the earliest attempts at examining these differences was done by McGowan and Nittrouer (1988). Speech samples of 8 children and 4 adults producing the fricatives /s/ and /ʃ/ were obtained and acoustically analysed. It was found that the back-cavity resonance is more prominent in children's fricative spectra than in adults. This suggested that an important anatomical difference may be related to the small size of fricative constriction and glottal opening in children relative to adults. Rapid advances in computer technology have allowed additional types of analysis. Thus, Nissen and Fox (2005) conducted a study to examine the acoustic structure of fricatives produced by adults and children. The acoustic structure of the fricatives (/f, θ, s, ʃ/) was described in terms of multiple acoustic parameters including duration, amplitude and spectral moments. It was found that the three measures; duration, amplitude and spectral moments varied significantly between the two groups. These findings highlight that the articulatory differences are developmental in nature, and it would be of significance to consider age when standardizing A_FS. For the purpose of this study, any AIPC factors that shows differences between stuttered/not stuttered comparisons for any age group indicate that that factor is potentially important to include in the design of A_FS. It can be seen that gemination and consonant by manner of articulation are factors that showed significant differences in both groups. Thus, words that carry any of these phonetic features are more likely to have an effect on fluency.

Further patterns in the data

Before presenting the proposed algorithms for counting syllables and counting disfluencies, we present some of our observations on the effects of syllabic and phonetic factors that appeared in the analysis. This information will be incorporated in the algorithms. Worked examples are also given to illustrate the application of those features in the A_FS algorithm.

The phonology of the definite article.

The Arabic definite article forms a syntactic word with the noun that follows it and defines it and a phonological word with the preceding word (Heselwood & Watson, 2013). This means that the definite article prefixes a noun or an adjective and renders it definite and the two (definite article + noun/adjective) form a syntactic unit. Depending on the context, the definite article may or may not be phonologically attached to the preceding word. The noun that follows the definite article may be a regular noun that can stand on its own (e.g. /ʔal.wa.lad/' the boy ') or it may be an irregular word that cannot be broken down further (e.g. /ʔal.laa.ti/ اللذي and /ʔal-la.ti/ التي). Depending on whether the consonant in the noun that follows the definite article is a coronal or a non-coronal consonant, the definite article has two underlying forms.

1) Clearance: When the definite article precedes a noun beginning with any of the following non-coronal consonants /b/, /ʤ/, /k/, /q/, /ʔ/, /f/, /χ/,/ħ/, /ʕ/, /h/, /m/, /w/ or /j/, it has the phonological form /ʔal/ or /ʔil/ (in some Arabic dialects) when the noun is utterance initial (e.g, /ʔal-kalb/ 'the dog'). Alternatively, when the noun follows a vowel-final word then the article has the phonological form /l/ (e. g, /ðay.lal.kalb/ ''the dog's tail).
2) Merging: When the definite article precedes a noun beginning with any of the following coronal consonants /t/, /tˤ/, /d/, /dˤ/, /θ/, /ð/, /ðˤ/, /sˤ/, /z/, /ʃ/ , /l/ , /n/ or /r/,

   then the consonant of the article /ʔaC/ must be the same as the consonant of the noun when the noun is utterance initial (e.g, /ʔaʃ-ʃams/). In this case, the coronal consonant is geminated resulting in an assimilated form. Alternatively, when the noun follows a vowel-final word then the article has the phonological form /C/ where C is the geminated coronal consonant (e.g. /no:raʃ.ʃams/ the sun's light/).

Linguistic properties and choice of word type

Analysis in this experiment was done at the word level where words were separated grammatically into three types; function words, content words and function-content words as described with examples in Table 3.1. This was in line with what was done in previous studies of stuttering in Arabic by Abdalla, Robb and Al-Shatti (2010) and Al-Tamimi et al. (2013). The reason for including the third word type (function-content) in addition to the two-word types in English (i.e. function words and content words) relates to the morphological nature of Arabic. That is, Arabic is a non-concatenative language with fusional features (Hakansson et al., 2003). It allows many morphemes to operate simultaneously within a word to encode functions such as gender and number. This illustrates that the content/function word dichotomy is not sufficient and justifies the addition of a new lexical category to assess fluency in Arabic. Our analysis showed that the morphological properties of function-content words increased difficulty and inclusion in analysis of this word type is important. Words in this lexical category were phonologically more complex than content words, which in turn were more complex than function words. This suggests that using this classification system provides an objective basis for dividing words into classes that differ in difficulty. There is still the need to divide words into classes according to other continuous measures such as frequency of word occurrence and neighbourhood density; all are linguistic variables that are expected to influence the fluency of produced speech (Anderson, 2007)

Formal algorithm for counting the number of syllables using A_FS

1) Divide words into syllables according to the following steps.
    1) Make vowels nuclei of syllables. The three short vowels /i/,/a/,/u/ and their counterpart long vowels /ii/, /aa/ and /uu/ always form the syllable nucleus. The number of syllables in an utterance is identical to the number of vowels (Al-Ani, 1970).
    2) The remaining components are referred to as "margin factors" and they are treated as follows.
        a. Segments that precede the nuclei form the syllable onset
        b. All other segments are syllabified as the form 'syllable final'.
    3) All consonants including /y/ and /w/ represent the margin morpheme in the syllable structure. These may indicate the initiation or termination of the syllable.
    4) Ignore the definite article when it precedes a noun that begins with a non-coronal consonant /b/, /ʤ/, /k/, /q/, /ʔ/, /f/, /χ/, /ʁ/, /ħ/, /ʕ/, /h/, /m/, /w/ or /j/.
    5) Geminated consonants are split into two components: the first is related to the first syllable and the second is related to the second syllable (Ibrahim, 2016). Figure 3.2 shows the syllable structure for the word /sabbaba/. The word has three syllables and a gemination

that occurs word medially. The first segment /b/ is related to syllable number 1 and the second /b/ is related to syllable number 2.

/sabbaba/ ‘he caused’

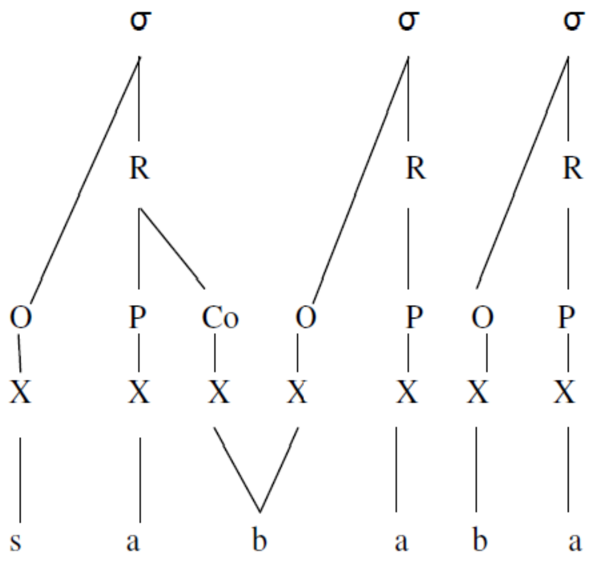

Figure 3.2: Syllable Structure for the word /sabbaba/ (Ibrahim,2016).

6) Make nuclei and word final segments into rhymes.

7) Join onsets and rhymes into syllables.

8) Continue the process until whole utterance has been completed.

On the basis of the description above, the syllable structure can be represented as indicated in Figure 3.3

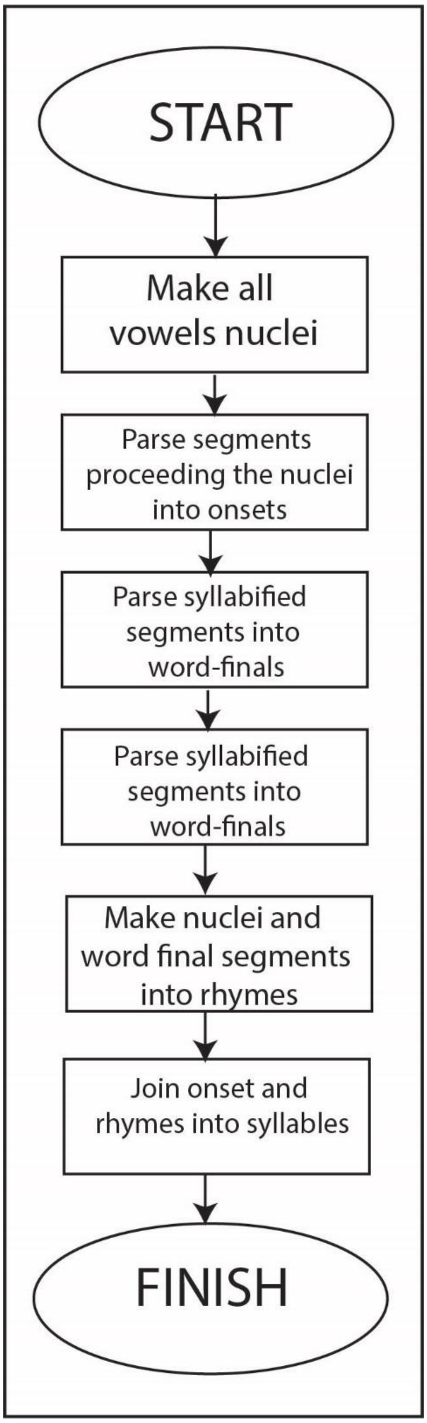

Figure 3.3: Flowchart showing the eight steps for counting the number of syllables

Issues in disfluency count in Arabic

Data from the above analyses showed stuttering instances with high duration when nouns were pronounced that began with the glottal stop /ʔ/ in the definite article /ʔal/. Most of those instances were multiple repetitions of the sound /ʔ/ followed by a vowel, but sometimes it involved a break within a word. In words where the noun itself begins with /ʔ/ and the definite article was a prefix, it became very difficult to produce this complex sequence of phonemes. For example, /ʔa-ʔa-ʔas-ʔa ʔas-ʔas-raar/. In this example, the child repeated the first sound of the definite article (which constitutes the function word within a function-content word) as well as the first syllable of the word /ʔasraar/ "secrets" (which constitutes the content word within a function-content word). This was still counted as one stuttering event because according to the SSI-4 manual each word can only be counted as disfluent once (Riley, 2009). In the proposed scheme, we propose a change in counting the number of stuttering events taking into account the word type to ensure that disfluency counts are not overestimated or underestimated. That is, we advise dividing words into one of three lexical categories that differ in morphological and phonological characteristics. Moreover, the statistical analysis revealed a relationship between phonological complexity and word types (Hypothesis 2). This provides further support for dividing stuttering loci into lexical categories when assessing fluency as this can provide a neat stuttering profile for children and adults (Tsai et al., 2011).

Another pattern that was also exhibited by all speakers, is stuttering that occurs in words where there is a geminated consonant, that raises issues associated with counting syllables. It is possible that the speaker is repeating a consonant phoneme to prepare for the process of transitioning from one phoneme to another (Wingate, 2012). This feature of Arabic is also accounted for by assigning a different weight to stuttering events on geminated consonants. Also, most participants showed stuttering events with long duration on words that began with, or included, complex sounds that are articulated in the back of the oral cavity like the velar /ɣ/ and /x/ and the laryngeal /ʔ/ and /h/ and pharyngeal /ħ/ and /ʕ/. Some of these phonemes occur in other semitic languages (e.g., /x/ in Hebrew), and they pose articulatory challenges. That is, they are pronounced with greater duration and possibly less accuracy even in typical speakers (Icht , & Ben-David, 2014).

Table 3.8 gives some examples. Support for different treatment for these two factors (geminated consonants and fricative sounds) comes from our statistical analysis where we assessed the importance of AIPC factors. Factor 2 (consonants by manner of articulation: fricatives) and Factor 9 (consonant by length) appeared the most important AIPC factors for children and adults in terms of their relationship with stuttering. Therefore, they will be assigned different weight when evaluating participants' stuttering.

Table 3.8: Examples of stuttering on complex phonemes.

Word Transcription	Translation	Stuttered Phoneme	Stuttering Type	Stuttering Location
ʕaʕaʕaʕaqlak	Your brain	/ʕ/	Prolongation – 1.84 seconds	Word initially
ʔʔʔalħħħa . ʤar	Stone	/ʔ/	Prolongation – 4.6 seconds	Word initially. The word also included another pharyngeal /ħ/
xxxxxam.xams.miy.yah	Five Hundreds	/x/	Prolongation – 3.8 seconds	Word initially. The word also included gemination.

This section finishes by a proposed algorithm for counting the number of disfluencies in Arabic.

Formal algorithm for counting disfluencies:

After counting the total number of syllables as described above, the percentage of stuttered syllables (%SS) needs to be obtained by dividing the number of stuttering events by the total number of syllables. All words should be analysed based on their lexical types and phonological features as is described below. Additionally, information derived from the disfluency analysis above using AIPC scale were incorporated to aid in quantifying disfluencies during scoring. With respect to speech behaviours to be counted as stuttering symptoms, we selected the three disfluency characteristics that are often cited in English (i.e., part-word repetition, prolongation and break). As mentioned earlier, these are also the symptoms that occur in Riley‘s (1994) SSI-3 manual. As this is the first attempt at establishing guidelines for assessing fluency in Arabic, we followed the most widely used syllable-based procedure, used the symptoms as a benchmark and then accommodated all necessary changes to come up with the new scheme. The same is true for the number of syllables to be counted, as we suggest obtaining 200 syllables. Support for this comes from work by Todd et al., (2014), which showed that a speech sample of 200 syllables is sufficient to compute a stuttering severity estimate. As well as counting the number of disfluencies accurately, it is equally important to quantify those disfluencies by assigning different weight to different words based on their phonological complexity. What has been learned from AIPC and the hypotheses we tested is that certain words with specific phonological characteristics pose more difficulty to the speakers and therefore attract more stuttering. The issue was how to find a valid formula that can combine lexical types, phonological complexity, and frequency of stuttering into an overall severity score. The study was limited by the small number of participants in both children and adults' groups; hence, this has made it even more difficult to construct a formula with limited data. Therefore, the focus has been on how to document the frequency of stuttered syllables, the specific stuttering symptom and the stuttering loci. This will provide detailed information on the nature of the disorder as it manifests in Arabic speakers and could also form the first stage of a long process of evaluating a new scheme to assess speech fluency which would extend from the current work. The scheme we propose involves a comprehensive analysis of speech based on a verbatim transcript. It also involves categorization of word type as well as phonological complexity and assigning these words different weight before integrating them in the stuttering severity formula.

The four steps for counting disfluencies are described next.

Step 1: Word type. Classify words into one of the three lexical types.

A) Function words form a closed class of words that have grammatical or functional roles but no lexical meaning and most of them are monosyllabic. This class includes pronouns (e.g. You for male /ʔant/ أنتَ) , prepositions (e.g. up /fawq/ فوق), articles (e.g. the definite article /ʔal/ أل), conjunctions (e.g. But /laakin/ لكن), clitics and auxiliary inflections linking verbs. This category can be divided into two types: (1) Free (munfasˤil), which means that the function word can stand on its own (e.g, /fii/ 'in' and /ʕan/ 'about); (2) Bound (muttasˤil), which means that the function word is attached to the word it precedes or follows (e.g./lawnoh/ 'its' color' the original word is colour /lawn/ but a morpheme is added to indicate that the colour belongs to a masculine second person.

B) Content words include nouns, adverbs and adjectives.

C) Function-content words include content words with one or more bound function words. A possible example is a noun with bound possessive, inflected verbs and definite articles with a noun. (e.g., /kaatib/ 'masculine writer' can be preceded by the bound function word /li/ (for) and /ʔal/ (the) and followed by the feminine plural suffix /aat/ to create the function-content word /li # l # kaatib # aat/ ‘for the writers (feminine form)’, where # indicates a morpheme boundary.

Note: When stuttering occurs on a function-content word and the participant for example stutters on the definite article (function word) and then the noun itself (content word), only the stuttering event(s) on the content word part is counted. For example, in a word like (ʔa-ʔal ba bawwaa.ba) there are four syllables (the definite article is ignored) and two stuttering events. The first stuttering event is the sound repetition of the first phoneme of the content part of the content function word /ba/ and the second one is the break before the last /ba/ syllable.

Step 2: Phonological complexity. For each word that has been lexically categorized, provide information on whether it is considered “easy” or difficult according to Table 3.9.

Table 3.9: Two factors selected from AIPC based on their importance (showing a significant relationship with stuttering).

Factor	Easy	Difficult
Consonant by manner of articulation [6]	The word does not contain any of the following Fricatives/ɣ/,/x/,/f/,/v/,/θ/,/ð/ /s/,/z/,/ʃ/ and /h/	The word contains any of the following Fricatives/ɣ/,/x/,/f/,/v/,/θ/,/ð/ /s/,/z/,/ʃ/ and /h/
Consonant by length: geminated consonants	The word does not contain a geminated consonant	The word contains a geminated consonant

Step 3: Sum up disfluencies.

Count the total number of disfluencies in each of the categories from step 1 and step 2 and divide by 200 (the total number of syllables). Repetition of whole function words should not be counted unless the function word itself had one of the above-mentioned stuttering symptoms (e.g., the function word /laakin/ is stuttered if the first syllable /laa/ is repeated, or if there is a prolongation or break within the word)

Step 4: Stuttering severity score. A total overall score that reflects stuttering severity needs to be obtained. To control for inflation that results from function-content words and the difficult words according to the two phonological complexity factors (Table 3.9), it is suggested that they are

[6] Fricative sounds in the consonant of articulation category occur in the IPC (Jakielski, 1998), but there are certain fricative sounds that exist in Arabic but not in English. These are the two velar sounds /ɣ/ and /x/ that are articulated at the back of the oral cavity. These two and all other fricative sounds that exist in Arabic (/f/, /v/. /θ/, /ð/, /s/, /z/, /ʃ/ and /h/) were included.

multiplied by .5 before adding them up in the formula. The rationale behind this adjustment is that due to their complexity, these word types may attract more stuttering and therefore this might have an influence on the overall stuttering score.

Severity Score = 1*(%Function words) + 1*(%Content words) + .5*% (Function-content words) + 1*(%easy words) + .5*% (difficult words) + .5*% (words with geminated consonant)

Figure 3.4 below summarized the steps for counting and weighing disfluencies in Arabic

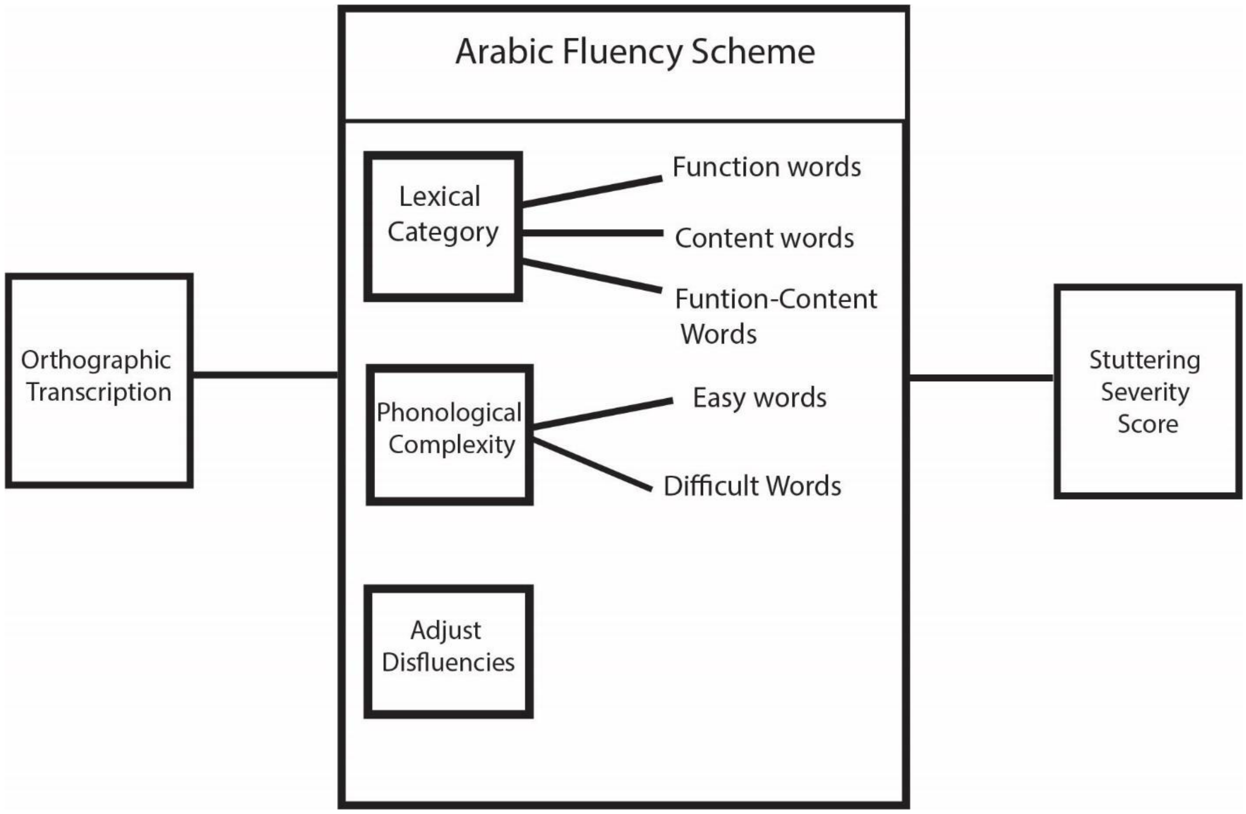

Figure 3.4: Summary of the steps to count disfluencies and obtain a stuttering severity score

General Discussion and Conclusion

The purpose of this chapter was to examine factors that affect syllable and disfluency count in Arabic and use them as a basis to develop the A_FS scheme. As a first step, two preliminary algorithms were proposed based on our investigation of syllabic, phonological and morphological features of the Arabic language. This gave the basis for the development of a formal scheme, which is intended to provide a framework for the characterization of syllabic and phonological structure of words in spontaneous samples. The empirical work then involved analysis of conversational speech samples of at least 200 syllables; these were obtained from Arabic adult and child speakers who stutter; of which seven of them had a confirmed stuttering diagnosis. Speech samples were transcribed using Arabic orthography and moments of disfluency were marked on the transcripts. The stuttering symptoms used were part-word repetition, prolongation or a break; these are equivalent to Riley's (1994) stuttering symptoms. The main goal of the analysis was to highlight areas where potential changes in counting the number of syllables and the number of disfluencies are needed to accommodate the requirements of the Arabic language. With respect to phonological factors, the analysis also takes some aspects of AIPC and incorporate those in designing the scheme. A commendable feature of AIPC is that it has a way of deriving a numerical value across all AIPC factors to characterize a word's difficulty. Consequently, two algorithms have been proposed: one for counting the number of syllables and one for counting disfluencies. These were presented with clear guidelines on how they should be applied. Overall, it seemed advisable to analyse words of different lexical categories (function, content and function-content words) separately as these word types tend to have different phonological characteristics and involve different types of stuttering. However, more research on Arabic is needed on the role of different lexical categories to see how it affects stuttering on different parts of speech (e.g. adjectives and adverbs which have specific inflectional structures). To explain, Arabic has unique inflectional structures on different parts of speech that potentially makes the link between word type and stuttering more complex (Vahab et al., 2013). Adjectives and adverbs in Arabic use a system of agreement on number and gender with the noun or pronoun that they modify. That is, this creates lexical flexibility in combining words of different forms, which in turn could have an impact on stuttering rate (Vahab et al., 2013). Furthermore, whilst English has a classic subject-verb object structure, the word order in Arabic is mostly verb-subject- object, but other forms are also acceptable such as the subject-verb-object (Watson, 2002). In fact, in many dialects of Arabic word order is usually dependent on factors such as the dynamism of the verb. Arabic also accommodates almost all patterns and word-forming processes that are used in inflectional languages, as well as ones that are specific to isolated languages (Vahab et al., 2013). There is obviously a much greater flexibility with respect to the position of the subject in Arabic, which necessitates further investigations on the effect of specific morphological variables on stuttering. It would be of interest also to examine the role of inflections in the forms of suffixes and prefixes on stuttering rate in Arabic.

A_FS is a new scheme that has not yet been examined for reliability or validity. As the next chapter concerns validating the AEN_NWR, work was also initiated to evaluate A_FS. Spontaneous speech samples are obtained from adults and children who stutter. The two formal algorithms were applied to count the number of syllables and to count disfluencies and obtain a severity score for each participant.

Limitations and Future Work

Probably the biggest limitation of this study is the small number of participants in both groups, adult and children. This led us to group content words with function-content words when testing certain hypotheses in order to conduct meaningful statistical analysis. For the same reason, it was necessary in one of the tests to limit statistical analysis to words of three syllables or more. Particularly, when testing hypothesis four (the association between stuttering and gemination), only words of three syllables or more were included. This is justified by the need of running meaningful analysis and the fact that longer stuttering is more likely to occur in longer words; particularly in word medial positions (Ghalib, 1984). It is worth noting here that work is needed on the frequency of occurrence of syllable forms in Arabic (Khattab & Al-Tamimi, 2013). Additionally, due to cultural issues, all participants were male, which made it impossible to examine gender-related differences. It is hoped, therefore, that future studies will include more adults and children. This will aid in increasing the power of statistical analysis.

Chapter 4 Towards Establishing Instruments to assess fluency in Arabic: Validation Study

"Those who stutter win, in the painful pauses of their demonstration that speech isn't entirely natural, a respectful attention, a tender alertness. Words are, we are reassured, precious". *John Updike, Getting the words out*

4.1 Introduction

Although the AEN_NWR was developed based on the same phonologically informed approach used with UNWR (Howell et al., 2017), which in itself has been established as an effective screening instrument to identify children who stutter, research on the AEN_NWR is still limited. Therefore, the aim of this chapter is to validate the claim made about AEN_NWR concerning its effectiveness at identifying fluency difficulties in a sample of PWS who speak Arabic as their first language. To provide support that the AEN_NWR is a reliable measure of phonological skills and speech fluency, the present work tested for a relation between AEN_NWR scores and the percentage of stuttered syllables (%SS). It was hypothesised that AEN_NWR scores would correlate with %SS in PWS, such that a lower AEN_NWR score would be associated with a higher %SS indicating higher levels of stuttering. Out of the three components of SSI-4, only the %SS was used in this study because Mirawdeli and Howell's (2016) reported that %SS alone, rather than the composite SSI score, was best for identifying children who stutter. Scores on the AEN_NWR were also compared against the A_FS. The goal was establishing the new A_FS as a framework for the characterization of syllabification and syllabic structure of words in spontaneous Arabic speech utterances; one that constitutes not only a disfluency measurement, but rather a reliable and clinically meaningful stuttering measurement procedure. Thus, we tested a second hypothesis that the A_FS would correlate negatively with the participants' total AEN_NWR score (number of trial items).

Additionally, a research question (research question 1) concerning changing in stuttering across ages has been examined; since data were collected from adults and children (See methods section).

4.2 Method

4.2.1 Participants.

There were two age groups in this study; a younger age group (age range from 6 years and 6 months to 17 years and 1 month; mean= 10 years 9 months, SD= 3 year 0 months; N=5; male=5) and an older age group (age ranges from 19 years and 2 months to 31 years and 1 months; mean= 24 years and 3 months, SD= 4years and 1 month; N=10; male=3). All participants were monolingual speakers of Arabic as a first language; and were categorised as displaying stuttering symptoms by an SLP.. Demographics on the children's age, gender and the spoken Arabic dialect were collected from the children's parents' and from the older group before starting the experiment. The general information of all participants are given in Table 4.1

Table 4.1: Participants demographic Information, %SS, scores on AEN_NWR and A_FS performance

	Group	**Gender**	**Age**	**Dialect**	**Treatment History**	**Total Syllables**	**%SS**	**AEN NWR**	**A_FS**
1	Group 1	Male	9:6	Hijazi	Previous	300	2.3	13	1.75
2		Male	11:6	Najdi	Previous	300	5.6	9	1.55
3		Male	11:5	Najdi	Previous	300	4	6	2
4		Male	6:6	Hijazi	Previous	300	10	0	1.70
5		Male	15:10	Omani	Previous	300	10.3	0	2
6		Female	17:1	Najdi	Previous	300	2.3	19	1.69
7	Group 2	Female	19:2	Najdi	Current	300	6.67	3	2
8		Male	22:0	Hijazi	Current	300	6.3	4	1.60
9		Male	25:2	Najdi	Previous	300	3	14	1.4
10		Female	25:0	Hijazi	Previous	300	6	3	1.3
11		Male	29:0	Hijazi	Previous	300	7.6	19	1.66
12		Female	23:4	Najdi	Previous	300	1	18	1.40
13		Female	25:1	Hijazi	Previous	300	7.3	2	2.4
14		Female	31:1	Bahraini	No treatment	300	4	17	4.5
15		Female	26:1	Hijazi	Previous	300	6.3	1	2.5

4.2.2 Procedure

Participants were tested in a quiet setting in one session which required approximately 20 minutes. Arabic was the language of instruction and conversation between the experimenter and participants. The experimenter conducted two tasks: (1) elicitation of spontaneous speech samples and (2) administration of the AEN_NWR. The study used a within-subjects design as participants completed both tasks; and the two tasks were conducted in randomized order. The complete session was recorded on a Sony DAT audio-recorder using a Sennheiser K6 microphone and audacity software. For the spontaneous speech, a sample of 300 syllables from all participants were obtained. These were elicited during a conversational speech with the researcher using topics of interest, such as school, travel, books and hobbies. When necessary, picture material from Riley (2009) were also used to elicit speech from children. For the AEN_NWR, the 28 non-words that were described in Chapter 2 were used as follows. In this experiment, all stimuli from the three sets (i.e., two syllables, three syllables and four syllables) were played to participants regardless of their performance. Our aim was to have as much data as we can on participants' performance on stimuli that vary in phonological complexity. Also, this would allow inter-reliability measures in the case that two researchers listen to the stimulus and provide their own scoring during off-line analysis. Participants were informed that they would hear made-up words; i.e. ones that do not exist either in Arabic or English. The examiner then gave the following instructions to the participant: "I am going to play some made-up words to you through the headphones and I want you to repeat them as right as you can. You will have to listen carefully because you will listen only once". Participants were allowed as much time as was necessary to respond. There were two practice items to make sure the output volume was appropriate, and the participant understood the nature of the task. The non-words were pre-recorded to ensure that factors such as differences in word stress patterns and accent would not affect the results. Recording of materials took place in an anechoic chamber and were obtained from a male professional phonetician who was phonetically trained in Arabic and English[7]

Data processing and analysis

AEN_NWR scoring

The non-word phonemic response transcriptions and target transcriptions for each non-word were aligned to identify how the two transcriptions differed. Differences in terms of alterations (i.e., deletions, substitutions, or insertion) of sounds were identified for scoring. Responses were scored as incorrect if they contained one or more phoneme errors (i.e., substitution, insertion or deletion of consonants). Responses from 20% of the participants, selected randomly, were scored independently by a second trained phonetician. Agreement on the number of correct stimuli was 88%. Data from the first author was used in statistical analysis.

Spontaneous speech

[7] The author would like to thank Christopher Lucas, SOAS University, for recording the non-word stimuli

The audio recordings of the participants' speech were orthographically transcribed after replaying them as many times as necessary by the researcher. Syllable counts were made according to the algorithm described in Chapter 3 (see Figure 3.3), and only the first 300 syllables were used in analysis. Todd et al. (2014) confirmed that a 200-syllable long speech sample was sufficient to obtain a reliable SSI score; however, due to the morphological nature of Arabic and in order to have sufficient syllable numbers to conduct analysis on, 300 syllables were obtained. Disfluency count was performed in two ways. First, the following guidelines from the original SSI-4 (Riley, 2009) were followed to score the first 300 syllables of all samples collected. Estimates were obtained by calculating the %SS (i.e., the percentage of stuttered syllables). %SS was calculated by determining the number of stuttered syllables and the total number of syllables. The number of stuttered syllables was then divided by the total number of syllables and multiplied by 100, resulting in the percentage of stuttered syllables %SS. A syllable is stuttered if the speaker exhibits one of the three disfluency characteristics. 1) repetition indicated by multiple repetition of a sound or a syllable that was not a word, 2) prolongation indicated by abnormal lengthening during the production of a phoneme, 3) break within a syllable.

Second, to obtain A_FS estimate, the proposed disfluency algorithm from Chapter 3 was applied. All stuttered words in the sample were classified into function, content or function-content words and were listed for each participant in an excel sheet. Single word utterances such as yes or no were excluded from the analysis. Next to each word, information was added on the phonological difficulty (i.e., whether or not the word contained a fricative) and the inclusion of gemination consonants. A severity score was then computed for all participants according to the formula in Chapter 3.

4.3 Results

A simple linear regression was used to assess whether AEN_NWR scores could be used to predict %SS scores. A correlation analysis using Pearson's r can be used since the variables were measured on a continuous scale. As expected, the negative correlation between AEN_NWR and the %SS component of SSI-4 was significant (r (15) = -.72), p <.000), indicating that a higher AEN_NWR score correlated with a lower %SS, making AEN_NWR scores a significant predictor of %SS. A scatter plot is given in Figure 4.1 with the %SS on the X axis and the AEN_NWR score on the Y axis.

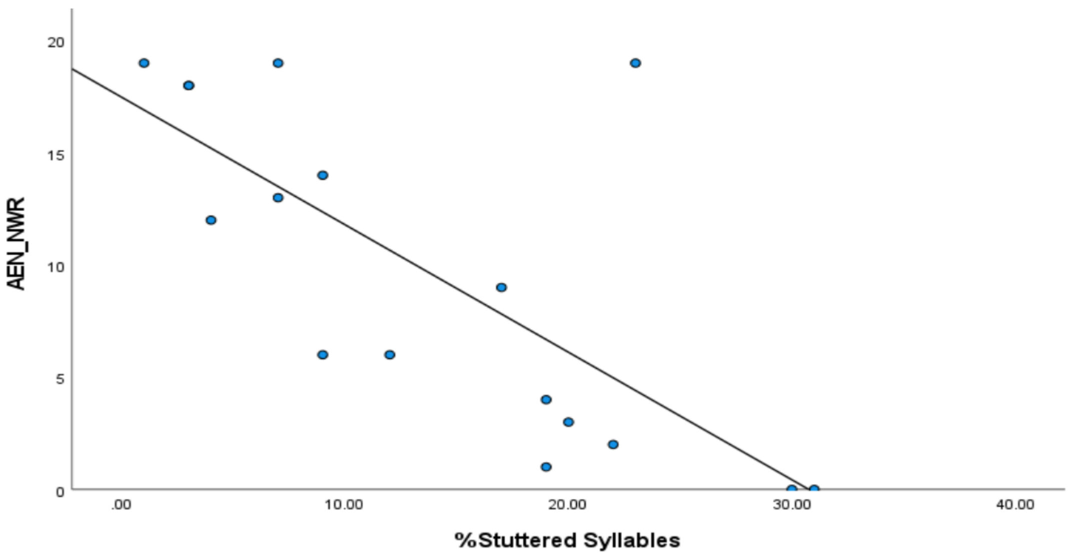

Figure 4.1: Scatter plot of raw %SS (frequency scores), (X-axis) against total AEN_NWR scores (Y-axis).

Participants' performance on A_FS was also compared with AEN_NWR scores. A correlation analysis using Pearson's r was used since the variables were measured on a continuous scale. A significant negative correlation between the AEN_NWR and A_FS scores was found using

Pearson's correlation coefficient (r (15) = -.60), p = .02), indicating that a higher AEN_NWR score correlates with a lower A_FS. The results are summarized in the scatterplot in Figure 4.2 showing scores on A_FS on the X axis and AEN_NWR scores on the Y axis.

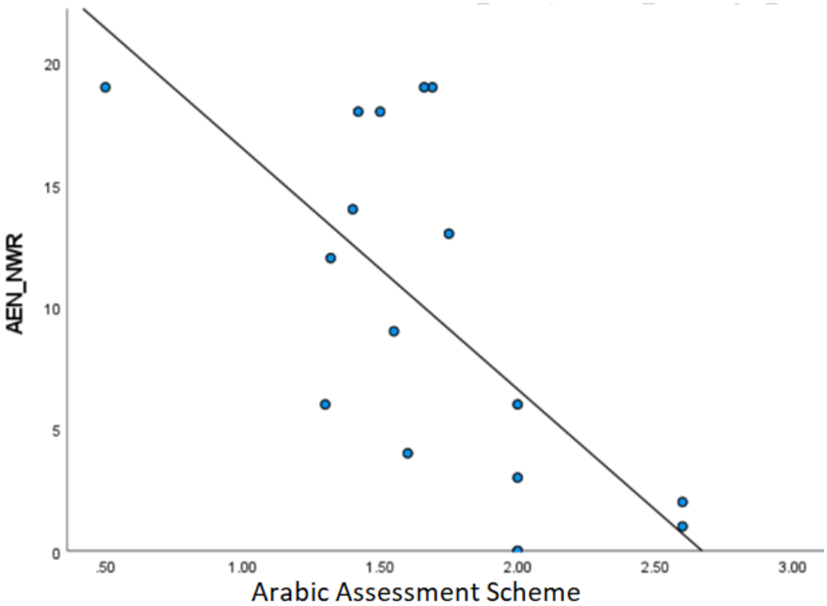

Figure 4.2: Scatter plot of estimates of A_FS (X-axis) against AEN_NWR scores (Y-axis).

Tests were conducted to examine developmental changes in stuttering for words of different lexical types (Research question 1). A one-way ANOVA was conducted with one between-group factor (adult vs children). The dependent variable was the percentage of disfluency on each of the three-word types. There was a highly significant difference in the percentage of stuttered function words for the children group compared to adults (F (1, 14) = 9.8, $p < .001$). For content words, there were no significant differences between children and adults on the percentage of stuttered words. With respect to function-content words, there was a significant difference on the percentage of stuttered words (F (1, 14) = 7.8, $p < .001$); adults stuttered significantly more than children on this lexical type.

4.4 Discussion

4.4.1 AEN_NWR criterion validity

One of the biggest aims of this thesis is to establish the AEN_NWR as an assessment tool that can identify WFD from stuttering in preschool children whose' first language is Arabic or English, including those who speak both languages to different levels. In this chapter, only a clinical sample of Arabic speakers who stutter were tested for their performance on the AEN_NWR, but future work will evaluate the performance of English speakers and compare the two groups (including people who are fluent or who stutter). Evidence was collected on the validity of the AEN_NWR as a reliable measure of phonological skills and hence of fluency of speech. It was hypothesized that AEN_NWR scores would correlate negatively with %SS (i.e., percentages of stuttered syllables, %SS). This specific task was carried out in order to obtain support that the AEN_NWR can act as a measure of speech production difficulty (which should be reflected in the %SS). This hypothesis was confirmed; it appears that participants who showed a higher percentage of stuttered syllables/dysfluent events scored lower on the AEN_NWR. These findings are consistent with the results by Howell et al. (2017) who showed a relationship between %SS and UNWR scores. Thus, this provides support of the AEN_NWR as an established measure of phonological skills for participants with stuttering symptoms. The high correlation with the %SS could be interpreted as an indication that the test has a high potential for identifying preschool children with speech disfluency. However, a conclusion cannot be made until the current results are compared with results of a control group of PWNS. This then would ensure that the AEN_NWR is a sensitive marker of fluency difficulty. WWRs were not included in the spontaneous speech samples as they are indicators of WFD and not a fluency difficulty (Howell et al., 2017).

The set of stimuli in the AEN_NWR conforms to accepted standards for NWR tasks including following language-specific phonotactic constraints (Arabic and English), avoiding later-developing consonants, and minimizing potential resemblance between real words and nonwords. Furthermore, the test does not require knowledge of lexical semantics for either of the two languages.

4.4.2 Evaluation of A_FS

This chapter also evaluated the Arabic speech-based assessment (A_FS) that was built on phonological principles of Arabic language properties. The aim was to investigate whether the newly designed instrument is an effective indicator of distinct disfluency types in spontaneous speech. Thus, A secondary hypothesis in this chapter examined the relationship between AEN_NWR scores and performance on A_FS. This hypothesis was confirmed; A_FS correlated negatively with AEN_NWR scores. This provides further confirmation on using A_FS as a framework that can contribute to effective assessment of stuttering in Arabic. Whilst AEN_NWR, in particular, has not yet been standardized to a large sample of Arabic and English speakers, for it to be used in external validation to compare A_FS against, it is based on the same phonologically informed approach used with UNWR.

The latter has been standardized in a large number of participants (n=96) and appeared to be a sensitive marker of fluency difficulty

It was interesting to see that the correlation between AEN_NWR and %SS outperforms the correlation between the AEN_NWR and A_FS. This could be attributed to the strict and clear guidelines in the SSI-4 manual, which was rigorously translated and adhered to in this thesis. That is, the SSI-4 manual provides a precise way to measure stuttering (Howell et al., 2013). As highlighted earlier in this thesis, only three types of speech symptom are counted are stutters. These are clearly defined to minimize any disagreement on which events should be counted; unlike other stuttering assessment scheme. Utterances such as rephrasing, phrases repetition and pausing are not counted as stuttering. More details on this are given in the next section.

4.4.3 Notes on guidelines to administering fluency tests in Arabic

Although SSI-data (as used in this chapter) were based on a translated version from English to Arabic; a commendable feature of SSI is that it provides clear description of events that are counted as stutters and what should not be counted as stutters. This was followed in this study as described earlier in the methods section. To the best of our knowledge, there is no published translated version of SSI to Arabic. Therefore, we translated the instructions and applied them to the extracts of speech samples we obtained. Out of the three components in SSI, only the %SS was used in data processing and analysis. Support for using this component comes from Mirawdeli and Howell's (2016) work that showed that it is not necessary to include all three components. The %SS has a similar performance to when all the three components of SSI are included. Hence, having these guidelines have probably aided in generating a translated SSI that can identify fluency in Arabic; although this measure still does not take into account the unique features and replicating the study with a larger sample may give different result. A_FS showed a high correlation with AEN_NWR; although this is a version that was tested for the first time and with a small number of participants. Thus, at this stage, it can act as a starting point to establish a viable instrument for assessing fluency in Arabic. The two algorithms proposed for counting syllables and disfluencies took into account the rich morphology of Arabic.

This was demonstrated by separating words into one of the three lexical categories: (1) function words; (2) content word; or (3) the combined function-content word category and then assigning weight to each lexical category. In the same way, words were assigned different weight depending on whether or not they contained a fricative sound. The idea was to mitigate, as much as possible, the effect of inflection when evaluating stuttering severity. It is important to note here that the main concern over the proposed scoring scheme was the arbitrary basis of weighting the parameters. While the AIPC provide a metric of phonetic complexity where some of the parameters here are used, it is an "all or nothing" approach (Howell, Au - Yeung, Yaruss & Eldridge, 2006). Furthermore, there is really no theoretical basis by which, for example, a geminated consonant would weigh 50% different than the same consonant when it is a singleton. The rationale behind this scoring was that words that have difficult phonemes or geminated consonants, and consequently are expected to impose difficulty, should weigh less to account for the difficulty in articulation due to phonological complexity. Additionally, there are differences in difficulty between radicals and empathic consonants and this is not accounted for in this model. In other words, individual factors may have different weighting and the level of phonetic difficulty need to be taken into account (Howell et al., 2013). This can be justified by the fact that this study is the first piece of work directly linking stuttering research with current phonological thinking in Arabic. Thus, it takes a step in moving away from arbitrary decision making to a more scientific investigation on stuttering. This should eventually lead to establishing a reliable and valid measure of fluency in Arabic. Backing up assumptions with more data must continue in order to allow for enhancing a rigorous valid instrument that can be used efficiently to assess fluency in speech for Arabic children at the age of (4+).

All in all, we now have two measures to assess fluency in Arabic; (1) the A_FS, which is a speech-based scheme to assess fluency that can provide details on symptoms that affect speech disfluency and (2) the AEN_NWR, which is broadly appropriate for the stuttering population tested. The design consideration behind the AEN_NWR involved eliminating any influences that language history could have on an individual's performance. The proposed set of stimuli conforms to accepted standards for NWR tasks including the following: language-specific phonotactic constraints (Arabic and English), avoiding later-developing consonants, and minimizing potential resemblance between real words and non-words.

4.4.4 Examination of developmental changes in disfluency symptoms

The classification of words into three types allowed examining the existence of an age-related shift in stuttering pattern. For the children group, there was a higher percentage of disfluencies on function words; whilst in the older group, disfluency decreased on function words and increased on function-content words. These findings are consistent with previous studies in English that showed that children stutter more on function words and adults on content words (e.g., Au-Yeung et al., 1998; Dworzynski et al., 2003; Dworzynski & Howell, 2004; Howell et al., 1999). Comparisons in English were made between function and content words as they form the two major lexical classes that differ in phonological complexity. EXPLAN theory offers an explanation as to why the pattern of disfluency change from function to content words for PWS who persist in stuttering during adulthood. Due to the complex phonological characteristics on content words, their speech planning takes longer compared to function words. What is worth noting here is that in Arabic, function-content words are phonologically more complex than function words and content words; given that they combine words from both categories into single complex words. This is one of the principal ways in which Arabic and English differ in a way that could affect stuttering. Work by Alqhazo and Al-Dennawi (2018) on Jordanian CWS investigated the influence of inflectional endings on the rate of stuttering. The results showed that CWS had significantly higher stuttering in words with inflectional endings. The findings were interpreted such that the extension of the word by adding inflectional suffixes changes the phonology of the word; and this change and makes it harder to be pronounce those words. As described earlier, the function-content lexical category is a dynamic one.

There are cases where two function words are attached to a content word: either two in initial position or one in initial position and one in final positions. For example, the content word (school مدرسة) can be attached to two prepositions to form one word that translates to (in the school بالمدرسة). The content word (glass كاس) can be attached to two function words; one proposition at the initial position and one at the final position to form the feminine plural suffix ات (by glassesبكاسات). There are also cases where a content word is attached to three function words; two at the initial word and one in the final word (e.g. for the teachers للمدرسات). In this case, the word is a function-content-function word. In this study, we limited the lexical categories to three types only; although it might be interesting to examine the impact of increase in phonological complexity as in the function-content-function word type on stuttering in future studies. Additionally, linguistic factors that are confounded with content and function words would be interesting to examine. Examples of lexical factors include word frequency (i.e. the number of times a certain word occurs in a language) and neighbourhood frequency (i.e., all other words that can be generated by changing one letter from the target word). Anderson (2007) showed that both lexical factors may influence the susceptibility of words to be stuttered by preschool CWS. Those factors may be accounted for in future studies by possibly considering them as covariates when conducting statistical analysis. Finally, studying each lexical category and even different lexical items (e.g., adverbs, nouns: derivation, compounding) have specific inflectional structures that is worth examining.

4.5 Conclusion, Limitation and Future Direction

This chapter makes a novel contribution to the subject area of disfluency as an aspect of language disorder, and to screening procedures of speech fluency. While the current data provide promising results, there are some limitations that should be emphasized here. First, all participants in this study have been diagnosed as displaying stuttering symptoms; however, different SLPs may use different guidelines to diagnose stuttering due to the lack of norms for assessing fluency in Arabic. The speech samples obtained in this study were in Arabic as the aim was to develop and standardize an instrument to assess fluency for Arabic preschool children; but for analysis of speech symptoms, the guidelines of SSI were followed with respect to what is counted as a stuttering symptom. Although a strong correlation was found between the %SS and the AEN_NWR, which might suggest that the guidelines can be generalized to Arabic, having guidelines designed specifically to Arabic could result in even a stronger correlation coefficient. To explain in more detail, the data used to evaluate SSI-4 statistically were collected from English speakers who stuttered, and it used English passages that had to be read as well as others that were spoken spontaneously. Consequently, the norms do not apply to the Arabic version as the standardization has not been conducted. Moreover, the test is not appropriate for assessment of Arabic in terms of the procedures used in SSI for counting the number of production units (i.e. syllables for English) as well as the specifications of stuttered events. Thus, in parallel to standardizing the AEN_NWR, future work should consider further improvement of the A_FS and turning it into a manual for assessing spontaneous speech sample in Arabic.

The second limitation involves the number of participants. Although the sample size is reasonable compared to other studies that examined the non-word repetition abilities in PWS (e.g., Hakim and Ratner (2004) had eight CWS and Sugathan and Maruthy (2020) had 17 CWS), it is desirable to replicate the results of this preliminary study using larger number of participants. This is particularly the case because participants in this study were from different age groups and there may be developmental differences between them. Having more participants could also permit assessing phonological performance at every non-word length to examine the relationship between repetition accuracy and phonological complexity. Several questions could be asked then concerning the effect of non-word length affects on the accuracy of performance in CWS. Finally, it would be of interest to examine matched groups of: (1) controls of Arabic speakers who do not stutter; (2) English speakers who stutter, and examine their performance on the AEN_NWR. Arabic speakers of dialects other than Saudi could also be examined; this would allow investigating the influence of dialect on AEN_NWR performance. Testing PWNS would also permit collecting further support on the fact that WWRs are indicator of WFD rather than speech disfluency.

With respect to A_FS estimates, similar to what was reported by Dworzynski and Howell (2004) when investigating phonetic complexity and stuttering in German language, the general scoring scheme of A_FS was based on AIPC factors that are not really independent of each other. Consequently, this makes it hard to isolate the unique effects of some factors as they may correlate with each other. For instance, in a word like /ʔas.saaʕa / ‘clock, there are four factors in this phonological unit that are contributing to the phonological complexity; 1) consonant by place of articulation (i.e. the radical /ʕ/), 2) consonant by manner of articulation (i.e. the fricatives /s/), 3) consonant by length because the word has a geminated consonant and 4) word length because the word has three syllables. Thus, results should be interpreted cautiously because the effect of a certain factor might not be high due to close interrelation with other factors. Additionally, the A_FS proposed algorithm was based on limited data collected in Chapter 3. This should be accounted for in future analyses by increasing the number of participants. With only ten speakers' data examined, it is unlikely that these results reflect characteristics of the entire speaking population.

Chapter 5 Treating Word-finding Difficulty in Arabic Children with English as an Additional Language

"In most people there is a settled place they speak from; in me it remains unsettled", *John Updike, Getting the words out*

5.1 Introduction

So far, we have outlined the AEN_NWR as an assessment tool that can potentially identify WFD from speech disfluency in preschool children whose' first language is Arabic or English. The question that may arise is when the two communication disorders are identified, what can be done to improve word-finding and speech fluency. Children who show stuttering symptoms need to be referred to SLT for full evaluation and intervention. There are also procedures involve training working memory to enhance fluency difficulty; these can be delivered in schools and should not preclude intervention administered by SLTs. For example, Howell et al., (2020) addressed disfluency using WM training. 232 reception class children from five primary schools were assessed by obtaining measures of their %SS and %WWR. 12 were at high-risk of fluency difficulty and received WM training over two weeks. The results showed marked improvements; children's %SS dropped from pre-test to post-test and these improvements lasted for at least a week after the intervention. WFD can also be addressed by giving them phonological or semantic training. In this study, the focus is on examining the impact of phonological training using specially designed material for treating WFD in Arabic children with EAL. As background, previous work into the nature of WFD is reviewed.

Language difficulties may result in significant and ongoing problems for children. However, this is not easy to rectify because, due to differing models of language and development, and the heterogeneous nature of language disorders, the study of language deficits challenge researchers and practitioners. To address the heterogeneity issue, researchers in different studies identified subgroups of children with specific types of language difficulties. Over the years, studies have generally examined lexical deficits, semantic deficits, phonological deficits and verbal dyspraxia (Conti-Ramsden & Botting, 1999). Previous research has shown the importance of studying the word-retrieval system in children; highlighting how it plays a central role in language development and cognitive processes that will be later used in the reading system (Wolf, 1991). However, there is little research on lexical access and its association with WFD. A population survey estimated the prevalence of WFD in children with speech and language difficulties in London and showed that 23% of those experienced WFD (Dockrell et al., 1998). Children in the study had received services from speech and language therapists, but no information was given on their language (i.e. whether they were children with English only (EO) or English as an additional language (EAL)). Hence, the identification of children at-risk for WFD is an important task for educators and clinicians.

5.1.1 Nature of WFD in children

The foundation of word-finding difficulty can be established by Wolf's (1980) model which included the following four stages: (1) stimulus perception; (2) recognizing and categorizing the perceptual information to a concept; (3) lexicon; searching for a level to give to the concept; and (4) motor system; which is the stage where the stimulus is produced. The cause of WFD is unclear; and there are several reasons why it arises within the language production system. As discussed in Chapter 1, EXPLAN theory accounted for two different ways in which fluency failures can affect normally fluent speech and make it stuttered. The work in EXPLAN separated WWR (an indication of WFD) from other stuttering symptoms (i.e., part-word repetition, prolongation and breaks). The last three symptoms have been agreed upon as characteristics of disfluent speech by many authors (e.g., Riley (1994)). WFD, on the other hand is often characterized by whole word repetition on function words and some content words that interrupt speech. Such disruption to fluent speech is often experienced by children with EAL (Howell et al., 2017). Children respond to difficulty in finding words usually by repeating function words or pausing prior to a content word that they are not able to retrieve. This speech interruption leads to a delay and gives children time to find an alternative way to deliver their speech Howell et al. (2017).

Previous research has generally focused on phonological and semantic deficiencies to explain the characteristics of WFD in children. The clinical pattern of WFD involves an inability to find the appropriate word and using alternative word forms to compensate for the difficulty in retrieving the target word (Messer & Dockrell 2006). In their analysis of children's word finding skills in discourse, German and Simon (1991) included WWR as a characteristic of WFD as well as the following ones: reformulation (i.e. replacing or changing a word); empty words which are words that do not add content or specificity like the word "well"; substitution by changing the target word for another word because the target is not retrievable such as saying tall instead of high; insertions which are words or phrases that the child adds to comment on the speech itself like saying : "He ate the … I don't know the name"; delays which are pauses that last about 6 seconds or more without vocalization; and time fillers which is a vocalization of vowel sound like uh, um that the child produces while retrieving the target word.

WFD can also occur in the speech of children with child onset fluency disorder. The fact that WWR occur in the speech of children who stutter and those with WFD might raise a problem in treatment as children with WFD but not fluency issues might be misdiagnosed as having stuttering and thus they would receive an inappropriate intervention. As an example, a child might exhibit a WWR and starts repeating the same word due to inability to retrieve the following word that does not exist in that child's vocabulary. In that case, the child is exhibiting a symptom of WFD and not stuttering. Thus, speech disfluencies and WFDs are thought to be independent of each other (Howell, 2013). The UNWR (Howell et al., 2017) that was introduced in Chapter 1 aims to distinguish these two groups: children with WFD and children with speech disfluency. Howell et al.'s (2017) procedure involve analysing spontaneous speech samples to obtain %SS (Riley 2009). Children whose scores are above a 3%SS criterion alone are designated as having fluency difficulty. %WWR is calculated by counting all repetitions of monosyllabic words; and then estimating the percentage out of all syllables. Any children not stuttering (%SS less than 3%) who have high rates of %WWR have WFD. This procedure provides an exclusive indicator of WFD (Howell et al., 2020).

5.1.2 WFD in children with EAL

Mirawdeli (2016) reported that children with EAL display WFD which can sometimes be mistaken for disfluency as children use WWR or other hesitation phenomena when they cannot find a word. Thus, we want to highlight that WFD occurs in the speech of children using an additional language, whether it is English or another additional language that is not their native one. In this section, previous studies that show how EAL children or bilinguals in general show WWR symptoms in their speech are reviewed.

Rad and Abdullah (2012) looked at the effect of context on hesitation strategies by EAL students whose first language was Iranian. Participants were university students at a public university in Malaysia and they all had the same level of language competence (IELTS scores of 5.5). It was found that background knowledge can affect hesitation. When speakers talk about themselves, the rate of hesitation is low because the speaker already has sufficient knowledge about the speech and possibly has many words in the lexicon to select from. When answering an unpredictable question, however, hesitation becomes more frequent because the speaker gets involved in the process of utterance planning and also, might not have adequate vocabulary in memory to draw from. Speakers of EAL require more time to plan their utterance due to their minimal experience with the additional language and they show less automatization in their additional language compared to their native language (Rad & Abdullah, 2012). A further study was conducted by Bada (2010) where he explored repetition in additional languages (English and French) by native Turkish speakers. A main emphasis of the study was investigating what element(s) of the target language a speaker might repeat.

Participants were recruited from English and French language teaching departments at Cukurova University. Their ages ranged from 19 to 21 years and they were all in their first year of studying an additional language. Repetitions were made in the speech of language learners irrespective of types of syntactic /lexical elements at a word or sentence level. The author also reported that repetition occurs in the speech of native speakers of different languages including German and French, which support the claim that repetition is a universal strategy that is not specific to one language. Participants may have produced WWR in their speech to gain time to produce the new utterance. For example: "Simba is a very..very pretty animal ". The age of participants in the two studies mentioned above is different from the samples in this study and it is possible that there may be some developmental changes; however, they still shed light on the nature of WWR that occur when the speaker cannot find a word in their lexicon. A number of studies have investigated this phenomenon in children of early school age. Differences in lexical access between children with EO and children with EAL were investigated by Yan and Nicoladis (2009). The first group of participants were 20 French children with EAL, and they ranged in age between 7;0 and 10;1. The control group had 25 children with EO and in the same age range of the experimental group. The results showed that children with EAL appeared to have difficulty accessing words compared to children who have EO. The results of this study are consistent with what have been reported in other studies, that children with EAL exhibit more WFD. Additionally, Howell et.al (2017) studied WFD in 4-5 years old children who attended UK schools and had diverse language backgrounds. The authors reported that children with EAL can have WFD in English that result in symptoms like WWR. Consistent with what was reported in other studies, this study also showed that children show word repetition to gain more time to use alternative ways when the target word cannot be easily retrieved.

5.1.3 Review of work on WFD intervention

Despite the negative consequences of WFD, there is a scarcity of well-controlled intervention studies for preschool children. Moreover, the available studies are inconsistent in their methodologies, including participant numbers, intervention intensity and its duration. All of this make it challenging to compare these studies or draw general conclusions. Furthermore, the current WFD interventions are not sensitive to children's specific demands in our work (heterogeneous language background) due to the materials being language-specific. The majority of research focuses on monolingual English speakers; and there is a need for interventions for children with different languages profiles (Ebbels, 2014). Current and novel treatment procedures for WFD must be rigorously designed to direct treatment practices for this population.

Recently, Best el., (2018) carried out a randomized control trial (n=20) study to demonstrate the effect of a WFD intervention with children with WFD in schools. The study compared phonological and semantic interventions and children were assessed three times before and once after the intervention. The intervention was carried out over six weeks and employed a word-web protocol where children were encouraged to generate semantic or phonological features of words. It was found to be effective in improving retrieval of treated items. Children in the experimental group gained on average four times as many items as the control group. This was a rigorously designed study that employed a clinically realistic intervention in terms of intensity and duration. Another important aspect of the study was that it took place at a mainstream primary school where WFD is a common problem. It should also be noted that this study has not targeted children with EAL. Children were English speakers; who either have been exposed to English at home from birth, or have been in an English speaking nursery at the age of 3 and continued to be exposed to English after that at home. Moreover, Best et al. (2018) pointed out a critical point concerning factors that may have affected the effectiveness. That is, the child background could have an influence on WFD which might affect their lexical retrieval; which in turn is likely to change over time. This reinforces the idea that WFD is a vocabulary problem (Howell et al., 2017) and that it should be treated separately from stuttering.

It is important to emphasize that there are only few studies on treating WFD, and the results of these studies differ with respect to the appropriateness of phonological vs semantic training (Wright et al., 1993). This indicates that caution must be exercised before selecting the most suitable treatment. Further support on the equivocal results come from work by Bragard et al. (2012) who stated that the issue about the effectiveness for semantic vs phonological intervention is open to debate. It is possible that both types of intervention are needed; or it could be that one type of treatment outperforms the other. In Wright's et al. (1993) study, eight children received semantic training and seven children received phonological training as a WFD intervention. It was found that the phonological treatment group made significant improvement post intervention in naming untrained pictures, whilst the semantic group did not. This study focused only on phonological training to treat WFD. This allowed to test whether phonological training is effective; and whether or not semantic training may also be needed.

5.1.4 The current study: Aim and Hypothesis

The aim of this experiment was to treat WFD in Arabic speaking reception class children with EAL. The study took place in mainstream schools, where WFD is a common issue, whilst there could be no child who exhibit stuttering symptoms. This study was based on previous work conducted by Howell and Sorger (2017) that showed educational impact of an intervention that employed semantic training tailored to specific language groups (English, Polish and Urdu). The current study aimed to extend Howell and Sorger's (2017) study by evaluating the effectiveness of a language-specific phonological training intervention for improving WFD in Arabic children with EAL. It is worth noting here that the screening AEN_NWR provides easy phonological materials for Arabic and English children with WFD as they are constructed using shared phonological properties between the two languages. On the other hand, the training NWR stimuli that are presented in this chapter are considered difficult because they include all of the English-specific features that were excluded when generating the AEN_NWR. Hence, the aim was to use the difficult nonwords in a phonological training instrument using the priming effect. The rationale behind studying phonological priming in typically developing children is that if the nonword prime reduces reaction time, then the prime processing must have activated phonological representations that are required for naming pictures

(Wilshire & Saffran, 2005). The priming effects are revealed when responses to the pictures are faster and more accurate when the pictures were preceded by phonologically corresponding nonword primes (e.g. the nonword prime kɛd preceded the target word bed).

The goal of this NWR was to improve WFD in Arabic children with EAL and then to determine whether this training transfers to real English words with similar sound structure as the non-word primes. As mentioned, the non-words are intended to be used to train on English phonemic sequences that are exercised infrequently, or probably not at all in Arabic. It was hypothesised that phonological training with such material should prime a child's use of unfamiliar phonological networks and improve lexical access. The non-words serve as primes and by voluntary repeating them, this activates internal nodes so that they become readily available for the processing of a related target (Glaser, 1992). In the next section, a description is given about the phonology of both languages highlighting the phonological differences and similarities between the two languages that lead to creating the training stimuli.

5.2 Phonological Training Materials

The non-words were specifically designed to include patterns not used in Arabic but are required for English. As will be highlighted in the methods section, the Arabic children (Experimental group) had been exposed to English in their school, but it was still not their primary language. Thus, they had less experience with the English phonological structures used to name the objects in the picture naming task. It was expected that intervention involving repeated exposure to the unfamiliar sequences would increase familiarity with such structures and improve performance (i.e., reduced latency in picture naming). The non-words include non-words for Arabic children with EAL. These materials incorporate a combination of consonant and vowel characteristics (monophthongs and diphthongs) in addition to consonant clusters that occur word-initially. Those were chosen because they represent patterns present in English that are not found in Arabic and will therefore be challenging for an Arabic-speaking child with EAL. The words were ramped up in difficulty over six 'lists'. The first two lists started with the least phonologically complex materials (i.e. vowels and diphthongs). All new vowels and diphthongs appeared in monosyllabic non-words to focus the child's attention on that phoneme as the target of the training and thus should facilitate pronunciation. The third list included new consonant phonemes that appear as the first consonant of each non-word; also to focus the child's attention on the target of the training. Training then extended to words with consonant clusters word-initially and then includes materials with complex clusters word-initially and word-finally. One way in which English is different than Arabic is in the structure of the syllables; English allows many more consonant clusters word-initially. This can present an articulatory challenge for Arabic children with EAL.

1. Vowel Materials:

Arabic has three short monophthong vowels /a/, /i/, /u/, and three counterpart long monophthong vowels /æ/,/i:/,/u:/. Long vowels contrast phonemically with their short vowel counterparts. For example, the two words /ka.tab/ 'He wrote' and /kæ.tab/ 'He messaged', contrast in that the vowel in the first word is short while the one in the second word is long. The words are contrastive as the two forms make two different words with two different meanings and thus, the vowel length in Arabic is phonemic. The English vowel system, however, contrasts long vowels and lax short vowels and the number of vowels in English is 12 although this varies depending on the accent or dialect. The distinction between specific English vowels, especially ones that do not have an equivalent or near equivalent is problematic for an Arabic child with EAL. For example, distinguishing between certain pairs of vowels like /ɪ/ and /ɛ/ in words like bit and bet will be challenging to Arabic children with EAL because of the lack of knowledge about the distinction between these two phonemes. The three vowels /ɛ /, /ʌ/ , /ɔ/ and the long vowel /ɑ/, that exist in English but do not have equivalent or near equivalents in Arabic, were selected because an Arabic child might make subtle changes in the duration of those vowels because of the unfamiliarity of that phoneme. These vowels, along with their lexical sets as described by Wells (1982) are given in Table 5.1 below.

Table 5.1: English vowels as described by Wells (1982)

Symbol	Definition	Lexical set
/ɛ/	Front mid unrounded lax	DRESS
/ʌ/	Central mid unrounded lax	STRUT
/ɔ/	Back mid rounded lax	LOT
/ɑ/	back low unrounded	START[8]

The following list of non-words in Table 5.2 were designed to train on the English vowels described above. Their corresponding real words are also given in Table 5.2. For simplicity, all the non-words are monosyllabic CVC and all consonants in the onset and coda exist in Arabic and English. The corresponding set of non-words represent real high frequency objects.

Table 5.2: List 1 of non-word training materials (the short vowels /ɛ /, /ʌ/, /ɔ/ and the long vowel /ɑ/)

Non-word Material		Real-word Material	
Orthographic Transcription	**IPA Transcription**	**Orthographic Transcription**	**IPA Transcription**
Ked	kɛd	Bed	bɛd
Heas	hɛs	Head	hɛd
Zuck	zʌk	Duck	dʌk
Shum	ʃʌm	Sun	sʌn
Nup	nʌp	Cup	kʌp
Sog	sɔg	Dog	dɔg
Tar	tɑ:	Car	kɑ:

2. Diphthong Materials:

Arabic has two diphthongs formed by vowels that are immediately followed by a glide in the same syllable. These two diphthongs were historically coalesced in dialects like Cairene and

[8] Materials are generated following the non-rhotic English accent because this part of training took place in London.

Sudanese (Watson, 2002). English has a phonemic inventory of the following seven diphthongs /ɪj/, /Ɛj/, /ɑj/,/oj/, /aw/,/əw/ and /ʉw/. This is another area where a child is expected to find it challenging to pronounce a word that has any of those diphthongs that do not occur in Arabic. The list of non-words and words in Table 5.3 thus focus on this aspect and include five diphthongs. For simplicity, these non-words are monosyllabic and only include a simple onset and word-final consonants that are found in Arabic and English.

Table 5.3: List 2 of the non-word training materials (Diphthongs)

Non-word Material		**Real-word Material**	
Orthographic Transcription	**IPA Transcription**	**Orthographic Transcription**	**IPA Transcription**
Doy	doj	Toy	toj
Noy	noj	Boy	toj
Fow	faw	Cow	kaw
mion	majən	Lion	lajən
moil	mojl	Oil	ɔjl
Naib	nƐjb	Face	fƐjs
Boam	bəwm	Home	həwm
Boud	bəwd	Bone	bəwn
Fouse	faws	House	haws
Hoat	həwt	Boat	bəwt
Bouth	bawθ	Mouth	mawθ

3. Consonant Phoneme Materials.

The following two lists of non-words and words were designed specifically to include consonants that are found in English but not in Arabic. These include /p/, /v/, /dʒ/, /g/, /tʃ/, /ŋ/, where some occur word initially and some occur word-medially and word-finally as follows. The first list of non-words (Table 5.4) begin with /p/, /v/, /dʒ/, /g/, /tʃ/ with all of them occurring word initially to focus the child's attention on the unfamiliar consonants. The phoneme /g/ was included although it occurs in Arabic but only as a replacement to the Arabic phoneme /q/. For example, in a word like /**qalam**/ 'pen', it will usually be pronounced /**galam**/ in most Saudi Arabic dialects.

Table 5.4: List 3_A of the non-word training materials (consonants word initially)

Non-word Material		**Real-word Material**	
Orthographic Transcription	**IPA Transcription**	**Orthographic Transcription**	**IPA Transcription**
pem	pɛm	Pen	pɛn
vad	vad	Van	van
vush	vʌʃ	Vase	vɑːz
jass	ʤas	Jam	ʤam
gub	gʌb	Gun	gʌn
goot	gʉ́wt	Goose	gʉ́ws

Table 5.5 includes non-words that end with /ŋ/ as this consonant phoneme does not exist in Arabic (it exists in some MSA words as an allophone of the phoneme /n/). In English, however, /n/ and /ŋ/ are two different phonemes and /n/ can occur in the middle of a word or word finally. This explains why a child with EAL finds it challenging to pronounce /ŋ/ because of unfamiliarity with this articulation. A child usually adds /g/ to a word that ends with /ŋ/. Table 5 displays two non-words, and their corresponding real words were designed with this influence in mind. Care was taken to include only CVC nonwords and selecting consonants that exist in Arabic to focus the child's attention on the target of the training.

Table 5.5: List 3_B of the non-word training materials (The consonant /ŋ/)

Non-word Material		Real-word Material	
Orthographic Transcription	**IPA Transcription**	**Orthographic Transcription**	**IPA Transcription**
Ding	dɪŋ	Ring	rɪŋ
Ling	lɪŋ	King	kɪŋ

The structure of the training materials then continues to advance in phonological complexity. The non-words above are considered simple because monosyllabic non-words were mostly used. Also, training targeted vowels, diphthongs and consonant phonemes that exist in English but not in Arabic. All non-words and their corresponding real words included only one phoneme that was the target of the training. Care was taken that all other phonemes are ones that exist in the child's phonemic inventory.

4. *CC cluster onset with consonants other than 's' in word initial consonant cluster.*

There has been a major controversy about initial consonant clusters in spoken Arabic across different dialects. It is agreed that MSA does not allow word-initial consonant clusters, therefore the syllable template CCVC does not occur in Arabic. As mentioned earlier in this chapter, however, MSA is used in formal schooling and in the media and a child is brought up listening to MSA on TV or other electronic devices. Thus, a child might have already been exposed to different patterns of legal consonant clusters by hearing Arabic in dialects that do allow consonant clusters like the Saudi Najdi for example. It is important to note that these might not be an underlying clusters; but rather a surface cluster resulting from a tendency to simplify clusters in normal speech (Algeo, 1978). The question remains as to whether a child is able to transfer the ability to pronounce consonant clusters to words in the child's additional language. A study by Elmahdi and Khan (2015) investigated pronunciation problems that face Saudi pupils as they are learning EAL in secondary schools. The results of their study demonstrated that the EAL students were unintentionally inserting a vowel sound in the words onsets of different English syllables, and it was interpreted that the reason for de-clustering is the influence of the participant's first language. This could be attributed to the fact that while in English, word-initial consonant clusters form a complex onset, the sequence of two consonants in Arabic word-initially can be an empty onset that contains a null vowel. As proposed in Kaye's (1990) analysis of Moroccan Arabic, an underlying empty nucleus must be realized in word-initial consonant clusters, which might be expressed phonetically sometimes when it occurs in such position. Also, as mentioned when creating the Arabic- English UNWR, there are certain dialects that do not allow consonant clusters word-initially as in English. For the reasons above, it seems justifiable to train an Arabic speaker with EAL on pronouncing consonant clusters when they occur word initially. Table 5.6 shows non-words that begin with a consonant cluster other than 's', along with their corresponding real words.

Table 5.6: CC cluster onset with initial consonant other than 's' (List 4)

Non-word Material		Real-word Material	
Orthographic Transcription	**IPA Transcription**	**Orthographic Transcription**	**IPA Transcription**
Trein	tríjn	Tree	t r íj
Flam	flám	Flag	f lág
Blum	blʌ́m	Blue	blʉ́w
Bloob	blʌ́b	Blood	b lʌ́d

The following list of nonwords was separated from the list above although they both have CC clusters because the first consonant here is 's' and this phoneme has a different sonority profile from other consonant phonemes (Harris, 1990). The list has non-words that begin with /s/ + a consonant that is in the Arabic inventory. The list also includes non-words that begin with s followed by a consonant that is not in the Arabic inventory and was targeted in the training materials above. The list also includes words with a consonant cluster and diphthong (e.g. snow) because diphthongs were targeted earlier in training. These non-words and their corresponding real words are given in Table 5.7

Table 5.7: CC cluster onset with 's' word initially (List 5)

Non-word material		Real-word material	
Orthographic Transcription	**IPA Transcription**	**Orthographic Transcription**	**IPA Transcription**
Stis	stɪ́s	star	stá:
Snid	snɪ́d	snow	snə́w
Spoos	spʉ́ws	spoon	spʉ́wn

5. Two consonant clusters word-initially and word-finally

Throughout the intervention and as the child develops in fluency, the non-words increase in phonological complexity. The set of non-words in Table 5.8 includes two consonants in the onset and word finally. Generally, there are minimal restrictions on the two consonants that occur word-finally in Arabic so a child should not face difficulty pronouncing them.

Table 5.8: (List 6) Two consonant clusters word-initially and word-finally

Non-word Material		**Real-word Material**	
Orthographic Transcription	**IPA Transcription**	**Orthographic Transcription**	**IPA Transcription**
Kridge	kríʤ	Bridge	bríʤ
Klidge	klíʤ	Fridge	fríʤ
Kriend	krέnd	Friend	frέnd

5.3 Method

5.3.1 Participants

Sixty-three reception class children (31 males, 32 females, mean age 5;1 at baseline assessment in September 2019) from two schools; one in Riyadh (N=33) and one in London (N=30) were tested. Children in Riyadh (referred to as the experimental group from now on) spoke Arabic predominantly but were exposed to English as an additional language in their school. The school in which they went required English as the primary form of communication between teachers, staff and the students. Children in the London school (control group) spoke English predominantly, although some had exposure to French through one of their parents. Testing took place at different times throughout the day in both schools from September 2019 until March 2020. The age of the children ranged from 4;9 to 6;1 at baseline assessment. The attrition rate for the experiment amounted to 3 children who only took part in the pre-assessment and were absent at all other assessments. These children's individual data were analysed in order to give feedback about the children's speech to their teachers, however, they were excluded from the statistical group analyses. All of the children except for one had normal hearing. Language history information was obtained from the children and verified by teachers to assay fluency in each language.

Children's language, literacy and phonological performance were assessed individually at pre-, post-, and follow up assessments. Therefore, this was a within-subject (repeated measures) design, since all children included in the analyses completed the entire set of tasks at all stages. However, the order in which the tasks were carried out was randomized across participants. Language skills were assessed using three tasks: a speech production task, a speech comprehension task and a picture-naming task. Phonological performance was measured using the AEN_NWR task. In total, there were five individual tasks that fell into three categories at the three assessment stages. The independent variables (IVs) included the assessment stages (3 levels: pre/post/follow-up), and the children group (experimental vs control).

5.3.2 Procedure

All tasks were administered individually. Children were taken out of class to a separate quiet room for the individual assessments. The testing in Riyadh took place throughout the weekdays, whereas children were only tested on Saturdays in London. The length of the assessment sessions varied between children, but usually took between 20-25 minutes per child. After finishing each session, every child received a sticker as an incentive. During the pre-test, all assessment tasks were administered apart from the picture-naming task. The intervention sessions were delivered one week after intervention. Intervention sessions started with the picture-naming task; followed by the intervention and then all other tasks again (post-test). Follow-up sessions were conducted one week later, and all five tasks were administered. All sessions were audio recorded using a Sennheiser SC 660 USB ML wired headset and the Audacity software for later analysis. Statistical analyses were made using IBM SPSS 26. Care was taken to ensure that the phonological training procedure did not conflict with any conventional SLT intervention children might subsequently receive, and that screening procedures were appropriate for children who speak Arabic and/or English. The intervention was designed to be efficient and easy for teachers and school staff to administer.

Tasks

Speech Production

The speech production task collected a spontaneous speech sample of at least 200 syllables. This was used to detect symptoms of speech disfluency and word-finding difficulty. Different picture materials were used to elicit speech across assessment stages. Drawings from German's (1991) TWFD manual were used. These depicted different scenes (e.g. people in the park) (see Appendix 2). At each assessment, one of the pictures was presented to the child on A4 paper, and the child was instructed to describe the picture. If the child did not respond, he or she was prompted to continue or was asked other general questions to elicit more speech; the experimenter made physical concomitant ratings (distracting sounds, facial grimaces, head movements, movements of extremities on a scale of 0 to 5 whereby 0 indicates 'none' and 5 indicates 'severe and painful looking' whilst the child was speaking (Riley, 2009) for later SSI analysis. Each child's speech was subsequently transcribed and analysed using three measures of fluency: Riley's 'Fluency Analyses for Stuttering Severity Instrument Score': the percentage of syllables indicating dysfluent symptoms such as repeated or prolonged sounds (=%SS). The percentage of syllables consisting of Whole-word repetitions (=%WWR) was calculated by subtracting the number of syllables that comprised WWRs from the total syllable count used for the %SS analysis. Then, the number of non-fluent events (multiple iterations of WWR were considered as a single non-fluent event) was divided by the adjusted total syllable count to obtain the %WWR. This provided an indication of word-finding difficulty in a similar way to which SSI indexes stuttering.

Speech Comprehension

The aim of this task was to measure a child's comprehension abilities after being narrated a story based on a picture they saw which they answered some questions about. Different material was used at each assessment stage and their presentation was randomized and counterbalanced. The pictures (printed out on A4 paper) and their corresponding stories and questions were adapted from German's TWFD (1991) and are given in Appendix 3. The experimenter instructed the child to listen carefully to a pre-recorded story about the picture that was presented as they needed to answer some questions about it afterwards. The experimenter then narrated the story and ensured that the child was paying attention. Finally, the specific questions were asked, and accuracy of answers (maximum score of 8 points) was recorded. Points were given even if the child could not name an item but pointed to the required one.

AEN_NWR

NWR measures a child's phonological skills based on how accurately they can repeat non-words of different syllable lengths. As described in Chapter2, the AEN_NWR takes the phonological constraints of Arabic and English into account to create non-words that are equally pronounceable by children from either language. The task was carried out using

MATLAB and *PsychToolbox v3* on a laptop. First, the experimenter instructed the child to carefully listen to "funny made up words" via headphones and to repeat them afterwards. After every response from the child, the experimenter would score the response by typing "1" (correct), "0" (incorrect), or "NR" (no response given) into MATLAB. Differences in vowel quality and stress pattern were not counted as errors. There were two practice items at the beginning followed by ten other non-words of two-syllable length, which were randomized in order of presentation by the software to cancel potential order effects. If a child correctly repeated a minimum of eight of these words, they progressed to the next syllable length where 10 non-words (+ two practice items) consisting of three syllables were presented, and so on. As described in Chapter 2, the difficulty of the task gradually increased from two-syllable non-words to four-syllable non-words. Each child's data was automatically saved in an excel file by MATLAB, which included the child's age, gender, first language (L1), their participant number, and 'total correct score' of all correctly repeated words. Audio recordings of the tasks were reanalysed later to check the scoring (number of accurately repeated non-words) and the latter were used in further analyses. A list of all the non-words used is given in Appendix 1.

Picture Naming task

The picture naming paradigm has been widely used in psycholinguistic research with children. The tasks involved all stages of language production and it gives the experimenter a higher degree of control (Masterson, Druks & Gallienne, 2008). In this task, children named pictures of English words with selected phonological structures that corresponded to those in the intervention materials. The aim was to assess whether phonological training can be transferred to real word objects using those pictures. Forty-eight pictures were selected from CBeebies word frequency list; a sub corpus of the SUBTLEX (Van Heuven at al., 2014) corpora. CBeebies is a British English database that was built using children's spoken language, which provides a better reflection of the language used compared to written texts (Chen & Meurers, 2018). Cbeebies has a word frequency Zip scale; a standardized measure that indicates the number of times every word has appeared in the corpus. The Zip scale is easy to understand using a logarithmic scale that goes from 1 (very low-frequency words) to 6 (very high-frequency content words). Only words with a Zip scale score higher than 4.60 were selected to minimize the number of errors due to not knowing the picture. The experimenter ran through every picture prior to the task to ensure participants named all items correctly before proceeding. A few pictures required correction across participants. Reaction time was calculated from the onset of the picture to the participant's voice onset. Pictures were presented on a computer screen with gray background situated 57 cm away from the child. The order of presentation of pictures was randomized. Different sets of pictures were used in pre-test and post-test sessions. At follow-up, pictures were randomly chosen from pre-and post-test pictures, ensuring that all phonological training targets were maintained. Psychopy software (Peirce, 2007) was used to design the experiment and display the pictures. Each picture onset was marked with an inaudible click, which was audio recorded together with the child's verbal response. The temporal gap between the onset of the click and the response was then measured manually. The pictures remained on the screen until the child provided verbal labels for them, or for up to 3000 milliseconds, whichever came first. A blank screen containing a fixation crosshatch (+) was presented between trials with an interstimulus interval of 1000 milliseconds.

Intervention

During the intervention, children repeated the English set of non-words that were designed to expose them to English phonotactic structures that are challenging for them because they represent patterns of English not found in Arabic. The stimuli were presented starting with the least phonologically complex materials (List 1) and progressed through complex clusters word initially and word finally until List 6. Children were told prior to listen to the non-words an repeat them. All stimuli were pre-recorded by a native British male speaker who was phonetically trained in English and Arabic. The aim was to provide phonological training using the created language specific non-word materials. The materials were predicted to benefit a child with EAL who speaks Arabic as well as well as a child with Arabic as an additional language who speaks English. In total, there were four individual tasks that fell into three categories (Figure 5.1) at the three assessment stages.

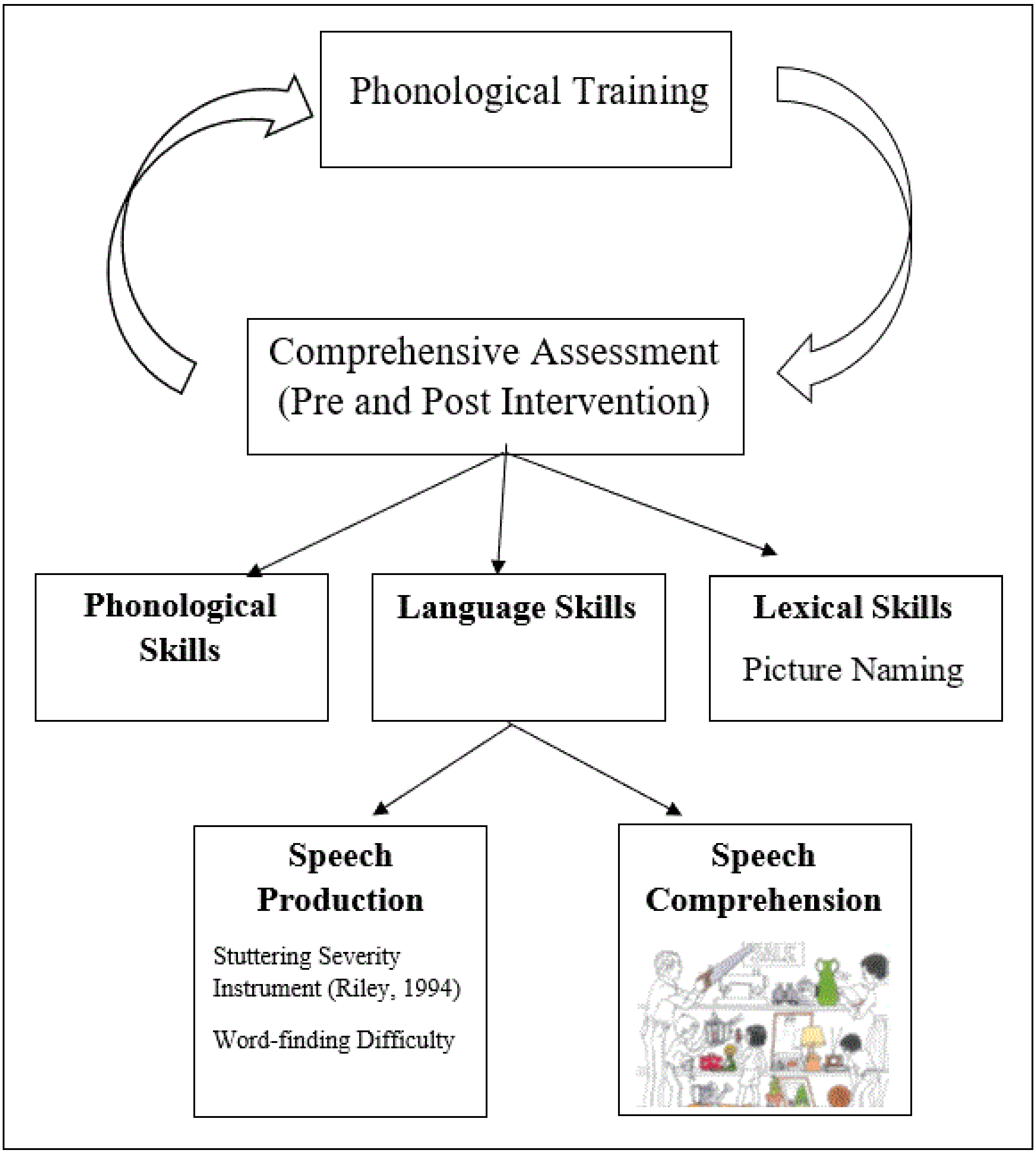

Figure 5.1 Study design outlining the different tasks at three assessment stages and the intervention.

5.4 Results.

Descriptive statistics

Descriptive statistics (means and standard deviations) are given in the figures below for all measures across the three assessment stages. Figure 5.2 displays measures of AEN_NWR scores. Speech comprehension scores are given in Figure 5.3. Fluency measures (%SS and %WWR) are given in Figure 5.4 and Figure 5.5 respectively. Reaction time measures are given in Figure 5.6.

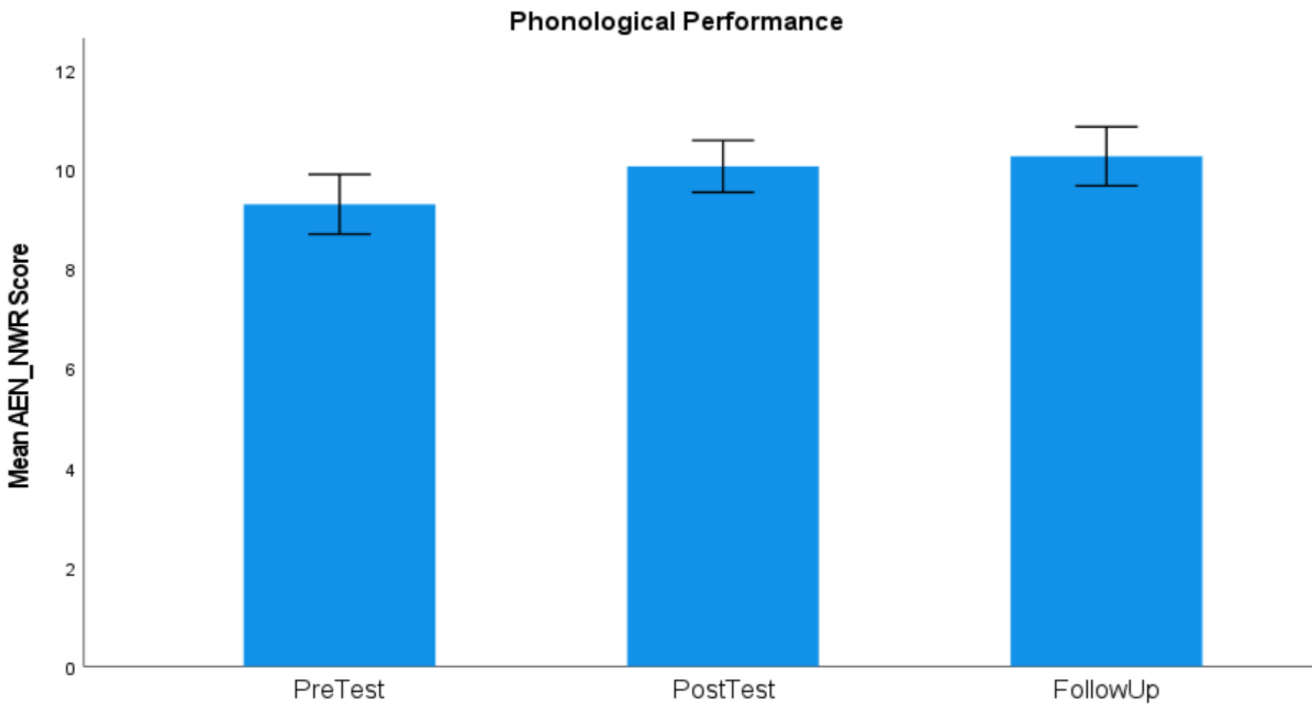

Figure 5.2 Mean scores and standard deviations on AEN_NWR task for all participants collectively across all assessment stages (Maximum score is 25)

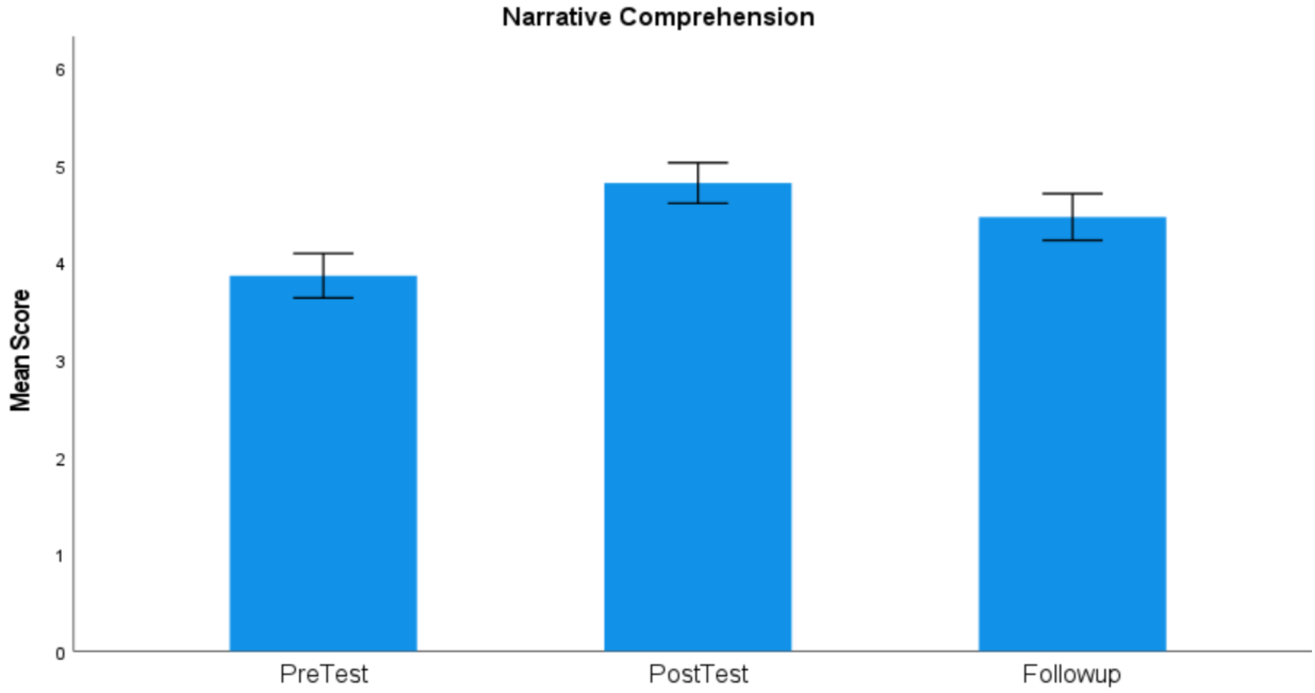

Figure 5.3: Mean scores and standard deviations for the speech comprehension task for all participants collectively across all assessment stages

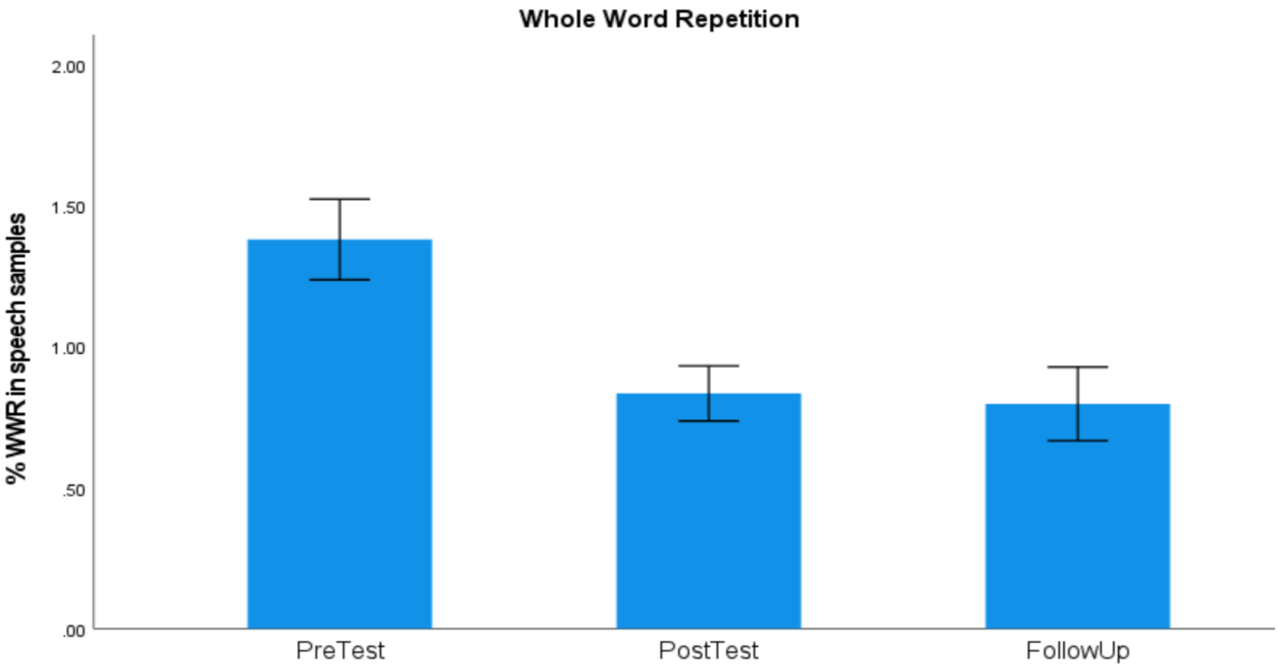

Figure 5.4: Mean scores and standard deviations for %WWR for all participants collectively across all assessment stages

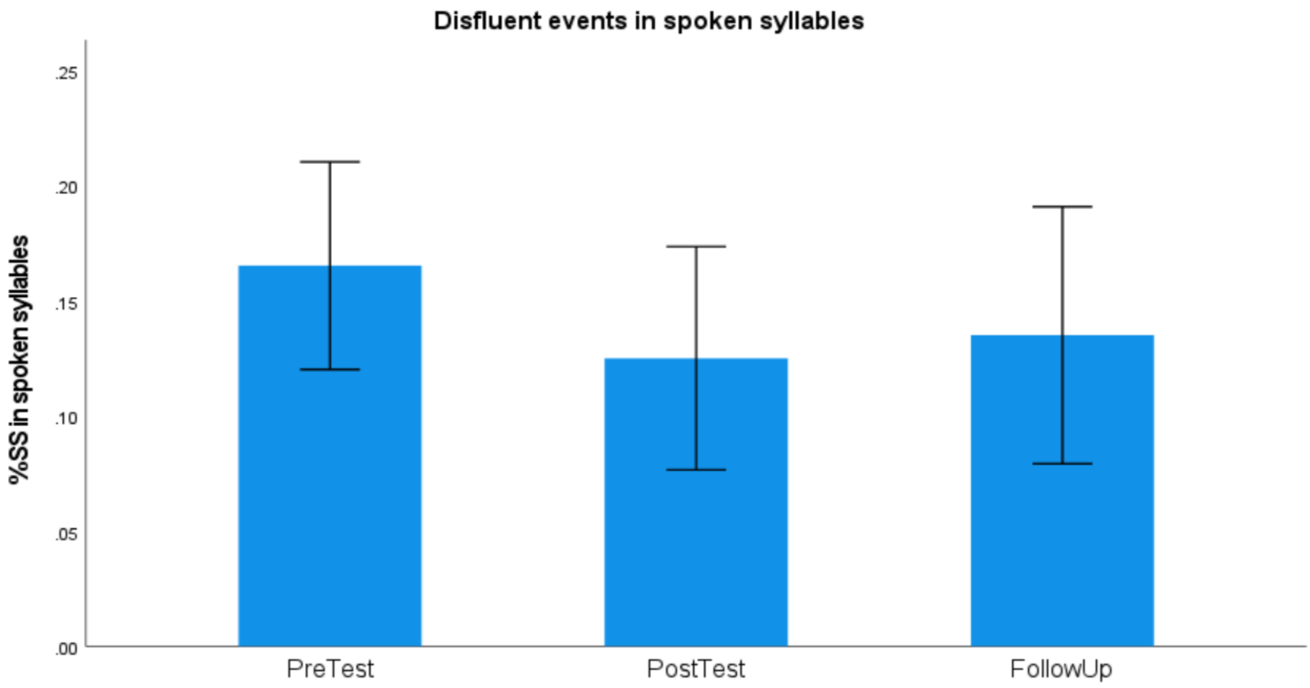

Figure 5.5: Mean scores and standard deviations for %SS for all participants collectively across all assessment stages

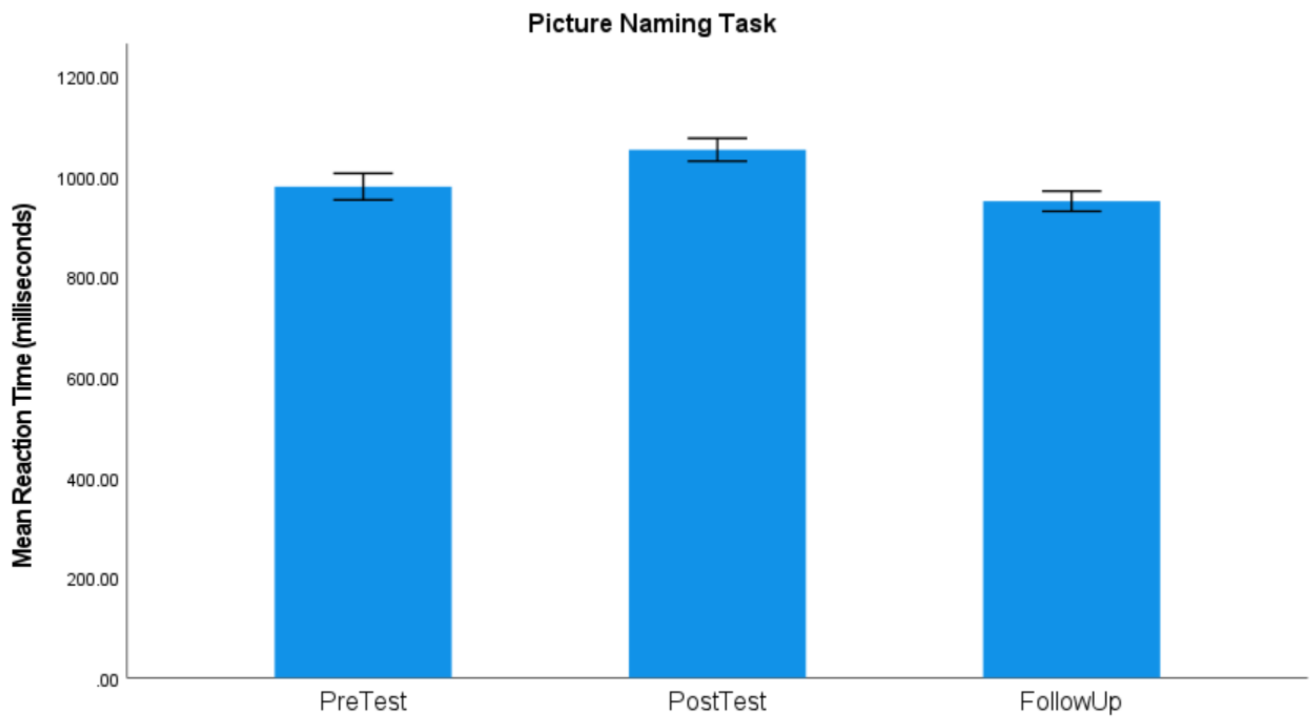

Figure 5.6: Mean scores and standard deviations of reaction time in milliseconds for all participants collectively across the three assessment stages.

Three-way ANOVAs with group (experimental vs control) as a between-subjects factor as well as time (pre, post and follow-up) as within subject factor were examined to determine the effect of the intervention. There were five dependent variables: (1) AEN_NWR scores as a measure of phonological performance; (2) speech comprehension; (3) reaction time in the picture naming task ; (4) %SS and (5) %WWR. Details of ANOVA results for each of the five tasks are summarized below.

Phonological Performance (AEN_NWR scores)

A significant interaction effect between phase and group was found for the AEN_NWR score ($F(1,57) = 18.0$ $P<.000$). This indicated that the intervention had different effects on phonological performance in the two groups. While no significant main effect of the intervention was found for all children as a whole group, the results indicated that there was a significant main effect of the intervention on AEN_NWR scores for the experimental group ($F(1,31) = 16.87$, $p < .001$). Bonferroni-corrected pairwise tests showed that the experimental group of significantly improved in their phonological performance from pre-test (M = 7.70, SEM=.61) to post test (M = 9.42 , SEM=.61) and also from post-test to follow-up (M = 11.00 , SEM=.70) as shown in Figure 5.7 below. On the other hand, the control group showed a slight decline on AEN_NWR from pre-test (M=11.22; SEM=.96) to post-test (M=11.07; SEM=.84). This group showed a further decline in their scores on follow-up (M=9.18; SEM=.98), but this decline was insignificant compared to pre-test and to post-test. No further significant effects were found.

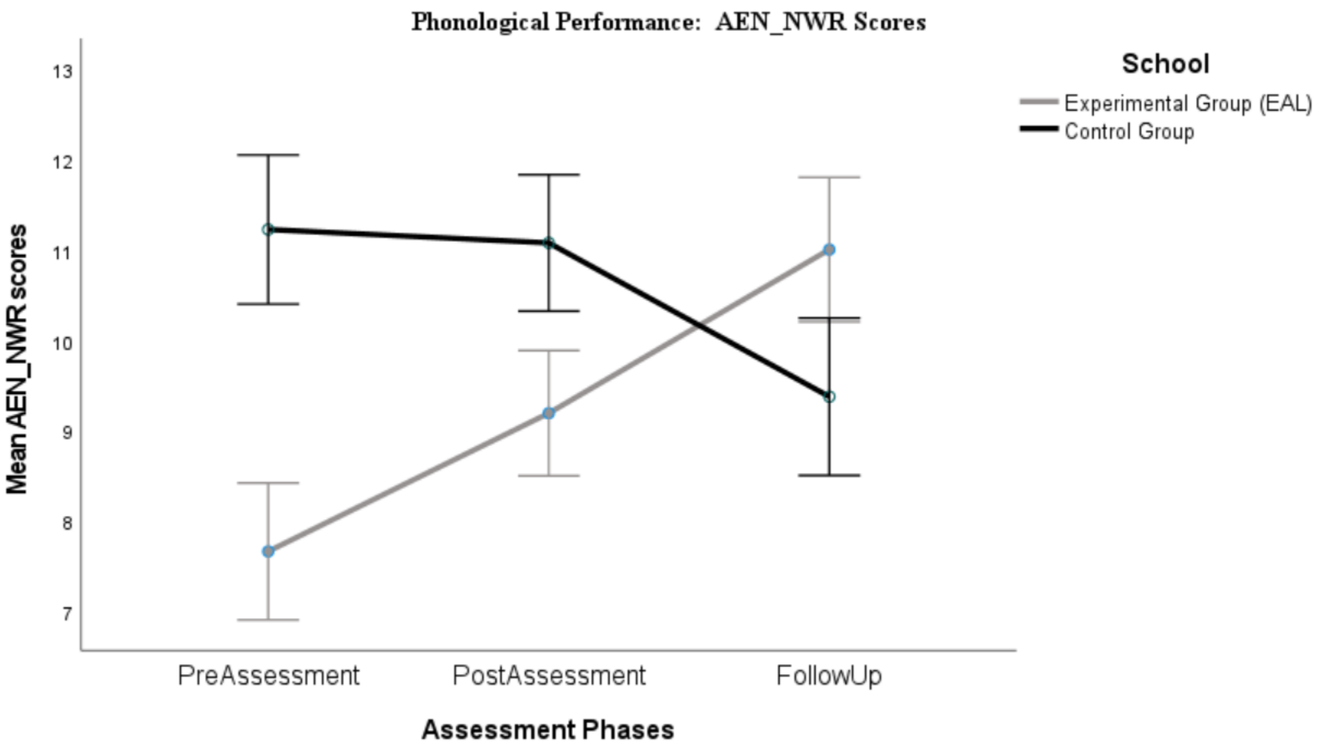

Figure 5.7: Graph of mean AEN_NWR for both experimental group and control group at each assessment phase. Bars are standard errors.

Speech Comprehension

Overall, there was a main effect of the intervention on the children's performance on narrative comprehension (F (2,61) = 5.2; $p < .001$). Bonferroni-corrected pairwise comparisons showed that while children's scores on this task significantly improved from pre-test (M=3.8; SEM=.23) to post-test(M=4.38; SEM=.2), and a slight increase from post-test to follow-up (M=4.5 ; SEM=.2), although this was not significant. Children in the experimental group showed an increase on their scores from pre-test (M =3.70; SEM=.32) to post-test (M=4.55; SEM=.33), although the increase was insignificant. In contrast with what was expected, there was a decline on this task during the follow-up session (M=3.85; SEM=,27). Children in the control group showed a significant improvement in their scores from pre-test (M=4.3; SEM=.33) to post-test (M=5.10; SEM=.32); they also improved slightly more in the follow-up (M=5.2 ; SEM =.36) but this was not significant. No interaction effect was found. Figure 5.8 summarizes the results.

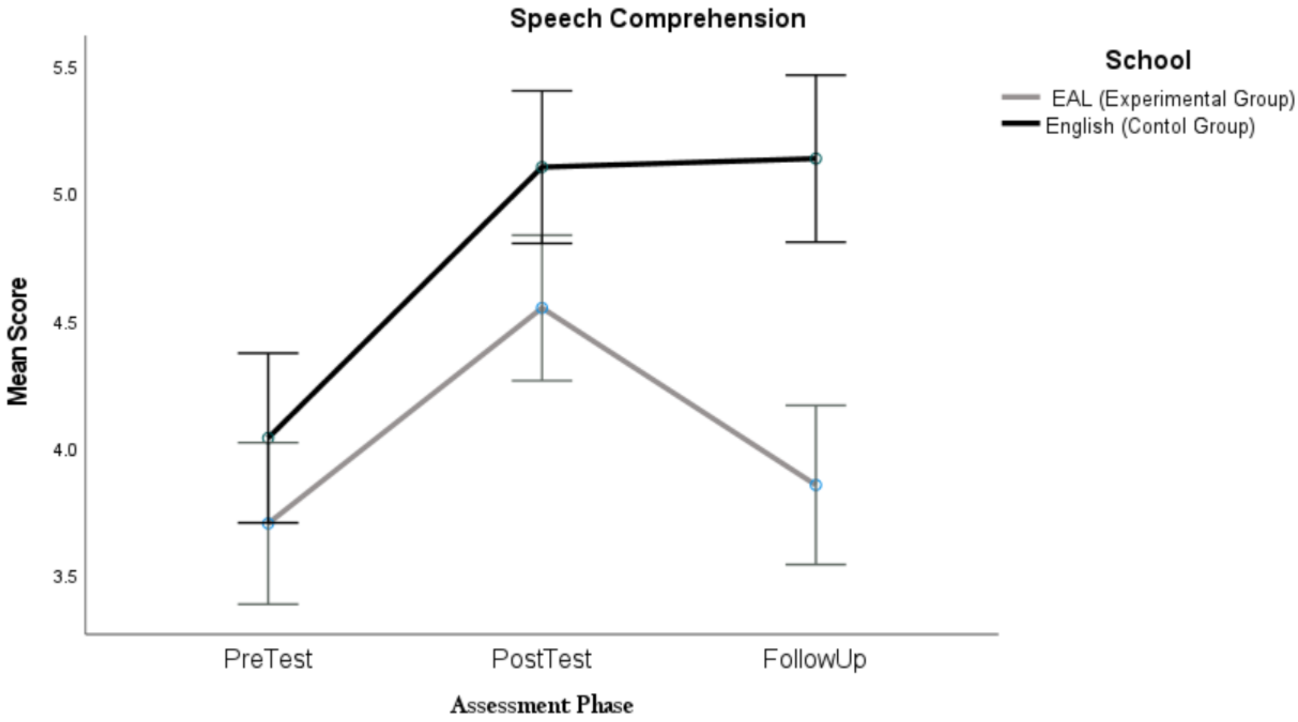

Figure 5.8: Graph of mean scores on speech comprehension task for both experimental group and control group at each assessment phase. Bars are standard errors.

WFD

There was a significant main effect of the intervention on %WWR. The results of the three by two ANOVA showed a significant main effect of time on the decline in WFD during the spontaneous speech sample ($F(1, 61) = 9.2$, $p < .001$). Overall, there was a decline in %WWR from pre-test

(M=1.40; SEN = .14) to Post-test (M=.83; SEM=.09) and this was significant. The %WWR even declined more in follow-up (M=.79; SEM=.13) but this was not significant compared to post-test. Bonferroni-corrected pairwise comparisons revealed that the main effect came from a decline in the WFD in post-assessment (M=.82; SEM=.13) compared to the pre-assessment baseline (M=1.45; SEM=.20) for the experimental group. There was even a decline in follow-up for this group (M=.56; SEM=.1) but this decline was not significant in comparison with the post-test. With respect to children in the control group, Bonferroni-corrected pairwise comparisons showed that the reduction in WFD at post-test (M=.85; SEM=.23) compared to the pre-assessment baseline (1.3; SEM= .21) was not significant. A significant interaction effect between group (experimental vs control) and assessment phase was found indicating differences between the groups in responding to the intervention (F (1,58) = 4.45, p=.03). The results are summarized in Figure 5.9.

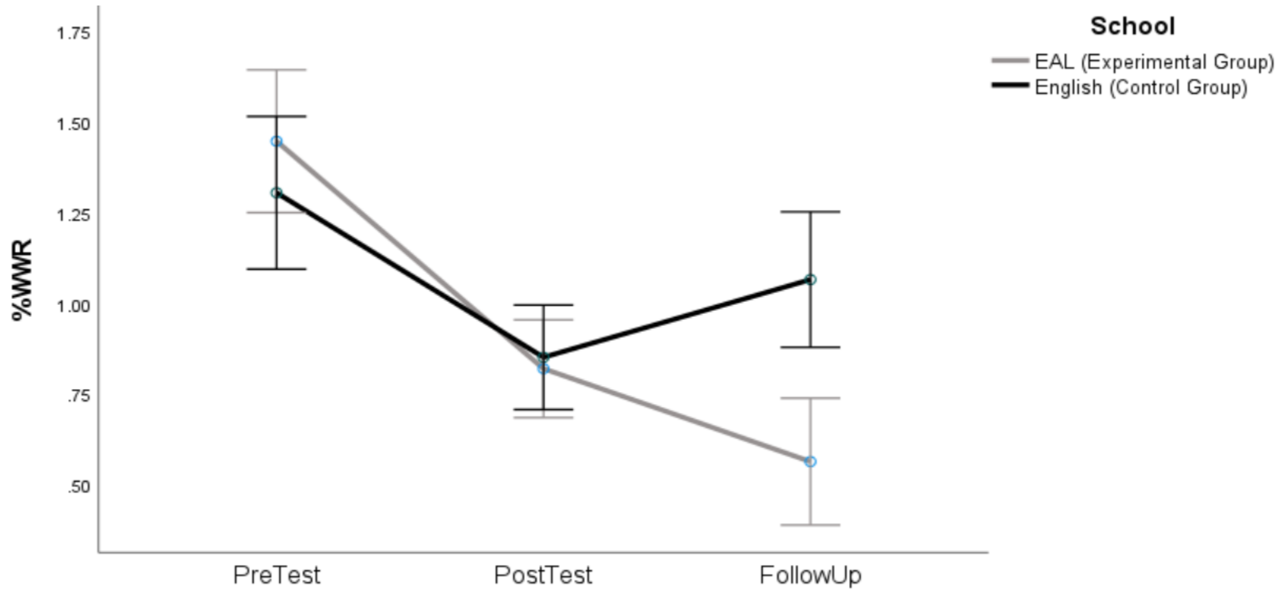

Figure 5.9: Graph of mean %WWR for experimental group and control group at each assessment phase. Bars are standard errors.

Because some recordings contained only around 100 syllables of speech, whereas others contained up to 300 syllables, a second analysis was conducted in which records where the number of syllables was less than 200 was excluded. This second analysis showed that there was a main effect of the intervention ($F (1,54) = 13.01$; $p < .001$). %WWR significantly declined from pre-test (M= 1.43; SEM = .2) to post-test (M=.87; SEM = .10) and the reduction was sustained during follow-up (M=.84; SEM=.136). An interaction between time and group was also significant ($F (1,54) = 4.93$; $p = .03$) indicating that the two groups responded differently to the intervention. As shown in Figure 5.10 below, the experimental group started with higher level of %WWR compared to the control group. This group showed a consistent decline in %WWR throughout the assessment phases (Pre-test: M=1.5; SEM=.20; Post-test: M =.85; SEM = .13 ; Follow-up: M =.56 ,SEM = .11). The control group showed a decline in their %WWR from pre-test (M=1.3; SEM=.23) to post-test (M=.87; SEM= .16). A slight increase in the %WWR occurred in follow-up for this group (M=1.12; SEM=.25). None of these changes for this group were significant. These findings indicate that the experimental group showed a greater improvement across assessment stages compared to the control group.

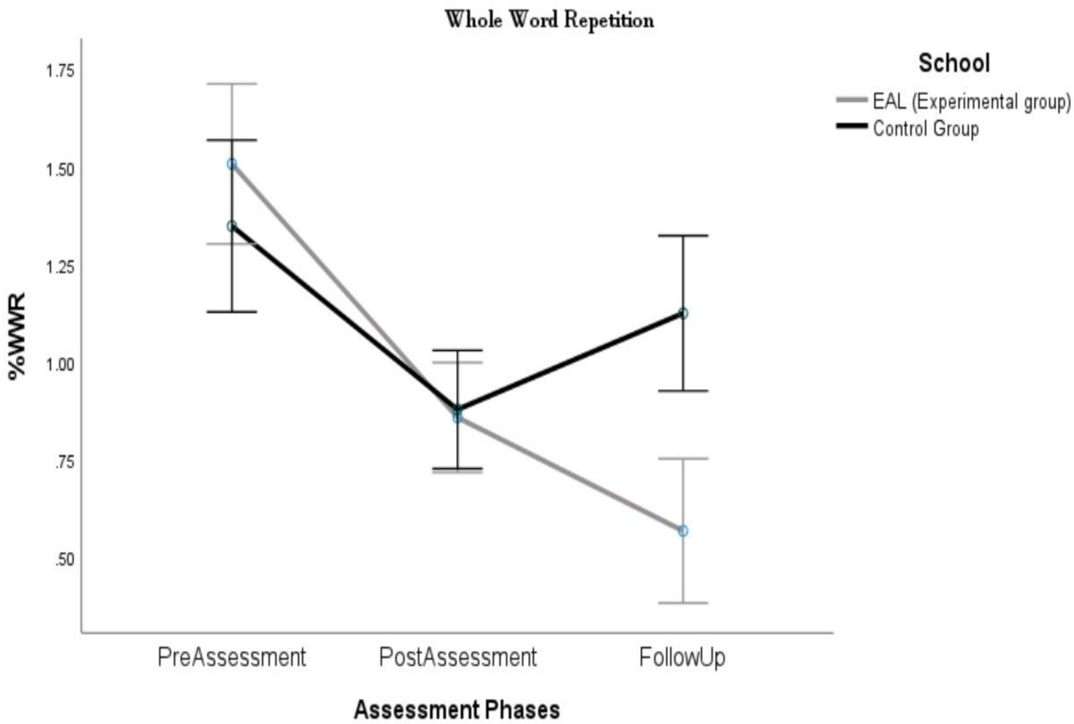

Figure 5.10: Graph of mean %WWR for both experimental group and control group at each assessment phase (excluding samples where syllables number was less than 200). Bars are standard errors.

Stuttered Syllables

There were changes in %SS after the intervention. Taking the average of all participants, both post-interventions mean %SS (M =.12, SEM = 0.05) and follow-up mean %SS (M = .14, SEM = 0.05) were lower than pre-intervention mean %SS (M .17, SEM = .04). However, the results of the three by two repeated-measure ANOVA did not show a significant main effect of time on %SS (F (1, 58) = .3, p > .05)). Bonferroni-corrected pairwise showed that children in the experimental group did not show a change in their %SS from pre-test (M=.17; SEM=.41) to post-test; although they showed a nonsignificant decline from post-test (M= .17; SEM=.46) to follow-up (M=.09; SEM= .44). Children in the control group school showed a reduction in their mean %SS post-intervention (M=.07; SEM=.22) compared to pre-intervention (M=.15; SEM=.26); however, a paired sampled t-test showed that the difference was not significant ($p > .05$). Additionally, their increase in %SS at follow-up test compared to post-test (M=.18; SEM=.41) was not significant ($p > .05$). These results are summarized in 5.11,

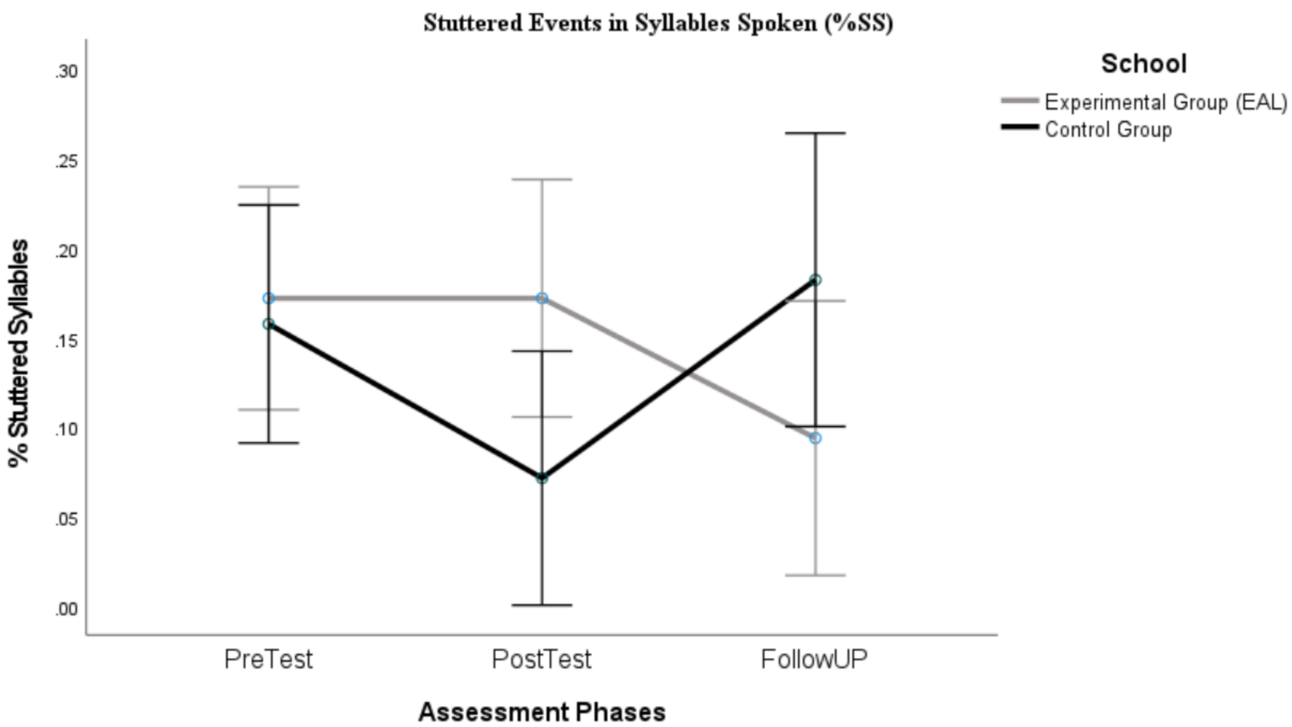

Figure 5.11: Graph of mean %SS for both experimental group and control group at each assessment phase. Bars are standard errors.

Again, due to differences in the number of syllables produced by children, a second analysis was conducted where a fixed number of syllables was used. No significant main effect was found for the intervention on %SS for both groups ($p > .05$). The experimental group's %SS remained the same during pre-test and post-test (M=.183; SEM= .07). A slight reduction occurred during follow-up (M= .1; SEM =.1) although this was not significant ($p > .05$). The intervention had no significant effect on the %SS for the control group either. The %SS declined slightly post-intervention and then increased slightly during follow-up, but none of these changes were significant.

Reaction Time

The main effect of the intervention on the mean Reaction time was significant ($F(1,61) = 25.71$; : $p <$ .001). Additionally, the interaction effect between groups and the intervention was significant ($F(2,61) =$ 4.6; $p = .01$). This indicates that there were differential changes in reaction time across the assessment stages between the two groups. Bonferroni-corrected pairwise showed that all children significantly improved in their response latency (decrease in reaction time) from pre-test (M = 1091.38, SEM=30.93) to post test (M=973.85, SEM=22.49) and also from post-test to follow-up (M = 910.40, SEM=.20.10) as shown in Figure 5.12 below. Inspection of the intervention effect on the mean RT in the two groups separately showed that the effect of the intervention on the RT for children in the experimental group was significant (($F(1,29) = 28.31$; $p < .001$). Children in this group showed a significant reduction in their RT at post-test (M=958.0; SEM=32.20) compared to pre-test (M=1141.51; SEM=44.24); they also showed further reduction at follow-up (M=892.80; SEM=22.93) compared to the mean RT at post-test; this was also a significant reduction. With respect to the control group, they also showed a decline in their reaction time from pre-test (M=1041.118; SEM=42.93) to pots-test (M=989.72; SEM=31.16); and a further decline during follow-up (M=927.98; SEM=33.80).

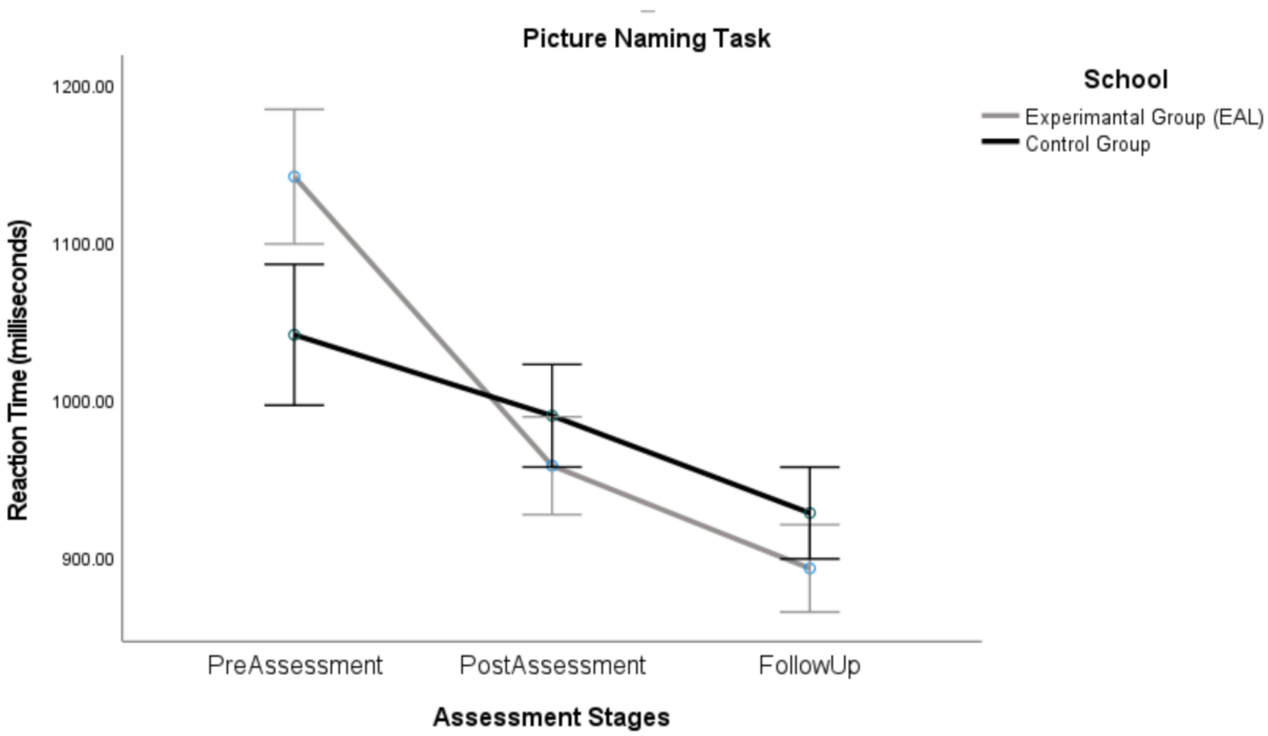

Figure 5.12: Graph of mean Reaction Time for both experimental group and control group at each assessment phase. Bars are standard errors.

Finally, an additional set of analysis were conducted to examine the effect of the intervention on all tasks; this time the complete group of children was divided based on whether or not they have been exposed to Arabic. As highlighted in the methods section, children in the experimental group spoke Arabic predominantly, but some of them were exposed to Arabic from birth and some were not as they were children of Expats who work in Saudi Arabia. In the same way, children in London spoke predominantly English as they were born and raised in the UK, but some of them had some exposure to Arabic by parents at home through one of their parents. This information was collected from teachers based on parents' reports when children registered in schools. Only ANE_NWR scores, an interaction effect was found between the language factor (exposure to Arabic) and the assessment across sessions ($F(1,57) = 4.1$; $p = .04$). Figure 5.13 summarizes the results.

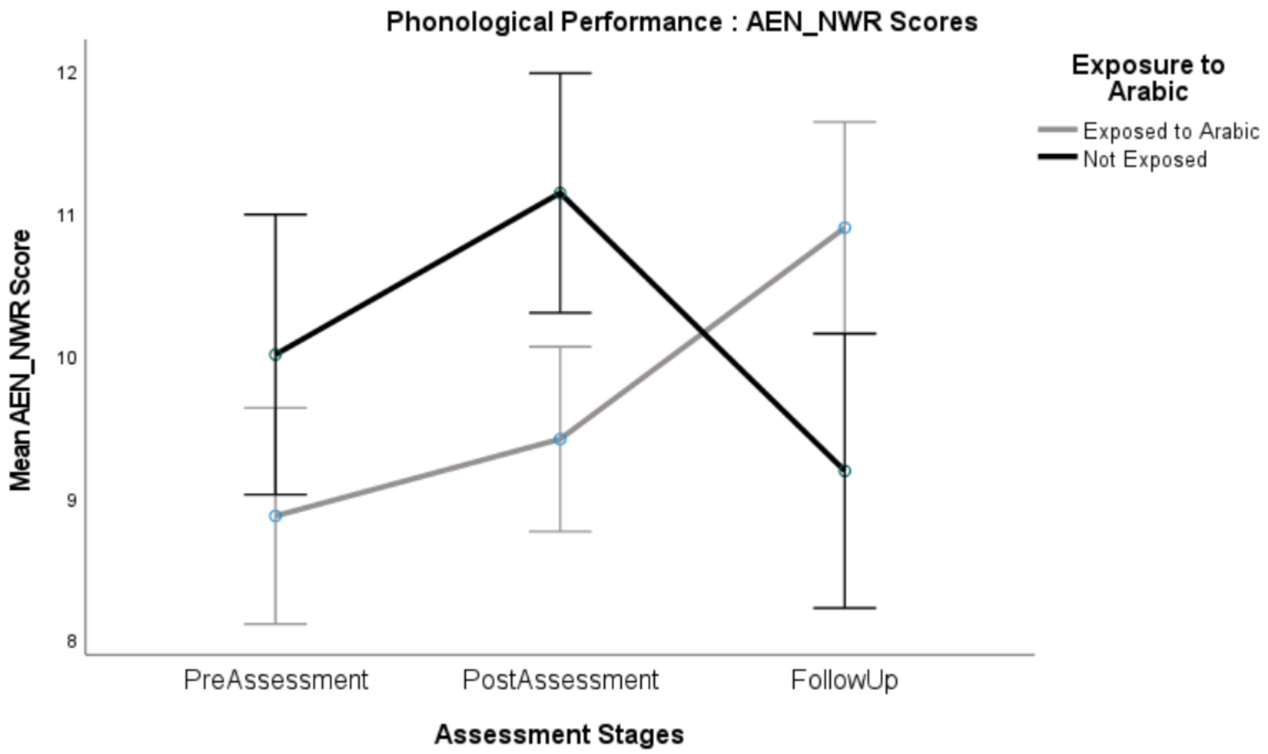

Figure 5.13: Graph of mean AEN_NWR for both experimental group and control group at each assessment phase. Bars are standard errors.

5.5 Discussion

The present study aimed to improve word-finding in reception-class Arabic children with EAL. Furthermore, the differences between a group of Arabic children with EAL and a control group of English children on speech comprehension, phonological performance, lexical access and two measures of speech fluency were investigated. Results must be interpreted with caution, as they are based on data for one intervention session only, but they provide some interesting pointers for further investigation.

In this study, the focus has been on phonological training for Arabic children with EAL; by teaching them aspects of English phonology that they lack. Statistical analyses of the data have shown that children in the experimental and control groups generally performed better post intervention: they showed higher accuracy in phonological performance, faster reaction times, and a decline in WFD (as indicated by WWR) ; and this was partly sustained at the follow-up session one week following the intervention. It is worth noting that the improvements seen in both groups could be interpreted as an indication that , overall, children started getting better with language overtime. This would make difficult to determine whether the effects are due to teaching input in schools rather than the intervention. Thus, in order to account for any maturity effects between groups of children, future work could consider a study with a larger sample in which children are sub-divided into groups according to age and months in 3-month intervals.

Some non-significant effects where significance was predicted might be non-significant because of low power. The scores for each task were analysed and compared between the two groups. The fact that there were differences in how the two groups reacted to the intervention in most of the tasks suggest that a language-specific training might be necessary. A summary of significant effects that were found can be seen in Table 5.9 below.

Table 5.9: A summary of significant main effects for the experimental and control groups

Task	Group/Factor	Effect found	Relation to hypothesis
Phonological performance	Experimental Group	Significant improvement in scores across the three assessment stages.	Consistent with the hypothesis
Phonological performance	Interaction	Significant effect for the experimental group compared to the control group	Consistent with the hypothesis
Speech Comprehension	Within -group	Overall significant main effect of the intervention for all (improvement in scores from pre-test to post-test)	Consistent with the hypothesis
Speech Comprehension	Control group	Significant improvement from pre-test to post-test	Consistent with the hypothesis
%WWR (First analysis)	Within-group	Overall significant main effect of the intervention (decline in %WWR from pre-test to post-test).	Consistent with the hypothesis
%WWR (First analysis)	Experimental group	Overall significant main effect of the intervention (decline in %WWR from pre-test to post-test)	Consistent with the hypothesis
%WWR (first analysis)	Interaction	Significant effect for the experimental group compared to the control group	Consistent with the hypothesis
%WWR (Second analysis)	Within-group	Overall significant main effect of the intervention (decline in %WWR from pre-test to post-test and from post-test to follow-up)	Consistent with the hypothesis
Reaction Time	Control group	Reduction from pre-test to post-test, and from post-test to follow-up	Consistent with the hypothesis
Reaction Time	Experimental Group	A significant reduction from pre-test to post-test, and from post-test to follow-up	Consistent with the hypothesis

With respect to phonological skills, the intervention led to a marked improvement in AEN_NWR scores for the experimental group. There was an initial improvement in AEN_NWR scores for the treatment group, as shown by the significant increase in their AEN_NWR scores between pre-treatment and post-treatment assessment phases. Furthermore, there was a significant increase in AEN_NWR scores between the post-treatment and one-week follow-up assessment. This demonstrated that the improvement in phonological performance was further enhanced after the intervention ended. Hence, it can be suggested that WFD training affects phonological ability for Arabic children with EAL. In contrast, the control group showed no significant changes, although they started at a significantly higher level compared to the experimental group. On the basis that phonological training should improve phonological performance for children with EAL, this was an expected result. The insignificant decline during follow-up could be attributed to practice effect; although it remains unclear why a decrement happened. Future studies should take this into account and administer a shorter non-word repetition task. This would necessitate several validation of the current AEN_NWE until achieving a sensitive test with minimal number of items.

When it comes to narrative comprehension, the control group started at a higher level compared to the experimental group, and they continued to outperform the experimental group in post-test and follow-up. This could be explained by to the fact that all stories were pre-recorded in English by a native British speaker. Hence, it is possible be that this posed more of a challenge to the Arabic children compared to children in the control group. Even if the Arabic children with EAL understand English, they are less familiar with the British accent compared to the control group of children who study in school in London.

In terms of disfluency characteristics (%SS and %WWR), two analyses were conducted as follows. In the first analysis, complete samples collected from all children were analysed despite differences in the number of syllables obtained. Results of this analysis showed that all children's scores on symptoms of disfluent speech (%SS and %WWR) tended to decline over time, but the effects were not always significant. Also, the control group showed an increase in %SS and %WWR; a result that was not expected. It was considered that the insignificant improvements for the two measures and the unexpected increase for the control group could be attributed to the fact that recordings varied in length; some recordings contained only around 100 syllables of speech, whereas others contained up to 300 syllables. Children spoke for longer durations especially at post and follow-up assessments, which might have affected pre- and post-intervention comparisons (children might show a proportionally higher number of WFD characteristics at post assessments due to a higher number of syllables spoken, which does not particularly allow a reliable comparison of the outcomes at the different assessment stages). Additionally, children got more used to the experimenter over time and this might affect their language output. Thus, a second analysis was conducted in which a fixed number of syllables was used to obtain measures of %SS and %WWR. In terms of %WWR, the overall group of children showed an immediate effect across pre and post training, and there were long-term effects (pre to follow-up). Children in the experimental group did not change immediately following the intervention, but they did show a decline in follow-up. The control group, however, did not show a clear patten of how the intervention had an effect on their %WWR. The unexpected increase in the %WWR could possibly be due to motivational effect; that is the intervention procedure was lengthy and some children might have lost interest towards the end of the study. This is of extreme importance to be accounted for in future studies in which testing sessions should be shorter. Additional time might also be needed between the assessment stages, but then care should be taken to ensure that the effects are due to teaching input rather than the intervention. Although it is unclear whether the intervention was a factor in reducing %WWR for the experimental group, their improvement in WFD over time supports the notion that WWR is not a stuttering symptom. It is a vocabulary problem that happens frequently when children speak EAL (Howell, 2017).

The results also showed that there were no significant changes in %SS over the assessment phases for both groups. This suggests that %SS was not affected by the WFD training, and that the intervention did not improve articulation, but rather it works at the phonological and lexical levels. Thus, WFD and speech disfluency are two communication difficulties, and each requires a different type of intervention. The WFD procedure suggested here could be offered in schools. A fluency difficulty training procedure needs to be developed to improve %SS. Howell et al. (2020) designed a procedure that uses working memory training on children at risk of fluency difficulty. The procedure was designed so that it is appropriate for children with English only and other children with different first languages who use EAL in UK schools and proved to be effective. Support for working memory training to improve speech fluency comes from numerous studies that show that WM capacity is related to speech fluency. For instance, Denman and Green (1986) investigated the role of WM in verbal fluency where the participants had to access the lexicon to produce words that are appropriate to the context. The results showed that verbal fluency is related to individual differences in WM capacity. A further study by Daneman (1991) investigated whether WM capacity could account for individual differences in verbal fluency. Several measures were taken in this study to assess verbal fluency including a speech generation task where participants were presented with a picture and asked to talk about it for one minute. The results supported the previous results showing that WM capacity is related to individual differences in verbal fluency during speech production. Participants with a greater capacity to hold and process information in the temporary storage were more fluent and also less prone to articulatory errors that may occur while speaking. Mota (1999) examined whether WM capacity correlated with fluency in speech production task involving an additional language. This study was based on Danemen (1991) mentioned earlier, and methodology was adapted. The test was applied to 16 EAL speakers with Portuguese as their first language. Participants' fluency in English was assessed by a speech generation task. There was a significant correlation between WM capacity, measured by the speaking span test, and fluency. This significant correlation corroborated the report by Danmen (1991) about the relation between an individual's WM capacity and fluency in speech. A similar study investigated whether engaging in WM tasks benefitted speech fluency of adults who do and do not stutter (Eichorn, Marton, Schwartz, Melara, & Pirutinsky, 2016).The results showed that all participants benefitted from WM training in enhancing their fluency. This suggests that WM capacity relates to fluency regardless of stuttering status. From the studies reported above, it can be concluded that there is a correlation between WM and fluency. However, this does not provide evidence about the cause of fluency difficulty as correlation does not imply causation in the relationship between WM and fluency. Howell et al. (2020) noted that although the specific contribution for WM on enhancing speech fluency has not yet been established, the two-week WM training improved fluency in twelve preschool children who have fluency difficulty. In the same way, a working memory training may be developed for Arabic children with EAL. A comprehensive package can then be created in which training procedures for fluency and WFD may be further developed and improved so that they can be used easily in schools.

Analyses of reaction time on the picture-naming task provide insight into the importance of phonological training as an intervention procedure. The findings of the assessments of picture-naming indicated that both groups showed faster RT following the intervention and improved even more at the follow-up. The control group started at a lower level prior to the intervention (i.e. they were faster than the experimental group). A possible reason this has happened is because some of the pictures were challenging to name for the experimental group particularly. This has happened despite conducting a familiarization procedure prior to beginning the task in which the experimenter named all pictures clearly. Nevertheless, it is possible that even this familiarization was not sufficient specially with some pictures that occur less frequently than others. For example, in response to seeing a picture of a “vase”, many of the children either did not respond or said “flower” referring to what was inside the vase. Additionally, there were some examples were the EAL children used an alternative name to the target word (e.g. saying “road” when the target word was “street”, saying “rock” when the target word was “stone”). Consequently, the children were not able to accurately name those items that formed part of the post-test and that contributed to their lower response latency during the pre-test. The fact that the experimental and the control groups improved over time is most probably due to familiarity with the task and training effects, but also suggests that the training material was appropriate and beneficial for all participants. Thus, results provide some support for the hypothesis that language-specific phonological training has an effect in improving word-finding for these children. This provides further insight into potential interventions for children with word-finding difficulty and associated fluency problems. This is also consistent with results from Howell (2015) in which the effectiveness of a WFD intervention in children was examined. A mixture of semantic-based training and phonetic cues was used in Howell’s study and it was shown that it had an impact on WFD as well as fluency.

Although no formal language proficiency measure was taken, information on exposure to Arabic was collected from parents and teachers to examine if it had an effect on one of the previous measures (i.e. fluency measures, narrative comprehension, phonological performance or reaction time). The significant interaction effect between exposure to language and the intervention for the AEN_NWR indicates that children in the two groups (London and Riyadh) responded differently to the intervention.. Future work should include standardized measures of proficiency of language. One task that would be appropriate is a lexical retrieval task which was used by Marshall et al, (2014). It involves asking children to name as many objects as they can of a category in all the languages they speak. This can then be analyzed to see whether there are differences in proficiency between languages a child speaks. Alternatively, a language history questionnaire can be used, in which parents can fill out information for their children.

Research and Clinical Implications

The outcomes of the study are valuable and useful in many ways, as they have educational implications. The intervention showed effects for Arabic children with EAL and for English children. Thus, the study provides a good approach to improving word-finding in preschool EAL children with other first languages, given that the right intervention materials are used, and the procedures could be easily and efficiently administered by teachers. The intervention could even be more beneficial for children if semantic training was included as part of the interventions to improve a range of skills further.

Limitations and Future Studies

The current study had some limitations that may have affected the results and the interpretation. Although the sample size was reasonable compared to other intervention studies, a larger sample size is needed to assess the clinical importance of the intervention in both groups. This would also allow additional detailed analyses to be conducted such as multilevel analysis. Moreover, future interventions should involve control materials that match the current ones in length but does not pose challenges for Arabic with EAL. This would be beneficial in determining whether improvements were really due to the phonological training. Another alternative would be to use nonwords that follows the phonological constraints of Arabic and use those to train a control group of English speakers. These materials were developed as part of this thesis; however, it was not feasible to recruit a comparable group of children. A list of these non-words can be found in Appendix 4.

The biggest limitation in this study probably was that children participated in one intervention session only. Future studies should consider multiple exposure to the intervention and an examination of the intervention effects following each session. In Best et al.'s (2018) WFD study that was mentioned above, children had six intervention sessions. Moreover, in order to account for any maturity effects between children, a study with a bigger sample could sub-divide the children into groups according to age and months in 3-month intervals. Finally, the selection of both groups was constrained by the need to conduct all testing in English to ensure consistency in conducting all tasks. Thus, for the experimental group, children knew some English through their schools or their caretakers; but they also came from a very high socioeconomic status (SES). In future studies, a matched control group mush be selected to eliminate any performance differences that may result in differences in SES.

Chapter 6 Self-report Measures for Assessing Speech Fluency in Adults and Children Arabic Speakers: Future Directions

"Thank you, I said bravely, dropping the syllables, cleanly, like marbles, and secretly full the most pathetic pride imaginable. I had spoken to strangers." *Alexis Hall,* Waiting for the Flood

6.1 Introduction

As part of the fluency assessment, it is important to have self-report instruments that can rapidly identify speech disfluency when large numbers of children need to be assessed. This chapter examines the importance of developing and standardizing self-report instruments for assessing fluency problems. The new changes in the definition of stuttering the DSM- V (2013) are discussed; along with how an instrument can be designed to accommodate these changes.

6.2 Classification and Diagnosis of Stuttering

In the updated version of the DSM-V (2013), stuttering is referred to as "Childhood-onset Fluency Disorder" to emphasize the predominance of onsets of developmental stuttering in early childhood; in which the average age of onset is from 2 to 7 years in more than 80% of cases.

Andrews and Harris (1964) reported that the lifetime incidence of stuttering is 1% in their study on stuttering in 1142 families in the United Kingdom with children born between May and June 1947. The study ended when the children were 15 years old, and established that the point prevalence of stuttering up to the age 15 was approximately 4.9%. Yairi and Ambrose (2005) confirmed that approximately 5% of pre-school age children exhibit episodes of stuttering. Prevalence varies as a function of age. Approximately 5% of pre-school age (3 to 4 years old) children are affected, but this percentage drops to 1% at the end of primary school (10 to 11 years old) and remains at this level through the lifespan (Yairi & Ambrose, 1999). Moreover, stuttering is characterized in DSM-V as "disturbance in the normal fluency and time patterning of speech that is inappropriate for the individual's age" (American Psychiatric Association, 2013, p.46). Disturbance to 'normal fluency' implies judgments are made relative to typical speakers (TS). DSM-V also indicates where there are differences between those not likely, versus likely, to stutter in domains other than speech symptoms (American Psychiatric Association, 2013). The domains include stress, quality of life and anxiety which also need to be assessed in ways that are appropriate for TS and PWS. The main takeaway message is that new instruments to assess speech fluency need to meet the following two requirements: (1) can be completed by both people who stutter (PWS) and typically fluent speakers (TS); and (2) include affective components of speech disfluency. As stuttering is not seen solely as a behavioural disorder characterized by observable disfluency symptoms, it seems appropriate to use self-report to access feelings of anxiety, sense of achievement, and other relevant emotions towards speech fluency. In the next section, current instruments to assess speech fluency in adults are reviewed to show that they do not meet DSM-5's requirement of (1) assessing speech disfluency on a continuum with normal speech and (2) including emotional consequences of stuttering.

Review of Clinical Instruments for Assessing Fluency and its Affective Concomitants

There exists a variety of tests to measure speech fluency and anxiety in both children and adults. With respect to the observable characteristics of speech, current measures of speech production difficulties include Riley's Stuttering Severity Instrument (SSI-4; Riley, 2009), and non-word tests such as

UNWR (Howell et al., 2017). Both measures have good reliability and validity supporting their use as a basis for designing a tool that can measure fluency. The SSI-4 is a widely used measure of stuttering. Information on this test has been discussed earlier in this thesis and therefore will not be repeated here. The main limitations of SSI-4 when the purpose is for assessing large samples from the general population stem from the fact that it does not take into account affective concomitants and that test administration and scoring are time consuming. To combat the problem of time required to assess children using SSI, the UNWR (Howell et al., 2017) makes a good alternative to speech-based procedures as it can identify fluency difficulty more quickly. It was shown that scores on the UNWR provide a sensitive marker of speech disfluency as children with fluency difficulty showed poor UNWR performance. Measures like UNWR, however, imply some practical issues: they are labour intensive, and require training for the person carrying out the assessments, and live scoring may not always be accurate. In short, valid measures that can identify speech disfluency are available; and they have many potential benefits for children. However, practicality considerations arise if those measures are administered in schools. Children have to be tested individually by a trained staff member or teacher. The children are required to speak a minimum of 200 syllables in order that a reliable estimate of stuttering severity can be provided (Riley, 1994; Todd, Mirawdeli, Costelloe, Cavenagh, Davis & Howell, 2014). And in the case of UNWR, training on live scoring is essential. This would all be time-consuming if teachers were required to assess all children in their classes. Moreover, these measures concern solely the observable characteristics of disfluent speech, and they do not extend across all the speech and affective components associated with dysfluent speech specified in DSM-V.

There are also several established and validated questionnaires which focus specifically on stuttering as a speech disorder. For example, the Wright and Ayre's Stuttering Rating Profile (WASSP, Wright & Ayre, 2009) is a 60-item instrument where the items are scored on 7-point scales. The self-report scale has subscales for measuring the social elements of stuttering that meet the criteria for the World Health Organization's International Classification of Functioning, Disability, and Health framework (ICF; WHO, 2001). WASSP items, however, do not apply to TS. For example, items on the 'stuttering behaviours' dimension such as 'uncontrollable stutters' only apply to PWS, and, therefore, are not appropriate for samples that include TS. Also, OASES (Yaruss & Quesal, 2006) is a 100-item instrument, which evaluates stuttering in relation to the life experience of PWS. The self-report scale is used as a broad outcome measure of stuttering. Like WASSP, OASES focuses specifically on stuttering, making some items inapplicable for the general population. An example is "I think about my stuttering all the time". A commendable feature of the Dutch version of OASES (Koedoot et al., 2011) is that re-standardization and validation were undertaken. After translation into Dutch, the psychometric performance was evaluated statistically as the English norms would not apply (Karimi et al., 2011). The target population for OASES is persons with fluency difficulty and a 100-item instrument would be appropriate for them. For instruments that target the general population of speakers including TS, a 100-item instrument would be too lengthy for them.

Our review now turns to self-report measures of social anxiety. The idea is to review measures of fluency and social anxiety; and then to propose one self-report measure that can objectively assess fluency whilst also including affective components.

Established self-report instruments for social anxiety that are appropriate for use with the general population such as the 'Social Interaction Anxiety Scale' (SIAS; Heimber et al., 1992) and the 'Liebowitz Social Anxiety Scale' (LSAS; Fresco et al., 2001; Liebowitz, 1987) are valid measures that have demonstrated good psychometric properties. Such tools examine anxiety in various socio-interactional settings that demand oral communication. For instance, LSAS includes items about initiating a conversation and speaking on the telephone in a public place. High social anxiety is associated with avoidance of such social situations (Schlenker & Leary, 1985). These several instruments could be used jointly to assist clinicians and researchers in assessing fluency. However, it would be prohibitively time-consuming to administer multiple tests. This is especially true in schools and clinics where time constraints are imposed. Thus, one questionnaire is required that assesses fluency and its associated psychological consequences efficiently. Ideally, this questionnaire should be very quick to administer and should not require training for staff to conduct. It can act as a check-point in which children who do not pass the screen could be then referred to lengthy evaluation. This will make it appropriate for assessing children and patients where time allowed for testing is restricted.

6.3 Shift to Continuous Measures in DSM-5

Continuity of fluency between TS and PWS was introduced into DSM-V to give clinicians more latitude when assessing disorders where severity varies (APA, 2013). The Schizotypy Personality questionnaire (SPQ; Raine, 1991) is an example of a continuum-based self-report scale that reflects the continuity view. The 79-item questionnaire was developed as an overall measure of schizotypal personality in the general population including non-psychiatric patients. SPQ was modelled closely on DSM criteria and assesses schizotypal personality disorder on nine subscales (Raine, 1991). A fluency questionnaire could be developed with a similar motivation in mind, i.e. to generate a scale that facilitates the assessment of fluency in the entire population. PWS and TS are placed on a fluency continuum where speakers with a clinical diagnosis of stuttering tend to be at one extreme of the continuum. Multiple subscales can be constructed to assess different aspects of fluency. The inclusion of these subscales enables individuals who may have sub-clinical levels of speech symptoms but high levels of social anxiety and other affective consequences of stuttering to be identified as PWS. This would be useful for research into individual profiles, or subtypes, shown by PWS.

6.4 Discussion and Conclusion

The aim of this chapter was to explore the importance of developing self-report measures to aid in the assessment of speech disfluency. The reviewed research in this chapter highlighted that there is an urgent need for a fluency questionnaire with multiple subscales to examine several fluency constructs. Such a questionnaire can potentially allow identification of dysfluent speakers based on self-reports of speech symptoms, or affective concomitants or mixtures of both; thereby advancing our understanding of manifestation of fluency. It should be emphasized that when a questionnaire is developed, it would not be a replacement for established measures used for assessment of fluency in clinical samples. SSI-4, for instance, provides a useful index of fluency difficulty and gives clinicians and researchers detailed information about the symptoms of dysfluent speech. The goal, however, is to complement established measures by developing an instrument to measure fluency on a continuum using a sufficient number of items to provide broad information that benefits clinicians' in tailoring their therapy to clients' specific needs.

Chapter 7 General Discussion

"That speech, and the sounds of the human voice, are invested with a kind of mysterious virtue, which makes them the natural and living expression of thought and feeling.." *Jonathan Rée", I see a voice*

7.1 Introduction

The overall aim of this thesis was to: (1) facilitate screening procedures to identify stuttering and WFD in school aged (4+) children who speak Arabic and/or English; and (2) to provide phonological training interventions to treat WFD in Arabic children who learn English as additional language. To do this, I developed a battery of fluency assessment procedures that takes into account the unique characteristics of the Arabic language. I collected data on these measures, which should contribute to the establishment of norms that could guide clinical diagnosis and intervention with Arabic preschool children.

In the first chapter, I reviewed psycholinguistic models of speech production. I presented alternative viewpoints on how speech disfluencies can affect normally fluent and stuttered speech. This paved the way to the detailed explanation of two specific types of fluency failures: stuttering and WFD. I concluded that fluent speech control is best explained by the EXPLAN model of fluency failure (Howell, 2002; Howell & Au-Yeung, 2001; 2002), which maintains that producing fluent spontaneous speech is associated with appropriate synchrony between language-planning and motor-execution factors. Although not comprehensive, EXPLAN addresses a wide range of topics related to disfluency including how different speech disfluency symptoms are associated with different types of fluency failures. Additionally, it is one of the few fluency models that paid attention to the role of phonological planning in explaining stuttering.

Whilst sensitive phonological performance measures exist such as the UNWR (Howell et al., 2017) that can identify fluency difficulty and WFD in samples of children who speak English as well as children who learn with EAL, this test does not apply to Arabic. It is surprising that no headway has been made towards developing instruments to screen for fluency that takes into account the structure and features of Arabic. In Chapter 2, I presented a brief description of the Arabic language's phonology specifying areas that aided in the generation of Arabic and English non-words. The outcome of this chapter was a list of twenty-seven non-words that forms the AEN_NWR. The stimuli conform to accepted standards for NWR tasks including the following: language-specific phonotactic constraints (Arabic and English), avoiding later-developing consonants, and minimizing potential resemblance between real words and nonwords. Furthermore, the test does not require knowledge of lexical semantics for either of the two languages.

The first empirical chapter (Chapter 3) revolved around establishing a speech-based instrument (Arabic fluency assessment) that took into account the phonological and morphological features of Arabic. In this chapter, I reported results of an empirical study in which the instrument (A_FS) was administered to Arabic speaking adults and children who stutter. I proposed two algorithms to assess fluency: one for counting the number of syllables and one for counting disfluencies.

In Chapter 4, I validated the AEN_NWR in a group of children and adult Arabic speakers who have been diagnosed as displaying stuttering symptoms by SLTs. I showed that participants who showed a higher percentage of stuttered syllables/dysfluent events, which indicates less fluent speech within their speech sample, scored lower on the AEN_NWR. This supports the idea that the AEN_NWR is a reliable measure of phonological skills and hence of fluency of speech.

In the third empirical chapter, Chapter 5, I presented an innovative form of intervention for Arabic children with EAL suffering from WFD. The technique involved providing phonological training with non-words, which focussed on the difficult structures of English, in which an Arabic child would be expected to know if they were to become proficient in English. In the first part of the chapter, I identified relevant language-specific factors and described how the non-words were designed and provided lists of the stimuli. In the second part of the chapter, I reported results of an experiment in which I administered the intervention to two groups of children. The hypothesis I tested was that phonological training with such material should prime a child's use of unfamiliar phonological networks and improve lexical access. This intervention was hoped to extend to vocabulary acquisition, literacy and to eventually build confidence with the wider educational community.

In Chapter 6, I explored the importance of developing self-report instruments to assess speech fluency; ones that meet the DSM-5 criterion. In the next section, I summarize the main findings, discuss theoretical and clinical implications of the results, and suggest directions for future research.

7.2 Report of main findings

Examining the relationship between stuttered episodes and different linguistic factors in Arabic is a complex task. The A_FS (see Chapter 3) was built on phonological principles of Arabic language properties. Results from analysing the spontaneous speech extracts yielded interesting findings that helped in learning about disfluency characteristics of Arabic that are associated with the language unique structure. Extracts of spontaneous speech samples were analysed using parameters of an Arabic index of phonemic complexity (AIPC) to learn about disfluency characteristics of Arabic that are associated with the language unique structure. Independent sample t-tests showed that function-content words have significantly higher AIPC scores than content words, which in turn have significantly higher AIPC scores than function words. When examining the significant effects of the AIPC factors on stuttering, the consonant by length factor represented in geminate consonants appeared to influence stuttering in both age groups; children and adults. Its impact on stuttered words in the child group was higher than in the adult groups. This finding validates previous finding by Al-Tamimi et al. (2013) on the impact of gemination on stuttering. Pronouncing geminate consonants requires a conscious effort, relative to singleton counterparts, to make them perceptible. This puts an extra challenge for PWS particularly when the geminate occurs word-medially. A geminated consonant is pronounced almost twice as long as the singleton length, which shows that there is a phonemic difference between the geminate and the corresponding single consonant (Hedia & Plag, 2017). The difficulty in producing those consonants might be attributed to moving and maintaining the articulators to retain a longer occlusion time. The act of holding the articulators might indicate the high articulatory effort it takes to produce a geminated consonant compared to its single counterpart (Al-Tamimi, 2004). Thus, this might explain why gemination attracted more disfluency in the form of repetition, prolongation and breaks. In all three forms, disfluency might occur on the geminated cluster itself or prior to it, as the speaker plans to produce the gemination.

It was wise to analyse word types for the three different categories separately in order not to mask predictors of either type when doing a combined analysis. This can be explained in terms of, for both age groups, function-content and content words appeared to behave differently in relation to stuttering. Additionally, this provides support to the interpretation that disfluencies on function words are qualitatively different that disfluencies on content words; and that the differences are not influenced by linguistic properties (Dworzynski, 2004). Most function words are either one- or two-syllable words and so when looking at the word length factor they are always considered 'easy' words. Also, these are limited in number and have high frequency in the language, which might affect the chance of them being stuttered.

Chapter four addressed the issue of validating the AEN_NWR as a measure of phonological skills for participants with stuttering symptoms and hence of fluency of speech. The AEN_NWR makes a good alternative to speech-based procedures for assessing children in schools; and it can equally test

Arabic and English children. All participants who were tested in this chapter were referrals from SLTs, and there was a significant association between their %SS and AEN_NWR scores. Performance on the AEN_NWR should not be affected in children with EAL who have WFD because they often have to produce untrained English phoneme sequences, which have the same demands as the non-word stimuli (Howell et al., 2017). This was not tested in this chapter, but it would be interesting to recruit a matched control group in future work to provide additional support on the validity of the AEN_NWR as a sensitive measures of speech disfluency.

Chapter four also reported preliminary efforts in starting to validate A_FS to establish it as a new measure that can objectively analyse connected speech to identify distinct types of speech symptoms as they occur in Arabic. The hypotheses that A_FS scores would correlate negatively with AEN_NWR scores was supported by the data. Although the relationship between A_FS and the AEN_NWR was statistically significant, it was not strong to demonstrate high validity levels. As mentioned in Chapter 4, the sample size of the study was small to draw general conclusions. Also, the English SSI in itself, although a widely used stuttering measure, has many documented limitations. Most of these limitations concern the manual's reporting of the instrument's inter-judge and intra-judge reliability across the subscores, total score, and final severity score (Davidow & Scott, 2017). The authors advise that future studies should consider collecting data

regarding whether or not additional re-playing of the recorded speech files improve reliability; and whether or not this may be important for enhancing the clinical and research value of the SSI-4

Thus, as the intention is to establish a new speech-based instrument to assess fluency in Arabic, our focus would be on establishing new procedures, data collection methods and clear guidelines. Davidow and Scott (2017) proposed the collection of speech naturalness data as an alternative to the final severity score produced by SSI. Speech naturalness would be performed on a 9-point scale where 1 is highly natural and 9 is highly unnatural. The authors argued that using the 9-point scale allows for a more fine-grained measurement compared to the stuttering severity score on SSI-4. Again, there is no data on the psychometric properties of this measure in Arabic, which is an area that merits investigation in future studies.

Statistical analyses of the data in the WFD treatment study (Chapter 5) have shown that children generally performed better post interventions; they improved in a range of skills over time (See Table 5.9 for a summary of the significant effects). Taking all findings together, it can be seen that phonological training is beneficial for all children but receiving language-specific material might even be more beneficial for children with EAL. However, it remains unclear whether language-specific materials are always needed to improve word-finding for preschool children. This gives a direction for future studies in which control materials should be used to examine the importance of using language-specific materials.

7.3 Summary and Implications

The implication from the first empirical chapter (Chapter 3) was that subtle linguistic differences in linguistic structure could have an effect on the pattern of disfluencies observed in spontaneous speech.

Furthermore, there were differences in the types of disfluencies on words of different grammatical classes. Phonetic difficulty of words from grammatical classes was also observed to affect stuttering; This highlights a complex structure of Arabic content and function-content words. These results can be accounted for by the EXPLAN model, assuming that lexical types share common characteristics that are associated with fluency failures across the languages. The EXPLAN model makes specific reference to developmental changes in disfluency patterns and specifically highlights points of difficulties where fluency is more likely to break down, which is not the case in other theoretical models that account for stuttering.

Another notable factor of EXPLAN is the implicit assumption of the continuity between stuttered and fluent speech. Despite that, it emphasizes that authors should not include the wrong disfluency symptoms when assessing stuttering (Howell, 2010). Whole-word repetition and hesitations are not a symptom of stuttering and children who exhibit this type of disfluency should not be classified as stuttering. In other words, WWR of function words is not problematic and it might not be necessary to refer children immediately into a treatment program when they show these disfluencies (Howell, 2004). Eventually, this pattern of disfluency changes over development and children who stall on function words (see Chapter 1 for detailed explanation) have a high chance of recovery. Children repeat whole words to buy time to generate plan of a successive word to produce fluent spontaneous speech. In the case of children with EAL, they have two sets of plans that they have to feed through the execution system. Work on older speakers with EAL has commenced

(Howell et., 2004) in Spanish. Stuttering rates of function and contents words in English and Spanish of a Spanish speaker with EAL were examined. The speaker showed more part-word repetitions on content words in his first language Spanish and more whole function word repetitions English. It would be of interest to replicate this study in Arabic and examine the exchange patterns of stuttering in Arabic speakers with EAL.

The continuity of fluency between TS and PWS was examined in Chapter 6 and a proposal was made regarding the development of self-report measures of speech disfluency. The idea was to have a screening instrument that can apply to the entire population; not just PWS. The literature review provided insights on the development of a fluency questionnaire, to bridge the gap of the lack of self-report instruments that meet the new DSM criterion. Online platforms to administer the questionnaire can be made available online; this makes them accessible to a wide range of PWS, ones who have been clinically diagnosed and ones who self-report stuttering symptoms as well as to the normal populations of speakers.

7.4 Problems Encountered and Future Directions

7.4.1 Problems encountered

In the first part of the thesis the main problem encountered was the fact that in many Arabic countries, including Saudi Arabia, there is less awareness about the stuttering, the importance of early diagnosis and intervention. Moreover, there is a stigma associated with stuttering which may lead some PWS to internalize their stuttering. This made it extremely difficult to obtain speech samples from PWS, particularly from female participants. Eventually research links with several stuttering support groups in Saudi Arabia and other Arabian Gulf countries as well as SLPs who work with PWS were established. This provided the main breakthrough and only then was it possible to recruit a larger sample size to validate the AEN_NWR (Chapter 4). The sample in Chapter 4 was also less gender-biased and included males and females as well as children and adults, which allowed examination of developmental changes in disfluency types.

The design of new tests for children with EAL brings with it a number of problems, too. This is especially true in a language where there is, unfortunately, paucity in assessment tools, published standardized tests or normative data all of which apply to Arabic. To date, there have been no studies on how the unique phonetic, phonological and morphological factors of Arabic are associated with stuttering apart from Al-Tamimi et al. (2013) and this thesis. This has led many researchers and clinicians to translate assessments that were developed and normed for English language and culture without taking into account linguistic and cultural differences. The issue is that features of the structure of Arabic affect syllable (and word) counts and stuttered events in English do not correspond directly with the patterns seen in Arabic. Therefore, a straight translation for application to Arabic before re-standardization is not appropriate. For an instrument to be employed in clinics and research, it first needs to be translated to the target language and the translation needs to be checked. Then the instrument has to be validated and standard scores need to be obtained (Karimi, Nilipour, Shafiei &

Howell, 2011). Furthermore, a limitation in administering the AEN_NWR in its current form is that it requires live scoring by humans who are – ideally – phonologically trained in Arabic and English to permit perceiving subtle differences in speech production. Although judgments in live scoring are made in a systematic fashion, they may not always reflect the true performance of an individual. This can be particularly harmful if the AEN_NWR was used as diagnostic tools, and misclassifications were made. This is a limitation that applies to the AEN_NWR as well as other NWR available at present that identify other communicative disorders. Further reliability and validity analyses are needed for the AEN_NWR, quality phonetic transcription also should be assured, and an even bigger dataset should be collected.

Treatment of WFD was limited by the lack of multiple intervention sessions, which was mainly due to time-constraints imposed by schools. During school days, both in Saudi Arabia and in the UK, it was not possible to conduct lengthy testing with preschool children for more than 20 minutes. It was even more difficult to keep children outside their classrooms for multiple testing to avoid them missing on important activities.

7.4.2 Future Directions

Additional work is needed on the linguistic patterns of stuttering that would test the current findings in more detail and expand on them. Ideally suited for this would be test cases of monolingual Arabic as well as Arabic speakers with EAL (adults and children) who stutter. It would be interesting to see whether individuals would show the same patterns that were examined here. Would these speakers show a pattern where they would stutter almost exclusively in Arabic on function-content words and content words, whereas in comparison also stutter on English function words, too. Additionally, if they have one dominant language what would be the pattern of stuttering in the two languages. More case studies are still needed to clarify this within the same individual; and to provide further understanding of how stuttering presents across Arabic and English. In relation to that, the literature on bilingualism and stuttering are generally limited; they are also inconsistent in terms of issues like participants' number, age groups and the languages spoken by participants (Van Borsel et al., 2001) This makes comparing studies or drawing a general conclusion on stuttering and bilingualism a difficult task. In a recent study, Gkalitsiou et al . (2017) investigated the loci of stuttering in young Spanish-English bilinguals. The authors reported more stuttering on function than content words in the participants' spontaneous speech samples. Examining the loci of stuttering in Arabic-English CWS warrant future investigations. It would be of interest then also to see if those bilinguals stutter more in their more proficient language.

Our results also indicate that there is still a lot more to find out about the processes that take place at the time of stuttering onset and what exactly leads to the developmental age changes in the symptoms of stuttering. Ideally suited to this question are longitudinal studies. It would be ideal to test a large group of children that would have a predisposition to stuttering, possibly due to familial history of stuttering (Dworzynski, 2004). Then, individual profiles can then be analysed to see whether there is anything that distinguishes children who develop stuttering from children who remain fluent.

Another interesting future direction in this research area would be the possibility to develop procedures to automate scoring of nonwords; this would be much more time- and labour efficient (Barrett, Hu & Howell, submitted). One way this can be started is by building an inventory of phonemes that children have difficulty in articulating, so that scoring of AEN_NWR can be improved, and reliable results can be obtained. It is worth noting that no known studies have looked at factors which might influence the live scoring of non-words in particular. The current data of AEN_NWR can also be re-analysed based on demographics. It would be interesting to observe if there are any differences in repetitions of non-words based on gender, language background or dialect.

References

Abdoh, E. M. A. (2011). A study of the phonological structure and representation of first words in Arabic (Doctoral dissertation, University of Leicester).

Abu-Chacra, F. (2007). Arabic: An essential grammar. Routledge.

Abu-Rabia, S. (2012). The role of morphology and short vowelization in reading morphological complex words in Arabic: Evidence for the domination of the morpheme/root-based theory in reading Arabic. *Creative Education*, *3* (04), 486.

Al Tamimi, Y. A., & Al Shboul, Y. (2013). Is the phonotactics of the Arabic complex coda sonority-based?. *Journal of King Saud University-Languages and Translation*, *25*(1), 21-33.

Al-Ani, S. H. (1970). An Acoustical and Physiological Investigation of the Arabic/E/.

Algeo, J. (1978). What consonant clusters are possible?. *Word*, *29*(3), 206-224.

Alm, P. A. (2005). *On the causal mechanisms of stuttering*. Lund University.

Alqattan, S. (2015). *Early phonological acquisition by Kuwaiti Arabic children* (Doctoral dissertation, Newcastle University).

Alqhazo, M., & Al-Dennawi, S. (2018). The linguistic aspects of the speech of Jordanian children who stutter. *International Journal of Pediatric Otorhinolaryngology*, 109, 174-179.

Alshangiti, WMM; (2015) Speech production and perception in adult Arabic learners of English: A comparative study of the role of production and perception training in the acquisition of British English vowels. Doctoral thesis , UCL (University College London).

Al-Tamimi, F. Y. (2004). An experimental phonetic study of intervocalic singleton and geminate sonorants in Jordanian Arabic. *al-'Arabiyya*, 37-52.

Al-Tamimi, F., Khamaiseh, Z., & Howell, P. (2013). Phonetic complexity and stuttering in Arabic. *Clinical Linguistics & Phonetics*, *27*(12), 874-887.

Amayreh, M. M. (2003). Completion of the consonant inventory of Arabic. *Journal of Speech, Language, and Hearing Research*, 46(3), 517-529.

Amayreh, M. M., & Dyson, A. T. (1998). The acquisition of Arabic consonants. *Journal of Speech, Language, and Hearing Research*, *41*(3), 642-653.

Ambrose, N. G., & Yairi, E. (1999). Normative disfluency data for early childhood stuttering. *Journal of Speech, Language, and Hearing Research*, *42*(4), 895-909.

Anderson, J. D. (2007).Phonological neighborhood and word frequency effects in the stuttered disfluencies of children who stutter. *Journal of Speech, Language, and Hearing Research*, 50, 229–247

Anderson, J. D., & Wagovich, S. A. (2010). Relationships among linguistic processing speed, phonological working memory, and attention in children who stutter. *Journal of fluency disorders*, *35*(3), 216-234.

Anderson, J. D., Wagovich, S. A., & Hall, N. E. (2006). Nonword repetition skills in young children who do and do not stutter. *Journal of Fluency Disorders*, *31*(3), 177-199.

Andrews, G., & Harris, M. (1964). *The syndrome of stuttering.* Spastics Society Medical Education.

Archibald, J. (1998). Second language phonology, phonetics, and typology. Studies in Second Language Acquisition, 189-211.

Archibald, L. M., & Gathercole, S. E. (2007). Nonword repetition in specific language impairment: More than a phonological short-term memory deficit. *Psychonomic Bulletin & Review*, *14*(5), 919-924.

Au-Yeung, Gomez, I. V., & Howell, P. (2003). Exchange of Disfluency With Age From Function Words to Content Words in Spanish Speakers Who Stutter. Journal of Speech, Language, and Hearing Research, 46(3), 754–765.

Baddeley, A. (1986). Working memory / Alan Baddeley (Oxford psychology series ; no. 11). Oxford: Clarendon.

Baddeley, A. (1992). Working memory. *Science*, *255* (5044), 556-559.

Baddeley, A. D. (1996). The concept of working memory. *Models of short-term memory*, 1-27.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47-89). Academic press.

Bakhtiar, M., Seifpanahi, S., Ansari, H., Ghanadzade, M., & Packman, A. (2010). Investigation of the reliability of the SSI-3 for preschool Persian-speaking children who stutter. *Journal of Fluency Disorders*, 35(2), 87-91.

Barrett, L. Hu, J. & Howell, P. (submitted). Systematic Review of Machine Learning Approaches for Detecting Developmental Stuttering. *IEEE/ACM Transactions on Audio, Speech and Language, Processing.*

Bercow, J. (2008). The Bercow Report: A Review of Services for Children and Young People (0--19) with Speech. *Language and Communication Needs. Nottingham: DCSF Publications.*

Bernhardt, B. M., Másdóttir, T., Stemberger, J. P., Leonhardt, L., & Hansson, G. Ó. (2015). Fricative acquisition in English-and Icelandic-speaking preschoolers with protracted phonological development. *Clinical Linguistics & Phonetics*, 29(8-10), 642-665.

Best, W. (2005). Investigation of a new intervention for children with word-finding problems. *International Journal of Language & Communication Disorders*, *40*(3), 279-318.

Best, W., Hughes, L. M., Masterson, J., Thomas, M., Fedor, A., Roncoli, S., & Kapikian, A. (2018). Intervention for children with word-finding difficulties: a parallel group randomised control trial. *International Journal of Speech-Language Pathology*, *20*(7), 708-719.

Bloodstein, O. (1970). Stuttering and normal nonfluency a continuity hypothesis. *British Journal of Disorders of Communication*, *5*(1), 30-39.

Bosshardt, H.-G. (1990). Subvocalization and reading rate differences between stuttering and nonstuttering children and adults. *Journal of Speech and Hearing Research*, *33*, 776–785.

Boudelaa, S., & Marslen-Wilson, W. D. (2005). Discontinuous morphology in time: Incremental masked priming in Arabic. *Language and Cognitive Processes*, *20*(1-2), 207-260.

Boudelaa, S., & Marslen-Wilson, W. D. (2010). Aralex: A lexical database for modern standard Arabic. *Behavior Research Methods*, 42(2), 481-487.

Boudelaa, S., & Marslen-Wilson, W. D. (2015). Structure, form, and meaning in the mental lexicon: Evidence from Arabic. *Language, Cognition and Neuroscience*, 30(8), 955-992.

Bragard, A., Schelstraete, M. A., Snyers, P., & James, D. G. (2012). Word-finding intervention for children with specific language impairment: A multiple single-case study.

Broom, Y. M., & Doctor, E. A. (1995). Developmental phonological dyslexia: A case study of the efficacy of a remediation programme. *Cognitive Neuropsychology*, *12*(7), 725-766.

Byrd, C. T., Vallely, M., Anderson, J. D., & Sussman, H. (2012). Nonword repetition and phoneme elision in adults who do and do not stutter. *Journal of Fluency Disorders, 37*(3), 188–201.

Caramazza, A., & Hillis, A. E. (1990). Where do semantic errors come from?. *Cortex*, *26*(1), 95-122.

Chen, X., & Meurers, D. (2018). Word frequency and readability: Predicting the text-level readability with a lexical-level attribute. *Journal of Research in Reading*, *41*(3), 486-510.

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.

Conti-Ramsden, G., & Botting, N. (1999). Classification of children with specific language impairment: Longitudinal considerations. Journal of Speech, Language, and Hearing Research, 42(5), 1195-1204.

Conture, E. G., Caruso, A. J., Rustin, L., Purser, H., & Rowley, D. (1987). Assessment and diagnosis of childhood disfluency. *Progress in the Treatment of Fluency Disorders*, 57-82.

Daland, R., Hayes, B., White, J., Garellek, M., Davis, A., & Norrmann, I. (2011). Explaining sonority projection effects. *Phonology*, *28*(2), 197-234.

Daneman, M. (1991). Working memory as a predictor of verbal fluency. *Journal of Psycholinguistic research*, *20* (6), 445-464.

Daneman, M., & Green, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language, 25*, 1–18.

Davidow, J. H., & Scott, K. A. (2017). Intrajudge and Interjudge Reliability of the Stuttering Severity Instrument–Fourth Edition. *American Journal of Speech-Language Pathology*, *26*(4), 1105-1119.

De Houwer, A. (2011). The speech of fluent child bilinguals. *In Multilingual aspects of fluency disorder*s (pp. 3-23). Multilingual Matters.

Dockrell, J. E., Messer, D., & George, R. (2001). Patterns of naming objects and actions in children with word finding difficulties. *Language and Cognitive processes*, *16*(2-3), 261-286.

Dockrell, J. E., Messer, D., George, R., & Wilson, G. (1998). Children with word-finding difficulties—Prevalence, presentation and naming problems. *International Journal of Language and Communication Disorders*, 33, 445–454.

Dworzynski, K. (2004). German and English comparison of fluency development and stuttering. University of London, University College London (United Kingdom).

Dworzynski, K., & Howell, P. (2004). Predicting stuttering from phonetic complexity in German. *Journal of Fluency Disorders*, *29* (2), 149-173.

Ebbels, S. (2014). Effectiveness of intervention for grammar in school-aged children with primary language impairments: A review of the evidence. *Child Language Teaching and Therapy*, *30*(1), 7-40.

Eichorn, N., Marton, K., Schwartz, R. G., Melara, R. D., & Pirutinsky, S. (2016). Does working memory enhance or interfere with speech fluency in adults who do and do not stutter? Evidence from a dual-task paradigm. *Journal of Speech, Language, and Hearing Research*, *59*(3), 415-429.

Elmahdi, O. E. H., & Khan, W. A. (2015). The pronunciation problems faced by Saudi EFL learners at secondary schools. Education and Linguistics Research, 1(2), 85-99.

Faust, M., Dimitrovsky, L., & Shacht, T. (2003). Naming Difficulties in Children with Dyslexia: Application of the Tip-of-the-Tongue Paradigm. Journal of Learning Disabilities, 36(3), 203–215.

Ferguson, C. A. (1959). Diglossia. word, 15(2), 325-340.

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, *27*(4), 513–543.

Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language*, *28*(2), 200-213.

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.

Gathercole, S.E. (2008). Working Memory. In J. Byrne (Ed.), Learning and Memory: A Comprehensive Reference, 2, 1st Edition, 33–51. Oxford: Academic Press.

German, D. J. (1991). Test of word finding in discourse. *Austin, TX: Pro-Ed.*

German, D. J. (2000). *TWF-3: Test of Word Finding*. Pro-Ed.

German, D. J. (2002). A phonologically based strategy to improve word-finding abilities in children. Communication Disorders Quarterly, 23(4), 177-190.

German, D. J., & Simon, E. (1991). Analysis of children's word-finding skills in discourse. *Journal of Speech, Language, and Hearing Research*, *34*(2), 309-316.

Gkalitsiou, Z., Byrd, C. T., Bedore, L. M., & Taliancich-Klinger, C. L. (2017). Stuttering on function words in bilingual children who stutter: A preliminary study. *Clinical linguistics & phonetics*, *31*(10), 791-805.

Glaser, W. R. (1992). Picture naming. *Cognition*, *42* (1-3), 61-105.

Gollan, T. H., & Silverberg, N. B. (2001). Tip-of-the-tongue states in Hebrew–English bilinguals. *Bilingualism: Language and Cognition*, *4*(1), 63-83.

Hakansson, G., Salameh, E. K., & Nettelbladt, U. (2003). Measuring language development in bilingual children: Swedish-Arabic children with and without language impairment. Linguistics, 41(2; ISSU 384), 255-288.

Hakim, H. B., and Ratner, N. B. (2004). Nonword repetition abilities of children who stutter: an exploratory study. *Journal of Fluency Disorders*. 29, 179–199.

Harley, T. A., & MacAndrew, S. B. (2001). Constraints upon word substitution speech errors. *Journal of Psycholinguistic Research*, *30*(4), 395-418..

Harris, J. (1990). Derived phonological contrasts. *Studies in the pronunciation of English: A commemorative volume in honour of AC Gimson*, *87*, 105.

Harris, J. (1994). English sound structure. Oxford, Blackwell.

Harris, J., & Gussmann, E. (2002). Word-final onsets. *UCL Working Papers in Linguistics*, *14*, 1-42.

Hassan, S. M., Sallam, Y. A. N., Khodeir, M. S., & Mahmoud, O. S. (2017). Correlation between stuttering severity and pragmatic development in Egyptian children who stutter. *The Egyptian Journal of Otolaryngology*, *33*(4), 670-678.

Hedia, S. B., & Plag, I. (2017). Gemination and degemination in English prefixation: Phonetic evidence for morphological organization. *Journal of Phonetics*, *62*, 34-49.

Homma, T., & Yamada, J. (2014). High Durational Variability of Consonant Geminates and Their Surrounding Vowels in Stuttering Japanese Speakers' Fluent Speech. *Canadian Journal of Speech-Language Pathology & Audiology*, *38*(1).

Howell P, Chua LY, Yoshikawa K, Tang H, Welmillage T, Harris J and Tang K (2020). Does Working-Memory Training Given to Reception-Class Children Improve the Speech of Children at Risk of Fluency Difficulty? *Frontiers in Psychology.* 11:568867.

Howell, P. (2004). Assessment of some contemporary theories of stuttering that apply to spontaneous speech. *Contemporary issues in communication science and disorders*, *31*(Spring), 123-140.

Howell, P. (2007). Signs of developmental stuttering up to age eight and at 12 plus. Clinical Psychology Review, 27(3), 287-306. doi: 10.1016/j.cpr.2006.08.005

Howell, P. (2010). Recovery from Stuttering. New York, NY: Psychology Press.

Howell, P., & Au-Yeung, J. (2002). The EXPLAN theory of fluency control applied to the diagnosis of stuttering. *AMSTERDAM STUDIES IN THE THEORY AND HISTORY OF LINGUISTIC SCIENCE SERIES 4*, 75-94.

Howell, P., & Rusbridge, S. (2011). The speech and language characteristics of developmental stuttering in English speakers. *Multilingual Aspects of Fluency Disorders*, *5*, 93.

Howell, P., Au-Yeung, J., & Sackin, S. (1999). Exchange of stuttering from function words to content words with age. *Journal of Speech, Language, and Hearing Research*, *42*(2), 345-354.

Howell, P., Davis, S., & Williams, R. (2009). The effects of bilingualism on stuttering during late childhood. *Archives of disease in childhood*, *94*(1), 42-46.

Howell, P., Ruffle, L., Fernandez-Zuniga, A., Gutiérrez, R., Fernandez, A. H., O'Brien, M. L., ... & Au-Yeung, J. (2004). Comparison of exchange patterns of stuttering in Spanish and English monolingual speakers and a bilingual Spanish-English speaker. *Theory, research and therapy in fluency disorders*, 415-422.

Howell, P., Soukup-Ascencao, T., Davis, S., & Rusbridge, S. (2011). Comparison of alternative methods for obtaining severity scores of the speech of people who stutter. *Clinical Linguistics & Phonetics*, *25*(5), 368-378.

Howell, P., Tang, K., Tuomainen, O., Chan, K., Beltran, K., Mirawdeli, A. & Harris, J. (2017). Identification of fluency and word-finding difficulty in samples of children with diverse language backgrounds. *International Journal of Language & Communication Disorders, 52,* 595-611.

Hubbard, C. P. (1998). Stuttering, stressed syllables, and word onsets. *Journal of Speech, Language, and Hearing Research*, *41*(4), 802-808.

Howell, P., Au-Yeung, J., Yaruss, S. J., & Eldridge, K. (2006). Phonetic difficulty and stuttering in English. *Clinical linguistics & phonetics*, *20*(9), 703-716.

Huthaily, K. (2003). Contrastive Phonological Analysis of Arabic and English [Master's Thesis]. University of Montana.

Ibrahim, M. A. (2016). The CV Treatment of Consonantal Gemintes in Arabic: A Non-Linear Approach. *The Buckingham Journal Of Language and Linguistics*, *9*, 21-31.

Ingham, R. J., & Cordes, A. K. (1997). Identifying the authoritative judgments of stuttering: Comparisons of self-judgments and observer judgments. *Journal of Speech, Language, and Hearing Research*, *40*(3), 581-594.

Icht, M., & Ben-David, B. M. (2014). Oral-diadochokinesis rates across languages: English and Hebrew norms. *Journal of Communication Disorders*, *48*, 27-37.

Iverach, L., & Rapee, R. M. (2014). Social anxiety disorder and stuttering: Current status and future directions. *Journal of Fluency Disorders*, 40, 69-82.

Jakielski, K. J. (1998). Motor organization in the acquisition of consonant clusters. PhD Thesis, University of Texas at Austin. Ann Arbor Michigan, UMI Dissertation services.

Jiang, J., Lu, C., Peng, D., Zhu, C., & Howell, P. (2012). Classification of types of stuttering symptoms based on brain activity. PloS one, 7(6), e39747.

Johnson, W. A. L. (1959). *The Onset of Stuttering. Research Findings and Implications.[By] W. Johnson and Associates*. University of Minnesota Press.

Kadri, Y., & Nie, J. Y. (2006). Effective stemming for Arabic information retrieval. In *Proceedings of the International Conference on the Challenge of Arabic for NLP/MT*.

Karimi, H., Nilipour, R., Shafiei, B., & Howell, P. (2011). Translation, assessment and deployment of stuttering instruments into different languages: Comments arising from Bakhtiar et al., "Investigation of the reliability of the SSI-3 for preschool Persian-speaking children who stutter" [J. Fluency Disord. 35 (2010) 87–91]. *Journal of Fluency Disorders*, 36, 246–248.

Kaye, J. (1990). 'Coda'licensing. *Phonology*, *7*(1), 301-330.

Kaye, J. (2013). *Phonology: A cognitive view*. Routledge.

Kristjansson, E. A., Desrochers, A., & Zumbo, B. (2003). Designer's Corner-Translating and Adapting Measurement Instruments for Cross-Linguistic and Cross-Cultural Research: A Guide for Practitioners. *Canadian Journal of Nursing Research Archive*, 127-142.

Levelt, W. J. (1992). Accessing words in speech production: Stages, processes and representations. Cognition, 42(1-3), 1-22.

Levelt, W. J. (1999). Models of word production. *Trends in cognitive sciences*, *3* (6), 223-232.

Levelt, W. J. (2001). Spoken word production: A theory of lexical access. Proceedings of the National Academy of Sciences, 98(23), 13464-13471.

Levelt, W. J. M. (1983). Monitoring in self-repair in speech. *Cognition, 14,* 41-104.

Mahr, G. C., & Torosian, T. (1999). Anxiety and social phobia in stuttering. *Journal of Fluency Disorders*, *24*(2), 119-126.

Mansson, H. (2000). Childhood stuttering: Incidence and development. *Journal of Fluency Disorders*, *25*(1), 47-57.

Masterson, J., Druks, J., & Gallienne, D. (2008). Object and action picture naming in three-and five-year-old children. *Journal of Child Language*, *35* (2), 373-402.

McGowan, R. S., & Nittrouer, S. (1988). Differences in fricative production between children and adults: evidence from an acoustic analysis of /ʃ/and/s. *The Journal of the Acoustical Society of America*, *83*(1), 229-236.

McGregor, K. K. (1997). The nature of word-finding errors of preschoolers with and without word-finding deficits. Journal of Speech, Language, and Hearing Research, 40(6), 1232-1244.

Messenger, M., Onslow, M., Packman, A., & Menzies, R. (2004). Social anxiety in stuttering: Measuring negative social expectancies. *Journal of Fluency Disorders*, *29*(3), 201-212.

Messer, D., & Dockrell, J. E. (2006). Children's naming and word-finding difficulties: Descriptions and explanations.

Migliaccio, R., Boutet, C., Valabregue, R., Ferrieux, S., Nogues, M., Lehéricy, S., & Teichmann, M. (2016). The brain network of naming: a lesson from primary progressive aphasia. PloS one, 11(2), e0148707.

Mines, M. A., Hanson, B. F., & Shoup, J. E. (1978). Frequency of occurrence of phonemes in conversational English. *Language and speech*, *21*(3), 221-241.

Mirawdeli, A. (2015). Identifying children who stutter or have other difficulties in speech production in school reception classes. Procedia-Social and Behavioral Sciences, 193, 192-201.

Mirawdeli, A., & Howell, P. (2016). Is it necessary to assess fluent symptoms, duration of dysfluent events, and physical concomitants when identifying children who have speech difficulties?. Clinical linguistics & phonetics, 30(9), 696-719. *of Language & Communication Disorders*, 33(4), 445-454.

Munson, B., Kurtz, B., & Windsor, J. (2005). The Influence of Vocabulary Size, Phonotactic Probability, and Wordlikeness on Nonword Repetitions of Children With and Without Specific Language Impairment. *Journal of Speech, Language, and Hearing Research*, 48(5), 1033-1047.

phonology. Wiley-Blackwell , Oxford , 2990 – 3019

Postma, A., & Kolk, H. (1993). The covert repair hypothesis: Prearticulatory repair processes in normal and stuttered disfluencies. *Journal of Speech, Language, and Hearing Research*, *36*(3), 472-487.

Rad, S. K., & Abdullah, A. N. (2012). Effect of context on types of hesitation strategies used by Iranian EFL learners in L2 oral language tests. *English language teaching*, *5*(7), 102.

Raine, A. (1991). The SPQ: a scale for the assessment of schizotypal personality based on DSM-III-R criteria. Schizophrenia bulletin, 17(4), 555-564.

Richels, C., Buhr, A., Conture, E., & Ntourou, K. (2010). Utterance complexity and stuttering on function words in preschool-age children who stutter. Journal of fluency disorders, 35(3), 314-331.

Riley, G. (1994). Stuttering severity instrument for young children (SSI-3). *Stuttering Severity Instrument for Young Children (SSI-3)*.

Riley, G. (2009). The stuttering severity instrument for adults and children (SSI-4). Austin, TX:

Riley, G. D. (1972). A stuttering severity instrument for children and adults. Journal of speech and hearing disorders, 37(3), 314-322.

Roberts, P. M. (2011). Methodology matters. In *Multilingual aspects of fluency disorders* (pp. 353-370). Multilingual Matters.

Rohrer, J. D., Knight, W. D., Warren, J. E., Fox, N. C., Rossor, M. N., & Warren, J. D. (2008). Word-finding difficulty: a clinical analysis of the progressive aphasias. Brain, 131(1), 8-38.

Rubin, H., & Liberman, I. Y. (1983). Exploring the oral and written language errors made by language disabled children. Annals of dyslexia, 33(1), 111-120.

Sander, E. K. (1972). When are speech sounds learned?. *Journal of Speech and Hearing Disorders*, 37(1), 55-63.

Sasisekaran, J., & Byrd, C. (2013). Nonword repetition and phoneme elision skills in school-age children who do and do not stutter. *International Journal of Language & Communication Disorders*, *48*(6), 625-639.

Schlenker, B. R., & Leary, M. R. (1985). Social anxiety and communication about the self. *Journal of Language and Social Psychology*, *4*(3-4), 171-192.

Selkirk, E. (1995). Sentence prosody: Intonation, stress, and phrasing. *The Handbook of Phonological Theory*, *1*, 550-569.

Shafiro, V., Levy, E. S., Khamis-Dakwar, R., & Kharkhurin, A. (2013). Perceptual confusions of American-English vowels and consonants by native Arabic bilinguals. *Language and speech*, *56*(2), 145-161.

Sheehan, J. G. (1953). Theory and treatment of stuttering as an approach-avoidance conflict. The Journal of Psychology, 36(1), 27-49.

Sievers, E. (1881). Grundzüge der phonetik (Fundamentals of phonetics). *Leipzig: Breitkopf und Hartel.*

Simonyan, K., Ackermann, H., Chang, E. F., & Greenlee, J. D. (2016). New developments in understanding the complexity of human speech production. *Journal of Neuroscience*, *36*(45), 11440-11448.

Smith, A., Sadagopan, N., Walsh, B., & Weber-Fox, C. (2010). Increasing phonological complexity reveals heightened instability in inter-articulatory coordination in adults who stutter. *Journal of Fluency Disorders*, 35, 1–18

Snow, P. (2014). Oral language competence and the transition to school: Socio-economic and behavioural factors that influence academic and social success. *International Journal on School Disaffection*, 11(1), 3-24.

Snowling, M. J. (1981). Phonemic deficits in developmental dyslexia. *Psychological research*, *43*(2), 219-234.

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

Sugathan, N., & Maruthy, S. (2020). Nonword repetition and identification skills in Kannada speaking school-aged children who do and do not stutter. *Journal of Fluency Disorders*, *63*, 105745.

Telser, E. B. (1971). *An assessment of word finding skills in stuttering and nonstuttering children*. Northwestern University.

Todd, H., Mirawdeli, A., Costelloe, S., Cavenagh, P., Davis, S.., & Howell, P. (2014). Scores on Riley's Stuttering Severity Instrument versions three and four for samples of different length and for different types of speech material. Clinical Linguistics and Phonetics, 28, 912-926.

Tsai, V. P. L., Brundage, S. B., & Ratner, N. B. (2011). Linguistic analysis of stuttering in bilinguals: Methodological challenges and solutions. Multilingual aspects of fluency disorders, 5, 308.

Tucker, M. A. (2011). The morphosyntax of the Arabic verb: Toward a unified syntax-prosody.

Vahab, M., Zandiyan, A., Falahi, M. H., & Howell, P. (2013). Lexical category influences in Persian children who stutter. *Clinical linguistics & phonetics*, *27*(12), 862-873.

Van Borsel, J., Maes, E., & Foulon, S. (2001). Stuttering and bilingualism: A review. *Journal of Fluency Disorders*, *26*(3), 179-205.

Van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176-1190.

Watson, D., & Friend, R. (1969). Measurement of social-evaluative anxiety. *Journal of consulting and clinical psychology*, *33*(4), 448.

Watson, J. C. (2002). *The phonology and morphology of Arabic*. Oxford University Press on Demand.

Watson, JCE (2011) Word stress in Arabic. In: The Blackwell companion to phonology.

Wells, J. C. (1982). Accents of English: Volume 1 (Vol. 1). Cambridge University Press.

Whiteside, S. P. (1996). Temporal-based acoustic-phonetic patterns in read speech: Some evidence for speaker sex differences. *Journal of the International Phonetic Association*, *26*(1), 23-40.

Wilshire, C. E., & Saffran, E. M. (2005). Contrasting effects of phonological priming in aphasic word production. *Cognition*, *95*(1), 31-71.

Winder, R. B., & Ziyadeh, F. J. (2019). *Introduction to Modern Arabic*. Princeton University Press.

Wingate, M. E. (1964). A standard definition of stuttering. Journal of Speech and Hearing Disorders, 29(4), 484-489.

Wingate, M. E. (2002). *Foundations of stuttering*. San Diego: Academic Press.

Wingate, M. E. (2012). The structure of stuttering: A psycholinguistic analysis. Springer Science & Business Media.

Wolf, M. (1991). The word-retrieval deficit hypothesis and developmental dyslexia. *Learning and Individual Differences*, *3*(3), 205-223.

Wright, S. H., Gorrie, B., Haynes, C., & Shipman, A. (1993). What's in a name? Comparative therapy for word-finding difficulties using semantic and phonological approaches. *Child Language Teaching and Therapy*, *9*(3) , 214–229. https://doi.org/10.1177/026565909300900304

Yairi, E. (1997). Disfluency characteristics of childhood stuttering. In R. F. Curlee & G. M. Siegel (Eds.), Nature and treatment of stuttering (2nd ed., pp. 49-78).

Yairi, E., & Ambrose, N. G. (1999). Early childhood stuttering I: Persistency and recovery rates. *Journal of Speech, Language, and Hearing Research*, *42*(5), 1097-1112.

Yairi, E., and Ambrose, N. G. (2005). Early Childhood Stuttering. Austin, TX: Pro-Ed.

Yan, S., & Nicoladis, E. (2009). Finding le mot juste: Differences between bilingual and monolingual children's lexical access in comprehension and production. *Bilingualism: Language and Cognition*, *12*(3), 323-335.

Yaruss, J. S. (2000). Converting between word and syllable counts in children's conversational speech samples. Journal of Fluency Disorders, 25(4), 305-316.

Zangl, R., & Mills, D. L. (2007). Increased brain activity to infant-directed speech in 6-and 13-month-old infants. *Infancy*, *11*(1), 31-62.

APPENDICES

Appendix 1

Orthographic transcription of the AEN_NWR stimuli

Syllable Length	Syllable Template	Orthographic Transcription
2	CV.CVC	sibad
		damif
		fibil
		manib
		tundan
	CVC.CVC	nastim
		bundaf
		nambik
		saftif
	CV.CVCC	takisk
		bamift
3	CV.CV.CVC	danibum
		sifakuf
	CV.CV.CVCC	natadulb

		sigadilk
		lazafusk
	CVC.CVC.CVC	risfundab
		mundaftis
		randintak
	CVC.CVC.CVCC	luntambilf
		rimbandusk
4	CV.CV.CV.CVC	lisakubam
	CVC.CV.CV.CVC	zimtakazum
	CV.CV.CV.CVCC	rifatanult
		dakanufast
	CV.CVC.CV.CVCC	kabaltakɪft
	CVC.CVC.CVC.CVCC	gandisfimbask

Appendix 2

Arabic non-words for native English speakers: Stimuli creation, lists of non-the non-words and corresponding real words.

Arabic non-words for native English speakers

1. Consonant Materials: phonemes that exist in Arabic but not in English

Based on previous studies on the phonological differences between English and Arabic, it seems that the most phonological patterns that are found in Arabic and absent from English are associated with consonant phonemes as well as consonant phonotactics. In this section, will begin by non-words that are designed specifically using phonemes that exist in Arabic but do not exist in English. This will be followed by non-words that have consonant phonemes existing in both English and Arabic but with different phonemic realization. Therefore, the focus will be on training English speakers on phonemic realizations that do are permissible in English. Other patterns of Arabic to be trained on include germination materials occurring word-medially and word-finally. Also, training will focus on Arabic consonant cluster phonotactics including word-final clusters with level and rising sonority.

This list of non-words have been designed specifically to include consonants that are found in Arabic but not in English, with all of them occurring word initially to focus the child's attention on the unfamiliar consonants, which are /tˤ/, /dˤ/ , /ðˤ/, /sˤ/, /x/ , /ɣ/ ,/q/, /ħ/, /ʔ/ and /ʕ/.

Also, all of the non-words are monosyllabic or disyllabic. The consonants that appear in all sequences will always be ones that exists in English and Arabic. The list of the corresponding word materials has some disyllabic words for the purpose of selecting objects that will be of high frequency for the child to recognise. All of the consonant phonemes in those real words, apart from the target phoneme, will be ones that exist in both Arabic and English.

Non-word Material		Real-word Material		English Meaning
Orthographic Transcription	IPA Transcription	Orthographic Transcription	IPA Transcription	
طمل	tˤ aml	طبل	tˤabl	Drum
طيم	tˤ i:m	طين	tˤi:n	Mud
ضمس	dˤims	ضرس	dˤirs	Tooth
قاز	qals	قلب	qalb	Heart
ظم	ðˤɪm	ظل	ðˤil	Shadow
صوبة	sˤu:ba	صورة	sˤu:ra	Picture
خسون	xasu:n	خروف	xaru:f	Sheep
غنسة	ɣan.sa	غرفة	ɣurfa	Room
غسومة	ɣas.su:ma	غسالة	ɣas.sa:la	Washer
حلين	ħali:n	حليب	ħali:b	Milk
عوس	ʕu:s	عود	ʕu:d	Stick
عاس	ʕa:s	عين	ʕayn	Eye
أز	ʔas	أب	ʔab	Father

2. Consonant Materials: phonemes that exist in Arabic and English but have different phonemic realization or different restrictions in English

/ʔ/ is a phoneme in Arabic that can occur word initially, in the middle and end of a syllable. In English, however, [ʔ] is an allophone that is produced by speakers of some English accents instead of [t] in words like beaten (/t/ followed by a nasal). Also, English speakers tend to produce a glottal stop when pronouncing words with vowel initial like am, on and if. Therefore, Training materials here will focus on placing [ʔ] word finally and word internally stressed syllables in disyllabic words. /h/ in Arabic is a voiceless glottal fricative phoneme that can occur word initially, medially and finally. In English, however, its occurrence is restricted the to onset of a stressed syllable (e.g. hat /hæt/ , behind /bɪˈhaɪnd/) while the /h/ in unstressed syllables is usually dropped. Therefore, non-words are designed to train English speakers to produce /h/ in the following two positions: 1) the onset of an unstressed syllable, 2) word-final positions.

Non-word Material		Real-word Material		English Meaning
Orthographic Transcription	IPA Transcription	Orthographic Transcription	IPA Transcription	
بَمه	bamh	وجه	wadʒh	face
فياه	fija:h	مياه	mija:h	Water
تفاه	tifa:h	شفاه	ʃifa:h	Lips
فلاه	fila:h	إله	ʔila:h	God
زاء	za:ʔ	ماء	ma:ʔ	Water
فَوَاء	fama:ʔ	سماء	sama:ʔ	Sky
تَواء	tawa:ʔ	هواء	hawæʔ	Air
فعاء	fiʕa:ʔ	وعاء	wiʕa:ʔ	Pot
مؤوس	moues	كؤوس	kuʔu:s	Glasses
نؤوس	nuʔu:s	رؤوس	ruʔu:s	Heads

3. Gemination Materials occurring word medially

Non-word Material		Real-word Material		English Meaning
Orthographic Transcription	IPA Transcription	Orthographic Transcription	IPA	
لبّاب	lab.bab	دباب	dab.bab	Motorcycle
لمام	lom.ma:n	حمّام	ħam.ma:m	Bathroom
بلاجة	bal.la:dʒa	ثلاجة	θal.ladʒa	Fridge

4. Gemination Materials occurring word finally

Non-word Material		Real-word Material		English Meaning
Orthographic Transcription	IPA Transcription	Orthographic Transcription	IPA Transcription	
نُب	nob.b	دب	dubb	Bear
در	dir.r	سرّ	sir.r	Secret
بم	bam.m	دم	dam.m	Blood

5. Word final cluster with level sonority

The following list of non-words focus on having a word final consonant cluster where the two consonant phonemes are of equal level of sonority. The first set of non-words represent final clusters that has even sonority and both consonants will be ones that exist in English and Arabic. The second set include one consonant phoneme (either the one in the onset or one of the two in the cluster) that exist in Arabic only, assuming the child has developed in fluency after receiving training on that phoneme.

Non-word Material		Real-word Material		English Meaning
Orthographic Transcription	IPA Transcription	Orthographic Transcription	IPA Transcription	
مفس	mafs	نفس	nafs	Self
مسف	masf	نسف	nasf	Destroy
تبد	tabd	كبد	kabd	Liver
لبت	labt	نبت	nabt	Plant
لمن	kamn	سمن	samn	Fat
فعش	faʕʃ	نعش	naʕʃ	Coffin
سفخ	safx	نفخ	nafx	Blow
دحش	daħʃ	وحش	waħʃ	Monster
كغص	kaɣs’	مغص	maɣs’	Colic
مخص	maxs’	شخص	ʃaxs’	Person

6. Word final cluster with rising sonority

A list of non-words that focus on having a word final consonant cluster involving a rising sonority. The first set of non-words represent final clusters that has rising sonority and both consonants will be ones that exist in English and Arabic. The second set include one consonant phoneme (either the one in the onset or one of the two in the cluster) that exist in Arabic only, assuming the child has developed in fluency after receiving training on that phoneme.

Non-word Material		Real-word Material		English Meaning
Orthographic Transcription	IPA Transcription	Orthographic Transcription	IPA Transcription	
فبل	ʃibl	شبل	ʃibl	Baby lion
فُبن	fobn	جبن	dʒobn	Cheese
كمل	kaml	نمل	naml	Bee
نجه	nadʒh	وجه	wadʒh	Face
كهر	kahr	نهر	nahr	River
فبس	fɪbs	لبس	lɪbs	Clothes
دسر	dasr	نسر	nasr	Eagle

مرف	marf	ظرف	ðˤarf	Envelope
لحر	laħr	بحر	baħr	Sea
نَطن	natˤn	بطن	batˤn	Stomach
سفل	sifl	طفل	tˤifl	Child
تَخل	taxl	نخل	naxl	Palm tree
كَهر	kaħr	سحر	siħr	Magic

Appendix 3

Consent forms and Information Sheets (English)

Parent/Guardian Information Sheet

Project Title: Non-word repetition tasks and their application to screening children for word finding and speech difficulty.

Principal Investigator: Professor Peter Howell
Division of Psychology and Language Sciences
Department of Experimental Psychology, University College London
26 Bedford way London WC1H 0AP.
Email: p.howell@ucl.ac.uk.

I am a doctoral student supervised by Professor Peter Howell. I am carrying out a research project that aims to identify children who have word-finding difficulty and to distinguish them from those who have other speech production problems. Earlier work in Professor Howell's speech research lab at UCL has allowed children with speech difficulties to receive treatment before problems become acute.
[School/ clinic name] has agreed to take part in our research study. I am contacting you as your child goes to that school/ clinic and we would like to ask your permission **in order for your child to participate in the study**. Participation in this study is entirely voluntary. Please take the time to read the following information. If you do not wish to have your child participate, please complete and return the form provided below. You do not need to return the form if you agree to your child's participation. If you have any further questions, please contact Professor Howell who will be happy to discuss the project with you.

Thank you for giving this your consideration,
Roa'a AlSulaiman
Doctoral student at University College London

Figure A5.1: English Information Sheet for Parents/Guardians of Children

What is the purpose of the study?
The study aims to identify children with word-finding difficulty and other speech and communication difficulties such as stuttering. This procedure is referred to as screening. Ideally, we want to screen all children in the reception classes in the schools and clinics with which we work. The procedures were designed to work with children who speak English as well as those whose first language is not English. A second aim concerns what teachers/ therapists can do when a child does not pass the screening. We have developed a type of treatment designed for children who have either word-finding difficulty or fluency difficulty.

Does my child have to take part?
No. Participation is voluntary and does not affect your child's education. The decision as to whether or not to take part is up to you and your child. **[If you do NOT wish your child to take part, please sign the attached consent form and return it to the school/clinic].** You can also withdraw your child at any time and all the data collected will be destroyed. Also, your child will only take part in the study if they are happy to do so when we visit (we will never put pressure on children to participate), and your child is free to end the study at any point.

What will my child be asked to do?
A researcher will come and visit your child at school/clinic for approximately 15 minutes. The researcher will speak with your child's teacher/therapist to ensure that the sessions do not clash with any important activities. Each child will be seen individually, will listen to a number of "nonsense" words one by one, and will be asked to repeat the word. Your child will be told that these are funny sounding made up word. The child will be audio-recorded while speaking to the researcher who has been cleared by the UK's Disclosure and Barring Service (a government scheme that checks the credentials of people who work with children). Each audio recording is then analysed anonymously at UCL for signs of fluency or word finding difficulty. All children occasionally show these signs and they are not usually a cause of concern. Parents are then invited to volunteer their child for the experimental fluency or word-finding training. This does not necessarily mean that they have fluency or word-finding problems. A separate information sheet with more details will then be given to you. We will also give you a link to a brief online questionnaire about your child's language that you can fill out at your own time on your phone or computer. No identifiable information is collected and you will be assigned a number that matches your child's participation number (the procedure is anonymised). Only the researchers will have access to the questionnaire. There is no obligation on parents who have been contacted to have their children participate in the training.

What are the benefits of taking part?
This research should help to improve policy and practice in how school and speech-language therapy services address the needs of children with speech production difficulties.

Are there any risks of taking part?
There are no risks beyond those that might arise during normal classroom activities. However, if your child does not enjoy taking part they can stop at any point.

Will details about my child be kept confidential?
Yes. The data collected will be stored anonymously by number. We will not store your child's name or other personal information, which could be used to identify them. All information will be treated as strictly confidential and handled in accordance with the provisions of the General Data Protection Regulation 2018.

What will happen to the results of the research study?
The data will be used as part of a PhD student report. Research publications may arise from the research and the findings may be reported in public presentations, such as at conferences, seminars, and in public engagement activities. The audiences at these events may include policy makers and professionals involved in child education. **No identifying information about any participating child will be disclosed in presentations at any of these outlets.** If you would like to receive a copy of the results once the study is complete, please contact Professor Howell at the address below.

What if I have further questions about the study or have a complaint?
We have not had any complaints in the past and children enjoy taking part. If you have any questions about any aspect of this study please contact: Peter Howell, Division of Psychology and Language Sciences, Department of Experimental Psychology, UCL 26 Bedford way London WC1H 0AP. Email: p.howell@ucl.ac. However, if you are unhappy and wish to complain formally, you can do this by UCL's data protection officer at data-protection@ucl.ac.uk.

ag

Figure A5.2: English Information Sheet for Parents/Guardians of Children

OPT OUT CONSENT FORM

I have read the above information and I **DO NOT** wish my child to take part in this study.

Name of child: ______________ ________________ _____________

Forename Surname Date of birth

_______________ _______________ _______________

Name of Parent Signature Date

Figure A5.3 Opt out form for parents/Gurdians of children

Appendix 4

Consent forms and Information Sheets (Arabic)

UCL

موضوع الدراسة: تطوير مقياس للتمييز بين التأتأة و صعوبة إيجاد الكلمات عند الأطفال

معلومات المشرف على الدراسة :

البروفيسور بيتر هاول
قسم علم النفس و علوم اللغة
كلية لندن الجامعية
البريد الإلكتروني Email: p.howell@ucl.ac.uk

إلى حضرة المسؤول /المسؤولة عن الطفل(ة)

السلام عليكم و رحمة الله ،

أنا طالبة دكتوراه ، و كجزء من دراستي أقوم بتصميم أداة سيتم استخدامها للفصل بين فئتين من الأطفال .
الفئة الأولى : الأطفال الذين يواجهون مشكلة صعوبة إيجاد الكلمات بسبب محدودية حصيلتهم اللغوية. تشيع هذه المشكلة بشكل أوسع عند الأطفال الذين قد يضطرون الى الانتقال الى بلد آخر يحتاجون فيه الى الحديث بلغة مختلفة عن لغتهم الأم .
الفئة الثانية : أطفال قد يواجهون فعلا مشاكل التأتأة و يحتاجون وقتها إلى رؤية أخصائي نطق و تخاطب لإجراء التدخل المناسب.
لتحقيق الهدف , قمنا بتصميم أدوات يتدرب من خلالها الطفل بطريقة ممتعة على أنماط الكلمات التي قد تعد جديدة عليه . لا يتم استخدام كلمات حقيقية هنا بل مجموعة من الحروف التي يجب أن تخضع لنظام لغة معين . الغرض هنا أن تكون هناك أدوات ثبتت فائدتها علميا , سهلة , و سريعة و تطبق على الأطفال بداية التحاقهم بالمدرسة من قبل مدرسيهم.
حصلت على موافقة مركز \مدرسة (مشكورين) لإشراك مراجعي المركز\المدرسة في الدراسة , و آمل الحصول على موافقتكم على مشاركة الطفل. بإمكانكم قراءة معلومات إضافية عن الدراسة و الإجراءات المتبعة خلالها أدناه. تسعدني الإجابة عل أي أسئلة أو استفسارات قد ترد إليكم عبر بريدي roaa.al-sulaiman.17@ucl.ac.uk أو جوال [redacted]

شاكرة و مقدرة تعاونكم

رؤى السليمان
طالبة دكتوراه في جامعة University College London
صفحة المعمل في الجامعة للاطلاع على بعض الدراسات
/https://www.ucl.ac.uk/speech-research-group

Figure A6.1: Arabic Information Sheet for Parents/Guardians of Children

أسئلة و أجوبة عن الدراسة

ما هو الغرض من من هذه الدراسة ؟

تصميم و تقنين أداة تساعد في التفريق بين التأتأة و اضطرابات التواصل الأخرى مثل صعوبة إيجاد الكلمات. الإجراء الذي نتبعه صُمم بحيث يناسب الأطفال المتحدثين باللغة الإنجليزية أو العربية كلغة أولى. الهدف الثاني هو القيام بتدريبات سريعة تحسن من الطلاقة للمساعدة في التغلب على صعوبة إيجاد الكلمات.

هل يجب على طفلي المشاركة ؟

لا. مشاركة الطفل اختيارية تماما ولا تؤثر على أي من الخدمات التي يتلقاها الطفل في المركز. قرار المشاركة يتوقف على رغبة الأهل و رغبة الطفل. أيضا، بإمكان الطفل الانسحاب في أي وقت دون ذكر أسباب و عندها ستقوم بإتلاف أي بيانات جمعت من الطفل. أيضا, لن نحاول إجبار الطفل على المشاركة في أي وقت.

ماذا سيطلب من الطفل؟

سيقوم الباحث بزيارة الطفل في العيادة\المدرسة لمدة 15 دقيقة تقريبا. في البداية , سيستمع الطفل إلى مجموعة من الحروف و سيطلب منه نطقها و سيتم تسجيل صوته. يقوم الطفل أيضا ببعض التدريبات المصممة بطريقة ممتعة و الهدف منها قياس أثر التدريب على الأنماط الجديدة. مثلا, يستمع الطفل إلى قصص مصورة وبعدها يجيب على أسئلة عليها . أيضا , تعرض للطفل مجموعة من الصور على جهاز الكمبيوتر و يطلب منه تسمية الصور و خلالها يتم تسجيل سرعة الاستجابة. تحفظ جميع التسجيلات الصوتية عن طريق أرقام بحيث لا يمكن التعرف على هوية المشارك كما لا يتم جمع أي معلومات شخصية.

مالفائدة من المشاركة في الدراسة ؟

نهدف من هذا البحث إلى تطوير الممارسات المتعلقة بالطرق المستخدمة من قبل المدارس وأخصائي التخاطب لتلبية احتياجات الأطفال الذين يواجهون صعوبات في إخراج الكلام بطلاقة.

هل توجد أي مخاطر من مشاركة الطفل في الدراسة ؟

لا لا توجد أي مخاطر و إذا لم يرغب الطفل في المشاركة ، فبالإمكان التوقف في أي وقت .

هل ستتم المحافظة على سرية المعلومات ؟

نعم، كافة المعلومات التي سيتم جمعها ستحفظ عن طريق الأرقام فقط . لن نحتفظ بأي معلومات عن اسم الطفل أو أي معلومات شخصية وسيتم التعامل مع المعلومات وفق لائحة أنظمة حماية البيانات للعام 2018

اين ستستخدم نتائج الدراسة ؟

ستستخدم النتائج كجزء من رسالة الدكتوراه . أيضا، قد تنشر بعض الأوراق العلمية من هذه الدراسة في مجلات علمية و مؤتمرات يحضرها بعض صانعي القرار المهتمين بتعليم الأطفال. لن يتم استخدام أي معلومات شخصية عن الأطفال في أي من هذه الجهات. بإمكانك التواصل معنا أيضا إذا رغبت في الاطلاع على نتائج الدراسة.

-

الموافقة

اطلعت على المعلومات أعلاه ولا مانع لدي من مشاركة الطفل

اسم الطفل

تاريخ الميلاد

توقيع ولي الأمر

Figure A6.2: Arabic Information Sheet for Parents/Guardians of Children